HANDBOOK OF FINANCE

VOLUME I

Financial Markets and Instruments

Frank J. Fabozzi
Editor

WILEY

John Wiley & Sons, Inc.

HANDBOOK OF FINANCE

VOLUME I

HANDBOOK OF FINANCE

VOLUME I

Financial Markets and Instruments

Frank J. Fabozzi
Editor

WILEY
John Wiley & Sons, Inc.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Handbook of finance / Frank J. Fabozzi, editor.
p. cm.
Includes index.
ISBN 978-0-470-04256-4 (3 v. set : cloth) – ISBN 978-0-470-07814-3 (v. 1) –
ISBN 978-0-470-07815-0 (v. 2) – ISBN 978-0-470-07816-7 (v. 3)
1. Finance–Handbooks, manuals, etc. I. Fabozzi, Frank J.
HG173.H327 2008
332–dc22 2008020129

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

About the Editor

Frank J. Fabozzi is Professor in the Practice of Finance and Becton Fellow in the Yale School of Management. Prior to joining the Yale faculty, he was a Visiting Professor of Finance in the Sloan School of Management at Massachusetts Institute of Technology. Professor Fabozzi is a Fellow of the International Center for Finance at Yale University and on the Advisory Council for the Department of Operations Research and Financial Engineering at Princeton University. He is an affiliated professor at the Institute of Statistics, Econometrics and Mathematical Finance at the University of Karlsruhe (Germany). He is the editor of the *Journal of Portfolio Management* and an associate editor of the *Journal of Fixed Income, Journal of Asset Management, Journal of Structured Finance,* and the *Review of Futures Markets*. He earned a doctorate in economics from the City University of New York in 1972. In 2002, Professor Fabozzi was inducted into the Fixed Income Analysts Society's Hall of Fame and is the 2007 recipient of the C. Stewart Sheppard Award given by the CFA Institute. He earned the designations of Chartered Financial Analyst and Certified Public Accountant. He has authored and edited numerous books on various topics in finance.

Contents

Contributors

Mark Adelson
Principal, Adelson & Jacob Consulting, LLC

Carol Alexander, PhD
Chair of Risk Management and Director of Research, ICMA Centre, Business School, The University of Reading

Roever W. Alexander, CFA
Managing Director, U.S. Fixed Income Strategy, JPMorgan Securities, Inc.

Andrew Alford, PhD
Managing Director, Quantitative Investment Strategies, Goldman Sachs Asset Management

Michele Allman-Ward
Managing Partner, Allman-Ward Associates

Noël Amenc, PhD
Professor of Finance, Edhec Graduate School of Business, Director, Edhec Risk and Asset Management Research Centre

Mark J. P. Anson, PhD, JD, CPA, CFA, CAIA
President and Executive Director of Nuveen Investment Services

Kodjo Apedjinou
Vice President, Lehman Brothers Inc.

Bala Arshanapalli, PhD
Professor of Finance, Indiana University Northwest

Enrique R. Arzac, PhD
Professor of Finance and Economics, Graduate School of Business, Columbia University

Erik Banks
Managing Director, Risk Advisory, Unicredit Group Europe

Chris Barr, CFA
Principal, Barclays Global Investors

Bülent Baygün, PhD
Head of Global Quantitative Strategy, Barclays Capital

Diana Berezina
Associate Director, Fitch Ratings

William S. Berliner
Executive Vice President, Countrywide Securities Corporation

Dan Bernstein
Director, Research, Bridgewater Associates

Peter L. Bernstein
President, Peter L. Bernstein Inc.

Anand K. Bhattacharya, PhD
Managing Director, Countrywide Securities Corporation

Harold Bierman Jr., PhD
The Nicholas H. Noyes Professor of Business Administration, Johnson Graduate School of Management, Cornell University

Jeremy Blatt
Managing Director, Deal Structuring and Execution, Finacity Corporation

Len Blum
Managing Partner, Westwood Capital, LLC

Ludovic Breger, PhD
Principal, Barclays Global Investors

Vineet Budhraja
Senior Research Analyst, Citi Alternative Investments

Gerald W. Buetow, Jr., PhD, CFA
President and Founder, BFRC Services, LLC

Rachel A. J. Campbell, PhD
Assistant Professor of Finance, Rotterdam School of Management, Erasmus University & Maastricht University, The Netherlands

Findlay M. Chapman, III
Principal, Findlay, Phillips and Associates

Ren-Raw Chen, PhD
Associate Professor of Finance, Rutgers University

Daniel L. Chesler, CMT, CTA
President, Chesler Analytics

Oren Cheyette, PhD
Principal, Barclays Global Investors

Anne Ching
Senior Consultant, Andrew Davidson & Co., Inc.

Moorad Choudhry, PhD
Head of Treasury, KBC Financial Products, London

Roger G. Clarke, PhD
Chairman, Analytic Investors, Inc.

Jeff Cohen
Securities Lending Manager, Susquehanna Intl Group, LLLP

Bruce Collins, PhD
Professor of Finance, Western Connecticut State University

Graham "Harry" Cross
Financial Tutor, 7City Learning

Christopher L. Culp
Senior Advisor, Lexecon Senior Fellow in Financial Regulation, Competitive Enterprise Institute Adjunct Professor of Finance, Graduate School of Business, University of Chicago

Alexandre Schutel Da Silva
Vice President, Quantitative Investments Group, Lehman Brothers Asset Management

Ray Dalio
President and Chief Investment Officer, Bridgewater Associates

Maurizio Dallocchio
Lehman Brothers Professor of Corporate Finance, Bocconi University

Aswath Damodaran, PhD
Professor of Finance and David Margolis Teaching Fellow, Stern School of Business, New York University

Andrew Davidson
President and Founder, Andrew Davidson & Co., Inc.

Henry A. Davis
Editor, *Journal of Structured Finance*

Rui J. P. de Figueiredo, Jr., PhD
Associate Professor, Haas School of Business, University of California at Berkeley

Harindra De Silva, PhD, CFA
Managing Director, Analytic Investors, Inc.

Chris P. Dialynas
Managing Director–Portfolio Management, Pacific Investment Management Company

Chris DiAngelo
Partner, Dewey Ballantine

Michael Dorigan, PhD
Senior Quantitative Analyst, PNC Capital Advisors

Kevin Dowd, PhD
Professor of Financial Risk Management, Centre for Risk and Insurance Studies, Nottingham University Business School

Pamela P. Drake, PhD, CFA
J. Gray Ferguson Professor of Finance and Department Head of Finance and Business Law, James Madison University

Robert Dubil, PhD
Associate Professor, Lecturer of Finance, David Eccles School of Business, University of Utah

Steven I. Dym, PhD
President, Mariner Capital Partners

Brian Eales, BA, MSc (Econ)
Academic Leader, London Metropolitan University

Mark Edwards
Vice President, ITG Solutions Network, Inc.

Mahmoud A. El-Gamal, PhD
Professor of Economics and Statistics, and Chair of Islamic Economics, Finance, and Management at Rice University

Robert F. Engle, PhD
Michael Armellino Professorship in the Management of Financial Services, Leonard N. Stern School of Business, New York University

Frank J. Fabozzi, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Dorsey D. Farr, PhD, CFA
Principal, French Wolf & Farr

Mark C. Faulkner
Managing Director, Spitalfields Advisors

Bruce J. Feibel, CFA
Global Director of Products, Eagle Investment Systems

Stanley Jay Feldman, PhD
Chairman, Axiom Valuation Solutions and Associate Professor of Finance, Bentley University

Sylvan G. Feldstein, PhD
Director, Investment Department, Guardian Life Insurance Company of America

Niall Ferguson
Manager of Client Service Analytics, Bridgewater Associates

John D. Finnerty, PhD
Professor of Finance and Director of the MS in Quantitative Finance Program, Fordham University Graduate School of Business Managing Principal, Finnerty Economic Consulting, LLC

Jeffrey D. Fisher, PhD
Dunn Professor of Real Estate, Indiana University

Sergio M. Focardi
Partner, Intertek Group

Reto Francioni, PhD
Chief Executive Officer, Deutsche Börse

Roland Füss, PhD
Professor of Finance, Union Investment Endowed Chair of Asset Management, European Business School (EBS), International University-Schloss Reichartshausen

Daniel E. Gallegos
Principal, Barclays Global Investors

Sunita Ganapati

James P. Garland, CFA
President, The Jeffrey Company

Gary L. Gastineau
Managing Director, ETF Consultants, LLC

David Geltner, PhD
George Macomber Professor of Real Estate Finance, MIT

Juby George
Valuation Risk Group, Credit Suisse, London

R. Philip Giles, PhD
Adjunct Professor of Finance and Economics, Columbia University Graduate School of Business and President, CBT Worldwide, Inc.

Jean-René Giraud
Director of Business Development, Edhec Risk and Asset Management Research Centre

Felix Goltz
Senior Research Engineer, Edhec Risk and Asset Management Research Centre

Laurie S. Goodman, PhD
Co-head of Global Fixed Income Research Manager of U.S. Securitized Products Research, UBS

James L. Grant, PhD
JLG Research and Professor of Finance, University of Massachusetts (Boston)

Les Gulko, PhD
Principal, Cove Island Ventures

Francis Gupta, PhD
Director, Research, Dow Jones Indexes

Brian K. Haendiges, FSA, CRC, CRA
Head, Institutional Defined Contribution Plans, ING Retirement Services

G. Timothy Haight, DBA
President, Menlo College and Chair of the Board, Board of Commonwealth Business Bank (Los Angeles)

Brett P. Hammond, PhD
Managing Director and Chief Investment Strategist, TIAA-CREF Asset Management

David Haushalter, PhD
Corporate Research and Educational Associate, Susquehanna Intl Group, LLLP

Markus Hoechstoetter, Dr. rer. pol.
Lecturer, School of Economics and Business Engineering, University of Karlsruhe

Susan Hudson-Wilson, CFA
Member, Boards of Hawkeye Partners, LLC, Property & Portfolio Research, Inc. and University of Vermont Endowment

Claus Huber, CFA, FRM, PhD
Chief Risk Officer, Credaris Portolio Management

Bruce I. Jacobs, PhD
Principal, Jacobs Levy Equity Management

Brian J. Jacobsen, PhD, CFA, CFP
Associate Professor of Business Administration, Wisconsin Lutheran College Chief Economist Partner, Capital Market Consultants, LLC

Henry G. Jarecki, MD
Chairman, The Falconwood Corporation

Teo Jasic, Dr. rer. pol.
Postdoctoral Research Fellow at the Chair of Statistics, Econometrics and Mathematical Finance at the University of Karlsruhe in the School of Economics and Business Engineering and a Partner of an International Management Consultancy Firm in Frankfurt, Germany

Robert Johnson, PhD, CFA
Deputy Chief Executive Officer, CFA Institute

David M. Jones, PhD
President, DMJ Advisors, LLC

Frank J. Jones, PhD
Professor, Accounting and Finance Department, San Jose State University

Robert Jones, CFA
Managing Director, and Chief Investment Officer for Quantitative Equity, Quantitative Investment Strategies, Goldman Sachs Asset Management

Steven L. Jones, PhD
Associate Professor of Finance, Indiana University, Kelley School of Business–Indianapolis

Dieter G. Kaiser, PhD
Director Alternative Investments, Institutional Advisors GmbH Research Fellow, Centre for Practical Quantitative Finance, Frankfurt School of Finance Management

Andrew Kalotay, PhD
President, Andrew Kalotay Associates

Ozgur Kan, PhD, CFA, FRM
Product Specialist, Moody's Investors Service

Adrian Katz
CEO, Finacity Corporation

George P. Kegler
President, Cassian Consultants

Jeroen Kerkhof, PhD
Vice President, Morgan Stanley

Janghoon Kim, CFA
Research Analyst, Citi Alternative Investments

Christoph Klein, CFA
Director, Portfolio Management, Deutsche Asset Management

Stanley Kogelman
Chief Investment Officer, Summer Hill Inc. & Partner, Advanced Portfolio Management

Petter N. Kolm, PhD
Clinical Associate Professor and Deputy Director of the Mathematics in Finance M.S. Program, Courant Institute, New York University

Vinod Kothari
Independent financial consultant and trainer on securitization, visiting Faculty, Indian Institute of Management, Kolkata, India

Steven K. Kreider, PhD
Managing Director, Morgan Stanley Investment Management

Mark P. Kritzman, CFA
President and CEO, Windham Capital Management, LLC

James Lam
President, James Lam & Associates Senior Research Fellow, Beijing University

Glen A. Larsen, Jr., PhD, CFA
Professor of Finance, Indiana University, Kelley School of Business–Indianapolis

Yann Le Fur
Professor of Corporate Finance, HEC Paris

Véronique Le Sourd
Senior Research Engineer, Edhec Risk and Asset Management Research Centre

Wai Lee, PhD
Managing Director, Head of Quantitative Investments Group, Lehman Brothers Asset Management

Yann LeFur
Professor of Corporate Finance, HEC, Paris

Martin L. Leibowitz, PhD
Managing Director, Morgan Stanley & Co. Inc.

Alexander Levin, PhD
Director of Valuation Modeling, Andrew Davidson and Co., Inc.

Kenneth N. Levy, CFA
Principal, Jacobs Levy Equity Management

Thomas K. Liaw, PhD
Professor of Finance and Chair, St. John's University

Terrence Lim, PhD, CFA
Managing Director, Quantitative Investment Strategies, Goldman Sachs Asset Management

Maria Mednikov Loucks, CFA
Senior Managing Director, Black River Asset Management

Douglas J. Lucas
Executive Director and Head of CDO Research, UBS

Steven V. Mann, PhD
Professor of Finance, Moore School of Business, University of South Carolina

Rebecca J. Manning
Vice President, Harbor Asset Management

James Manzi, CFA
Consultant

Harry M. Markowitz, PhD
Consultant

Terrence F. Martell, PhD
Saxe Distinguished Professor of Finance, Zicklin School of Business, Baruch College/CUNY

Lionel Martellini, PhD
Professor of Finance, Edhec Graduate School of Business, Scientific Director, Edhec Risk and Asset Management Research Centre

Greg M. McMurran
Chief Investment Officer, Analytic Investors, Inc.

Christian Menn, Dr. rer. pol.
Associate, Sal. Oppenheim Jr. & Cie, Frankfurt, Germany

Ryan Meredith, CFA, FFA
Senior Research Analyst, Citi Alternative Investments

William L. Messmore
Vice President, Lehman Brothers

Steven Miller
Managing Director, Standard & Poor's LCD

Banikanta Mishra, PhD
Visiting Professor of Finance, University of Michigan and Professor of Finance XIM-Bhubaneswar, India

Priya Misra
Senior Vice President, Lehman Brothers Inc.

Stefan Mittnik, PhD
Professor of Financial Econometrics at the University of Munich, Germany, and Research Director at the Ifo Institute for Economic Research in Munich

William T. Moore, PhD
David & Esther Berlinberg Professor, Vice Provost for Academic Affairs, University of South Carolina

William Nelson, PhD
Professor of Finance, Indiana University Northwest

Thomas J. O'Brien, PhD
Professor, University of Connecticut

Dominic O'Kane, PhD
Affiliated Professor of Finance, EDHEC Business School, Nice, France

Dessislava A. Pachamanova, PhD
Assistant Professor of Operations Research, Babson College

Anthony F. L. Pecore
Research Analyst, Franklin Templeton Investments

John A. Penicook, CFA
Managing Director, UBS Global Asset Management

Bobby Pornrojnangkool, PhD
Vice President, Quantitative Investments Group, Lehman Brothers Asset Management

Anshul Pradhan
Associate, Lehman Brothers Inc.

Kuntara Pukthuanthong-Le, PhD
Assistant Professor, San Diego State University

Paul Puleo

Pascal Quiry
Professor of Corporate Finance, HEC Paris

Svetlozar T. Rachev, PhD, DrSci
Chair-Professor, Chair of Econometrics, Statistics and Mathematical Finance, School of Economics and Business Engineering, University of Karlsruhe and Department of Statistics and Applied Probability, University of California, Santa Barbara

Ellen J. Rachlin
Managing Director–Portfolio & Risk Manager, Mariner Investment Group, Inc.

Mahmud Rahman, PhD
Professor of Finance, Eastern Michigan University

Shrikant Ramamurthy
Managing Director, RBS Greenwich Capital

Srichander Ramaswamy, PhD
Senior Economist, Bank for International Settlements, Basel, Switzerland

Donald M. Raymond, PhD, CFA
Senior Vice President, Public Market Investments, Canada Pension Plan Investment Board

Adam V. Reed, PhD
Assistant Professor of Finance, University of North Carolina at Chapel Hill

Robert R. Reitano, PhD, FSA
Professor of the Practice in Finance, Brandeis University, International Business School

Mark Retik

Victor Ricciardi
Assistant Professor of Finance, Kentucky State University, and Editor, Social Science Research Network Behavioral & Experimental Finance, eJournal

Scott F. Richard, PhD
Managing Director, Morgan Stanley Investment Management

W. Alexander Roever, CFA
Managing Director, U.S. Fixed Income Strategy, JPMorgan Securities, Inc.

Paul Ross
Investment Associate, Bridgewater Associates

James Sagner
Managing Principal, Sagner/Marks and Associate Professor in the School of Business of Metropolitan College of New York

Antonio Salvi, PhD
Professor of Corporate Finance, EM Lyon and Bocconi University

Paul A. Samuelson, PhD
Institute Professor Emeritus, Massachusetts Institute of Technology

Anthony Sanders, PhD
John W. Galbraith Chair and Professor of Finance, Ohio State University

Bernd Scherer, PhD
Managing Director and Global Head of Quantitative Structured Products, Morgan Stanley Investment Management

Uwe Schillhorn, CFA
Executive Director, UBS Global Asset Management

Robert A. Schwartz, PhD
Marvin M. Speiser Professor of Finance and University Distinguished Professor, Zicklin School of Business, Baruch College, CUNY

Shani Shamah
Consultant, E J Consultants

William F. Sharpe, PhD
STANCO 25 Professor of Finance, Emeritus at Stanford University's Graduate School of Business

Koray D. Simsek, PhD
Assistant Professor, Faculty of Management, Sabanci University

Daniel D. Singer, PhD
Professor of Finance, Towson University

Eric H. Sorensen, PhD
President and CEO, PanAgora Asset Management

Russell Sparkes
Chief Investment Officer, Central Finance Board of the Methodist Church (UK)

Larry Speidell, CFA
General Partner, Ondine Asset Management, LLC

Beth Starr
Managing Director, Lehman Brothers

State Street Corporation

Meir Statman, PhD
Glenn Klimek Professor of Finance, Santa Clara University

Filippo Stefanini
Deputy Chief Investment Officer, Aletti Gestielle Alternative SGR Professor of Risk Management, Faculty of Engineering at Bergamo University in Italy

David M. Stein, PhD
Managing Director, Parametric Portfolio Associates

Stoyan V. Stoyanov, PhD
Chief Financial Researcher, FinAnalytica Inc.

Lee R. Thomas III, PhD
Managing Partner and CEO, Flint Rock Capital Management

James R. Thompson, PhD
Noah Harding Professor of Statistics, Rice University

Radu Tunaru, PhD
Senior Lecturer in Financial Mathematics, City University, CASS Business School

Robert Tzucker
Inflation Trading, Barclays Capital

Raman Vardharaj, CFA
Senior Quantitative Analyst, RS Investments

Antonio Villarroya, MD
Global Head of Rates Derivatives Strategy, Merrill Lynch

Wayne H. Wagner
Senior Advisor, ITG Inc., Principal OM/NI, A Trading Consultancy

Karen Weaver, CFA
Managing Director, Global Head of Securitization Research and Regional Research Head—the Americas, Deutsche Bank Securities, Inc.

Bruce W. Weber, PhD
Professor of Information Management, London Business School

Robert Whaley, PhD
Valere Blair Potter Professor of Management, The Owen Graduate School of Management, Vanderbilt University

Shane Whelan, FSAI, FSA, FFA, PhD
Lecturer in Actuarial Science, School of Mathematical Sciences, University College Dublin, Ireland

Jarrod W. Wilcox, PhD, CFA
President, Wilcox Investment Inc.

Edward E. Williams, PhD
Henry Gardiner Symonds Professor of Management, Rice University

Lan-Ling Wolff
Fredell & Co. Structured Finance Ltd.

Eugene Xu, PhD
Director, Global Securitization Research, Deutsche Bank Securities, Inc.

Preface

Over the past two decades, financial professionals have had available to them excellent reference books on specialty areas in finance. There are handbooks on corporate financial management, financial instruments, portfolio strategies, structured finance, capital budgeting, derivatives, and the list goes on. But to truly understand financial markets throughout the world, it is necessary to understand how financial decision makers—such as corporate treasurers, chief financial officers, portfolio managers, traders, and security analysts—make decisions and the tools that they employ in doing so. From that perspective, the idea for this handbook was conceived.

Finance is the application of economic principles and concepts to business decision making and problem solving. The field of finance can be considered to comprise three broad categories: financial markets and instruments, financial management, and investment management.

The field of *financial markets and instruments* deals with the role of financial markets in an economy, the structure and organization of financial markets, the efficiency of markets, the role of the various players in financial markets (i.e., governments, regulators, financial institutions, investment banks and securities firms, and institutional and retail investors), and the determinants of asset pricing and interest rates.

Financial management, sometimes referred to as *business finance,* is the specialized field in finance that is concerned primarily with financial decision-making within a business entity and encompasses many different types of decisions. (While financial management is sometimes referred to as *corporate finance,* the principles are applied to the management of municipalities and nonprofit profit entities.) We can classify financial management decisions into two groups: investment decisions and financing decisions. *Investment decisions* are concerned with the use of funds—the buying, holding, or selling of all types of assets. Basically, the types of assets acquired are either working capital, such as inventory and receivables, or long-term assets. Decisions involving the former are called *working capital decisions* and those involving the latter are called *capital budgeting decisions.* Financing decisions are concerned with the acquisition of funds to be used for investing and financing day-to-day operations. Basically, this involves the selection of the firm's capital structure—that is, the combination of equity and debt used to finance the firm—and is referred to as the *capital structure decision.* The financing decision also involves the determination of how much of the company's earnings to retain and how much to distribute to shareholders in the form of dividends. This decision is referred to as the *dividend decision.* Whether a financial decision involves investing or financing, the core of the decision will rest on two specific factors: expected return and risk. *Expected return* is the difference between potential benefits and potential costs. *Risk* is the degree of uncertainty associated with the expected returns.

Investment management is the area of finance that focuses on the management of portfolios of assets for institutional investors and individuals. The activities involved in investment management, also referred to as *asset management,* include working with clients to set investment objectives and an investment policy to accomplish those objectives, the selection a portfolio strategy consistent with the investment objectives and investment policy, and the construction of the specific assets to include in a portfolio based on the portfolio strategy. Investment management begins with the decision as to how to allocate funds across the major asset classes (e.g., stocks, bonds, real estate, alternative investments). This decision, referred to as the *asset allocation decision,* requires a thorough understanding of the expected returns and risks associated with investing in a specific asset class. Again, we see the importance of understanding expected return and risk. The investment strategy employed can be classified as either active or passive and the decision as to which type to follow depends on the client's view of the efficiency (i.e., the difficulty of obtaining superior returns) of the market for the asset class. The portfolio construction phase involves assembling the best portfolio given the client's investment objectives, given the investment constraints set forth in the investment policy, and the estimated expected return and risk of the individual assets that are potential candidates for inclusion in the portfolio.

These three general areas use theories and analytical tools developed in other disciplines. For example, theories about the pricing of assets and the determination of interest rates draw from theories in economics. In fact, many academics refer to finance as *financial economics.* There are investment management strategies that utilize theories and concepts that draw from the field of psychology, giving rise to the specialized field in finance known as *behavioral finance.* The complex nature of financial markets requires a finance professional to draw from the fields of statistics and econometrics in order to describe the movement in asset prices and returns, as well as to obtain meaningful measures of risk. The field of *financial risk management,* used both in financial management and investment management, employ these tools. These same tools are used by investment managers in formulating and testing potential strategies and in the valuation (pricing) of complex financial instruments known as derivatives. Investment managers and financial managers utilize sophisticated mathematical models developed in the area of operations research/management science to aid in making optimal allocation decisions such as in portfolio construction and the selection of capital projects. Managers also use simulation models, a tool of operations research, in a variety of activities that involve corporate and investment decisions. *Financial engineering,* sometimes referred to as *mathematical finance,* is the relatively new specialized field in finance that uses statistical and mathematical tools to deal with problems in all areas of finance and risk management.

This multivolume reference provides a bird's-eye view of finance that will help the reader appreciate the wide range of topics that the discipline of "finance" encompasses. While there are handbooks that address specialized areas within finance, the purpose of this three-volume handbook is to cover all of the areas mentioned above and is intended for professionals involved in finance, as well as the student of finance.

This three-volume handbook offers coverage of both established and cutting-edge theories and developments in finance. It contains chapters from global experts in industry and academia, and offers the following unique features:

- The handbook was written by more than 190 experts from around the world. This diverse collection of expertise has created the most definitive coverage of established and cutting-edge financial theories, applications, and tools in this ever-evolving field.
- The series emphasizes both technical and managerial issues. This approach provides researchers, educators, students, and practitioners with a balanced understanding of the topics and the necessary background to deal with issues related to finance.
- Each chapter follows a format that includes the author, chapter abstract, keywords, introduction, body, summary, and references. This enables readers to pick and choose among various sections of a chapter and creates consistency throughout the entire handbook.
- Each chapter provides extensive references for additional readings, enabling readers to further enrich their understanding of a given topic.
- Numerous illustrations and tables throughout the work highlight complex topics and assist further understanding.
- Each chapter provides cross-references within the body of the chapter. This helps readers identify other chapters within the handbook related to a particular topic, which provides a one-stop knowledge base for a given topic.
- Each volume includes a complete table of contents and index for easy access to various parts of the handbook.

TOPIC CATEGORIES

The allocation of the topics among the three volumes of the handbook required a good deal of time, with more than two dozen restructurings of the table of contents for each volume before reaching what I believe to be the most useful allocation for readers. There was no simple formula. The decision involved feedback from practitioners, academics, and graduate students. The final allocation to the three volumes was as follows.

Volume I (*Financial Markets and Instruments*) covers the general characteristics of the different asset classes, derivative instruments, the markets in which financial instrument trade, and the players in the market. Topics include:

- Market Players and Markets
- Common Stock
- Fixed Income Instruments
- Real Estate
- Alternative Investments
- Investment Companies, Exchange-Traded Funds, and Life Insurance Products
- Foreign Exchange
- Inflation-Hedging Products
- Securities Finance

Volume II (*Investment Management and Financial Management*) covers the theories, issues, decisions, and implementation for both investment management and financial management. Topics include:

- Investment Management
- Equity Portfolio Management
- Fixed Income Portfolio Management
- Alternative Investments
- Corporate Finance

The analytical tools, the measurement of risk, and the techniques for valuation are the subject of Volume III (*Valuation, Financial Modeling, and Quantitative Tools*). Topics include:

- Risk Management
- Interest Rate Modeling
- Credit Risk Modeling and Analysis
- Valuation
- Mathematical Tools and Techniques for Financial Modeling and Analysis

The chapters can serve as material for a wide spectrum of courses, such as the following:

- Financial markets
- Principles of finance
- Investment and portfolio management
- Corporate finance
- Derivative instruments and their applications
- Financial mathematics
- Financial engineering

Frank J. Fabozzi
Editor, *Handbook of Finance*

Guide to the *Handbook of Finance*

The *Handbook of Finance* is a comprehensive overview of the field of finance. This reference work consists of three separate volumes and 229 chapters. Each chapter provides a comprehensive overview of the selected topic intended to inform a broad spectrum of readers ranging from finance professionals to academicians to students to the general business community.

To derive the greatest possible benefit from the *Handbook of Finance*, we have provided this guide. It explains how the information within the handbook can be located.

ORGANIZATION

The *Handbook of Finance* is organized to provide maximum ease of use for its readers. The material is broken down into three distinct volumes:

- Volume I (*Financial Markets and Instruments*) covers the general characteristics of the different asset classes, derivative instruments, the markets in which financial instrument trade, and the players in the market.
- Volume II (*Investment Management and Financial Management*) covers the theories, issues, decisions, and implementation for both investment management and financial management.
- Volume III (*Valuation, Financial Modeling, and Quantitative Tools*) tackles the analytical tools, the measurement of risk, and the techniques for valuation.

TABLE OF CONTENTS

A complete table of contents for the entire handbook appears in the front of each volume. This list of titles represents topics that have been carefully selected by the editor, Frank J. Fabozzi. The Preface includes a more detailed description of the volumes and parts the chapters are grouped under.

INDEX

A Subject Index for the entire handbook is located at the end of each volume. The subjects in the index are listed alphabetically and indicate the volume and page number where information on this topic can be found.

CHAPTERS

Each chapter in the *Handbook of Finance* begins on a new page, so that the reader may quickly locate it. The author's name and affiliation are displayed at the beginning of the chapter.

All chapters in the handbook are organized according to a standard format, as follows:

- Title and author
- Outline
- Abstract
- Keywords
- Introduction
- Body
- Summary
- References

Outline

Each chapter begins with an outline indicating the content to come. The outline is intended as an overview and thus lists only the major headings of the chapter. Lower-level headings also may be found within the chapter.

Abstract

The abstract for each chapter gives an overview of the topic, but not necessarily the content of the chapter. This is designed to put the topic in the context of the entire handbook, rather than give an overview of the specific chapter content.

Keywords

The keywords section contains terms that are important to an understanding of the chapter.

Introduction

The text of each chapter begins with an introductory section that defines the topic under discussion and summarizes the content. By reading this section, the reader gets a general idea about the content of a specific chapter.

Body
The body of each chapter discusses the items that were listed in the outline section.

Summary
The summary section provides a review of the materials discussed in each chapter. It imparts to the reader the most important issues and concepts discussed.

References
The references section lists both publications cited in the chapter and secondary sources to aid the reader in locating more detailed or technical information. Review articles and research papers that are important to an understanding of the topic are also listed. The references provide direction for further research on the given topic.

HANDBOOK OF FINANCE

VOLUME I

PART 1

Market Players and Markets

CHAPTER 1

Overview of Financial Instruments and Financial Markets

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: Broadly speaking, an asset is any possession that has value in an exchange. Assets can be classified as tangible or intangible. A tangible asset is one whose value depends on particular physical properties—examples are buildings, land, and machinery. Assets, by contrast, represent legal claims to some future benefit. Their value bears no relation to the form, physical or otherwise, in which these claims are recorded. Financial assets, also referred to as financial instruments, are intangible assets. For financial assets, the typical benefit or value is a claim to future cash. Financial markets are classified as cash/spot markets and derivatives markets. Financial markets play a key role in the financial system of all economies. In most economies financial instruments are created and subsequently traded in some type of financial market.

Keywords: financial assets, financial instruments, issuer, investor, debt instrument, equity instrument, fixed income instruments, maturity, coupon rate, floating-rate securities, amortizing instrument, call provision, put provision, prepayment, search costs, liquidity, price discovery process, capital market, secondary market, primary market, over-the-counter market, derivatives markets, derivative instruments, futures contract, option contract

Participants in financial markets must understand the wide range of financial instruments and the role of financial markets. In this chapter, an overview of the instruments (both cash and derivative instruments), issuers, and investors is provided. The role of financial assets and financial markets are also explained.

ISSUERS AND INVESTORS

The entity that has agreed to make future cash payments is called the *issuer* of the financial instrument; the owner of the financial instrument is referred to as the *investor*. Here are seven examples of financial instruments:

1. A loan by Bank of America (investor/commercial bank) to an individual (issuer/borrower) to purchase a car.
2. A bond issued by the U.S. Department of the Treasury.
3. A bond issued by Nike Inc.
4. A bond issued by the city of San Francisco.
5. A bond issued by the government of Australia.
6. A share of common stock issued by Caterpillar, Inc., an American company.
7. A share of common stock issued by Toyota Motor Corporation, a Japanese company.

In the case of the car loan by Bank of America, the terms of the loan establish that the borrower must make specified payments to the commercial bank over time. The payments include repayment of the amount borrowed plus interest. The cash flow for this asset is made up of the specified payments that the borrower must make.

In the case of a U.S. Treasury bond, the U.S. government (the issuer) agrees to pay the holder or the investor the interest payments every six months until the bond matures, then at the maturity date repay the amount borrowed. The same is true for the bonds issued by Nike Inc., the city of San Francisco, and the government of Australia. In the case of Nike, Inc. the issuer is a corporation, not a government entity. In the case of the city of San Francisco, the issuer is a municipal government. The issuer of the Australian government bond is a central government.

The common stock of Caterpillar, Inc. entitles the investor to receive dividends distributed by the company. The investor in this case also has a claim to a pro rata share of the net asset value of the company in case of liquidation of the company. The same is true of the common stock of Toyota Motor Corporation.

DEBT VERSUS EQUITY INSTRUMENTS

Financial instruments can be classified by the type of claim that the holder has on the issuer. When the contractual arrangement is one in which the issuer agrees to pay interest and repay the amount borrowed, the financial instrument is said to be a *debt instrument*. The car loan, the U.S. Treasury bond, the Nike Inc. bond, the city of San Francisco bond, and the Australian government bond are examples of debt instruments requiring fixed payments.

In contrast to a debt obligation, an *equity instrument* obligates the issuer of the financial instrument to pay the holder an amount based on earnings, if any, after the holders of debt instruments have been paid. Common stock is an example of an equity claim. A partnership share in a business is another example.

Some securities fall into both categories in terms of their attributes. Preferred stock, for example, is an equity instrument that entitles the investor to receive a fixed amount. This payment is contingent, however, and due only after payments to debt instrument holders are made. Another "combination" instrument is a convertible bond, which allows the investor to convert debt into equity under certain circumstances. Both debt instruments and preferred stock are called *fixed-income instruments*.

CHARACTERISTICS OF DEBT INSTRUMENTS

There are a good number of debt instruments available to investors. Debt instruments include loans, money market instruments, bonds, mortgage-backed securities, and asset-backed securities. In the chapters that follow, each will be described. There are features of debt instruments that are common to all debt instruments and they are described below. In later chapters, there will be a further discussion of these features as they pertain to debt instruments of particular issuers.

Maturity

The term to maturity of a debt obligation is the number of years over which the issuer has promised to meet the conditions of the obligation. At the maturity date, the issuer will pay off any amount of the debt obligation outstanding. The convention is to refer to the "term to maturity" as simply its "maturity" or "term." As we explain later, there may be provisions that allow either the issuer or holder of the debt instrument to alter the term to maturity.

The market for debt instruments is classified in terms of the time remaining to its maturity. A money market instrument is a debt instrument which has one year or less remaining to maturity. Debt instruments with a maturity greater than one year are referred to as a capital market debt instrument.

Par Value

The par value of a bond is the amount that the issuer agrees to repay the holder of the debt instrument by the maturity date. This amount is also referred to as the principal, face value, or maturity value. Bonds can have any par value.

Because debt instruments can have a different par value, the practice is to quote the price of a debt instrument as a percentage of its par value. A value of 100 means 100% of par value. So, for example, if a debt instrument has a par value of $1,000 and is selling for $900, it would be said to be selling at 90. If a debt instrument with a par value of $5,000 is selling for $5,500, it is said to be selling for 110.

Coupon Rate

The *coupon rate*, also called the nominal rate or the contract rate, is the interest rate that the issuer/borrower agrees to pay each year. The dollar amount of the payment, referred to as the coupon interest payment or simply *interest payment*, is determined by multiplying the coupon rate by the par value of the debt instrument. For example, the interest payment for a debt instrument with a 7% coupon rate and a par value of $1,000 is $70 (7% times $1,000).

The frequency of interest payments varies by the type of debt instrument. In the United States, the usual practice for bonds is for the issuer to pay the coupon interest in two semiannual installments. Mortgage-backed securities and asset-backed securities typically pay interest monthly. For bonds issued in some markets outside the United States, coupon payments are made only once per year. Loan interest payments can be customized in any manner.

Zero-Coupon Bonds

Not all debt obligations make periodic coupon interest payments. Debt instruments that are not contracted to make periodic coupon payments are called zero-coupon instruments. The holder of a zero-coupon instrument realizes interest income by buying it substantially below its par value. Interest then is paid at the maturity date, with the interest earned by the investor being the difference between the par value and the price paid for the debt instrument. So, for example, if an investor purchases a zero-coupon instrument for 70, the interest realized at the maturity date is 30. This is the difference between the par value (100) and the price paid (70).

There are bonds that are issued as zero-coupon instruments. Moreover, in the money market there are several types of debt instruments that are issued as discount instruments.

There is another type of debt obligation that does not pay interest until the maturity date. This type has contractual coupon payments, but those payments are accrued and distributed along with the maturity value at the maturity date. These instruments are called accrued coupon instruments or accrual securities or compound interest securities.

Floating-Rate Securities

The coupon rate on a debt instrument need not be fixed over its life. *Floating-rate securities*, sometimes called floaters or variable-rate securities, have coupon payments that reset periodically according to some reference rate. The typical formula for the coupon rate on the dates when the coupon rate is reset is:

$$\text{Reference rate} \pm \text{Quoted margin}$$

The quoted margin is the additional amount that the issuer agrees to pay above the reference rate (if the quoted margin is positive) or the amount less than the reference rate (if the quoted margin is negative). The quoted margin is expressed in terms of basis points. A basis point is equal to 0.0001 or 0.01%. Thus, 100 basis points are equal to 1%.

To illustrate a coupon reset formula, suppose that the reference rate is the 1-month London Interbank Offered Rate (LIBOR). Suppose that the quoted margin is 150 basis points. Then the coupon reset formula is:

$$\text{1-month LIBOR} + 150\,\text{basis points}$$

So, if 1-month LIBOR on the coupon reset date is 5.5%, the coupon rate is reset for that period at 7% (5% plus 150 basis points).

The reference rate for most floating-rate securities is an interest rate or an interest rate index. There are some issues where this is not the case. Instead, the reference rate is the rate of return on some financial index such as one of the stock market indexes. There are debt obligations whose coupon reset formula is tied to an inflation index.

Typically, the coupon reset formula on floating-rate securities is such that the coupon rate increases when the reference rate increases, and decreases when the reference rate decreases. There are issues whose coupon rate moves in the opposite direction from the change in the reference rate. Such issues are called inverse floaters or reverse floaters.

A floating-rate debt instrument may have a restriction on the maximum coupon rate that will be paid at a reset date. The maximum coupon rate is called a cap.

Because a cap restricts the coupon rate from increasing, a cap is an unattractive feature for the investor. In contrast, there could be a minimum coupon rate specified for a floating-rate security. The minimum coupon rate is called a floor. If the coupon reset formula produces a coupon rate that is below the floor, the floor is paid instead. Thus, a floor is an attractive feature for the investor.

Provisions for Paying off Debt Instruments

The issuer/borrower of a debt instrument agrees to repay the principal by the stated maturity date. The issuer/borrower can agree to repay the entire amount borrowed in one lump sum payment at the maturity date. That is, the issuer/borrower is not required to make any principal repayments prior to the maturity date. Such bonds are said to have a bullet maturity. An issuer may be required to retire a specified portion of an issue each year. This is referred to as a sinking fund requirement.

There are loans that have a schedule of principal repayments that are made prior to the final maturity of the instrument. Such debt instruments are said to be amortizing instruments. The same is true for mortgage-backed and most asset-backed securities because they are backed by pools of loans.

There are debt instruments that have a call provision. This provision grants the issuer/borrower an option to retire all or part of the issue prior to the stated maturity date. Some issues specify that the issuer must retire a predetermined amount of the issue periodically. Various types of call provisions are discussed below.

Call and Refunding Provisions

A borrower generally wants the right to retire a debt instrument prior to the stated maturity date because it recognizes that at some time in the future the general level of interest rates may fall sufficiently below the coupon rate so that redeeming the issue and replacing it with another debt instrument with a lower coupon rate would be economically beneficial. This right is a disadvantage to the investor since proceeds received must be reinvested at a lower interest rate. As a result, a borrower who wants to include this right as part of a debt instrument must compensate the investor when the issue is sold by offering a higher coupon rate.

The right of the borrower to retire the issue prior to the stated maturity date is referred to as a "call option." If the borrower exercises this right, the issuer is said to "call" the debt instrument. The price that the borrower must pay to retire the issue is referred to as the call price.

Prepayments

For amortizing instruments—such as loans and securities that are backed by loans—there is a schedule of principal repayments but individual borrowers typically have the option to pay off all or part of their loan prior to the scheduled date. Any principal repayment prior to the scheduled date is called a prepayment. The right of borrowers to prepay is called the prepayment option. Basically, the prepayment option is the same as a call option.

Options Granted to Bondholders

There are provisions in debt instruments that give either the investor and/or the issuer an option to take some action against the other party. The most common type of embedded option is a call feature, which was discussed earlier. This option is granted to the issuer. There are two options that can be granted to the owner of the debt instrument: the right to put the issue and the right to convert the issue.

A debt instrument with a put provision grants the investor the right to sell the issue back to the issuer at a specified price on designated dates. The specified price is called the put price. The advantage of the put provision to the investor is that if after the issuance date of the debt instrument market interest rates rise above the debt instrument's coupon rate, the investor can force the borrower to redeem the bond at the put price and then reinvest the proceeds at the prevailing higher rate.

A convertible debt instrument is one that grants the investor the right to convert or exchange the debt instrument for a specified number of shares of common stock. Such a feature allows the investor to take advantage of favorable movements in the price of the borrower's common stock or equity and is referred to as a *conversion provision*.

FINANCIAL MARKETS

A financial market is a market where financial instruments are exchanged (that is, traded). Although the existence of a financial market is not a necessary condition for the creation and exchange of a financial instrument, in most economies financial instruments are created and subsequently traded in some type of financial market. The market in which a financial asset trades for immediate delivery is called the spot market or cash market. The other type of financial market is called a derivatives market.

Role of Financial Markets

Financial markets provide three major economic functions. First, the interactions of buyers and sellers in a financial market determine the price of the traded asset. Or, equivalently, they determine the required return on a financial instrument. Because the inducement for firms to acquire funds depends on the required return that investors demand, it is this feature of financial markets that signals how the funds in the financial market should be allocated among financial instruments. This is called the price discovery process.

Second, financial markets provide a mechanism for an investor to sell a financial instrument. Because of this feature, it is said that a financial market offers "liquidity," an attractive feature when circumstances either force or motivate an investor to sell. If there were not liquidity, the owner would be forced to hold a financial instrument until the issuer initially contracted to make the final payment (that is, until the debt instrument matures) and an equity instrument until the company is either voluntarily or involuntarily liquidated. While all financial markets provide some form of liquidity, the degree of liquidity is one of the factors that characterize different markets.

The third economic function of a financial market is that it reduces the cost of transacting. There are two costs associated with transacting: search costs and information costs. *Search costs* represent explicit costs, such as the money spent to advertise one's intention to sell or purchase a financial instrument, and implicit costs, such as the value of time spent in locating a counterparty. The presence of some form of organized financial market reduces search costs. Information costs are costs associated with assessing the investment merits of a financial instrument, that is, the amount and the likelihood of the cash flow expected to be generated. In a price efficient market, prices reflect the aggregate information collected by all market participants.

Classification of Financial Markets

There are many ways to classify financial markets. One way is by the type of financial claim, such as debt markets and equity markets. Another is by the maturity of the claim. For example, the money market is a financial market for short-term debt instruments; the market for debt instruments with a maturity greater than one year and equity instruments is called the capital market.

Financial markets can be categorized as those dealing with financial claims that are newly issued, called the *primary market*, and those for exchanging financial claims previously issued, called the *secondary market* or the market for seasoned instruments.

Markets are classified as either cash markets or *derivative markets*. The latter is described later in this chapter. A market can be classified by its organizational structure: It may be an auction market or an *over-the-counter market*.

DERIVATIVE MARKETS

So far we have focused on the cash market for financial instruments. With some financial instruments, the contract holder has either the obligation or the choice to buy or sell a financial instrument at some future time. The price of any such contract derives its value from the value of the underlying financial instrument, financial index, or interest rate. Consequently, these contracts are called *derivative instruments*.

The primary role of derivative instruments is to provide an inexpensive way of protecting against various types of risk encountered by investors and issuers. Unfortunately, derivative instruments are too often viewed by the general public—and sometimes regulators and legislative bodies—as vehicles for pure speculation (that is, legalized gambling). Without derivative instruments and the markets in which they trade, the financial systems throughout the world would not be as efficient or integrated as they are today.

A May 1994 report published by the U.S. General Accounting Office (GAO) titled *Financial Derivatives: Actions Needed to Protect the Financial System* recognized the importance of derivatives for market participants. Page 6 of the report states:

> Derivatives serve an important function of the global financial marketplace, providing end-users with opportunities to better manage financial risks associated with their business transactions. The rapid growth and increasing complexity of derivatives reflect both the increased demand from end-users for better ways to manage their financial risks and the innovative capacity of the financial services industry to respond to market demands.

Types of Derivative Instruments

The two basic types of derivative instruments are futures/forward contracts and options contracts. A *futures contract* or forward contract is an agreement whereby two parties agree to transact with respect to some financial instrument at a predetermined price at a specified future date. One party agrees to buy the financial instrument; the other agrees to sell the financial instrument. Both parties are obligated to perform, and neither party charges a fee.

An *option contract* gives the owner of the contract the right, but not the obligation, to buy (or sell) a financial instrument at a specified price from (or to) another party. The buyer of the contract must pay the seller a fee, which is called the option price. When the option grants the owner of the option the right to buy a financial instrument from the other party, the option is called a call option. If, instead, the option grants the owner of the option the right to sell a financial instrument to the other party, the option is called a put option.

Derivative instruments are not limited to financial instruments. In this handbook we will describe derivative instruments where the underlying asset is a financial asset, or some financial benchmark such as a stock index or an interest rate, or a credit spread. Moreover, there are other types of derivative instruments that are basically "packages" of either forward contracts or option contracts. These include swaps, caps, and floors.

SUMMARY

Financial instruments can be classified by the type of claim that the holder has on the issuer (debt and equity) and cash and derivative instruments. With debt instruments there is an interest rate that is specified by contract. It could be a fixed interest rate or a floating interest rate. Other characteristics of debt instruments are that they have a maturity value and provisions for paying off the principal borrowed. Some debt instruments may have call, put or conversion provisions. An equity instrument obligates the issuer of the financial instrument to pay the holder an amount based on earnings, if any, after the holders of debt instruments have been paid.

Financial markets provide three major economic functions: (1) the determination of the price of the traded asset (price discovery); (2) a mechanism for an investor to sell a financial instrument (liquidity); and (3) reduction in the cost of transacting (search cost and information costs).

Financial markets are classified as cash (spot) markets and derivative markets. Derivative instruments include future/forwards contracts and options. The primary role of derivative instruments is to provide investors and issuers a vehicle for hedging/controlling different types of risk that they encounter when operating in the financial market.

REFERENCES

Brynjolfsson, J., and Fabozzi, F. J. (eds.) (1999). *Handbook of Inflation Indexed Bonds*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., Ramsey, C., and Marz, M. (eds.) (2000). *The Handbook of Nonagency Mortgage-Backed Securities*, 2nd edition. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (ed.) (2000). *Investing in Asset-Backed Securities*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (ed.) (2001). *Investing in Commercial Mortgage-Backed Securities*. Hoboken: NJ: John Wiley & Sons.

Fabozzi, F. J. (ed.) (2002). *The Handbook of Financial Instruments*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (ed.) (2005). *The Handbook of Fixed Income Securities*, 7th edition. New York: McGraw-Hill.

Fabozzi, F. J. (ed.) (2006). *The Handbook of Mortgage-Backed Securities*, 6th edition. New York: McGraw-Hill.

Fabozzi, F. J., and Choudhry, M. (eds.) (2004a). *The Handbook of European Fixed Income Securities*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., and Choudhry, M. (eds.) (2004b). *The Handbook of European Structured Financial Products*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., and Jacob, D. (eds.) (1999). *The Handbook of Commercial Mortgage-Backed Securities*, 2nd edition. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., and Modigliani, F. (2002). *Capital Markets: Institutions and Instruments*, 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Fabozzi, F. J., Modigliani, F., Jones, F. J., and Ferri, M. (2002). *Foundations of Financial Markets and Institutions*, 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Fabozzi, F. J., and Pilarinu, E. (eds.) (2002). *Investing in Emerging Fixed Income Markets*. Hoboken, NJ: John Wiley & Sons.

CHAPTER 2

Fundamentals of Investing

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: The investment management process involves five steps: setting investment objectives, establishing an investment policy, selecting a portfolio strategy, constructing a portfolio, and evaluating performance. The investment process involves the analysis of the investment objectives of the entity whose funds are being invested. Given the investment objectives, an investor must then establish policy guidelines to satisfy the investment objectives. This phase begins with the decision as to how to allocate funds across the major asset classes and requires a thorough understanding of the risks associated with investing in each asset class. After establishing the investment objectives and the investment policy, the investor must develop a portfolio strategy. Portfolio strategies can be classified as either active or passive. The next step is to construct the portfolio by selecting the specific financial instruments to be included in the portfolio. Periodically, the investor must evaluate the performance of the portfolio and therefore the portfolio strategy. This step begins with the calculation of the investment return and then evaluates that return relative to the portfolio risk.

Keywords: individual investors, institutional investors, asset classes, mutual fund, systematic risk, unsystematic risk, inflation risk, credit risk, interest rate risk, duration, liquidity risk, exchange rate risk, reinvestment risk, call risk, prepayment risk, active portfolio strategy, passive portfolio strategy, efficient portfolio, performance evaluation

In this chapter the fundamentals of investing will be reviewed. We will explain these fundamentals in terms of the steps that are involved in investing. These steps include setting investment objectives, establishing an investment policy, selecting a portfolio strategy, constructing a portfolio, and evaluating performance.

SETTING INVESTMENT OBJECTIVES

The investment process begins with a thorough analysis of the investment objectives of the entity whose funds are being invested. These entities can be classified as *individual investors* and *institutional investors*.

The objectives of an individual investor may be to accumulate funds to purchase a home or other major acquisition, to have sufficient funds to be able to retire at a specified age, or to accumulate funds to pay for college tuition for children.

Institutional investors include:

- Pension funds.
- Depository institutions (commercial banks, savings and loan associations, and credit unions).
- Insurance companies (life insurance companies, property and casualty insurance companies, and health insurance companies).
- Regulated investment companies (mutual funds).
- Endowments and foundations.

- Treasury departments of corporations, municipal governments, and government agencies.

In general we can classify institutional investors into two broad categories—those that must meet contractually specified liabilities and those that do not. We refer to those in the first category as institutions with "liability-driven objectives" and those in the second category as institutions with "non-liability-driven objectives." Some institutions have a wide range of investment products that they offer investors, some of which are liability-driven and others that are non-liability-driven.

ESTABLISHING AN INVESTMENT POLICY

Once the investment objectives are identified, an investor must then establish policy guidelines to satisfy the investment objectives. Setting policy begins with the asset allocation decision. That is, a decision must be made as to how the investor's funds should be distributed among asset classes. In making the asset allocation decision, investors will look at the risk and return characteristics of the asset classes in which they may invest and the correlation between the returns of each asset class. We define what is meant by an asset class and the notion of risk in the sections to follow.

The asset allocation will take into consideration any investment constraints or restrictions. Asset allocation models are commercially available for assisting those individuals responsible for making this decision.

In the development of investment policies, the following factors must be considered:

- Client constraints
- Regulatory constraints
- Accounting and tax issues

Asset Classes

From the perspective of a U.S. investor, the convention today is to refer to the following as traditional *asset classes*:

U.S. common stocks
Non-U.S. (or foreign) common stocks
U.S. bonds
Non-U.S. (or foreign) bonds
Cash equivalents
Real estate

Cash equivalents are defined as short-term debt obligations that have little price volatility.

Common stock and bonds are further divided into other asset classes. For U.S. common stocks, the following are classified as asset classes:

Large-capitalization stocks
Mid-capitalization stocks
Small-capitalization stocks
Growth stocks
Value stocks

"Capitalization" means the market capitalization of the company's common stock. It is equal to the total market value of all of the common stock outstanding for that company. For example, suppose that a company has 100 million shares of common stock outstanding and each share has a market value of $10. Then the market capitalization of this company is $1 billion (100 million shares times $10 per share). The market capitalization of a company is commonly referred to as its "market cap" or simply "cap."

While the market cap of a company is easy to determine given the market price per share and the number of shares outstanding, how does one define "value" and "growth" stocks? We'll see how that is done in Chapter 30 of Volume II.

For U.S. bonds, the following are classified as asset classes:

U.S. government bonds
Investment-grade corporate bonds
High-yield corporate bonds
U.S. municipal bonds (that is, state and local bonds)
Mortgage-backed securities
Asset-backed securities

All of these securities are described in later chapters, where what is meant by "investment grade" and "high yield" is also explained. Sometimes, the first three bond asset classes listed above are further divided into "long term" and "short term."

The following asset classes are classified for the non-U.S. common stock and bond asset classes:

Developed market foreign stocks
Emerging market foreign stocks
Developed market foreign bonds
Emerging market foreign bonds

In addition to the traditional asset classes listed above, there are asset classes commonly referred to as alternative asset classes. Some of the more popular ones include hedge funds, private equity, venture capital, and managed futures.

How does one define an asset class? One highly respected investment manager, Mark Kritzman (1959, p. 79), describes how this is done as follows:

> . . . [S]ome investments take on the status of an asset class simply because the managers of these assets promote them as an asset class. They believe that investors will be more inclined to allocate funds to their products if they are viewed as an asset class rather than merely as an investment strategy.

He then goes on to propose criteria for determining asset class status, although we won't review the criteria he proposed here.

Along with the designation of an investment as an asset class comes a barometer to be able to quantify performance—the risk, return, and the correlation of the return of the asset class with that of other asset classes. The barometer is called a "benchmark index" or simply "index." Listed in Table 2.1 are benchmark indexes for the various asset classes that cover common stocks.

If an investor wants exposure to a particular asset class, he or she must be able to buy a sufficient number of the individual securities comprising the asset class. This means that if an investor wants exposure to the U.S. large-cap

Table 2.1 Benchmark Indexes for Common Stock Asset Classes

Asset Class	Benchmark Index
U.S. Large-Cap Equity	Standard & Poor's (S&P) 500
U.S. Large-Cap Value	Frank Russell 1000 Value, S&P/Barra 500 Value
U.S. Large-Cap Growth	Frank Russell 1000 Growth, S&P/Barra 500 Growth
U.S. Mid-Cap Equity	Frank Russell Mid Cap
U.S. Small-Cap Equity	Frank Russell 2000
U.S. Small-Cap Value	Frank Russell 2000 Value
U.S. Small-Cap Growth	Frank Russell 2000 Growth
International Equity	Morgan Stanley Capital International (MSCI) EAFE, Salomon Smith Barney International, MSCI All Country World (ACW1) ex U.S.
Emerging Markets	MSCI Emerging Markets

equity market and the S&P 500 is the index (consisting of 500 companies) representing that asset class, then the investor cannot simply buy the shares of a handful of companies and hope to acquire the expected exposure to that asset class. For institutional investors, acquiring a sufficient number of individual securities comprising an asset class is often not a problem. However, for individual investors, obtaining exposure to an asset class by buying a sufficient number of individual securities is not simple. How can individual investors accomplish this?

Fortunately, there is an investment vehicle that can be used to obtain exposure to asset classes in a cost-effective manner. The vehicle is an investment company, more popularly referred to as a *mutual fund*. This investment vehicle is the subject of Chapter 60 in Volume I. For now, what is important to understand is that there are mutual funds that invest primarily in specific asset classes. Such mutual funds offer investors the opportunity to gain exposure to asset classes without having expertise in the management of the individual securities in that asset class and by investing a sum of money that, in the absence of a mutual fund, would not allow the investor to acquire a sufficient number of individual assets to obtain the desired exposure.

Risks Associated with Investing

There are various measures of risk. We will describe each of them here.

Total Risk

The dictionary defines risk as "hazard, peril, exposure to loss or injury." With respect to investments, investors have used a variety of definitions to describe risk. Today, the most commonly accepted definition of risk is one that involves a well-known statistical measure known as the variance. Specifically, investors quantify risk in terms of the variance of an asset's expected return. The variance of a random variable is a measure of the dispersion of the possible outcomes around the expected value. In the case of an asset's return, the variance is a measure of the dispersion of the possible outcomes for the return around the expected return.

There are two criticisms of the use of the variance as a measure of risk. The first criticism is that since the variance measures the dispersion of an asset's return around its expected value, it considers the possibility of returns above the expected return and below the expected return. Investors, however, do not view possible returns above the expected return as an unfavorable outcome. In fact, such outcomes are favorable. Because of this, some researchers have argued that measures of risk should not consider the possible returns above the expected return. Various measures of downside risk, such as risk of loss and value at risk, are currently being used by practitioners. The second criticism is that the variance is only one measure of how the returns vary around the expected return. When a probability distribution is not symmetrical around its expected return, then a statistical measure of the skewness of a distribution should be used in addition to the variance.

One way of reducing the risk associated with holding an individual security is by diversifying. Often, one hears investors talking about diversifying their portfolio. By this an investor means constructing a portfolio in such a way as to reduce portfolio risk without sacrificing return. This is certainly a goal that investors should seek. However, the question is, how does one do this in practice?

Some investors would say that a portfolio can be diversified by including assets across all asset classes. For example, one investor might argue that a portfolio should be diversified by investing in stocks, bonds, and real estate. While that might be reasonable, two questions must be addressed in order to construct a diversified portfolio. First, how much should be invested in each asset class? Should 40% of the portfolio be in stocks, 50% in bonds, and 10% in real estate, or is some other allocation more appropriate? Second, given the allocation, which specific stocks, bonds, and real estate should the investor select?

Some investors who focus only on one asset class such as common stock argue that such portfolios should also be diversified. By this they mean that an investor should not place all funds in the stock of one company, but rather should include stocks of many companies. Here, too, several questions must be answered in order to construct a diversified portfolio. First, which companies should be represented in the portfolio? Second, how much of the portfolio should be allocated to the stocks of each company?

Prior to the development of portfolio theory by Harry Markowitz (1952), while investors often talked about diversification in these general terms, they never provided the analytical tools by which to answer the questions posed here. Markowitz demonstrated that a diversification strategy should take into account the degree of covariance or correlation between asset returns in a portfolio. (The covariance or correlation of asset returns is a measure of the degree to which the returns on two assets vary or change together.) Indeed, a key contribution of what is now popularly referred to as "Markowitz diversification" or "mean-variance diversification" is the formulation of a security's risk in terms of a portfolio of securities, rather than the risk of an individual security. Markowitz diversification seeks to combine securities in a portfolio with returns that are less than perfectly positively correlated in an effort to lower portfolio risk (variance) without sacrificing return. It is the concern for maintaining return, while lowering risk through an analysis of the covariance between security returns, that separates Markowitz diversification from other approaches suggested for diversification and makes it more effective.

The principle of Markowitz diversification states that as the correlation (covariance) between the returns for assets that are combined in a portfolio decreases, so does the variance of the return for that portfolio. The good news is that investors can maintain expected portfolio return and lower portfolio risk by combining assets with lower (and preferably negative) correlations. However, the bad news is that very few assets have small to negative correlations with other assets. The problem, then, becomes one of searching among a large number of assets in an effort to discover the portfolio with the minimum risk at a given level of expected return or, equivalently, the highest expected return at a given level of risk.

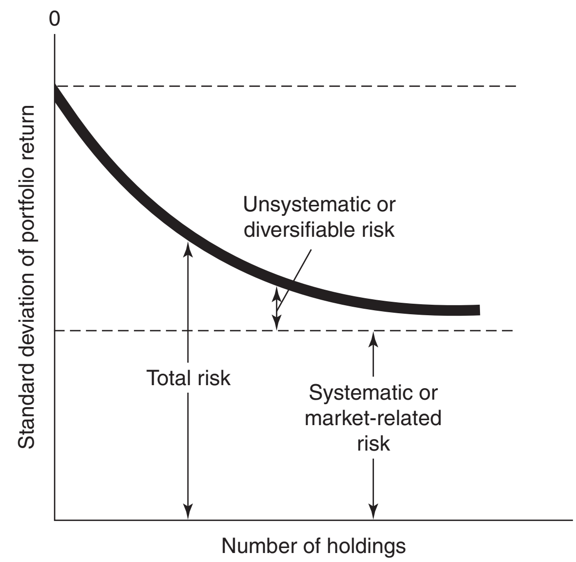


Figure 2.1 The Capital Market Line

Systematic versus Unsystematic Risk

The total risk of an asset or a portfolio can be divided into two types of risk: systematic risk and unsystematic risk. William Sharpe (1963) defined *systematic risk* as the portion of an asset's variability that can be attributed to a common factor. It is also called undiversifiable risk or market risk. Systematic risk is the minimum level of risk that can be attained for a portfolio by means of diversification across a large number of randomly chosen assets. As such, systematic risk is that which results from general market and economic conditions that cannot be diversified away.

Sharpe defined the portion of an asset's variability that can be diversified away as unsystematic risk. It is also called diversifiable risk, unique risk, residual risk, idiosyncratic risk, or company-specific risk. This is the risk that is unique to a company, such as a strike, the outcome of unfavorable litigation, or a natural catastrophe.

How diversification reduces unsystematic risk for portfolios is illustrated in Figure 2.1. The vertical axis shows the variance of a portfolio's return. This variance represents the total risk for the portfolio (systematic plus unsystematic). The horizontal axis shows the number of holdings of different assets (e.g., the number of common stock held of different issuers). As can be seen, as the number of asset holdings increases, the level of unsystematic risk is almost completely eliminated (that is, diversified away). Studies of different asset classes support this. For example, for common stock, several studies suggest that a portfolio size of about 20 randomly selected companies will completely eliminate unsystematic risk leaving only systematic risk. (The first study of this type was by Wagner and Lau 1971.) In the case of corporate bonds, generally less than 40 corporate issues are needed to eliminate unsystematic risk.

The relationship between the movement in the price of an asset and the market can be estimated statistically. There are two products of the estimated relationship that investors use. The first is the beta of an asset. Beta measures the sensitivity of an asset's return to changes in the market's return. Hence, beta is referred to as an index of systematic risk due to general market conditions that cannot be diversified away. For example, if an asset has a beta of 1.5, it means that, on average, if the market changes by 1%, the asset's return changes by about 1.5%. The beta for the market is 1. A beta greater than 1 means that the systematic risk is greater than that of the market; a beta less than 1 means that the systematic risk is less than that of the market. Brokerage firms, vendors such as Bloomberg, and online Internet services provide information on beta for common stock.

The second product is the ratio of the amount of systematic risk relative to the total risk. This ratio is called the coefficient of determination or R-squared. This ratio varies from 0 to 1. A value of 0.8 for a portfolio means that 80% of the variation in the return of the portfolio is explained by movements in the market. For individual assets, this ratio is typically low because there is a good deal of unsystematic risk. However, through diversification the ratio increases as unsystematic risk is reduced (see Figure 2.1).

Inflation or Purchasing Power Risk

Inflation risk, or purchasing power risk, arises because of the variation in the value of an asset's cash flows due to inflation, as measured in terms of purchasing power. For example, if an investor purchases an asset that produces an annual return of 5% and the rate of inflation is 3%, the purchasing power of the investor has not increased by 5%. Instead, the investor's purchasing power has increased by 2%. Inflation risk is the risk that the investor's return from the investment in an asset will be less than the rate of inflation.

Common stock is viewed by some as having little inflation risk. For all but inflation protection bonds, an investor is exposed to inflation risk by investing in fixed-rate bonds because the interest rate the issuer promises to make is fixed for the life of the issue.

Credit Risk

An investor who purchases a security not guaranteed by the U.S. government is viewed as being exposed to credit risk. There are several forms of credit risk: default risk, downgrade risk, and spread risk.

Default risk is defined as the risk that the issuer will fail to satisfy the terms of the obligation with respect to the timely payment of interest and repayment of the amount borrowed thereby forcing the issuer into bankruptcy. All investors in a bankrupt entity (common stockholders and bondholders) will realize a decline in the value of their security as a result of bankruptcy.

In the case of bonds, investors gauge the credit risk of an entity by looking at the credit ratings assigned to issues by rating companies, popularly referred to as rating agencies. There are three rating agencies in the United States: Moody's Investors Service, Inc., Standard & Poor's Corporation, and Fitch. These ratings are discussed in Chapter 24 of Volume III. When the credit rating of a bond is lowered by a rating agency, this action by a rating agency is referred to as the downgrading of a bond. The risk that a bond will be downgraded is called downgrade risk.

Credit spread risk is the risk that credit spreads in the market will increase resulting in poor performance of the bonds owned.

Liquidity Risk

When an investor wants to sell an asset, he or she is concerned whether the price that can be obtained from dealers is close to the true value of the asset. For example, if recent trades in the market for a particular asset have been between \$40 and \$40.50 and market conditions have not changed, an investor would expect to sell the asset in that range.

Liquidity risk is the risk that the investor will have to sell an asset below its true value where the true value is indicated by a recent transaction. The primary measure of liquidity is the size of the spread between the bid price (the price at which a dealer is willing to buy an asset) and the ask price (the price at which a dealer is willing to sell an asset). The wider the bid-ask spread, the greater the liquidity risk.

Liquidity risk is also important for portfolio managers that must mark to market positions periodically. For example, the manager of a mutual fund is required to report the market value of each holding at the end of each business day. This means accurate price information must be available. Some assets do not trade frequently and are therefore difficult to price.

Exchange Rate or Currency Risk

An asset whose payments are not in the domestic currency of the investor has unknown cash flows in the domestic currency. The cash flows in the investor's domestic currency are dependent on the exchange rate at the time the payments are received from the asset. For example, suppose an investor's domestic currency is the U.S. dollar and that the investor purchases an asset whose payments are in euros. If the euro depreciates relative to the U.S. dollar at the time a euro payment is received, then fewer U.S. dollars will be received.

The risk of receiving less of the domestic currency than is expected at the time of purchase when an asset makes payments in a currency other than the investor's domestic currency is called *exchange rate risk* or currency risk.

Risks for Bonds

There are systematic risks that affect bond returns in addition to those described above. They include interest rate risk, call/prepayment risk, and reinvestment risk.

Interest Rate Risk

The price of a bond changes as interest rates change. Specifically, price moves in the opposite direction to the change in interest rates. That is, if interest rates increase, the price of a bond will decline; if interest rates decrease, the price of a bond will increase. This is the reason a bond will sell above its par value (that is, sell at a premium) or below its par value (that is, sell at a discount). The risk that the price of a bond or bond portfolio will decline when interest rates increase is called *interest rate risk*.

The sensitivity of the price of a bond to changes in interest rates depends on the following factors:

- The bond's coupon rate
- The bond's maturity
- The level of interest rates

Specifically, the following relationships hold:

- All other factors being constant, the lower the coupon rate, the greater the price sensitivity of a bond for a given change in interest rates.
- All other factors being constant, the longer the maturity, the greater the price sensitivity of a bond for a given change in interest rates.
- All other factors being constant, the lower the level of interest rates, the greater the price volatility of a bond for a given change in interest rates.

Consequently, the price of a zero-coupon bond with a long maturity is highly sensitive to changes in interest rates. The price sensitivity is even greater in a low interest rate environment than in a high interest rate environment. For money market instruments, since their maturity is less than one year, the price is not very sensitive to changes in interest rates.

The price sensitivity of a bond to changes in interest rates can be estimated. This measure is called the *duration* of a bond. Duration is the approximate percentage change in the price of a bond for a 100-basis-point change in interest rates. For example, if a bond has a duration of 8, this means that for a 100-basis-point change in interest rates, the price will change by approximately 8%. For a 50-basis-point change in interest rates, the price of this bond would change by approximately 4%.

Given the price of a bond and its duration, the dollar price change can be estimated. For example if our bond with a duration of 8 has a price of \$90,000, the price will change by about 8% for a 100-basis-point change in interest rates and therefore the dollar price change will be about \$7,200 (8% times \$90,000). For a 50-basis-point change, the price would change by about \$3,600.

The concept of duration applies to a bond portfolio also. For example, if an investor has a bond portfolio with a duration of 6 and the market value of the portfolio is \$1 million, this means that a change in interest rates of 100 basis points will change the value of the portfolio by approximately 6% and therefore the value of the portfolio will change by approximately \$60,000. For a 25-basis-point change in interest rates, the portfolio's value will change by approximately 1.5% and the portfolio's value will change by approximately \$15,000.

How is duration computed? First, two prices are computed. One is based on an increase in interest rates and the second is based on a decrease in interest rates. Duration is then computed as follows:

$$\text{Duration} = \frac{\text{Price rates decrease} - \text{Price rates increase}}{2 \times \text{Initial price} \times \text{Change in price in decimal form}}$$

Typically, interest rates fluctuate up and down by an amount less than 50 basis points. But regardless of the rate change used, the interpretation is still that it is the approximate percentage price change for a 100-basis-point change in rates.

There are limitations of duration that the investor should recognize. First, in calculating duration or using the duration provided by financial consultants or fund managers, it is assumed that the prices calculated in the numerator are done properly. This is not a problem for simple bonds. However, there are bonds where if interest rates are changed the estimated price must be estimated by complex pricing models. In turn, those models are based on several assumptions. So, for example, it is not surprising that two brokers providing information on duration for a complex bond could have materially different estimates. One broker could report a duration of four while another a duration of six! Moreover, mutual fund managers who manage a portfolio containing a large allocation to complex bonds could report a duration that is significantly different than the true price sensitivity of the fund to changes in interest rates due to improperly calculating the duration of the complex bonds.

The second limitation of duration is that it is a good approximation for small changes in interest rates (e.g., 50-basis-point change in rates) but the approximation is poorer for a larger change in interest rates. This does not mean that it is not useful for giving the investor a feel for the price sensitivity of a bond or a portfolio.

The third limitation has to do with the duration of a portfolio. In computing the duration of the portfolio, first the duration of each bond in the portfolio is computed. Then a weighted average of the duration of the bonds in the portfolio is computed to get the portfolio duration. The limitation comes about because it is assumed that the interest rate for all maturities change by the same number of basis points. So, if a portfolio has a 2-year, a 10-year, and a 20-year bond, when using a portfolio's duration it is assumed that the 2-year, 10-year, and 20-year bonds change by the same number of basis points. This assumption is commonly referred to as the "parallel yield curve assumption."

Call/Prepayment Risk

A bond may include a provision that allows the issuer to retire or call all or part of the issue before the maturity date. From the investor's perspective, there are three disadvantages to call provisions. First, the cash flow pattern of a callable bond is not known with certainty because it is not known when the bond will be called. Second, because the issuer is likely to call the bonds when interest rates have dropped below the bond's coupon rate, the investor is exposed to *reinvestment risk*; this is risk that the investor will have to reinvest the proceeds when the bond is called at interest rates lower than the bond's coupon rate. Finally, the price appreciation potential of a bond will be reduced relative to an otherwise comparable bond without a call provision. Because of these three disadvantages faced by the investor, a callable bond is said to expose the investor to *call risk*. The same disadvantages apply to mortgage-backed and asset-backed securities where the borrower can prepay. In this case the risk is referred to as *prepayment risk*.

Reinvestment Risk

Reinvestment risk is the risk that proceeds available for reinvestment must be reinvested at a lower interest rate than the instrument that generated the proceeds. In addition to reinvestment risk when investing in a callable or prepayable bond, reinvestment risk occurs when an investor purchases a bond and relies on the yield of that bond as a measure of return potential. This point we be discussed later.

SELECTING A PORTFOLIO STRATEGY

Given the investment objectives and the investment policy, the investor must then develop a portfolio strategy. Portfolio strategies can be classified as either active or passive.

An *active portfolio strategy* uses available information and forecasting techniques to seek a better performance than a portfolio that is simply diversified broadly. Essential to all active strategies are expectations about the factors that influence the performance of an asset class. For example, with active common stock strategies this may include forecasts of future earnings, dividends, or price-to-earnings ratios. With bond portfolios that are actively managed, expectations may involve forecasts of future interest rates and sector spreads. Active portfolio strategies involving foreign securities may require forecasts of local interest rates and exchange rates.

A *passive portfolio strategy* involves minimal expectational input, and instead relies on diversification to match the performance of some index. In effect, a passive strategy assumes that the marketplace will reflect all available information in the price paid for securities. Between these extremes of active and passive strategies, new strategies have sprung up that have elements of both. For example, the core of a portfolio may be passively managed with the balance actively managed.

Given the choice among active or passive management, which should be selected? The answer depends on the investor's view of how "price-efficient" the market is and the investor's risk tolerance. By marketplace price efficiency we mean how difficult it would be to earn a greater return than passive management after adjusting for the risk associated with a strategy and the transaction costs associated with implementing that strategy. If an asset class is highly price efficient, the investor would want to pursue a passive strategy.

The most common passive strategy is indexing. In indexing, the investor designs a portfolio so that it will replicate the performance of the index.

CONSTRUCTING THE PORTFOLIO

Once a portfolio strategy is selected, the next step is to select the specific financial instruments to be included in the portfolio. (In the discussion to follow, we will refer to financial instruments as "securities.") This requires an evaluation of each security and the creation of an efficient portfolio. An *efficient portfolio* is one that provides the greatest expected return for a given level of risk, or equivalently, the lowest risk for a given expected return.

Constructing an Indexed Portfolio

As just mentioned, an investor who pursues the most popular form of a passive strategy, indexing, will assemble a portfolio that attempts to match the performance of the index. In theory, it is quite simple to do. An investor can purchase every security in the index. The amount purchased of a particular security should be equal to the percentage of that security in the index.

From a practical perspective, it may be difficult to buy all the securities comprising an index for several reasons. First, transaction costs from buying and rebalancing the indexed portfolio may be too expensive, resulting in the underperformance of the indexed portfolio relative to the index. Second, the amount to be invested may be such that all of the securities comprising the index cannot be acquired. For example, if an investor has $10,000 to invest in the stock market, the stock of only a few companies could be acquired. Finally, in some indexes not all of the securities can be acquired without great difficulty. For example, in the case of indexing to match the performance of a bond index, some of the bond issues included in the index may not trade frequently and are difficult to acquire.

For individuals, index replication is typically not accomplished by buying individual securities. Rather, if available, a mutual fund that has as its objective the creation of a portfolio to replicate an index can be purchased. This overcomes the problems of the individual investor creating the indexed portfolio. Managers of mutual funds have a larger amount to invest and therefore can acquire a large number of securities in the index and can do so minimizing transaction costs. A good example is the common stock indexed mutual funds.

For institutional investors, even with a large amount of funds to invest, the portfolio manager still faces the problem of transaction costs and unavailability of certain securities. There are trading arrangements that have been developed in some markets that allow for more efficient execution of trades so as to minimize transaction costs and therefore the likelihood that the indexed portfolio will underperform the index. For common stock, these trading arrangements are described in Chapter 11 of Volume I and Chapter 28 of Volume II. In the case of unavailable securities or a universe of securities so large that it is impractical to acquire all the securities in the index, there are methodologies that can be used to minimize the risk of not matching the index. We'll discuss this further below.

Constructing an Active Portfolio

In an active strategy, an investor is seeking to outperform the index or, in the case of liability-driven institutional investors, earn a higher return than a liability that it must pay. The construction of an active portfolio begins with an analysis of the factors that have historically determined the return on the index. Once these factors are identified, then the index can be decomposed into these factors or, more specifically, a risk profile of the index can be identified based on these factors.

Active management involves a deliberate decision by the portfolio manager to create a portfolio that departs from the risk profile of the index by accepting a larger or smaller exposure to one or more factors. Departures from the risk profile of the index represents bets on these factors. For example, consider common stock. One of the important factors that determines the risk profile of a common stock index such as the S&P 500 is the composition of the index in terms of industry sectors. Suppose that a portfolio manager believes that he or she can select industry sectors that can outperform and underperform. Then the portfolio manager will deliberately overweight the industry sectors that are expected to outperform and underweight those that are expected to underperform.

For an indexing strategy, in contrast, this approach involves creating a portfolio with a profile that matches the risk profile (that is, matching the factors) of the index. This mitigates the problem mentioned earlier of having to buy all the securities in the index.

EVALUATING PERFORMANCE

Periodically the investor must assess the performance of the portfolio and therefore the portfolio manager. This process begins with calculating the return realized over the investment period. The realized return is then compared to the return on the benchmark. The benchmark can be a market index or a minimum return established by a liability. The comparison will allow the investor to determine whether the portfolio outperformed, matched, or underperformed the benchmark.

However, the process does not stop there. It is common to compare the performance relative to the risk accepted—a reward-to-risk ratio. The most common measure used is the Sharpe ratio. The numerator of the Sharpe ratio is the return over the risk-free rate. The risk of the portfolio is measured by the standard deviation of the portfolio. The Sharpe ratio is thus:

$$\frac{\text{Portfolio return} - \text{Risk-free rate}}{\text{Standard deviation of portfolio}}$$

The Sharpe ratio is therefore is a measure of the excess return relative to the total variability of the portfolio.

For institutional investors, more elaborate techniques to assess performance are employed. The most common is the use of multifactor asset pricing models. While these models can be used to construct a portfolio, they can also be used to identify the reasons for underperformance or outperformance. These models do so by allowing the investor to determine the factor exposures that resulted in better or worse performance than the benchmark index.

SUMMARY

The investment management process involves five steps: setting investment objectives, establishing the investment policy, selecting the portfolio strategy, constructing the portfolio, and evaluating performance. Policy guidelines are established in order to satisfy the investment objectives and begin with the asset allocation decision. The policy guidelines must take into consideration client-imposed constraints, regulatory constraints (if applicable), and accounting and tax factors. After establishing the investment objectives and guidelines, the next step is to formulate a portfolio strategy. In doing so, a decision must be made as to whether an active or passive portfolio strategy is to be pursued. The choice depends on the investor's view of the efficiency of the market for the asset classes in which funds are to be allocated. Given the portfolio strategy, the specific securities to be held in the portfolio must then be selected. After the portfolio has been assembled, performance must be periodically evaluated. This involves comparing the performance of the portfolio to the established benchmark.

REFERENCES

Ammann, M., Kessler, S., and Tobler, J. (2006). Analyzing active investment strategies. *Journal of Portfolio Management* 32, 1: 56–67.

Anson, M. (2004). Strategic versus tactical asset allocation. *Journal of Portfolio Management* 29, 1: 8–22.

Ellis, C. D. (2003). The winner's game. *Journal of Portfolio Management* 28, 3: 27–34.

Fabozzi, F. J. (1999). *Investment Management*, 2nd edition, New Jersey, Prentice Hall.

Farr, D. D. (2006). Exploring the dimensions of active management 32, 1: 31–36.

Jacobs, B. I., and Levy, K. N. (2006). Enhanced active equity strategies. *Journal of Portfolio Management* 32, 3: 45–55.

Kritzman, M. (1999). Toward defining an asset class. *Journal of Alternative Investments* 2, 1: 79–82.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance* 7, 1: 77–91.

Markowitz, H. M. (1959). *Portfolio Selection*. Cowles Foundation Monograph 16. New York: John Wiley & Sons.

Nanda, S., and Peters, D. (2006). A very long-term buy-and-hold portfolio. *Journal of Portfolio Management* 31, 2: 65–73.

Sharpe, W. F. (1964). Capital asset prices. *Journal of Finance* 19, 3: 425–442.

Wagner, W. H., and Lau, S. (1971). The effect of diversification on risks. *Financial Analysts Journal* 27, 6: 48–52.

CHAPTER 3

The American Banking System

R. PHILIP GILES, PhD
Adjunct Professor of Finance and Economics, Columbia University Graduate School of Business and President, CBT Worldwide, Inc.

Abstract: After a 50-year hiatus, the American banking system is currently evolving from one characterized by severe geographic and product limitations into one with nationwide branching with banks offering a virtually complete array of financial services. This transformation was ushered in by a 10-year period of extensive deregulation at the state and federal levels, caused by a variety of external and technological factors. But although the banking system has been transformed, the complex web of overlapping banking regulatory authorities has not. Over the past two decades there has been substantial consolidation in the number of banking charters—with the number of banks falling by over half. However, we expect the dual banking system to survive, with a large number of individual banks compared to any other country in the world. One unanswered question at this time is whether the historic depository-nonfinancial business separation will continue.

Keywords: American banking system, structural transformation, post-Depression banking system, global banking constants, depository institutions, private bank, private banking, depository-nonfinancial separation, industrial loan companies, regulatory authorities, payments system, dual banking, Office of the Comptroller of the Currency (OCC), Federal Reserve System, member bank, Federal Deposit Insurance Corporation (FDIC), bank holding companies, board of governors of the Federal Reserve System (BGFRS), reserve requirements, relationships banking, disintermediation, money market mutual funds (MMMFs), bank failures, deregulation and reregulation, risk-based capital standards, off-balance-sheet (OBS) activities, technological progress, federal preemption, banking consolidation, nationwide branching, diseconomies of distance, universal banking

While the *American banking system* is unique to the world, it is neither emulated nor admired. This statement accurately describes the relationship of bank system differences between the United States and the rest of the world from colonial times until the present.

Although the Great Depression of the 1930s created unprecedented turbulence to the U.S. banking system, it was remarkably stable for the next 50 years. Giles (1983) details the structure of the banking system on the brink of change. Over the past two decades the American banking system has undergone tremendous structural transformation, moving closer to that found in other developed nations. According to Berger, Kashyap, and Scalise (1995) the 1980s and the first half of the 1990s were the most turbulent period for U.S. banking since the Great Depression. And, although structural change is still in progress, the ultimate form of the future banking system in the United States can be envisioned with some degree of certainty.

This chapter is organized into four parts. The first section describes the post-Depression banking system, including the regulatory structure, allowable bank products, and the scope and variety of banking organizations. This is followed by a description of the factors transforming the system into its current (evolving) form. The next section describes today's banking system and the last section summarizes the current structure and trends.

GLOBAL BANKING CONSTANTS

The banking system is a vital part of any economy. In brief:

- Banks pool and absorb risks for depositors and provide investment and working capital for nonfinancial industries.
- Banks are a particularly important source of funds for small borrowers.
- With the central bank's discount window, banks provide a backup source of liquidity for any sector in temporary difficulty.
- The central bank transmits monetary policy through banks.
- Banks provide the payment media.

The combination of a fractional reserve banking system and banks' role in providing the payments media means that problems in the banking sector can propel the entire economic system into a tailspin. This is due to the fact that in times of economic stress, currency can be hoarded (that is, kept under the mattress or buried in the backyard), causing large contractions in the money supply.

For this reason all banks are heavily regulated and deposit insurance is often provided to protect small depositors who are the heaviest users of currency. And, due to heavy regulation, we find there is extensive data on the banking system facilitating comprehensive research in the industry.

Despite the differences among banking systems, there are some universal truths to be found between large and small banks in a given country and among banking systems in a wide array of national economies. These constants are due to the universal needs of the general public:

- A safe place to store liquid wealth.
- A convenient means of making day-to-day payments.
- A source of credit.

There are global balance sheet constants which are listed below.

Bank assets can be placed into four slots, listed in decreasing level of liquidity:

- Cash
- Securities
- Loans
- Fixed assets

Two types of liabilities are found for any bank, with a third one for large banks in developed markets:

- Core deposits—small saving and time deposits and checking deposits of customers.
- Purchased funds—large negotiable CDs and interbank borrowing (for large banks in developed economies).
- Capital—including common equity plus other potential items.

In the United States as well as the rest of the world, most bank assets are expressed at book value in contrast to market value as found in securities firms and investment management companies. However as banks evolve from the "buy-and-hold" policy to more asset trading, the proportion of bank assets valued at the market has increased and this trend will continue.

For American banks, the loan-to-asset ratio is typically around 60%, while Asian banks usually have 70%. For any bank, fixed assets are usually around 1% of total assets. Finally, equity is usually between 5% and 10% of total assets.

From the above we can draw the following conclusions about banks around the world:

- Financial assets typically comprise 99% of total assets.
- Equity does not serve as a funding source, but plays the role of loss absorption.

STRUCTURE OF THE POST-DEPRESSION BANKING SECTOR

In this chapter the term "banking" refers to depository institutions. There are three major types of depositories:

1. Commercial banks.
2. "Thrifts," which comprise savings and loan associations and savings banks.
3. Credit unions.

A fourth category, state-chartered *industrial loan companies* (ILCs), have, until recently, played a niche role in the financial system.

Features of the "Old Structure" Still in Place

Although the U.S. banking system is converging to structures found elsewhere, some unique features are likely to remain. The following paragraphs point out these anomalies.

The Separation of Depositories from Nonfinancial Businesses

Today, virtually all U.S. depository institutions are chartered by a government authority. In the past it was possible to accept deposits as a "private bank," which is an unincorporated firm without a charter that accepts deposits, but as Spong (2000) points out, it is no longer possible to establish a new depository without a charter. Furthermore, the number of private banks has declined over the past decades so that few remain at this time. The most prominent remaining private bank in the United States is Brown Brothers Harriman & Co. in New York.

The term "private bank" should be differentiated from the function of "private banking" which involves managing the assets of high-net-worth clients. Many chartered banks have "private banking" departments, while overseas banking centers such as Geneva and Zurich have a number of prominent private banks that play a significant role in global private banking.

Historically, in the United States there has always been a separation between depositories and nonfinancial firms but it is unclear at this time if this separation will continue in the future. In the past the main exception to this separation has been the previously mentioned industrial loan companies which can be owned by nonfinancial firms.

Bank Charters and Regulatory Authorities

We will see that the U.S. regulatory structure is complex with many overlapping authorities. The following paragraphs detail the evolution of the U.S. regulatory system, emphasizing that structural change usually occurs only after an economic crisis. Table 3.1 presents a comprehensive list of relevant federal banking legislation (see www.fdic.gov/regulations/laws/important/index.htm).

Prior to the Civil War, the federal government had virtually no role in banking. Thus, two types of banks existed—state-chartered and private banks. The safety and quality of the banking system depended upon the level and enforcement of banking regulations. With bank regulation only at the state level, the quality of regulation varied widely. Because the predominant form of payment was specie and representative paper money, the quality of the payments system was poor. This presented a major deterrent to development of business relationships outside of the local area.

The economic crisis caused by the Civil War provided an opportunity to change the system. The National Bank Act of 1863 authorized federal chartering of depositories. This created the concept of "dual banking" for commercial banks. Today, dual banking exists for all major types of depository sectors listed earlier (except ILCs, which are

Table 3.1 Summary of Major Federal Banking Legislation

Year	Name	General Provisions
1863	National Bank Act	Established national banks; Created the Office of the Comptroller of the Currency (OCC); prohibited state banks from issuing currency
1913	Federal Reserve Act	Created the Federal Reserve System
1927	McFadden Act	Forced national banks to conform to branching laws of home state
1933	Glass-Steagall Act	Created the Federal Deposit Insurance Corporation; separated investment banking from commercial banking; placed interest rate ceilings on bank deposits
1956	Bank Holding Company Act	Defined a bank as any institution that simultaneously accepts demand deposits and offers commercial loans; extended Federal Reserve supervision to any company that controls two or more banks; prohibited BHCs from purchasing a bank in another state; nonbank business must be "closely-related to banking"
1970	Douglas Amendment to the Bank Holding Company Act	Gave the Federal Reserve System authority over formation and activities of one-bank holding companies
1978	International Banking Act	Brought foreign banks in the U.S. under federal regulatory framework; required deposit insurance for branches of foreign banks accepting retail deposits in the United States
1980	Depository Institution Deregulation and Monetary Control Act of 1980 (DIDMCA)	Lowered and unified reserve requirements; required all depositories offering transactions accounts to maintain reserves with the Fed; phased out interest rate ceilings on bank deposits; raised FDIC insurance ceiling from $40,000 to $100,000; opened the Fed's discount window to all chartered depositories
1994	Riegle-Neal Interstate Banking and Branching Efficiency Act (RNIBBEA)	Permitted bank holding companies to purchase banks in any state, subject to concentration limits; allowed interstate mergers between banks, subject to concentration limits and state laws
1999	Gramm-Leach-Bliley Financial Services Modernization Act	Securities firms and insurance companies allowed to purchase banks; banks allowed to underwrite insurance and securities and engage in real estate activities; allowed the creation of financial holding companies (FHCs) with wider array of activities than allowed for bank holding companies

only state-chartered), resulting in state and national banks, thrifts, and credit unions.

Although the writers of the National Bank Act would have preferred to transfer all bank regulatory and chartering authority to the federal government, it was not possible. Thus, to this day each of the 50 states maintains a state banking department that charters and regulates state-chartered banks. The *Office of the Comptroller of the Currency (OCC)* is responsible for chartering and regulating national banks.

The economic crisis caused by the Panic of 1908 demonstrated the necessity of a lender of last resort, resulting in the passage of the Federal Reserve Act in 1913. While all national banks were required to be members, membership was not mandatory for state-chartered banks and most state banks chose to be nonmembers. The term "*member bank*" originally meant that a bank belonging to the Federal Reserve was entitled to all Fed services, including the discount window, but was also subject to all Fed regulations, including reserve requirements. This resulted in three types of chartered commercial banks—national member banks, state member banks, and state nonmember banks.

The Great Depression began after the stock market crash of 1929 and in the next four years the United States lost over half of all banks through failure, mainly caused by currency drains by depositors (recall that a majority of banks were state nonmembers without discount window access). This banking panic demonstrated the need for deposit insurance. The Federal Deposit Insurance Corporation (FDIC) was formed as part of the Glass-Steagall Act of 1933. Again, universal coverage was not mandatory, although all member banks were required to have insurance. This created a fourth banking sector, state-chartered nonmember, uninsured bank. However, over time, deposit insurance has come to be viewed as an economic requirement, so that the few uninsured banks in the United States as of 2006 hold less than 0.2% of total deposits.

As of the end of 2006, about 65% of all banks were state-chartered, nonmember banks and about 25% of all banks had a national charter. However, national banks controlled about 65% of all banking assets while state nonmember banks comprised about 20% of the total.

Due to severe geographic and product restrictions, banks began to reorganize, forming bank holding companies (BHCs). A one-bank holding company (OBHC) allows banking organizations to own separate subsidiaries that offer services beyond that allowed for banks and outside of the single bank's market area. Multibank holding companies (MBHCs) provide a means for a given banking organization to operate outside the geographic area allowed for a single bank. The Bank Holding Company Act of 1956 gave the *Board of Governors of the Federal Reserve System (BGFRS)* the authority to regulate the nonbanking operations of bank holding companies (BHCs), and financial holding companies (FHCs) which are described below. Note that because ILCs did not offer demand deposit accounts, they did not fit the definition of "bank" in the BHC Act and thus escaped BGFRS oversight.

Table 3.2 presents the types of commercial banking organizations and the regulatory authorities to which they are subject. In this table, "X" under the regulation columns indicates primary regulatory authority, while the "X*" denotes secondary authority. Note that in addition to the

Table 3.2 Summary of U.S. Bank Regulatory Structure

Regulatory Authority		**OCC**			**50 State Banking Departments**			**12 Federal Reserve Banks**		**FDIC**		**BGFRS**
Regulatory Function		**Charter**	**Examine**	**Regulate**	**Charter**	**Examine**	**Regulate**	**Examine**	**Regulate**	**Examine**	**Regulate**	**Regulate**
Type of Banking Organization	National banks Fed member Insured	X	X	X			X*		X*		X*	
	State banks Fed member Insured				X	X	X	X	X		X*	
	State banks Nonmember Insured				X	X	X			X	X*	
	State banks Nonmember Uninsured				X	X	X					
	State ILCs Nonmember Uninsured				X	X	X			X	X*	
	BHCs & FHCs											X

examining staff of each of the 50 states, both the OCC and the Federal Reserve maintain a nationwide staff of examiners. Kohn (2004) points out that the quality of federal examination and supervision seems to vary noting that banks the Fed examines had the lowest failure rate, while the banks that OCC examined had the highest failure rate.

Product Constraints in American Banking

Originally, state regulations on allowable bank activities varied widely among the states, with some states allowing universal banking. But the 1933 Glass-Steagall Act forced the separation between banking and most securities activities so that banks were thereafter mostly confined to deposit acceptance and credit extension. Exceptions allowed bank dealing in all government bonds and underwriting of general obligation local government bonds. While brokerage activities were not specifically disallowed, discount brokerage was deemed to be less likely to be challenged, because no investment advice was offered.

When banks attempted to circumvent these restrictions through the bank holding company, the Bank Holding Company Act limited the products of their nonbanking subsidiaries to areas "closely related to banking." Thus, from 1933 until the mid-1980s there was almost a complete separation between the banking and securities industries.

Geographic Constraints in American Banking

The most unique aspect of American banking is the vast number of separately chartered depositories—banks, thrifts, and credit unions. Because the U.S. banking system was initially under individual state control, banks in a given state, as well as the state's banking authority, sought to protect their local markets from "foreign" banks, that is, any bank originating outside of the given state. Thus, all forms of interstate banking were prohibited.

The intrastate branching environments that existed until recently are listed below with the number of states as of 1975 given in parentheses, as detailed by Jayaratne and Strahan (1997).

- Statewide branching (14).
- Limited branching, usually within a county or metropolitan area (24).
- Unit banking, that is, no branching (12).

Although in the distant past bank customers would be content to conduct all of their banking business at a single physical branch, increased mobility created a need for geographic expansion of a given banking organization. With limited or no branching permitted, a new bank location often required a separately chartered bank. This led to the proliferation in the number of banks.

While state restrictions could only be applied to state-chartered banks, the 1927 McFadden Act placed national banks under the same geographic rules as that of the bank's home state. At its peak in the 1920s the United States had over 31,000 separately chartered banks. By 1933 this number had fallen to a little over 14,000 and remained remarkably stable for the next 60 years, peaking at 14,500 in 1984. (It is interesting to note that Texas, a unit-banking state, had almost 2,000 separately chartered banks in 1986.)

Geographic restraints created a large number of small, undiversified banks. Calomiris (2000, Chapter 1) demonstrates that the system was vulnerable to bank runs and portfolio shocks while Jayaratne and Strahan (1997) illustrate the system was also plagued with high costs due to inefficiencies.

When banks attempted to circumvent interstate banking limits by establishing multibank holding companies, Congress passed the Douglas Amendment to the Bank Holding Company Act of 1956 which gave the target state the authority to deny entry and this denial was universal.

Interest Rate Ceilings on Bank Deposits

One of the most unenlightened features of the 1933 Glass-Steagall Act placed interest rate limits on bank deposits. No interest was allowed on consumer and business checking accounts while "Regulation Q" limited the amount of interest that could be paid on small time and savings deposits as well as large time and certificates of deposit.

Initially, with a weak economy and no inflation these ceilings were not binding. But by the 1960s with accelerating economic growth and rising inflation, market interest rates exceeded rates that banks were allowed to pay.

Partial deregulation was required in the mid-1970s. Penn Central Railroad failed and defaulted on its commercial paper (CP) obligations causing the CP market to lock up. This meant that issuers were unable to roll over maturing obligations, but banks could not issue large certificates of deposits (CDs) when their maximum rate was below the market. Thus, interest rate ceilings on large CDs were eliminated.

Reserve Requirements

Reserve requirements force a bank to maintain a certain level of assets as a proportion of reservable liabilities. Originally, each state maintained separate reserve requirements. After 1914 the Fed had its own reserve requirements which were usually higher than that of the states. This was the reason a majority of state chartered banks chose to be nonmembers.

Originally, the Fed's reserve requirements were applied to most deposit liabilities including retail savings and time deposits, wholesale time deposits and all demand deposits, with the latter increasing with bank size. The highest reserve requirement exceeded 16% for checking accounts at large banks.

For all domestic time deposits, for example, large bank CDs, a bank faced three costs—the interest rate (i_{CD}), reserve requirements (RR_{CD}), and the FDIC insurance fee (C_{FDIC}). However, the FDIC only insured domestic deposits and reserve requirements could not be imposed on balances held out of the United States, such as Eurodollar deposits. A given bank's cost of accepting time deposits within the United States exceeded the cost of a virtually identical dollar-denominated deposits accepted in a foreign branch. The total cost of a domestic time deposit rate was labeled the "effective domestic cost" or "EDC" which was computed as $\text{EDC} = (i_{CD} + C_{\text{FDIC}})/(1 - RR_{CD})$. With LIBID representing the bid rate on Eurodollar deposits, any bank accepting time deposits both in the U.S. and London would be no worse off in London as long as LIBID $\leq$ EDC. In an environment with high interest rates the domestic regulatory burden substantially widened the spread between LIBID and domestic CD rates, reaching a peak of 269 basis points in September 1971. Because of this, and the natural reluctance of many dollar holders to avoid potential U.S. control over their deposits, the Eurodollar market grew rapidly at the expense of U.S. banking business and jobs.

Relationship Banking

Banks rely on external data from bond rating agencies for credit extension and loan pricing decisions for large borrowers. This is referred to as the "hard data" approach. There is a national market for large corporate loans and lending rates are very competitive.

As Fields, Fraser, Berry, and Byers (2006) point out, banks maintain a personal relationship with consumers and small and medium-sized business for their credit decisions. This information, called "soft data," is not made public. With the need for face-to-face communication, this is typically viewed as a local market.

FORCES FOR CHANGE IN AMERICAN BANKING IN THE 1980s AND 1990s

Starting in the 1980s many factors converged to substantially transform the U.S. banking system to a form more in line with that of other nations. Loss of market share due to increased external competition, increased volatility, and failure rates coupled with technological change led to substantial regulatory reform. Berger, Kashyap, and Scalise (1995) provide a comprehensive discussion of the early stages of this structural transformation.

External Competition

Banks lost dominance on both sides of their balance sheet to other firms. The following paragraphs describe these developments.

Disintermediation

Originally, banks provided the main source of borrowed funds for consumers, a major source of funds for small and middle-sized enterprises and working capital funds for large firms. However, external competitors provided an increasing slice of funds to each of these borrowing classes.

In the corporate sector after 1960 the commercial paper market began to replace banks in providing short-term funds for large highly regarded firms. With the development of the original-issue high-yield bond market, banks also began losing market share to middle market and lower-rated large firms.

From flow-of-funds data for the nonfarm-nonfinancial sector the combination of bank loans, commercial paper, and corporate bonds has traditionally comprised 40% to 50% of total liabilities. And, until 1985 the bank loan portion of these three sources ranged between 30% and 40%, reaching a peak of 42% in 1984. However, it now totals around 17% of these three sources, less than half of the historical share.

Banks traditionally provided consumers the main source of nonmortgage credit, but credit cards began to supplant bank loans from about 1975 on. In 1974, credit cards provided about 10% of the total while today this source provides over 50%.

Deposits as a Source of Funds

In 1960 the average American bank received 90% of total funds from deposits, with 55% comprised of consumer and business checking accounts paying no interest and the other accounts under Reg Q ceilings. This cheap and stable source of funds has fallen dramatically over the decades for a variety of reasons.

As consumers switched from currency and checks to credit cards for their daily purchases, their checking account balances fell accordingly. Similarly, with noncompetitive interest rate ceilings in a high inflation environment money market rates greatly exceeded maximum rates allowed by banks on consumer deposits.

Money market mutual funds (MMMFs) were first created in 1972. Assets consisted of Treasury bills, commercial paper, and large bank CDs with most of this yield passed on to shareholders. Because MMMFs give shareholders limited check writing abilities, banks lost market share. Today, the average bank receives less than 10% of its total funds from wholesale and retail checking accounts.

Volatility, Risk, and Failure

The banking structure that prevailed in the decade of the 1960s and 1970s became increasingly risky. With restricted branching, bank loan portfolios were undiversified geographically and by industry. Furthermore, as Berger, Kashyap, and Scalise (1995) point out, capital regulation was mainly ad hoc, did not differentiate by asset risk and did not include off-balance sheet activities. Over the 40-year span from the late 1960s to the present, the composite equity-asset ratio of American banks reached a low point of under 6% between 1978 and 1982.

With the twin bouts of inflation in the 1970s, extreme market volatility brought on bank stress and failures not seen since the Great Depression. Commodity and raw material prices rose sharply in the late 1970s, but, with a change in monetary policy in August 1979, inflation fell rapidly, carrying commodity prices down as well. This fall in prices was devastating for the agriculture and energy sectors as well as emerging market exporters.

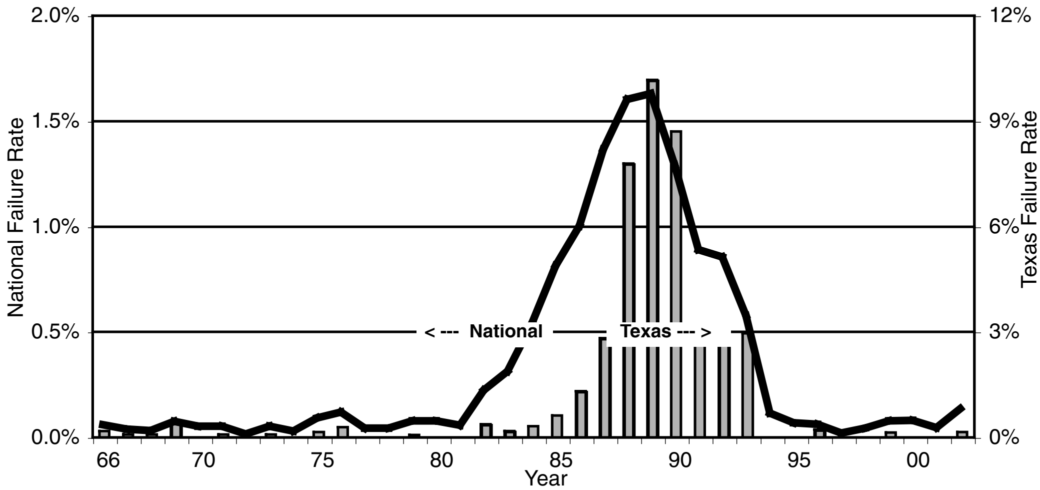


Figure 3.1 Bank Failure Rate

Between 1966 and 2002, the national failure rate for all banks averaged 0.3%. But during the 1982 to 1993 period the national failure rate tripled to an average of 0.9%, peaking at 1.6% during 1988 and 1989.

As bad as that sounds, it was much worse for Texas banks with energy, agriculture, and Latin American emerging market loans. Over the 1966 to 2002 period the Texas failure rate was four times the national rate at 1.2%, rising to 3.4% for the 1982 to 1993 period with a peak of 134 bank failures or 10.2% in 1989. As a means of reducing the cost of paying off the depositors of the insured depositors out-of-state MBHCs were allowed to buy failed Texas banks. Figure 3.1 presents the national and Texas failure rates between 1966 and 2006 illustrating the developing crisis of the late 1980s.

DEREGULATION, REREGULATION, AND TODAY'S EVOLVING BANKING SYSTEM

Regulatory and technological developments have substantially changed the banking system described above.

Deposit Deregulation

The movement from rigid regulation to recognition of market forces began with the gradual relaxing of rules on bank deposits.

Reserve Requirements

Partly due to the loss of dollar-based corporate banking business to offshore centers and the loss of retail bank deposits to MMMFs, the Depository Institutions Deregulation and Monetary Control Act (DIDMCA) was passed in 1980 and gradually reduced reserve requirements. By 1990, all reserve requirements were removed from retail and wholesale time and savings deposits and the reserve requirement for checking deposits was reduced to a flat 10%.

One result of this was to lower the domestic cost of large bank CDs so that since the mid 1990s rates on large domestic CDs, commercial paper, and LIBID trade within about 5 basis points.

Interest Rate Ceilings

The 1980 DIDMCA also removed interest rate ceilings from all savings and time deposits. And it allowed banks to offer interest-bearing checking accounts to consumers, giving them the ability to compete directly with MMMFs if they choose to do so.

Risk-Based Capital Standards

At the height of the post-Depression bank failure rate, the U.S. bank regulators met with their counterparts from the Group of Ten at the Bank for International Settlements (BIS), located in Basel, Switzerland, to design an international, risk-based capital system, now termed "Basel I."

To avoid regulatory arbitrage, the central banks of each participating country agreed to enforce the new standards on all of their "internationally active" banks. This standard is applied to all U.S. banks as well as BHCs.

Risk-Adjusted Assets

All assets are placed in one of four "buckets" depending on risk with different levels of capital required for each bucket. These buckets are presented in Table 3.3.

Off-Balance-Sheet Activities

Due to the lack of capital requirements and the development of derivative products, large banks began to derive an increasing amount of revenue from *off-balance-sheet (OBS) activities* in the form of fee income. As Berger, Kashyap, and Scalise (1995) point out, the ratio of

Table 3.3 Risk Weights for Credit Exposures under Basel I

Bucket	Allowable Assets	Risk Weight
Riskless	Central government securities of OECD countries	0%
	Cash and central bank deposits	
Money market risk	Local government bonds Deposits due from OECD banks	20%
Moderate risk	Residential mortgages Equivalent assets of interest rate swaps	50%
Standard risk	All other tangible assets	100%

noninterest income to total income for all U.S. banks rose from 7.0% in 1979 to 20.9% by 1994. Furthermore, from FDIC data we find this ratio in late 2006 is 29.0% for all banks.

A two-step process was defined in Basel I to account for OBS activity risks. Each activity is converted into an "equivalent asset" or "credit equivalent" by a "credit conversion factor." These factors are described in Table 3.4. Each credit equivalent is then placed into one of the four asset buckets described in Table 3.3.

Several points should be noted regarding Basel I. First, market risk is not addressed, only credit risk. This can be illustrated by noting that only options bought by the bank are covered, not options written by the bank. Note also that credit conversion factors are independent of the client. Thus, if a bank holds OBS products with an Organisation for Economic Co-operation and Development (OECD) central government, the risk weight would be 0%.

Uniform Definition of Capital

Capital was defined uniformly for all banks and was divided into two layers or "tiers." The minimum capital ratio for Tier 1 is 4% and the minimum total capital level for Tier 1 and Tier 2 is 8% of risk-adjusted assets.

The first tier consists of common equity and perpetual preferred stock without cumulative dividends, less intangible assets such as goodwill which cannot be liquidated to repay depositors. Tier 2 mainly consists of subordinated debt, other preferred stock, and a portion of a bank's reserve for loan losses.

Common equity is the most important part of capital since it has all desirable properties of capital, namely:

- Long-term stable funding, because equity is perpetual.
- Loss absorption, because common equity protects all other funds providers.
- Incentive for good management, since shareholders have voting rights.

As a result of Basel I, the equity/asset ratio of all U.S. banks by 2006 was at an all-time high of over 10%.

Table 3.4 OBS Credit Conversion Factors

Credit Products	
Commercial L/Cs	20%
Loan commitments	50%
Performance bonds	50%
Financial guarantees	100%
FX Risk Products	
Forward contracts	5%
Currency swaps	5%
FX options bought by bank	5%
Interest Rate Risk Products	
Interest rate swaps	0.50%
Forward rate agreements	0.50%
IR options bought by the bank	0.50%

Technological Progress and Bank Developments

Banks can offer new or improved products to their customers; they can lower operating costs and reduce risk through technological developments.

Front-Office Benefits

Automated teller machine (ATM) networks allow consumers to perform simple banking transactions at any time without waiting in line during banking hours. Improved funds transfer systems and automatic debits/credits reduce the need for paper-based payments. Improvements in automated clearing house functions allow near real-time transfer of funds accompanied with a substantial drop in transaction costs.

Internet banking, at the very least, provides access to relevant information for a customer from any location at any time. Some banks also allow transactions over the internet.

Back-Office Benefits

Cost shifts in payments processing have been substantial as Berger, Kashyap, and Scalise (1995) point out with two examples. Between 1979 and 1994 the real cost of processing a paper check rose from 1.99 cents to 2.53 cents while that for processing an electronic deposit fell from 9.10 cents to 1.38 cents.

Individual credit bureaus provide banks with instantaneous access to financial information on millions of individuals. Credit scoring systems allow banks to either augment their soft data or completely replace soft data with hard data on credit extension and loan pricing decisions to individuals.

Later developments fostered the creation of small business credit scoring systems which applies techniques for individual credit decisions to owners of small businesses, typically for loans up to $100,000. (See Berger, Frame, and Miller, 2005.)

Product Deregulation

Although the Glass-Steagall Act separated banking activities from securities activities, this separation began to erode in the 1980s. Securities firms offered MMMFs which had most features of bank checking and savings accounts (except for deposit insurance).

Furthermore, as bank loans began losing market share to fixed income securities, the Federal Reserve began loosening rules restricting securities activities of BHC nonbanking subsidiaries. Eventually, in 1989 the BGFRS authorized J. P. Morgan to underwrite corporate debt securities and later to underwrite stocks.

The Gramm-Leach-Bliley Financial Services Modernization Act of 1999 allowed securities firms and insurance companies to purchase banks, and allowed banks to underwrite insurance and securities. Mishkin (2004) provides a good description of the deregulation of banking activities.

Geographic Deregulation

Deregulation occurred between 1984 and 1994. Geographic deregulation will eventually create nationwide branching.

Changes in State Regulations

Prior to changes in federal legislation, state deregulation took place at two levels. In the mid-1980s individual states began to loosen regulations on intrastate branching, often moving from unit banking to limited branching and then to statewide branching.

At the same time, states began to form regional reciprocal relationships allowing MBHCs from neighboring states to buy in-state banks. The latter action created "superregional" banks in which one MBHC owned banks and branches in more than one state.

Federal Preemption of State Restrictions

At the federal level, the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994 (RNIBBEA) replaced 50 state entry laws over interstate banking. This is a rare example of federal banking laws preempting state banking laws.

This act initially allowed banking organizations to acquire banks in any state unless state law specifically prohibits. Hawaii was the only state to do so. Furthermore, under RNIBBEA BHCs could merge the banks they acquired into one bank with branches in different states. This has created interstate branching in 47 states. Texas, Montana, and Minnesota are the only states to opt out of this provision.

Results of Geographic Deregulation

Upheaval of the post-Depression banking system has been profound, including the number of bank charters, bank offices, and banking efficiency.

Banking Consolidation

One result of geographic deregulation has been to sharply reduce the number of separately chartered banks while at the same time increase the number of bank branches. Since 1984 the number of bank charters has fallen by 50% from 14,500 to around 7,000. At the same time, the number of bank branches has increased dramatically. For example, the number of banking offices in 1984 was 56,376 while in late 2006 it was 80,473.

Consolidation has also reduced the number of cities hosting large banks. As DeYoung and Klier (2004) point out, the corporate headquarters of merged banks are usually found in locations with the least geographic restrictions in former years, abandoning former-unit banking locales such as Chicago, Dallas, and Houston.

Bank Performance

Jayaratne and Strahan (1997) indicate that branching deregulation has improved bank performance while simultaneously benefiting bank customers, as diversification has decreased loan losses as well as loan pricing. Schuermann (2004) concludes that banks are less vulnerable to the business cycle with increased diversification through geographic expansion.

The Prospects for Nationwide Branching

While any bank now has the ability to form a nationwide branch network with a single bank charter, it may not necessarily be economically feasible to do so, given the vast geographic area. This same question is being asked within the European Union with their "Single Market Programme" potentially creating continent-wide banking organizations, as well as in Eastern Europe and Latin America.

One problem relates to the ability of senior management, located at the lead bank headquarters, to control junior managers at affiliates. That is, can "best practices" at the lead bank be transmitted to affiliates? A second question refers to the agency cost of distance, that is, does the lead bank have the ability to monitor far-flung affiliates?

Berger and DeYoung (2006) demonstrate that there are economic benefits to geographic expansion through their study of the MBHCs over the 1985 to 1998 time period. With technological developments, lead bank managers are indeed able to pass down efficiency improvements to acquired affiliates with this control increasing over time after acquisition. Furthermore, with technological improvements, diseconomies of distance have decreased.

Berger (2003) also demonstrates that technological progress has diminished the role of relationship banking. As cited earlier, retail lending no longer requires a local relationship and small business credit scoring has allowed "soft data" of relationship banking to be replaced by "hard data" which does not require a local presence for small business loans up to $100,000.

Concentration into Top Banks

FDIC data indicate that total deposits of all banks with (nominal) assets over $1 billion have increased from 68% in 2002 to 85% by late 2006. The top five banking organizations now control 35% of total domestic deposits. As of 2007, Bank of America had an 11% national share and JPMorgan-Chase has an 8.4% total.

One obstacle to further consolidation with the largest banking organizations is the RNIBBEA restriction that no BHC can control more than 30% of deposits in a single state and no more than 10% nationally, except for acquisition of failed banks. Thus, it appears that Bank of America has already reached its concentration limit.

Table 3.5 Risk Weights for Corporate Credit Exposures Under Basle II

Credit Assessment	AAA to AA−	A+ to A−	BBB+ to BB−	Below BB−	Unrated
Risk weight	20%	50%	100%	150%	100%

THE AMERICAN BANKING SYSTEM WILL CONTINUE TO BE UNIQUE

Although the U.S. banking system has witnessed profound structural changes in the two decades since the mid-1980s, it is likely to continue to be unique. Whether in the future the system will be emulated or admired is an open question.

The Outlook for Dual Banking

We should expect to see bank concentration continue through BHC mergers. We should also expect to see further consolidation as MHBCs continue to convert their acquired affiliates into branches.

Extrapolating current trends such as OCC preeminence and concentration of assets in national banks into the future one might conclude that we are heading towards the ultimate demise of dual banking, finally ending with a handful of nationally chartered banks. However, this is not the likely outcome.

One unique feature that we should expect to continue is the existence of a relatively large number of small, single market banks. While we have observed that the number of bank charters has fallen by 50% percent from 1984 to 2005, this trend appears to be slowing. Although relationship banking has declined in importance, it is still viable for small and middle-sized firms that do not have a bond rating but seek credit over a given limit. As Berger, Frame, and Miller (2005) illustrate the use of credit scoring for approve/deny decisions stops at the $100,000 level. Thus, relationship banking will continue for the middle markets sector, at least for the foreseeable future, and the dual banking structure will persist.

Continued Technological Advances in Risk Control

The original risk-based capital rules of Basel I were vital, but changes are needed due to financial system developments over the past two decades. The Basel Committee on Bank Supervision (2003) provides some of the features of Basel II.

Basel II incorporates market risk into capital requirements because banks have moved from their "buy-and-hold" philosophy into more investment banking activities. Basel II also incorporates ratings agency data in setting capital ratios for corporate credit exposures. Instead of lumping all corporate credits into the Standard Risk bucket with a 100% weight, there is now a wider spectrum of risk weights, ranging from low weights for investment-grade firms to high weights for speculative grade firms. Table 3.5 illustrates this refinement.

Basel II also encourages banks to use their own internal systems for capital requirements. It recognizes that individual banks may have a better grasp of the risks they face than regulators. This allows banks to use an internal ratings–based approach, which should lower capital requirements. Furthermore, Basel II allows large banks to use an advanced internal ratings–based approach, replacing more standard "one bank fits all" data with internal data, including credit scoring data.

It is apparent the Basel II approach will continue as financial engineering progresses beyond current derivative and structured finance products. While these products can indeed transfer risk, they can also concentrate risk in ways that have not been tested or observed at this time.

The Future U.S. Regulatory Structure

Mishkin (2004) describes the U.S. regulatory system as a "crazy quilt of multiple regulatory agencies with overlapping jurisdictions." As illustrated above we are unique in having dual banking as well as having three federal agencies with nationwide examining functions. However, no agency will voluntarily cede examination and regulatory authority. Bank chartering agencies claim they need to know the condition of a troubled institution in order to make the decision to pull the charter. Similarly, the FDIC needs to know the condition of the bank when it insures deposits. Kohn (2004) notes that when Chase switched from a national to a state charter in 1995, the OCC lost 2% of its budget.

But it may not be as redundant as it appears in Table 3.2, in that state bank examiners often share tasks with federal examiners on a coordinated examination. To streamline the system, it would appear that the examining function of the Federal Reserve Banks could be jettisoned, but it doesn't seem likely at the present time.

The Continued Separation of Depositories from Nonfinancial Firms

One unanswered question at this time is the continued separation of depositories from nonfinancial businesses. Although industrial loan companies have historically played an insignificant role in the U.S. financial system Ergungor and Thompson (2006) point out that between 1987 and 2005 assets have risen from $3.8 billion to $140 billion with the largest ILC now having assets exceeding $66 billion. Furthermore, although FDIC insurance was not originally available to ILCs, the Garn–St. Germain Act of 1982 extended insurance to this sector. Today, seven states charter ILCs and the ILC charter is effectively the only vehicle by which nonfinancial firms can enter banking. Recent nonfinancial-depository combinations include Toyota, GMAC, Target, and Home Depot, which own large ILCs offering a variety of financial services. As of 2005, the 10 largest ILCs include Merrill Lynch Bank USA, UBS Bank USA, American Express Centurion Bank, Fremont Investment & Loan, Morgan Stanley Bank, USAA Savings Bank, GMAC Commercial Mortgage Bank, GMAC Automotive Bank, Beal Savings Bank, and Lehman Brothers Commercial Bank.

Federal bank regulators are at odds over this topic with the FDIC supporting ILCs in the current form and BGFRS arguing for a change in the 1956 definition of "bank" to bring ILCs and their nonfinancial parent under the board's supervision.

SUMMARY

Over a very short time, the American banking system has progressed from one that was segmented by product line and geography into one allowing universal banking and nationwide branching. With risk-based capital rules, product and geographic deregulation and universal access to the Fed's discount window, today's banking system has lower costs and less risk than the system of the past.

REFERENCES

Basel Committee on Bank Supervision. (2003). Quantitative Impact Statement 3. [Available at www.bis.org/bcbs/qis/qis3.htm.]

Berger, A. N. (2003). The economic effects of technological progress: Evidence from the banking industry. *Journal of Money, Credit and Banking* 35, 2: 141–176.

Berger, A. N., and DeYoung, R. (2006). Technological progress and the geographic expansion of the banking industry. *Journal of Money, Credit and Banking* 38, 6: 1483–1513.

Berger, A. N., Dick, A. A., Goldberg, L. G., and White, L. J. (2005). Competition from large, multimarket firms on the performance of small, single-market firms: evidence from the banking industry. *Finance and Economics Discussion Series*, 2005-15, Board of Governors of the Federal Reserve System.

Berger, A. N., Kashyap, A. K., and Scalise, J. M. (1995). The transformation of the U.S. banking industry: What a long, strange trip it's been. *Brookings Papers on Economic Activity* 2: 55–218.

Berger, A. N., Frame, W. S., and Miller, N. H. (2005). Credit scoring and the availability, price and risk of small business credit. *Journal of Money, Credit, and Banking* 37, 2: 191–222.

Brevoort, K. P., and Hannan, T. H. (2006). Commercial lending and distance: evidence from community reinvestment act data. *Journal of Money, Credit and Banking* 38, 8: 1991–2012.

Calomiris, C. W. (2000). *U.S. Bank Deregulation in Historical Perspective*. New York: Cambridge University Press.

Davis, E., and Rice, T. (2006). Federal preemption of state bank regulation: a conference panel summary. Chicago Fed Letter, Federal Reserve Bank of Chicago.

DeYoung, R., and Klier, T. (2004). Why Bank One left Chicago: One piece in a bigger puzzle. *Chicago Fed Letter*, April.

Ergungor, O. E., and Thomson, J. B. (2006). *Industrial Loan Companies. Economic Commentary*. Federal Reserve Bank of Cleveland.

Fields, L. P., Fraser, D. R., Berry, T. L., and Byers, S. (2006). Do bank loan relationships still matter? *Journal of Money, Credit and Banking* 38, 5: 1195–1209.

Frame, W. S., Srinivasan, A., and Woosley, L. (2001). The effect of credit scoring on small-business lending. *Journal of Money, Credit, and Banking* 33, 3: 813–825.

Giles, R. P. (1983). A Model of the U.S. Correspondent Banking System, Unpublished PhD thesis, Columbia University.

Hannan, T. H., and Prager, R. A. (2006). The profitability of small, single-market banks in an era of multimarket banking. Finance and Economics Discussion Series, 2006-41, Board of Governors of the Federal Reserve System.

Hirtle, B., Levonian, M., Saidenberg, M., Walter, S., and Wright, D. (2001). Using credit risk models for regulatory capital: Issues and options. *Economic Policy Review,* Federal Reserve Bank of New York 7, 1: 21–35.

Jayaratne, J., and Strahan, P. E. (1997). The benefits of branching deregulation. *Economic Policy Review,* Federal Reserve Bank of New York (December).

Janicki, H. P., and Prescott, E. S. (2006). Changes in the size distribution of U.S. banks: 1960–2005. *Economic Quarterly,* Federal Reserve Bank of Richmond 92, 4: 291–316.

Jones, K. D., and Crichfield, T. (2005). Consolidation in the U.S. banking industry: Is the "long, strange trip" about to end? *FDIC Banking Review* 17, 4: 31–61.

Kohn, M. (2004). *Financial Institutions and Markets*, 2nd edition. New York: Oxford University Press.

Mishkin, F. (2004). *The Economics of Money, Banking and Financial Markets*, 7th edition. New York: Pearson.

Petersen, M. A., and Rajan, R. G. (2002). Does distance still matter? The information revolution in small business lending. *Journal of Finance* 57, 6: 2533–2570.

Radecki, L. J. (1998). The expanding geographic reach of retail banking markets. *Economic Policy Review,* Federal Reserve Bank of New York.

Schuermann, T. (2004). Why were banks better off in the 2001 recession? *Current Issues in Economics and Finance,* Federal Reserve Bank of New York 10, 1.

Spong, K. (2000). *Bank Regulation: Its Purposes, Implementation, and Effects*, 5th edition. Division of Supervision and Risk Management, Federal Reserve Bank of Kansas City.

Tirtiroglu, D., Daniels, K. N., and Tirtiroglu, E. (2005). Deregulation, intensity of competition, industry evolution, and the productivity growth of U.S. commercial banks. *Journal of Money, Credit and Banking* 37, 2: 339–360.

Tracey, W. F., and Carey, M. (2000). Credit risk rating systems at large U.S. banks. *Journal of Banking and Finance* 24: 167–201.

Walter, J. R. (2006). Not your father's credit union. *Economic Quarterly,* Federal Reserve Bank of Richmond 92, 4: 353–377.

CHAPTER 4

Monetary Policy: How the Fed Sets, Implements, and Measures Policy Choices

DAVID M. JONES, PhD
President, DMJ Advisors, LLC

ELLEN J. RACHLIN
Managing Director, Portfolio Management, Mariner Investment Group, Inc.

Abstract: The Federal Reserve's twin goals are to achieve sustainable economic growth and stable prices. The Federal Reserve strives also to contain inflationary expectations. The Fed seeks to influence the U.S. economy in a manner which allows it to grow at a pace which will employ all available resources but not so fast a pace as to fuel inflation. The Federal Reserve's effectiveness in achieving the dual goals of price stability and sustainable economic growth is limited to their direct influence on short-term interest rates and indirect effect on long-term interest rates. Historically, the Federal Reserve under different chairmen has favored targeting the Federal funds rate most of the time. The Federal Reserve is challenged in achieving the desired results due to its indirect influence over global capital flows. At times, global capital flows may work against what the Federal Reserve is trying to achieve. An investor who can anticipate Federal Reserve policy shifts can more accurately anticipate changes in asset valuations. The Federal Reserve, in setting short-term interest rates, ultimately affects asset repricings and valuations. An investor needs to understand the dynamics of Federal Reserve policy shifts and pay attention to key economic indicators that the Fed watches.

Keywords: price stability, monetary policy, open market operations, FOMC Meetings, discount rate changes, federal funds rate target, "leaning against the wind" policy, monetary policy transmission process, yield curve, borrowed reserves, sustainable economic growth

A significant element of competitive and successful equity and fixed income portfolio management is to understand and anticipate the effect of interest rate changes on asset prices. This chapter outlines the key components of the interest rate policy process undertaken by the Federal Reserve, the policy-making body that sets short-term interest rates. Although the Federal Reserve publicly announces its policy decisions, it is extremely useful for the portfolio manager to anticipate policy shifts. The portfolio manager who can anticipate policy shifts can more accurately anticipate changes in asset valuations. In order to anticipate policy shifts, the portfolio manager must not only understand the dynamics of the Fed's decision-making process but must watch the key economic indicators that the Fed watches.

Monetary policy is the U.S. government's most flexible policy tool. It is controlled by the Federal Reserve, which acts independently of government interference. The

Federal Reserve, since early 1994, immediately announces policy decisions and communicates forthcoming policy intentions through venues such as Congressional testimony and public speeches. Other government policy instruments that influence economic activity include fiscal policy (taxes and spending), trade, foreign exchange, and other regulatory practices. But none of these government policy tools are as flexible as monetary policy.

Since the Federal Reserve's inception in 1913, it has had the primary task of ensuring that financial conditions are supportive of sustainable, noninflationary economic growth. The Fed has several tools it can employ to influence aggregate demand, output growth and price behavior. But, as will be discussed later, all of the Fed policy tools directly influence short-term interest rates and only indirectly influence long-term rates.

The effectiveness of any Fed policy on achieving the goal of price stability is limited to the influence of that policy on both short- and long-term interest rates, real and nominal. Price stability is the primary prerequisite for steady long-term economic growth. Low inflation rates enable businesses to increase their investment in infrastructure, including new machinery and high-tech equipment. Therefore, a low inflationary environment brighten the prospects for future increases in productivity and an improved standard of living.

In the short term, the Fed must juggle the simultaneous objectives of stable prices and maximum employment (sustainable growth). Although the president appoints the Fed chairman by legislative decree, the Federal Reserve is an independent agency and is accountable to the public only through the legislative branch of the U.S. government. If the Fed were not an independent agency, it could be subject to political influences promoting economic growth over price stability.

Fed-induced *price stability* and the absence of consumer and business inflationary expectations are essential to containing speculation and allowing the capital markets to efficiently allocate funds in support of sustainable growth. Generally speaking, capital markets efficiently allocate funds to the sectors of the economy promising the highest risk-adjusted returns. This process is absolutely crucial to the wealth-creating success of modern capitalism. But, as we will discuss later, at times excesses of capital allocation can occur. Capital markets may allocate capital to countries where the risks of debt default or likely debt downgrades appears quite high, such as Mexico in 1995 and southeast Asia in 1997, as these participants have confidence central bankers will successfully stave off defaults. This exaggerated if not misplaced faith that somebody will bail them out of bad investment or lending decisions, called the moral hazard of central banking, is one of the few downside effects to the central bankers' role of lender of last resort.

KEY ECONOMIC INFLUENCES ON FED POLICY

Monetary policy is more art than science. In essence, Fed policy is a process of trial and error, observation and adjustment. The Fed's policies are often countercyclical to the business cycle. At best, Fed policy makers can hope to smooth the peaks and troughs of business cycles. In pursuit of this countercyclical policy approach, when output is excessive relative to the economy's sustainable potential and is potentially inflationary, Fed officials will lean in the direction of more restraint in their policy stance. They will tend to increase interest rates, eventually slowing aggregate demand and output growth to a more sustainable and potentially less inflationary pace. Conversely, when output falters and falls below the economy's sustainable potential, recession threatens. Accordingly, Fed officials will tilt their policy stance in the direction of greater ease, and lower interest rates. Lower interest rates serve to boost aggregate demand and output growth thereby lessening the threat of recession. A word of caution: Fed officials must feel their way along after implementing policy shifts because the effects from policy shifts are long, variable, and sometimes difficult to predict. The Fed enacts policy shifts based on economic forecasts. Economic forecasting is often an uncertain exercise.

Table 4.1 Significant Economic Releases

Payroll employment	Monthly—1st Friday
Housing starts	Monthly—3rd–4th week
Industrial production	Monthly—3rd week
ISM (supplier deliveries)	Monthly—1st business day
Motor vehicle sales	Monthly—1st–3rd business day
Durable goods orders	Monthly—4th week
Employment cost index	Quarterly
Nonfarm productivity growth	Quarterly
Commodity prices	Continuously released

One can develop an idea of the Fed's next policy objective by paying careful attention to various indicators of current economic activity. The key economic releases which serve as the Fed's intermediate policy indicators and that market participants should follow carefully include: nonfarm payrolls, ISM supplier deliveries, industrial production, housing starts, motor vehicle sales, durable good orders, labor compensation, productivity growth, and commodity prices. Table 4.1 gives the release cycles of these economic indicators. Consistent and meaningful changes in these economic indicators will signal changes in the business cycle and in Fed policy.

- Nonfarm payrolls are released monthly and detail the previous month's changes in the complexion of the workforce including numbers employed, hourly pay changes and hours worked. Supplier deliveries are part of the Institute for Supply Management's monthly survey. This report reflects survey results of the purchasing managers of hundreds of industrial corporations. The survey reports on the lead time between orders placed with suppliers and delivery of those orders. The greater the lead time, the stronger the economy and the lesser the lead time, the weaker the economy.
- Industrial production, released monthly, measures the collective output of factories, utilities, and mines. If final

demand is high and inventory stockpiles are rapidly shrinking, future industrial production, employment, and income will be boosted as inventories are restocked, thereby stimulating economic activity. If, in contrast, final demand growth is slowing and inventory growth is excessive, future industrial production, employment, and income will weaken as inventories are trimmed, thereby depressing economic activity.

- Housing starts, published monthly, are the number of new single- and multi-family housing units begun for construction in the previous month. Housing starts, which are financed, are highly sensitive to interest rate changes. If housing starts slow dramatically, this signals that interest rates are high enough in the current economic environment to choke off demand. Conversely, if housing starts are increasing, this signals that interest rates are low enough in the current economic environment to promote demand.
- Motor vehicle sales, released monthly, are a key reflection of consumer confidence and income. Motor vehicle sales are strongly positively correlated to both income levels and consumer confidence.
- Durable goods orders, released monthly, are new orders placed by consumers with manufacturers of "large ticket" consumer goods, expected to last three or more years. These items may include appliances or business machinery.
- Commodity prices, for which the market receives continual input, are important indicators of future price rises in both producer and subsequently consumer prices. The most influential prices are those of raw goods and materials such as oil, lumber, metals, and agricultural commodities. Consistent, sympathetic, and significant price increases in these raw goods and materials will signal higher future prices in finished consumer goods.
- Employment cost index which includes workers' wages, salaries and benefits is released quarterly.
- Nonfarm-productivity growth which is defined as output per hour and is released quarterly, is closely followed by Fed officials. In order to estimate the economy's sustainable potential, it is necessary to add productivity growth plus labor force growth. These are supply-side factors.

It is extremely difficult to recognize meaningful and consistent changes in these economic variables: non-farm payrolls, industrial production, housing starts, motor vehicle sales, durable good orders, commodity prices employment cost index, and productivity growth. Even if changes in these variables appear consistent and meaningful, it is difficult to predict whether the changes in the economic variables are temporary or if left unchecked will be longer lasting.

If the Fed believes the changes in key economic variables are consistent and potentially longer lasting, they will take measures to influence the availability of credit in the economy which in turn influences aggregate demand and output growth. The Fed's most frequently employed policy tools include open market operations and changes to the discount rate. Less frequently, the Fed will employ changes in bank reserve requirements or verbal persuasion aimed at influencing bank behavior and capital market conditions with respect to the supply of credit to consumer and business borrowers, and even more rarely, the Fed may change margin requirements on stocks.

Through open market operations, the purchase or sale of U.S. government securities, the Fed either adds liquidity or funds into the market or subtracts liquidity or funds from the system. By changing the discount rate, the Fed changes the rate it charges depository institutions for the privilege of borrowing funds at the discount window. In January 2003, the Fed acted to tie the discount rate to the Federal funds rate. For financially sound member banks, the discount rate on primary borrowings at the discount window exceeds the Federal funds rate by 100 basis points. For secondary borrowings by less financially sound banks, the discount rate exceeds the Federal funds rate by 150 basis points. The combination of these tools can either make the cost of funds, that is, interest rates, cheaper or dearer. *Open market operations* work on the principles of supply and demand while changes in the discount rate directly alter the interest charged on funds. Discount rate changes are proposed by the board of directors of one or more district reserve banks for the approval by the board of governors of the Federal Reserve. Open market operations are conducted by the Federal Open Market Committee (FOMC) in a manner consistent with decisions made at the periodic *FOMC meetings*. The FOMC consists of the seven members of the board of governors plus five voting bank presidents.

Two policy tools, changes in the reserve requirements and verbal persuasion, are tools infrequently used to reinforce stated Fed policy aims and they are used to complement policy changes already enacted through open market operations. *Discount rate changes* are more commonly used to put into effect Fed policy aims implemented through open market operations. These tools are employed by the Federal Reserve board of governors to underscore a policy of easing or tightening. Figure 4.1 attempts to simplify the decision making and policy implementation process of the FOMC.

Historically, the Federal Reserve, under different chairmen, has introduced two contrasting techniques for implementing open market operations (see Figure 4.1). Initially, the Fed has used as its operating procedures (guidelines) a rigid *federal fund rate target*, generally in effect from the late 1920s through the late 1970s. More recently, Fed officials have introduced a more flexible Federal funds rate target. When the Fed uses a rigid Federal funds rate target, Fed open market operations tend to have procyclical results. That is, during economic expansions, the Fed's use of a rigid *federal funds rate target*, in the face of increasing money and credit demands, would result in the full accommodation of these demands, thereby triggering an acceleration in money and credit growth, excessive real growth, and the mounting threat of inflation. Conversely, during economic downturns, the Fed's use of a rigid Federal funds rate target, in the face of declining money and credit demands, would result in weakening money and credit growth, slowing real growth, and lessening inflation pressures. Fed chairman William McChesney Martin

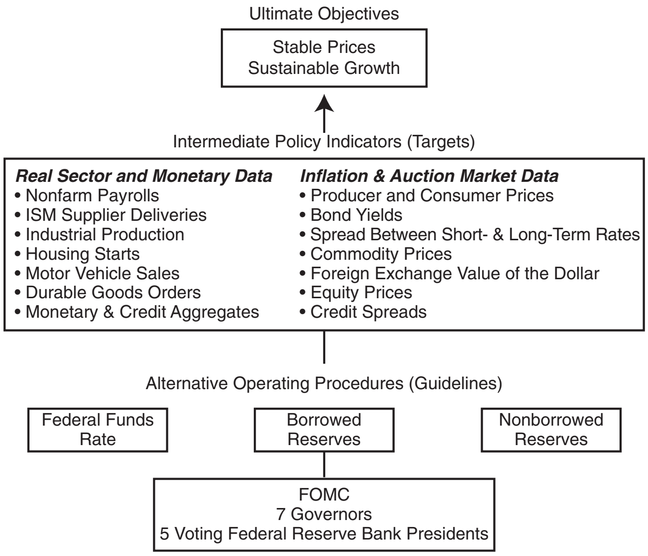


Figure 4.1 Fed Policy Objectives, Intermediate Indicators and Alternative Open Market Operating Procedures

Jr., who was Fed chairman from 1951 to 1970, started the transition to a more flexible Federal funds rate target. He sought to achieve countercyclical effects when he introduced his "*leaning against the wind*" policy approach. Under this approach, if economic growth appears too strong relative to the economy's sustainable potential and consequently, potentially inflationary, the Fed would tighten its policy stance and increase its Federal funds rate target in order to restrain money and credit growth with the aim of slowing aggregate demand and output growth, thereby, lessening inflationary pressures. Conversely, if economic growth weakens, the Fed would "lean" toward an easier policy stance and lower its Federal funds target in order to stimulate economic growth. Under Fed chairmen Paul Volcker (1979–1987) and Alan Greenspan (1987–2006) still greater flexibility was introduced into the Fed's *federal funds rate target* in order to enhance countercyclical policy actions.

Regarding the intermediate policy indicators in Figure 4.1, Fed chairman Volcker tended to place primary policy emphasis on curtailing money and credit growth. In Volcker's own words in a statement before the Joint Economic Committee of the U.S. Congress on June 15, 1982, "[a] basic premise of monetary policy is that inflation cannot persist without excessive monetary growth, and it is our view that appropriately restrained growth of money and credit over the longer run is critical to achieving the ultimate objectives of reasonably stable prices and *sustainable economic growth*." Subsequently, however, Chairman Greenspan found it necessary to lessen the Fed's emphasis on monetary and credit growth in favor of greater policy emphasis on a wider range of intermediate indicators of the real sector, inflation, and auction (financial) markets. Greenspan feared that owing to globalization, securitization, and, most importantly, financial product innovation such as hedge funds, money and credit growth was no longer a reliable predictor of economic activity and inflation. Greenspan also placed more emphasis on transparency and verbal persuasion in seeking to increase the effectiveness of monetary policy.

IMPLEMENTING MONETARY POLICY: THE TRANSMISSION PROCESS

The *monetary policy transmission process* has always been a long and variable one. In the past the banking system, the conduit for monetary policy, was the dominant source of credit for consumers and businesses. Typically, it has taken from six to twelve months for a shift in monetary policy to work its way through the banking system and capital markets to impact aggregate demand and output. It takes even longer for a given policy shift to influence price behavior. Complicating this process in today's world, a declining share of credit is supplied through the banking system and a rising share of credit is supplied through globally integrated capital markets. As a result, the Federal Reserve today more than in the past, must be highly attuned to financial market participants' perceptions of Fed intentions and potential market impact of the Fed's perceived intentions. The banking system remains the point of contact for the Fed when it initiates shifts in policy stance. However, Fed intentions and related market expectations of their intentions remain a critical concern in the transmission of Fed policy shifts. This process results in capital market asset price and interest rate adjustments that ultimately influence changes in aggregate demand, output growth, and inflation.

Fed authorities began in February 1994 to immediately announce policy decisions. (See Figure 4.2 for policy statements following FOMC meetings.) Today's Fed monetary policy transmission process is a transparent one. Monetary policy transparency easily conveys Fed policy intentions. Typically, Fed officials, through speeches, interviews, and congressional testimony, will seek to prepare financial market participants for any policy shift that may be in store in upcoming policy meetings. Clear information on current Fed policy helps the monetary policy transmission process operate more effectively. Under former Fed chairman Greenspan, the Fed sought to be more transparent, and refined its methods of communication.

Historically, the Fed policy transmission process has worked largely by manipulating the cost of credit as supplied by the banking system. Specifically, to effect a policy shift, the Fed has traditionally begun by changing the composition of bank reserves. For example, more Fed restraint means the Fed manipulates a rising share of borrowed to total reserves, resulting in an increase in the cost of reserves. The increased cost of reserves is reflected in a higher Federal funds rate. (The Federal funds rate is the rate on bank reserve balances held at the Fed that are loaned and borrowed among banks, usually overnight.) Conversely, less Fed restraint (more ease) means the Fed manipulates a declining share of *borrowed reserves* to total reserves. This action results in a declining cost of reserves

Friday, February 4, 1994 at 11:05 a.m.—FOMC meeting
"Chairman Alan Greenspan announced today that/the Federal Open Market Committee decided to increase slightly the degree of pressure on reserve positions. The action is expected to be associated with a small increase in short-term money market interest rates.

The decision was taken to move toward a less accommodative stance in monetary policy in order to sustain and enhance the economic expansion.

Chairman Greenspan decided to announce this action immediately so as to avoid any misunderstanding of the committee's purposes, given the fact that this is the first firming of reserve market conditions by the committee since early 1989."

Tuesday, March 22, 1994 at 2:20 p.m.—FOMC meeting
"Chairman Alan Greenspan announced today that the Federal Open Market Committee decided to increase slightly the degree of pressure on reserve positions. This action is expected to be associated with a small increase in short-term money market interest rates."

Monday, April 18, 1994 at 10:06 a.m.—FOMC telephone conference, call
"Chairman Alan Greenspan announced today that the Federal Reserve will increase slightly the degree of pressure on reserve positions. This action is expected to be associated with a small increase in short-term, money market interest rates."

Tuesday, May 17, 1994 at 2:26 p.m.—FOMC meeting
"The Federal Reserve today announced two actions designed to maintain favorable trends in inflation and thereby sustain the economic expansion.

The Board approved an increase in the discount rate from 3% to 3.5%, effective immediately, and the Federal Open Market Committee agreed that this increase should be allowed to show through completely into interest rates in reserve markets.

These actions, combined with the three adjustments initiated earlier this year by the FOMC, substantially remove the degree of monetary accommodation which prevailed throughout 1993. As always, the Federal Reserve will continue to monitor economic and financial developments to judge the appropriate stance of monetary policy.

In taking the discount action, the Board approved requests submitted by the Boards of Directors of eleven Federal Reserve Banks—Boston, New York, Philadelphia, Richmond, Atlanta, Chicago, St. Louis, Minneapolis, Kansas City, Dallas and San Francisco. The discount rate is the interest rate that is charged depository institutions when they borrow from their district Federal Reserve Bank."

Wednesday, July 6, 1994 at 2:18 p.m. —FOMC meeting
"The meeting of the FOMC ended at 12:35 pm and there will be no further announcement."

Figure 4.2 Sampling of the Federal Reserve's Official Statements of FOMC Actions, First Half of 1994

that is reflected in a lower Federal funds rate. Borrowed reserves are those reserves that banks borrow temporarily at the Fed's discount window for purposes of adjusting their reserve positions. Banks traditionally try to avoid borrowing at the Fed discount window. There is a perception that such borrowings are a sign of financial weakness. Banks that are forced to borrow temporarily at the discount window will, generally, first turn to other sources of loanable funds such as Federal funds or repo borrowings. Figure 4.1 describes the importance of bank reserves to Fed policy implementation.

Banks, when faced with greater Fed restraint and a rising cost of loanable funds, find their net interest margins narrowing or their profits declining. In that case, the *yield curve* typically flattens or, when the Fed is tightening aggressively, inverts as short-term rates are pushed above longer term rates. Under these circumstances, banks have less incentive to increase their investments and loans. This results in a decline in the availability of funds and an increase in the cost of bank credit to consumers and businesses. Therefore, consumers and businesses will cut back on their borrowing and spending. This in turn results in a declining rate of increase in real economic growth and eventually, a moderation in inflation pressures. Conversely, a Fed move toward an easier policy posture reduces banks' cost of funds. Banks find their net interest margins widening or profits increasing because the fed funds rate is far more elastic than long-term interest rates and the yield curve will steepen. Banks' incentive to increase the availability and reduce the cost of credit increases. This stimulates consumer and business borrowing and spending, thereby spurring real economic growth and eventually triggering a rise in inflationary pressures. The only exception to our converse case is in the environment of an inverted yield curve such as the U.S. government bond curve in the early 1980s. Despite the Fed's efforts to ease short-term rates in the initial stages, the reduction in the cost of funds to banks may not have a significant impact on potential profit margins if the yield curve is inverted enough. Long-term rates may be too low on a relative basis to short-term rates to make bank or other financial institutions' extensions of long-term credit profitable, at least initially.

To view this monetary transmission process from the investment side, Fed policy shifts set off a chain reaction. For example, in the case of a Fed shift towards a more restrictive policy posture, investors who hold short-term credit market instruments such as Treasury bills or money market mutual funds will find interest rates on their short-term investments moving up to higher and more attractive levels relative to yields on longer-term bonds. Accordingly, investors will shift their investments down the yield curve. They will sell longer-term bonds and place the proceeds in shorter-term money market investments. This process will result in rising longer-term interest rates. Rising longer-term interest rates will, in turn, make the returns on bonds more attractive relative to the returns on stocks. As a result, investors will sell stocks, place the proceeds in bonds, and stock prices will decline. As capital market expectations of future Fed restrictive intentions are formed, these portfolio asset adjustments between money market investments, bonds, and stocks will be hastened and intensified.

THE IMPACT OF MONETARY POLICY: ITS DECLINING DIRECT INFLUENCE

Since the mid-1970s, there has been a sharp decline in the proportion of bank credit to total credit available. The bank share of total credit continues to fall. In mid-1970, it was 55%. By 2006, the banks' share of total credit was reduced to 25%. The main factors contributing to the declining bank share of total credit have been globalization of credit resources, securitization, and financial product innovation. The result has been a rising share of credit extended directly through the capital markets to consumers and businesses. Among the major new nonbank institutional suppliers of credit through the capital markets are mutual funds, hedge funds, pension funds, finance companies, and insurance companies. Currently, with the advent of the information revolution, these nonbank lenders are virtually in as good a position as bank lenders to assess market and credit risks.

Today, the Fed's policy transmission process works increasingly to a greater extent through capital market asset price adjustments and interest rates (that is, bonds, stocks, etc.) than through the availability of funds. As in the past, the Fed initiates a policy shift by changing the composition of bank reserves. As we have previously explained, there is a resulting change in the cost of funds as reflected in a change in the Federal funds rate. The Federal funds rate prompts positively correlated changes in short-term market rates. These changing costs of short-term credit include bank loans made at the prime rate and funds raised in the commercial paper market. The impact on capital market price adjustments work in the following manner: as short-term borrowing costs rise, borrowers find longer-term borrowing rates relatively more attractive. Eventually, corporate bond and fixed-rate mortgage offerings increase, driving up longer-term interest rates. This impact of Fed policy shifts on short-term and long-term market interest rates is magnified as Fed intentions are recognized by capital market participants. The participants form expectations of further Fed tightening (easing) moves, thereby affecting longer-term interest rates. Longer-term interest rates are influenced by the average of expected short-term rates plus a term premium that includes inflation expectations. The effect of capital market participants is reflected in the changing shapes of the yield curve as the Fed funds rate changes.

Rising longer-term interest rates and declining stock prices will increase the cost of capital. Increasing the cost of capital decreases business investment. Also, higher longer-term rates depress housing activity. In addition, declining financial asset prices depress consumer wealth and consumer spending, resulting in a decline in the pace of real economic activity. This process serves to moderate inflationary pressures. Commodity prices are likely to be falling in such an environment. Moreover, increasing interest rates will generally cause the value of the U.S. dollar to appreciate in the foreign exchange markets relative to other currencies. A stronger U.S. dollar will cause a decline in exports and a rise in imports, all other factors being equal. This rising trade deficit also serves to dampen economic activity.

GLOBAL CREDIBILITY: THE CENTRAL BANKER'S RESPONSIBILITY

Important influences on the global financial environment include: market deregulation or regulation, financial innovation, integrated global financial markets, and advanced information processing and communications technology. There is a massive pool of mobile capital that relentlessly seeks out countries where business activity generates the highest possible return for a given amount of risk. In order to compete effectively for capital from global investors, countries must pursue disciplined macroeconomic policies and pro-business microeconomic policies including deregulation and privatization. Countries competing for capital must aim for balanced and sustained noninflationary growth.

A more sobering lesson for modern-day central bankers is their reduced effectiveness in controlling massive global capital flows and related financial asset price bubbles. At times this has been manifested in capital market participants' overly optimistic view of central bankers' abilities and desire to stave off debt defaults. This may be particularly true in the case of staving off sovereign and quasi-sovereign debt where there is a history of central bankers providing meaningful amounts of liquidity. The legacy was underscored in the Mexican financial crisis in 1995. The U.S. government provided amounts up to US$50 billion to the Mexican government, staving off a debt default. The benefits of staving off the Mexican default were not without drawbacks. This lesson can be found in the Asian financial crisis that began in mid-1997. The Asian financial turmoil began in the rapidly growing economy of Thailand and spread to the other Southeastern Asian countries of Malaysia, Indonesia, and the Philippines. These developing countries had benefited from an abundance of foreign liquidity. But the heavy capital inflows eventually resulted in excessive growth, mounting trade deficits, and speculative financial bubbles typically manifested in frenzied local bank-financed speculation in equities and real estate. The currency crisis in these Southeast Asian countries was triggered as escalating trade deficits scared away global money managers, triggering a rapid depreciation in their currencies, with interest rates rising sharply in response. As the bubble burst, real estate and equity prices plummeted. This unforeseen instability posed a major threat to the affected countries' banking systems, as bad debts mounted.

It was not until equity market selling pressures spread to Hong Kong that the rest of the world began to take

Table 4.2 Changing Values in Asian Equities, Bonds, and Currencies

Country	Currency Levels		Equity Index—6-Month Return Local Currency Terms	Fixed Income Yield Benchmark	
	7/07/97	1/07/98	7/07/97–1/07/98	7/07/97	1/07/98
Malaysia	2.53	4.06	−50.87	T10+63	T10+260
Indonesia	2432.00	8000.00	−46.58	T10+118	T10+650
South Korea	883.00	1650.00	−47.7	T10+86	T10+525
Philippines	26.41	45.00	−36.53	T30+221	T30+440
Thailand	28.63	52.88	−41.5	T10+82	T10+500
Hong Kong	7.74	7.74	−35.8	T10+73	T10+160
Japan	112.78	131.73	−23.74	T10−371	T10−364

serious notice. With the return of Hong Kong to Mainland China, the Chinese government kept the Hong Kong dollar pegged to the U.S. dollar as a matter of political principle. Nevertheless, speculators continued to attack the Hong Kong dollar on the assumption that it had to fall in line with other southeastern Asian currencies in order for Hong Kong to remain competitive. In its effort to fight off the speculative attack on the Hong Kong dollar, the Hong Kong Monetary Authority was forced to sharply increase interest rates, thereby weakening the Hong Kong real estate market and threatening Hong Kong banks with mounting bad debts.

Next, the Asian currency crisis spread to the larger South Korean economy, where the heavily indebted financial system was vulnerable, and ultimately to the huge Japanese economy which was still attempting to recover from the bursting of its own 1980s financial bubble. Then, like a rapidly spreading contagion, the Asian currency depreciation and equity market plunge spread to Latin America and even eastern Europe and Russia where previously high-performing debt and equity markets registered extremely disorderly declines, and ultimately to declines in the western European and U.S. stock markets. Table 4.2 illustrates the magnitude of these Asian market declines.

The importance of the Asian financial crisis is that it illustrates the lessening influence that central bankers have on today's globally integrated capital market flows, apart from serving as last-resort lenders of liquidity. The role of last-resort lender, however, should not be minimized. The central bank and supra-led package of loans to Mexico in 1995 staved off a dramatic currency crisis that could have led to a debt default. With the stunning advances in information processing and communications technology, global money managers can move capital around the world at virtually the speed of light. This capital, as already noted, seeks out opportunities offering the highest risk-adjusted returns, but it flees from turmoil. The point is that the increasingly efficient global capital markets are linked more tightly than ever before. Apart from maintaining anti-inflation credibility and serving as lenders of last resort, central bankers, including the U.S. Fed chairman, may in the future have only a marginal influence on these massive global capital market flows and related financial asset price bubbles.

Moreover, and perhaps, since the Asian currency turmoil, it is the stark power of the global capital markets themselves rather than domestic politicians or central bankers that are forcing major financial system changes in the affected countries, including the desirable privatization of public corporations and large scale banking reform. The only means by which governments (or the IMF) can stabilize market forces is to respond by offering larger or more effective financial reform packages than global capital market participants expect. For example, global money managers are demanding that bank reform include provisions for allowing insolvent banks to fail and for the weaker banks to be acquired by healthy domestic or foreign financial institutions. In addition, taxpayer funds, along with deposit insurance, must be used to pay off depositors in failed banks. Also, most importantly, bank reform must make provision for transparency, including full disclosure of bad loans and off-balance-sheet items by banks and securities firms.

Huge, global pools of mobile capital may serve to actually discipline national and global macroeconomics policies. If, for example, any developed or emerging country tries to boost growth through overly stimulative macroeconomic policies that are potentially inflationary for political reasons, its trade deficit will worsen and its currency will depreciate. Global institutional investors and money managers will become fearful of the increased inflationary threat and sell bonds, thus pushing long-term interest rates higher and helping to choke off growth in that developing or emerging country.

Former Fed chairman Greenspan was faced by a "conundrum" when he and his fellow policy makers began a prolonged series of rate firming actions mid-2004. Specifically, despite an impressive series of short-term rate-hikes, long-term rates actually declined, reducing the effects of Fed's firming actions. Eventually, Fed officials concluded that this atypical situation in which Fed policy was becoming less accommodative as the capital markets were becoming more accommodative, reflected a unique combination of low inflation expectations, a global savings glut, heavy global carry trades by hedge funds and other large institutional investors, and currency interventions, especially by Japan and China.

Accordingly, the best that any country can do for its citizens is to create a favorable economic climate for participation in the world economy. There are many important economic building blocks for positive participation in the world economy. These building blocks include deregulation, privatization, free markets, minimal government interference, longer-term productivity-enhancing measures (investment in education, job training, research as well as the implementation of technological innovations, and rewards for savings and investment), and, above all, central bank anti-inflation credibility and consistency. Longer-term price stability creates steady, predictable levels of economic growth. These are the rewards of pursuing a monetary policy that seeks price stability and, thereby, sets the stage for enhanced productivity.

In sum, while central bankers still play a key stabilization role in the effort to ensure that financial conditions are supportive of sustainable economic growth, the ability of central banks to influence massive global flows of mobile capital and related asset price bubbles is diminishing. This raises the spectre of additional currency crises from time to time, not unlike those in Mexico in 1995 and in Southeast Asia in 1997. To be sure, central bankers can help limit the private sector's speculative tendencies by maintaining a high level of anti-inflation credibility. Moreover, central banks can help contain the damage when asset price bubbles break by serving as last-resort lenders of liquidity. But in the final analysis, these central bank influences are marginal compared with today's sheer power of global capital market forces.

SUMMARY

In general, central bankers have the mandate to create an investment environment which results in attractive risk-adjusted return on capital and stable capital flows. The Federal Reserve's policy transmission process influences the expectations of capital markets participants. While the Federal Reserve alters short-term borrowing rates, all lenders of capital adjust their cost of capital in response.

The Federal Reserve's policy shifts are countercyclical and affect the forward economic environment. Evident several months after the fact, the impact of these policy shifts is designed to enhance the economic environment. A positive economic environment attracts stable capital flows which ultimately aid stable, noninflationary economic growth. This is the ultimate goal of a central banker.

REFERENCES

Chang, K. H. (2003). *Appointing Central Bankers: The Politics of Monetary Policy in the United States and the European Monetary Union* (Political Economy of Institutions and Decisions). Cambridge: Cambridge University Press.

Chappell, H. W. Jr., McGregor, R. R., and Vermilyea, T. A. (2005). *Committee Decisions on Monetary Policy: Evidence from Historical Records of the Federal Open Market Committee*. Boston: MIT Press.

Federal Reserve Bank of Minneapolis. (2001). *A Prescription for Monetary Policy: Proceedings from a Seminar Series*. Toronto: Books for Business.

Friedman, M., and Schwartz, A. J. (1971). *Monetary History of the United States, 1867–1960*. Princeton: Princeton University Press.

Gruben, W. (1997). *Exchange Rates, Capital Flows, and Monetary Policy in a Changing World Economy (The Federal Reserve Bank of Dallas).* Germany: Springer.

Greenspan, A. (2004). *Monetary Policy and Uncertainty: Adapting to a Changing Economy. Federal Reserve Bank of Kansas City.* University Press of the Pacific.

Jones, D. M. (1989). *Fed Watching and Interest Rate Projections: A Practical Guide*. New York: New York Institute of Finance.

Jones, D. M. (1992). The only game in town (monetary policy of the Federal Reserve System under Chairman Alan Greenspan). *Mortgage Banking* 53, 1: 35–38.

Jones, D. M. (1996). *The Buck Stops Here: How the Federal Reserve Can Make or Break Your Financial Future*. New York: Prentice Hall Trade.

Jones, D. M. (2002). *Unlocking the Secrets of the Fed: How Monetary Policy Affects the Economy and Your Wealth-Creation Potential*. Hoboken, NJ: Wiley.

Kettl, D. (1986). *Leadership at the Fed*. New Haven, CT: Yale University Press.

Kindleberger, C. P. (1978). *Manics, Panics and Crashes*. New York: John Wiley & Sons.

Meltzer, A. H. (2004). *History of the Federal Reserve, Volume 1: 1913–1951*. Chicago: University of Chicago Press.

Meyer, L. H. (2004). *A Term at the Fed: An Insider's View*. New York: Harper Collins.

Wooley, J. T. (1986). *Monetary Politics: The Federal Reserve and the Politics of Monetary Policy*. London: Cambridge University Press.

CHAPTER 5

Institutional Aspects of the Securities Markets

JAMES R. THOMPSON, PhD
Noah Harding Professor of Statistics, Rice University

EDWARD E. WILLIAMS, PhD
Henry Gardiner Symonds Professor of Management, Rice University

M. CHAPMAN FINDLAY, III
Principal, Findlay, Phillips and Associates

Abstract: Every individual who has more money than required for current consumption is potentially an investor. Whether a person places his or her surplus funds in the bank at a guaranteed rate of interest or speculates by purchasing raw land near a growing metropolis, he or she has made an investment decision. The intelligent investor will seek a rational, consistent approach to personal money management. The best method for some is simply to turn their funds over to someone else for management. A significant number of investors do indeed follow this policy, and it is quite likely the correct decision for many. Others, however, manage their own money or even become professionals who manage other people's funds. The first step is understanding the institutional aspects of the securities markets in order to be a successful participant in them. The various investments media and the environment in which they trade are important elements in this regard.

Keywords: stock market efficiency, efficient market hypothesis (EMH), allocational efficiency, transactional efficiency, organized markets, bid price, ask price, initial public offering (IPO), seasoned equity offering (SEO), primary market, limit order, stop order, Securities Act of 1933, Securities Exchange Act of 1934, margin purchases, stock market indicators

In this chapter, we shall not unveil any mysteries about getting rich quickly. Indeed, if such secrets existed, it is doubtful that the authors would be willing to reveal them. Nevertheless, there are systematic procedures for making investment decisions that can enable the rational investor to maximize his or her economic position given whatever constraints he or she wishes to impose. Economic position is tacitly assumed to be the primary goal of the investor, although there may well be those who have other central goals. The purchaser of an art collection, for example, may be more interested in the aesthetic aspects of his investment than the financial appreciation that might be realized. There is nothing irrational about this, and it is not difficult to construct optimal decision models for such an individual. Similarly, another person may be strongly concerned about pollution, or human rights. Such a person may refuse to buy shares in companies that pollute or that do business in countries where ethnic cleansing is practiced. Again, this can be perfectly reasonable behavior, but these investors should at least be aware of the opportunity costs of their decisions. That is, they should realize that such an investment policy may have economic costs. The polluter may be a very profitable company whose stock could produce exceptional returns for its holders.

Maximizing economic position cannot usually be taken as the only objective of the investor. There is some correlation between the returns one expects from an investment and the amount of risk that must be borne. Thus, decisions must be made that reflect the ability and desire of the individual to assume risk. The literature in finance is very specific in both theoretical and practical terms about risk bearing and the optimal portfolio for the investor.

Although intelligence is about the only important personal attribute requisite for any kind of decision making, there are other traits that may be helpful to the investor. In particular, a certain amount of scientific curiosity may be very important to successful investors. By scientific curiosity we do not mean knowledge or even interest in disciplines generally considered "science," such as biology or chemistry, although the scientifically trained analyst may have an advantage in scrutinizing the stocks of high-technology companies. Rather, scientific curiosity refers to the systematic pursuit of understanding. An investor should be willing to take the time and spend the energy to know him or her and the investing environment.

It is unfortunately true that many otherwise successful people make poor investors simply because they do not have a logical investment policy. They have only vague objectives about what they want (such as "capital appreciation" or "safety of principal"), and they often substitute general impressions for solid fact gathering. How many highly competent doctors, for example, go beyond the recommendations of their brokers (friends, parents, relatives, or drug company salespeople) when selecting a security? How many businesspersons take the time to familiarize themselves with the income statements and balance sheets of the firms in which they hold stock? How many professional portfolio managers make purchases based on a well-researched, documented effort to uncover those securities that others have passed over? Even in the case of portfolio managers, the number may be surprising. Of course, it could be reasoned that the doctor may not have the time or knowledge to make a thorough investigation of his or her investments and that the businessperson is too occupied with his or her own company to do a systematic search for information. If this is the case, then both doctor and businessperson should seek expert counsel.

Although knowledge of what other managers are doing is important and an experienced person's market "feel" may be superior to any professor's theoretical model, too often even the professional tends to substitute rumor and hunch for sound analysis and thorough investigation. In addition to intelligence and scientific curiosity, the modern investor needs to be reasonably versed in mathematics and statistics.

There is any number of investment possibilities that the investor may consider. The simplest is the commercial bank savings account, or certificate of deposit insured by the U.S. government. Next, in terms of simplicity, are the U.S. Treasury bills issued by the federal government in maturities of one year or less. These media provide safety of principal, liquidity, and yields that are not unattractive by historical standards. Nevertheless, they require little analysis as an investment vehicle, and any discussion of them must perforce be brief. There is a place for such investments in the portfolio of most investors, however, and the role of liquidity in investment strategy is a focal point in the portfolio management.

At the other end of the investments spectrum are such highly illiquid assets as real estate, oil well interests, paintings, coins, stamps, antiques, and even ownership of business enterprises. These investments require a very specialized analysis, and anyone who is contemplating the purchase of such assets should be even more careful than the investor in securities.

In between the savings account, certificate of deposit, or U.S. Treasury bill and the illiquid assets mentioned above, are a host of investments that can generally be described as securities. A security is an instrument signifying either ownership or indebtedness that is negotiable and that may or may not be marketable. Securities are by far the most popular form of semiliquid investment (that is, investment that is marketable but that may not be salable near the price at which the asset was purchased), and they can be analyzed in a systematic, consistent fashion.

It was mentioned before that the investor should be well aware of the investment environment before he or she makes a decision. The environment for securities includes such important variables as the general state of the economy, the expected risk and return levels from purchasing a specific security, and the economic position of the investor. It also includes the more specific aspects of securities regulations, financial information flows, the securities markets, and general measures of security price performance (such as the Dow Jones averages). There are entire chapters in this handbook series devoted to each of these topics, and we will not purport to examine any of them in much detail. Nevertheless, the more important elements of these subjects will be discussed later in this chapter.

THE STOCK MARKET EFFICIENCY QUESTION

Of all the forms of securities, common stocks (and derivatives of common stocks) are the most romantic. Although the bond markets are quite important to both issuing corporations and many investors (pension fund money, life insurance reserves, bank portfolio funds, and so on in the aggregate are more heavily invested in bonds than equities), it is the stock market that engenders the interest of most investors. This is undoubtedly true because the rewards (and penalties) of stock market investment well exceed those obtainable in the bond market. Furthermore, equity analysis is more complicated than bond appraisal, and greater skill is required in selecting common stocks than fixed income securities. This is not to say that bond analysis is simple or even uninteresting. Indeed, some of the more sophisticated minds in the investments business are engaged in the bond market. Nevertheless, few people spend their lunch hours talking about the bond market, and the future performance of a bond rarely comes up in bridge table or golf course discussions. It is common stocks that entice most investors, and some investors have been known to feel a greater empathy for their stocks than their spouses. Thus, common stocks (and derivatives of common stocks) will be our focal point in the discussion in this chapter.

There is a school of thought that maintains that only insiders and those privileged to have information not known to the rest of us can make large profits in the stock market. These people subscribe to a theory of stock prices called the *efficient market hypothesis* (*EMH*). EMH advocates argue that the current price of a stock contains all available information possessed by investors and only new information can change equity returns. Because new information becomes available randomly, there should be no reason to expect any systematic movements in stock returns. Advocates of the EMH feel that the stock market is perfectly efficient and the cost of research and investigation would not be justified by any "bargains" (that is, undervalued stocks) found.

The EMH has been tested by a number of scholars. These researchers have considered various hypotheses about the behavior of the stock market, from notions that past stock prices can be used to forecast future prices (the belief held by stock market chartists or "technicians") to reasoned opinions that stocks exist that are undervalued by the market and that these stocks can be uncovered by a thorough investigation of such fundamental variables as reported earnings, sales, price-to-earnings (P/E) multiples, and other pieces of economic or accounting data. The latter view of the market has long been held by most investors, and the whole profession of security analysis depends upon it. From the early days of the first edition of Graham and Dodd (1934) down to the present, analysts have been taught that there are overpriced stocks and underpriced stocks and it is the job of the analyst to determine which are which. The EMH advocates have found, however, that the presence of so many individuals trying to find bargains (and overpriced stocks to sell short) makes it impossible for any one of them to outperform the general market consistently. Thus, as the economy grows and earnings increase, it is possible to make money in the stock market, but it is impossible to expect more than "average" returns. This is true, they say, because there are many buyers and many sellers in the market who have a great deal of similar information about stocks. If any one stock were "worth" more than the price for which it was currently selling, sharp analysts would recommend buying until its price rose to the point at which it was no longer a bargain. Similarly, if a stock were selling for more than its intrinsic value, analysts would recommend selling. The price of the security would fall until it was no longer overpriced.

The EMH has gained great currency in many quarters, particularly among academic economists, beginning in the early 1970s (see Fama, 1970). Nevertheless, it has not convinced too many practitioners, and many financial economists today no longer accept the EMH unequivocally (see Fama, 1996, and especially Haugen, 1999). This may be for two reasons. In the first place, if the EMH were believed, it would be hard for professionals to justify the salaries that they are paid to find better-than-average performers. Second, many analysts have suggested that their very presence is required for the EMH to work. If they could not find undervalued stocks, they would not come to their desks each day; and if they did not appear, there would no longer be that vast army of competitors to make the stock market efficient and competitive! Moreover, many analysts point out that there are substantial differences of opinion over the same information. Thus, although every investor may have available similar information, some see favorable signs where others see unfavorable ones. Furthermore, various analysts can do different things with the same data. Some may be able to forecast future earnings, for example, far more accurately than others simply because they employ a better analytical and more systematic approach. It is these differences of opinion and analytical abilities that make a horse race, and most practitioners (and an increasing number of financial economists) believe that this is what the market is all about.

SOME HISTORY

Long before the EMH began to gain advocates, many economists (and almost all practitioners) believed that the stock market was neither competitive nor efficient (see Williams and Findlay, 1974). These individuals viewed the market historically as an expression of the whim and fancy of the select few, a large gambling casino for the rich, so to speak. It has been observed that securities speculation in the past has been far from scientific and that emotion rather than reason has often guided the path of stock prices. Inefficiency proponents believed that people are governed principally by their emotions and that bull and bear markets are merely reflections of the optimism or pessimism of the day. They argued that economics plays a slight role in the market and that investor psychology is more important. This view traces back over 130 years. Charles Mackay (1869), in a famous book published in 1869, entitled *Memoirs of Extraordinary Popular Delusions and the Madness of Crowds*, argued:

> In reading the history of nations, we find that, like individuals, they have their whims and their peculiarities, their seasons of excitement and recklessness, when they care not what they do. We find that whole communities suddenly fix their minds upon one object, and go mad in its pursuit; that millions of people become

> simultaneously impressed with one delusion, and run after it, till their attention is caught by some new folly more captivating than the first. We see one nation suddenly seized, from its highest to its lowest members, with a fierce desire of military glory; another as suddenly becoming crazed upon a religious scruple; and neither of them recovering its senses until it has shed rivers of blood and sowed a harvest of groans and tears, to be reaped by its posterity. . . . Money, again, has often been a cause of the delusion of multitudes. Sober nations have all at once become desperate gamblers, and risked almost their existence upon the turn of a piece of paper. (pp. vii–viii)

Mackay's fascinating story details some of the most unbelievable financial events in history:

- John Law's Mississippi scheme, which sold shares to the French public in a company that was to have a monopoly of trade in the province of Louisiana. Mississippi shares were eagerly bought up by French investors who knew that this "growth stock" could not help but make them rich. After all, it was common knowledge that Louisiana abounded in precious metals.
- The South Sea Bubble, which induced Englishmen to speculate on a trading monopoly in an area (the South Atlantic) owned by a foreign power (Spain) that had no intention of allowing the English into the area for free trading purposes. The fevers produced by the South Sea spilled over into other "bubbles," one of which proposed to build cannons capable of discharging square and round cannonballs ("guaranteed to revolutionize the art of war") and another that sought share subscribers to "a company for carrying on an undertaking of great advantage, but nobody to know what it is" (Mackay, 1869, p. 53).
- The Tulipomania, which engulfed seventeenth-century Holland. Fortunes were made (and later lost) on the belief that every rich man would wish to possess a fine tulip garden (and many did, for a while at least). Tulip bulb prices reached astronomical levels, as one speculator bought bulbs to sell at higher prices to a second speculator, who purchased to sell at even higher prices to yet another speculator.

In fact, as Mackay was writing, Jay Gould and Jim Fisk were busily manipulating the value of the shares of the Erie Railroad in the United States [see Adams and Adams (1956)]. It was common practice for directors in many companies to issue information causing the price of their firm's stock to rise. They then sold their shares at inflated prices to the unsuspecting public. Some months later, they would release discouraging information about the company's prospects, in the meanwhile selling short the shares of their company. When the new information drove the price of the shares down, the directors would cover their short positions, again reaping nice profits at the expense of the unaware.

These practices continued on into the 1920s, an era when everybody believed that the life style of a J. P. Morgan or a Harvey Firestone could be within his reach. As Frederick Lewis Allen has pointed out in his wonderfully nostalgic yet penetrating Only Yesterday, it was a time when "the abounding confidence engendered by Coolidge Prosperity . . . persuaded the $40,000 a year salesman that in some magical way he too might tomorrow be able to buy a fine house and all the good things of earth"(Allen, 1931, p. 11). A speculative binge started in 1924 with the Florida land boom (where "investors" paid large sums of money for plots that turned out in many cases to be undeveloped swampland) and continued on throughout most of the rest of the decade. As historian David Kennedy has pointed out, "Theory has it that the bond and equity markets reflect and even anticipate the underlying realities of making and marketing goods and services, but by 1928 the American stock markets had slipped the bonds of surly reality. They catapulted into a phantasmagorical realm where the laws of rational economic behavior went unpromulgated and prices had no discernible relation to values. While business activity steadily subsided, stock prices levitated giddily" (Kennedy, 1999, p. 35). All this came to an end with the stock market crash (and Great Depression that followed) in October, 1929.

From the Gilded Age to the 1920s, stock values were based on dividends and book value (that is, net asset value per share). In other words, stocks were valued much like bonds, based on collateral and yield. Since it was hard to "fake" a dividend, manipulators often resorted to "watering the balance sheet" (writing up asset values to unreasonable levels so as to raise the book value). This discussion seems quaint today with the Securities and Exchange Commission (SEC) requiring an accounting change (eliminating pooling on acquisitions), which will result in "watering" balance sheets in much the same way as in days of yore.

Earnings and even earnings growth as a basis for value were touted during the 1920s, especially to justify the ever higher prices. The book often cited for providing intellectual justification for the stock market excesses of the 1920s was Edgar Lawrence Smith's, *Common Stocks as Long Term Investments* (Smith, 1924). This is really an unfair assessment. What Smith did show, going back a century over periods of severe inflation and deflation, was that a diversified portfolio of stocks held over long periods might expect to earn a current yield of the short (e.g., commercial paper) rate and appreciate (from the retention of earnings) at about 2.5%, which can probably be interpreted as a "real" (inflation adjusted) return. At current values, this would translate to a shareholder rational required return of about 10% for the average stock. Not only is that estimate quite reasonable but it also is consistent with studies of long-run equity returns since the 1920s by the firm of Ibbotson Associates. Furthermore, it contrasts with the arguments of the Dow Jones 40,000 crowd who contended, at the beginning of the year 2000 at least, that the average stock should currently be selling at 100 times earnings (see Glassman and Hassett, 1999). The market correction that began in March 2000 has pretty much laid this notion to rest.

With the 1930s came the first edition of Graham and Dodd's classic *Security Analysis*. Graham and Dodd (1934) devoted much of their attention to adjustments to make financial statements more conservative. They would allow conservative P/E multiples on demonstrated, historical earning power, with considerable attention again paid to book value and dividends. Finally, a much neglected book, J. B. Williams's *The Theory of Investment Value*, also appeared (Williams, 1938). In it he developed most of the valuation formulations of financial mathematics. Along the way, he demonstrated that share price could be expressed as the discounted present value of the future dividend stream. Such currently trendy notions as free cash flow analysis and economic value added flow from his work.

Thus, in normal markets, the valuation approaches basically argue that investors should pay for what they can see (net assets, dividends) or reasonably expect to continue (demonstrated earning power). As markets boom, more and more optimistic future scenarios need to be factored into rational valuation models to obtain existing prices. Beyond some point the assumptions become so extreme that explanations other than rational valuation suggest themselves.

Although the EMH was yet decades away, economists in the 1920s and 1930s did advance the notion that markets should conform to some rationality which, with its great rise and collapse, the stock market seemed not to obey. John Maynard Keynes was one of the more perceptive observers of this phenomenon; and in 1936, he wrote *The General Theory of Employment, Interest, and Money* (Keynes, 1936). He believed that much of man's social activities (including the stock market) were better explained by disciplines other than economics. Keynes felt that "animal spirits" exercised greater influence on many economic decisions than the "invisible hand." His reasoning was based partly on a very acute understanding of human nature. It also depended on Keynes' lack of faith in the credibility of many of the inputs that go into economic decision making. He argued, ". . . our existing knowledge does not provide a sufficient basis for a calculated mathematical expectation. In point of fact, all sorts of considerations enter into the market valuation which are in no way relevant to the prospective yield" (Keynes, 1936, p. 152). He argued further that ". . . the assumption of arithmetically equal probabilities based on a state of ignorance leads to absurdities. "

To Keynes, the market was more a battle of wits like a game of Snap, Old Maid, or musical chairs than a serious means of resource allocation. One of the most often quoted metaphors in the *General Theory* tells us:

> . . . [P]rofessional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. (Keynes, 1936, p. 156)

In the stock exchange, pretty girls are replaced by equities that one speculator believes will appeal to other speculators. Thus,

> A conventional valuation which is established as the outcome of the mass psychology of a large number of ignorant individuals is liable to change violently as the result of a sudden fluctuation of opinion due to factors which do not really make much difference to the prospective yield; since there will be no strong roots of conviction to hold it steady. (Keynes, 1936, p. 154)

During this period, economists would discuss efficiency in two ways—neither of which is related to the efficiency in the EMH. *Allocational efficiency* related to transmitting saving into investment and also investment funds to their highest and best use. Such issues arose in the debate over central planning versus market economies. *Transactional efficiency* related to the costs of trading (e.g., commissions and taxes, bid-ask spread, round-trip costs, etc.). Keynes favored high trading costs, both to keep the poor out of the casino and to force a longer-term commitment to investments. He was also not optimistic about the ability of the market to allocate capital well ("when the investment activity of a nation becomes the byproduct of a casino, the job is likely to be ill done.")

Keynes's view was generally adopted by economists at least until the 1950s. A leading post-Keynesian, John Kenneth Galbraith, argued along Keynesian lines in his book *The Great Crash* (Galbraith, 1955) that stock market instability (inefficiency) had been important in economic cycles in the United States since the War between the States and that the 1929 collapse was a major factor leading to the Great Depression of the 1930s. Attitudes about the nature of the stock market began to change in the 1960s, and many financial economists began to interpret stock price movements as a "random walk."

In sum, the 1929 crash and Great Depression were still very recent memories as the United States emerged from World War II. The Securities Acts (see the following discussion) passed during the 1930s were predicated upon the view, rightly or wrongly, that fraud pervaded the markets. Certain financial institutions were either prohibited or generally severely restricted in their ability to hold common shares. A "prudent man's" portfolio would still be mostly (if not all) in bonds. The stock market was, at worst, a disreputable place and, at best, a place where one not only cut the cards but also distrusted the dealer.

Viewed in this context, the EMH becomes one of the most remarkable examples of image rehabilitation in history. From a very humble start, within a decade or two it had converted numerous academics. By the end of its third decade, it had converted a plurality of the U.S. Supreme Court! Whether these people (especially the latter) know it or not, they have adopted the position that an investor, having undertaken no analysis, can place a market buy order for any stock at any time and expect to pay a price which equals the true value of the shares. In other words, they now have so much trust in the dealer that they do not even bother to cut the cards!

THE ROLE OF FINANCIAL INFORMATION IN THE MARKET EFFICIENCY QUESTION

A fundamental postulation of the efficient market hypothesis is that investors (except insiders) have similar information with which to work. Indeed, the entire foundation of the EMH is based on the presumption that all data have been digested by the market and that the current price of a security reflects all available information. Opponents of the EMH believe that information is not perfectly disseminated among investors and that investors may tend to interpret information differently. Because the ability to make above-average returns in the market depends upon differences in the flow and understanding of information, it is very important that the purchaser of securities appreciate all the possible sources of financial data.

A large quantity of information is generated by agencies and services that put together reports for the investing public (institutional and personal). These reports vary from general economic prognostications to very concrete analyses of industry and corporate prospects. Moody's, Standard & Poor's, and FitchRatings supply numerous bulletins and reports on a daily, weekly, and monthly basis. The Value Line Investment Survey publishes reports on hundreds of companies and ranks stocks in terms of quality, potential short-term and long-term price performance, and yield. Brokerage and investment banking houses also publish numerous reports that analyze and evaluate individual companies and securities. The bigger firms maintain substantial staffs of analysts, and it is not unusual for major entities to have analysts who cover the securities in only one industry.

In addition to the services listed above, there are a number of private investment letters that are distributed to a clientele of paid subscribers. These letters cost varying amounts, from a few dollars per year to thousands of dollars per year, depending on the nature of the information provided and the previous track record of the publishers of the letter. Some investment letters have been prepared for years and are widely respected.

An unfortunate feature of the aforementioned reports, however, is that the information that they contain is available to a large audience. Thus, if one of the leading brokerage houses recommends a particular stock, that knowledge is immediately transmitted to all participants and incorporated into stock prices. If only these data existed, there would be a good reason to accept the EMH on *a priori* grounds. There are other pieces of information, however, that may not be easily transmitted or understood. Important data are not always available to the general public, and even when widely disseminated some information is not easily interpreted. Data in the latter category often appear in the financial press and require specialized knowledge for proper understanding. Although many of the articles that are found in such publications as the *Wall Street Journal*, *Barron's*, *Forbes*, *Fortune*, and *BusinessWeek* are very specific and provide obvious input in the appraisal of a firm, frequently it is not easy to make sense of an isolated fact that may be reported about a particular company. For example, suppose a firm reports its third quarter earnings and the statement is immediately carried by the *Wall Street Journal*. Suppose further that reported earnings are up significantly from the previous quarter and from the same period (quarter) in the previous year. What does this mean? The average reader might assume that the report is bullish and that the firm has done well. However, he looks at the price of the firm's stock and finds that it went down with the publication of the information. How can this be explained? One possibility is that professional analysts who had been scrutinizing the company very carefully for years had expected the firm to do even better than the reported figures and were disappointed by the result. Another possibility is that the market had discounted the good news (that is, the improvement was expected and the price of the stock had been previously bid up accordingly). When the news was not quite as good as expected, the market realized that it had overanticipated the result, and the price thus had to fall.

Information of this sort can often bewilder the "small" investor, and even the seasoned analyst is sometimes surprised by the way the market reacts to certain new inputs. Nevertheless, it is situations such as these that make possible above-average stock market performance. The investor who went against the crowd with the belief that the firm was not going to do as well as others expected and sold his shares (or established a "short" position) would have profited. Here, of course, superior forecasting or a better understanding of the situation would have enabled the shrewd investor to realize better-than-average returns. Hence, it is clear that the appropriate evaluation of financial information is the key to long-run investment success, and the trained analyst will have a decided advantage. Because the published investigations of others will very likely be generally available, and hence already included in the current price of a security, it should be obvious that the portfolio manager or private investor who really desires above-normal profits will have to make independent evaluations. The inputs that go into independent appraisals may include publicly available information, such as a corporation's annual report and Form 10K, but these data will be uniquely interpreted. Moreover, the professional analyst may have access to certain inputs that are not generally known to the public. Competent securities analysts spend many hours each week interviewing corporate officers and employees of the firms that they follow. The capable analyst learns how to take clues from what managers say and use these clues to good advantage. This is not quite as good as inside information, but the really top analysts can sometimes deduce facts from evidence garnered that equals that possessed by management. These people are rare, however, and they are very highly paid. Moreover, this source is becoming increasingly limited by the SEC Rule FD, which requires companies to specify certain officers (employees) who can communicate with the investing public and requires immediate public dissemination in the event of inadvertent disclosure of material information.

Computers are used to perform tasks in a matter of seconds that previously required thousands of man hours, and a whole new industry (financial information technology) has come into existence that has offered its own stock market high fliers in recent years. Significant amounts of data are now "online," including "First Call" earnings estimates of analysts, "ProVestor Plus" company reports, Standard & Poor's Stock Reports, Vickers "Insider Trading Chronology," and virtually all press releases made by every publicly-held company. SEC filings are immediately available from EDGAR (electronic data gathering and research) On Line (edgaronline.com), which also provides a "People Search" with information on directors and officers, including salary and stock option data. Other similar information is provided by Yahoo! (finance.yahoo.com). Other useful web sites include the following: smartmoney.com; premierinvestor.com; moneycentral.msn.com; fool.com; CBS.marketwatch.com; bloomberg.com; Kiplinger.com; bigcharts.com; and Morningstar.com. Also, most publicly held companies maintain web sites that post all major information about the company. Of course, one should not overlook the various "chat" room sources (which are mostly "gripe" sessions from disgruntled employees—some stockholders as well), the information content of which have only recently become the subject of academic research.

THE ROLE OF ORGANIZED MARKETS IN THE MARKET EFFICIENCY QUESTION

There is another prerequisite even more important than widely available information to the efficient market hypothesis: the existence of large, well-behaved securities markets. From economics, it will be recalled that a perfectly competitive market is one in which: (1) there are many buyers and sellers, no one of which can influence price; (2) there exists perfect information that is costless and equally available to all participants in the market; (3) there is a homogeneous commodity traded in the market; (4) there is free entry of sellers and buyers into the market (no barriers to entry); and (5) there is an absence of taxes and other transaction costs (e.g., brokerage commissions). The price observed in such a perfect market would be not only efficient but also in equilibrium. Clearly, however, no market meets this sufficient condition for the EMH. Advocates of the EMH would contend that the price could behave "as though" it were efficient if none of the above conditions were even approximated. (For a further discussion, see Chapter 5 in Thompson, Williams, and Findlay [2003].)

The discussion earlier on the role of financial information in the market efficiency question, as well as the discussion later on the role of securities market regulation in the market efficiency question, deals with the descriptive validity of the information availability assumption. This section and the next deal with the question of which "price" is assumed to be efficient. For example, if a stock is quoted \$5 bid, \$10 asked, what is the price which supposedly unbiasedly estimates the fully informed value? Likewise, if uninformed trading pushes the price up or down from the last trade, which is the right price? Finally, later in this chapter our attention turns to trying to identify the market return that one cannot expect to beat.

For a market to be competitive, it usually must be sufficiently large so that any buyer (or seller) can purchase (sell) whatever quantity he or she wishes at the "going" price (that is, the price set through the negotiations of all buyers and sellers together). The securities markets generally satisfy the size requirement in that literally billions of dollars worth of stocks and bonds are traded daily just in the United States. This does not mean that there is a good market for every single stock or bond in the hands of the public, however. If there is sufficient trading depth in a particular security, it will be possible to trade at a price very near the most recent past transaction price. As many investors can testify, however, there are numerous stocks (and many more bonds) that trade in very thin markets. That is, the number of participants in the market is so small that one may have to bid well above the last price to buy or ask well under that price to sell. Such a market is clearly not even nearly perfect.

To a large extent, whether or not a particular stock or bond is traded in a broad market depends on the "floating supply" of the issue outstanding. A stock with only a million shares in the hands of the public that "turns over" only 10% of the supply annually (that is, 100,000 shares annually, or less than 500 shares on average each business day) will probably not trade often enough or in sufficient quantity to have market depth. Such a security may show rather substantial price volatility from one transaction to the next since the time span between transactions may be several hours or even days. Thus, no one buyer could accumulate more than a few shares at any one time without driving the price higher. Similarly, no seller could liquidate much of a position without pushing the price down.

One way to be sure that a stock does have a reasonably large floating supply and regular transaction patterns is to make certain that it is traded in an organized market, and organized securities markets have existed for centuries. Records show that securities were trading as early as 1602 in Antwerp and that an organized exchange existed in Amsterdam by 1611. Today, most of the leading capitalist countries have at least one major exchange, and the United States boasts several. In North America, the New York Stock Exchange (NYSE) is still the most important market, but the Nasdaq (a computer-based automatic quotation service) list competes with many newer, "high-tech" stocks. Securities listed on the American Stock Exchange and the regional exchanges do not generally have the market depth (or the earnings and assets) of stocks listed on the NYSE. Options on common stocks are traded on the Chicago Board Options Exchange (CBOE) and on the American, Pacific, and Philadelphia stock exchanges.

Many securities are traded on more than one exchange. An advantage of dual listing is that extra trading hours may be secured for a firm's stock. Thus, a company's shares listed on the NYSE and the London and Tokyo Stock Exchanges could be traded almost around the clock. In fact, there is almost continuous trading in most major stocks held worldwide even when an organized exchange is not open. A disadvantage of this extended trading may be increased volatility. (For a further discussion, see Chapter 5 in Thompson, Williams, and Findlay [2003].)

On an exchange, there is a record of each transaction, and an investor can observe the last price at which the security was traded. The investor may call his or her broker (or check an online trading account) and find out at what price a particular stock opened (its first price for the day), find the high and low for the day, and obtain a current quotation with market size. A *bid price* is the amount that is offered for the purchase of a security and the *ask price* is the amount demanded by a seller. There will customarily be a "spread" between the bid and ask prices that serves to compensate those who make a market in the security (a specialist on most exchanges and various dealers on the Nasdaq).

The primary market is the first sale or new issue market. When a firm (or government) sells its securities to the public, it is a primary transaction. The first public sale of stock by a firm is the *initial public offering* (*IPO*); a subsequent sale by the firm is called a *seasoned equity offering* (SEO). After a bond or share is in the hands of the public, any trading in the security is said to be in the secondary market. Purchases of securities in the primary markets are usually made through investment bankers who originate, underwrite, and sell new issues to the public. In the usual case, a firm will arrange to "float" an issue through its bankers. The firm will be given a price for its securities that reflects market conditions, yields on equivalent securities (either the company's or those of similar concerns), and the costs involved to the investment bankers to distribute the stocks or bonds. Title to the securities is customarily taken by the underwriting syndicate (several banking houses), although small corporations usually have to settle for a best efforts distribution in which the bankers merely act as selling agents for the company.

The *primary market* (new issue market) has been quite popular for speculative investors. The reason for this popularity is the fantastic price movements experienced by many stocks after initial sale (IPOs). In the late 1990s, it was not unusual for newly public stocks to double or even quadruple in the first day of trading. Some "high tech" stocks went up by a factor of ten or more within days or weeks of their IPO. It was no wonder that just the mention of a new issue was often enough to get investors clamoring for "a piece of the action." It is interesting to note that this is not a new phenomena, and nearly all bull (rising price) markets for decades (actually centuries) have featured startling performers that rose to unbelievable levels even though these were brand new (or, in any case, not very seasoned) companies. For a while in 1999 and early 2000, just about any company with "dot com" or "e" or "i" in its name seemed to be able to go public and have the stock price skyrocket within hours or, at most, days. Companies that never made money (and some that had never made a sale!) were accorded *market capitalizations* (number of shares outstanding times price per share) that often exceed those of old-line companies that have been in business for decades.

One "tech" stock that played the game with a vengeance was Aether Systems which provides "wireless data services and software enabling people to use handheld devices for mobile data communications and realtime transactions" (Aether Systems, Inc. Form 10K for 1999, p. 2). Aether went public at $16 per share in October 1999. On March 9, 2000, the stock closed at $315! During the "tech crash" in April, 2000, the stock fell to $65. It rebounded in only a few weeks to well over $100, but subsequently fell to below $5. History suggests that these "highfliers" would eventually collapse in price (with many going bankrupt) when more sober market conditions reappear (as they always must, and, post 2000, did). Many have offered these examples as prima facie evidence of market inefficiency. Peter Bernstein provides numerous other historical examples of people paying ridiculous prices for "growth" in his book *Against the Gods: The Remarkable Story of Risk* (Bernstein, 1996, pp. 108–109).

THE ROLE OF TRADING IN THE MARKET EFFICIENCY QUESTION

In economic theory, traders play an important role in effecting a market equilibrating process. That is, if the price of A is "low" relative to its "real value" and the price of B is "high" relative to its real value, traders will buy A and sell B until "equilibrium" values are established. Of course, there must be some common agreement on just what "real value" means and how such is determined, and the economics literature has searched for this answer for over 200 years. Suppose Sam can buy a pound of chocolate on Fifth Street for $5 and sell it on Sixth Street for $6. He would be advised to buy all the chocolate he could on Fifth, run over to Sixth and sell all he could. Now the very process of Sam's engaging in this activity causes the price on Fifth Street to rise (excess demand) and the price on Sixth Street to fall (excess supply). Indeed, in theory at least, others besides Sam should enter this market, and the price of chocolate should eventually settle (say at $5.50) unless there were transactions costs involved in moving chocolate from Fifth Street to Sixth Street (say, Fifth Street is in New York and Sixth Street is in Houston). Economists have debated at length on how long this process should take (in theory, quite rapidly), and under what conditions others would join Sam in this endeavor. Suppose Sam is the only one who knows chocolate can be bought on Fifth for $5 and sold on Sixth for $6 with virtually no transactions costs. The existence of imperfect knowledge may provide Sam with quite an opportunity. Of course, information then takes on its own value. Issues such as "Why does Sam know about this opportunity?" and "Why don't others?" come into play. Also, suppose it takes equipment to move chocolate from Fifth to Sixth. The requirement of having capital investment may create a barrier to entry (and impute an "opportunity cost" for the alternative use of the equipment) which may prevent others from joining the market. Economists are generally suspicious of "free lunches" and usually search for reasons why the world looks the way it does. Suppose the chocolate on Fifth Street is actually inferior to that for sale on Sixth Street This may well explain the price difference, and it may mean that Sam does not have such a good opportunity after all. Thus, the existence of fairly homogeneous products may be required for this arbitrage opportunity to really exist.

Now let us return from chocolate to stock. In order to evaluate whether opportunities may exist, one should know something about just how trading takes place. Just as Sam had to know how to find Fifth Street and Sixth Street and buy and sell chocolate and judge the quality of chocolate, so must the intelligent investor know about how the stock market functions and who the players are. We have already established some of this above, but it would be wise to identify a few more important elements. First, a securities dealer maintains an inventory of securities in which he or she makes a market by purchasing and selling from his or her own account as a principal. A broker acts as an agent for his or her customers in purchasing or selling securities. On the floor of an exchange this is done through a specialist (a trader charged by the exchange with maintaining a market in the stock). A broker may also act by buying from and selling to dealers. As an agent, the broker is expected to obtain the best available price for the customer and is paid a commission for this service.

The simplest order to a broker to buy or sell a security at the best available price is called a market order. In the dealer market (sometimes called the over-the-counter, or OTC, market), the broker would check dealer quotes and execute the order at the best available price. On the floor of the NYSE, the floor broker for the customer's firm would walk (or transmit electronically) to the post where the security is traded. From the assembled group of other floor brokers, floor traders (who trade for their own account), and the specialist, the customer's broker would determine the best available price and execute the order. There are, however, other types of orders. A *limit order* includes a maximum (minimum) price that the customer is willing to pay (receive) to buy (sell) the stock. A *stop order* contains a price above (below) the current market price of the stock that, if reached by the market, the customer desires to trigger a buy (sell) market order. Since it is quite likely that neither stop nor limit orders could be executed immediately, the floor broker would instruct the specialist to enter them in his "book" for execution when their terms were met. A stop limit order performs like a stop order with one major exception. Once the order is activated (by the stock trading at or "through" the stop price), it does not become a market order. Instead, it becomes a limit order with a limit price equal to the former stop price. For example, Smith places a stop limit order to sell stock with a stop price of $50 a share. As with the stop order, once the stock trades at $50, the order is triggered. However, the broker cannot sell it below $50 a share no matter what happens. The advantage of this order is that the buyer sets a minimum price at which the order can be filled. The disadvantage is that the buyer's order may not be filled in certain fast market conditions. In this case, if the stock keeps moving down, Smith will keep losing money.

After a stock split or a dividend payout, the price on all buy limit orders and the stop price on sell stop and sell stop limit orders is adjusted. For example, if Jones places an order to buy 100 shares of XYZ Corporation at $100 a share and the stock splits 2 for 1, the order will automatically be adjusted to show that he wants to buy 200 shares of XYZ at $50, reflecting the split. Other restricted orders include the following:

- Good-until-canceled (GTC) orders remain in effect until they are filled, canceled, or until the last day of the month following their placement. For example, a GTC order placed on March 12th, left unfillled and uncancelled, would be canceled automatically on April 30th.
- A day order is a limit order that will expire if it is not filled by the end of the trading day. If one wants the same order the next day, it must be placed again.
- "All or none" is an optional instruction that indicates that one does not wish to complete just a portion of a trade if all of the shares are not available.
- "Do not reduce" means that the order price should not be adjusted in the case of a stock split or a dividend payout.
- "Fill or kill" is an instruction to either fill the entire order at the limit price given or better or cancel it.

The major functions of the specialist (mentioned above) are to execute the orders in his book and to buy and sell for his or her own account in order to maintain an orderly market. To limit possible abuses, the specialist is required to give the orders in the book priority over trades for his or her own account and to engage in the latter in a stabilizing manner. The larger blocks of stock being traded on the exchanges in recent years have caused the capital requirements for specialists to be increased, and rule violations (such as destabilizing trading) are investigated. Even the EMH advocates, however, agree that the specialist's book constitutes "inside information "and this group can earn above-normal profits.

In the past, NYSE designated commissions were charged by member firms on a 100 share (round lot) basis. No discounts were given. Since May 1, 1975 (known as "May Day" to many retail brokers), discounting has been allowed in a deregulated environment. For large transactions, as much as 75% of a commission might be discounted by the larger retail brokers (such as Merrill Lynch). Even larger discounts are now provided by firms that call themselves discount brokers (such as Charles Schwab), and some deep discount brokers are charging as little as $5 per transaction for market order trades done over the Internet. Some are even free if a large enough balance is kept in the brokerage account; and for larger individual accounts (e.g., $100,000), many brokers are now allowing unlimited free trading for an annual fee of 1% to 1.5% of the portfolio value.

Reduced revenues resulting from negotiated commissions coupled with the higher costs of doing business altered the structure of the brokerage industry. A number of less efficient houses collapsed, were merged with stronger concerns, or undertook voluntary liquidation. During the 1970s, several large houses failed and millions of dollars in customer accounts were jeopardized. In order to prevent loss of investor confidence, the Securities Investor Protection Corporation (SIPC) was established to protect the customers of SIPC member firms. The SIPC provides $500,000 ($100,000 cash) protection per account in the event of failure. This arrangement does not protect against trading losses but rather in the event of failure of brokers to satisfy their agency responsibilities. Suppose Mr. X is a customer of ABC & Co. which is a member of the SIPC. Suppose further that X keeps his stocks in custody with ABC (that is, "street name") and the firm goes bankrupt. X would be able to look to the SIPC to make good on the value of his investments up to $500,000. Many brokerage firms purchase insurance to provide protection above $500,000 and the capital requirement for firms has been increased. Thus, a more concentrated, hopefully stronger, and more efficient brokerage community has emerged over the past three decades. This has made it possible for the transactions costs to be reduced tremendously; and this, in turn, should have made markets relatively more efficient than they were, say, 30 years ago.

A caveat should be noted here: Substantially lower commissions and Internet trading have inevitably led to the phenomenon of the undercapitalized "day trader." These are individuals who may have as little as $5,000 who speculate on small price movements in a particular stock within a single day's trading. This phenomenon was not possible when commissions were large, but with $5 trades almost anyone with a computer and Internet access can play the game. The evidence suggests that most of these traders get wiped out, or suffer large percentage losses to their portfolios, within months of initiation of trading. Thus, even in the biggest bull market in history (ending in the year 2000), there were traders who lost most of their money by treating the stock market like a computerized Las Vegas. Interestingly, there are economists who contend that this added "liquidity" has actually made the markets more efficient! (See Malkiel [1999] for an advocate of such reasoning.)

A final note on the mechanics of securities trading: From the beginning of the NYSE in the late eighteenth century, stocks (originally U.S. government securities) were traded in eighths, quarters and halves. Some stock even traded in fraction of eighths, but the basic trading unit was the $^1/_8$, which was one-eighth of a dollar or $0.125. This peculiarity was a result of having the old Spanish dollar (which was divided into eighths and thus called "pieces of eight") being a major currency during the U.S. colonial era. After the decimal U.S. currency system was effected, securities continued to trade in eighths, first as a matter of convenience and later because it increased bid/ask spreads where dealers make most of their money (buying at 11^7/_8$ and selling at $12 is much more profitable than buying at $11.99 and selling at $12). In late 1999 and early 2000, there was a movement initiated by the SEC and adopted by the stock exchanges (and the Nasdaq) to change trading to decimal units. Thus, we no longer buy (and sell) stocks at prices such as 16^3/_8$ or 30^1/_8$; rather, we may buy or sell at the more sensible $16.37 (or $16.38) or $30.12 (or $30.13). This may not seem like a big change, but has made the arithmetic of trading much simpler. (Quick Mental Check: XYZ goes from 10^5/_{64}$ to 10^{13}/_{32}$. What is your profit? How much easier is it to calculate an advance from $10.08 to $10.41!) Also, the greater competition has reduced spreads such that dealer margins have been reduced in favor of investors.

THE ROLE OF SECURITIES MARKET REGULATION IN THE MARKET EFFICIENCY QUESTION

Many people feel that a major element contributing to the efficiency of the U.S. securities markets is the regulation of those markets by the federal government. Before 1933, there were no laws governing stock-exchange or investment-house activities, and widespread manipulation and questionable practices abounded. Corporations were not required to provide information to investors, and fraudulent statements (or no statements at all) were issued by any number of companies. As securities speculator Joseph P. Kennedy (father of future President John F. Kennedy) once remarked to one of his partners, "It's easy to make money in this market. . . . We'd better get in before they pass a law against it" (Kennedy, 1999, p. 367). The excesses of the 1920s were attributed in part to the lack of comprehensive legislation regulating the securities industry, and with the coming of the New Deal a number of laws were indeed passed to prevent a recurrence of the events that led to the 1929 crash.

The *Securities Act of 1933* (the '33 Act) requires full and complete disclosure of all important information about a firm that plans to sell securities in interstate commerce. Issues of securities exceeding certain dollar limits, and all issues sold in more than one state, must be registered. A prospectus must be prepared by the issuing company and distributed to anyone who is solicited to buy the securities in question. The prospectus must include all pertinent facts about the company, such as recent financial reports, current position, a statement about what will be done with the funds raised, and a history of the company. Details about the officers and directors of the company are also required.

The *Securities Exchange Act of 1934* (the '34 Act) established the Securities and Exchange Commission (the "SEC" or the "Commission"). It also regulates the securities markets and institutional participants in the market, such as brokers and dealers. All exchanges are required to register with the Commission, although much of the supervision of individual exchanges is left up to the governing bodies of each exchange. Amendments to the act (e.g., the Maloney Act of 1938) now also include the OTC markets, although broker-dealer associations are accorded the same self-regulatory authority as the exchanges enjoy (see discussion of self-regulation below.) The '34 Act also calls for the continual reporting of financial information by firms that have "gone public" and a major part of the financial (and other) information flow from reporting companies to the public is done pursuant to this act and amendments to it. Interestingly, President Franklin Roosevelt appointed Joseph P. Kennedy (the speculator mentioned above) to be the first Chairman of the SEC. ("a choice often compared to putting the fox in the henhouse or setting a thief to catch a thief " (Kennedy, 1999, p 367)!

The Investment Company Act of 1940 regulates the management and disclosure policies of companies that invest in the securities of other firms. Under the act, investment companies may be organized as unit trusts, face-amount certificate companies, or management investment companies. Only the last classification is currently significant, and it is further subdivided into open-end and closed-end management investment companies. Closed-end companies sell a fixed number of shares, and these shares then trade (often at a discount to net asset value) just like other shares. Many closed-end funds are listed on the NYSE. Some even issue preferred stock (or income shares) and borrow money. Open-end companies are better known as mutual funds and are required to redeem their shares at net asset value upon demand; because of this requirement, they may not issue long-term debt or preferred stock. If a fund registers with the SEC, agrees to abide by the above rules, invests no more than 5% of its assets in the securities of any one issuer, holds no more than 10% of the securities of any issuer, pays at least 90% of its income out to fund shareholders, and meets other rules, it may pay taxes only on earnings retained. Capital gains and dividends (interest) earned are paid out to the holders of the fund's shares who pay taxes at their respective individual or institutional rates.

Other acts that are important include the Public Utility Holding Company Act of 1935, which regulates the operations and financial structure of gas and electric holding companies; the Trust Indenture Act of 1939, which requires that bonds (and similar forms of indebtedness) be issued under an indenture that specifies the obligations of the issuer to the lender and that names an independent trustee to look after the interests of lenders; and the Investment Advisors Act of 1940, which requires the registration of investment counselors and others who propose to advise the public on securities investment.

One of the major excesses of the pre-1933 era was the practice of buying stocks on low margin. *Margin purchases* are those for which the investor does not advance the full value of the securities that he or she buys. Thus, if an investor bought one hundred shares of General Motors at 60 on 50% margin, he would only put up $3,000. The remaining $3,000 would be borrowed either from his broker or a bank. Margin purchases may increase the rate of return earned by an investor. If General Motors went up 10% to 66, the investor in the above example would have made $600/$3,000 = 20%. They also increase the degree of risk exposure. If GM went down 10%, the loss would be 20%. During the late 1920s, investors were buying stocks on less than 10% margin. Large rates of return were earned by everyone so long as stock prices were advancing. When prices began to skid in late 1929, however, many people were wiped out in a matter of days. Investors tended to build their margin positions as prices rose by buying more shares with the profits earned. Thus, a man might have put $1,000 into stock worth $10,000 in January 1929. As prices advanced by 10%, say, in February, he might have used his profit to buy another $9,000 worth of stock. His commitment was still $1,000, but he controlled $20,000 in stock. As prices rose further, he might have increased his position to $25,000. But suppose prices fell just a little, say 4%. This decline would be enough to wipe out his investment completely! Such a decline began during October 1929, and many investors were sold out of their stocks as prices fell. The process of liquidating shares as prices went below margin levels caused further price deterioration that, in turn, forced more liquidations. The snowballing effects of this phenomenon produced the major crash of October 29, 1929, and contributed to the subsequent collapse of both the stock market and the American economy.

Because of the problems directly traceable to margin purchases, the Securities Exchange Act of 1934 gave the board of governors of the Federal Reserve System the power to set margin requirements for all stocks and bonds. Since 1934, margins have been allowed as low as 40% but have also been as high as 100% (no borrowing permitted). To some extent, the sobering experiences of 1929 caused a natural reaction against margin purchases in subsequent years. Nevertheless, most participants in the market today have only read about 1929 and would, if given the chance, follow their forefathers down the same speculative path. To protect them and society from such excesses, extremely low margins are no longer permitted.

Another practice that caused problems prior to 1933 was the short sale. When one sells a security he or she does not own but borrows from someone else to make delivery, he or she is said to sell that security short. The device has been used for years and can be defended on economic grounds even though it does sound a bit immoral to be selling something one does not own. In practice, the short sale is consummated by specialists (who have responsibility for maintaining an orderly market on the NYSE) and dealers far more frequently than by the investing public. The average investor might consider selling short a security if she believed its price was going to decline. She would simply call her broker and ask to sell so many shares of such and such company short at a given price. If the broker could find the shares for the customer to borrow (usually from securities held by the broker in his own account or securities margined with the broker), the transaction could be effected. Because short selling can tend to exacerbate downward movements in stock prices, it is easy to see how speculative excesses could occur through unregulated use of the device. Thus, the Securities Exchange Act of 1934 allows the SEC to set rules for short selling. There are several regulations in effect now governing the practice, the most important being the "uptick" requirement. This rule prevents a short sale while a stock is falling in price. Thus, if a stock sells at $40, then $39.50, then $39, no short sale could be effected until there is an advance above $39.

Since the average securities firm functions as investment banker (representing the issuing firm), broker (representing the customer), and dealer (representing itself) simultaneously, the potential for conflict of interest is great. Many of the laws previously discussed were passed to protect the general public when such conflicts arise. These laws, in turn, provide for substantial self-regulation. This is manifested by exchange regulations for member firms and NASD (National Association of Securities Dealers) rules for most others, who subject themselves to such rules in order to obtain securities from other NASD firms at less than the price to the general public. NYSE members must restrict all transactions in listed stocks to the floor of the exchange, even though the larger firms could merely match buy and sell orders in their own offices. NASD firms may only trade with their customers if their price is the best obtainable and must reveal if they acted as principal on the other side of the transaction. Any research recommendations by broker-dealers must indicate if the firm holds a position in the stock. Other regulations call for ethical behavior on the part of members by prohibiting such practices as: (1) the spreading of rumors; (2) the recommending of securities clearly inappropriate for a given customer; and (3) the encouraging of excessive numbers of transactions (called "churning") in a given account. Although many of these rules have protected the public, others are clearly designed to protect the economic position of the broker-dealer community itself.

The Securities Exchange Act of 1934 defines officers, directors, and holders of more than 5% of the shares of a firm as insiders. Such persons are required to file a statement of their holdings of the firm's stock and any changes in such holdings (within a month) with the SEC. Profits made by insiders on shares held less than six months must also be reported and may be legally recovered by the firm (through a shareholders' derivative suit if necessary); in addition, malfeasance suits could be filed for other injuries to shareholder interests. Over the years, holdings of related persons have come to be included in the determination of whether the 5% rule were met and persons related to insiders were also considered to be insiders for the purpose of the law. In general, the principle was established that insiders and their relatives should not gain a special benefit over other shareholders by virtue of the information about the firm they possess. The Texas Gulf Sulphur case of the mid-1960s, in which corporate insiders withheld information about a minerals discovery until they could obtain stock, clearly reestablished this point through both civil and criminal action.

Several other cases expanded the concept of insider information. In the cases of Douglas Aircraft and Penn Central in the 1970s, brokerage houses obtained insider information (of bad earnings and impending bankruptcy, respectively) and informed selected institutional investors before the general public. Subsequent suits and exchange disciplinary actions against the brokerage houses involved, and suits against the institutions, indicate that second- and third-hand possessors of inside information may also be classed as insiders. A securities analyst was charged in the Equity Funding case some years ago for providing information to selected investors that did not even originate from the company itself but rather from former employees. We clearly have moved in the direction of classifying insiders more on the basis of the information they possess than the position they hold in regard to the firm. Although such a situation would tend to validate the EMH by default, its long-run implications for investigative research analysis and individualistic-portfolio management are not encouraging.

THE ROLE OF STOCK MARKET INDICATORS IN THE MARKET EFFICIENCY QUESTION

When the EMH postulates that only "normal" returns can be earned in the stock market, an implicit assumption is made that there is some sort of average that summarizes stock market performance in general. In fact, it is extremely difficult to calculate measures of this sort. Perhaps the most widely used average is the Dow Jones Industrial Average (DJIA) that appears in the *Wall Street Journal* each day. The DJIA is computed by taking the price of each of 30 selected blue-chip stocks, adding them, and dividing by a divisor. The divisor initially was the number of stocks in the average (originally 12), but because of the obvious biases of stock splits (a two-for-one split would tend to cause the price of a share to fall by one-half), the divisor was adjusted downward for each split. The divisor now is well below one, which, in reality, makes it a multiplier. In addition to the DJIA, there is a Dow Jones Transportation (formerly rail) Average of 20 stocks, a Dow Jones Utility Average of 15 stocks, and a composite average of all 65 stocks.

Dow Jones also calculates market indices for a number of foreign markets, an Asia/Pacific Index and two World Indices (one with U.S. stocks and another without). Each is computed in the same manner as the DJIA. For many investors, the Dow Jones averages are the market. When an investor calls his broker to ask what the market is doing, he is very likely to get a response such as "down 56.75." The broker means, of course, that the DJIA is down 56.75 points. This information may have very little to do with what the investor really wants to know (that is, how are my stocks doing?). The DJIA is not an overall indicator of market performance, although many use it as if it were. In fact, only blue-chip stocks are included in the average. The thousands of other stocks that are not blue chips are not represented. Moreover, the DJIA has been criticized by many even as a measure of blue-chip performance. Because the DJIA merely adds the prices of all included stocks before applying the divisor, a stock that sells for a higher price receives a larger weight in the measurement.

The difficulties associated with the Dow Jones averages have led to the development of other stock price averages and indexes. Standard & Poor's computes an industrial index, a transportation index, a utility index, and a composite index (500 stocks) that include both the price per share of each security and the number of shares outstanding. These figures thus reflect the total market value of all the stocks in each index. The aggregate number is expressed as a percentage of the average value existing during 1941 to 1943, and the percentage is divided by 10. The S&P indexes are better overall measures of stock market performance than the Dow Jones averages because more securities are included. Furthermore, the statistical computation of the S&P indexes is superior to the Dow Jones method.

There are a number of other indexes that are also prepared. Both the New York and American Stock Exchanges compute measures that include all their respective stocks. The NYSE Common Stock Index multiplies the market value of each NYSE common stock by the number of shares listed in that issue. The summation of this computation is indexed, given the summation value as of December 31, 1965. The American Stock Exchange average simply adds all positive and negative changes of all shares and divides by the number of shares listed. The result is added to or subtracted from the previous close. The NASD, with its automated quotation service, computes a composite index based on the market value of over 5,000

OTC stocks plus indices for six categories representing industrials, banks, insurance, other finance, transportation, and utilities. The broadest index is calculated by Wilshire Associates. Their Wilshire 5000 Equity Index is based on all stocks listed on the New York and American Stock Exchanges plus the most actively traded OTC stocks. Although there is no single perfect indicator of average performance, many analysts are tending to favor the Wilshire Index as the most broadly indicative. Nevertheless, most observers do not ignore the Dow Jones averages because so many investors are influenced by them. Fortunately (at least for measurement purposes), there is a high positive correlation between the price movements of all stocks. Thus, if most stocks are going up (or down), almost any stock price measure will indicate this.

SUMMARY

It is important to understand the institutional aspects of the securities markets in order to be a successful participant in them. The various investments media and the environment in which they trade are important elements in this regard. Of all the forms of securities, common stock and derivatives of common stocks are the most romantic but are also the most difficult to analyze.

There is a school of thought that maintains that the current price of a stock contains all available information about that stock and only new information can change equity returns. This theory of stock market behavior is called the efficient market hypothesis (EMH). The EMH has been tested over the years by a number of scholars. It was generally endorsed by financial economists in the 1970s, but most practitioners never accepted the theory. Today, even many financial economists no longer accept the EMH, at least without qualification.

A crucial assumption of the efficient market hypothesis is that stock prices reflect all public information. A major form of such information is that supplied by research agencies and services. Reports periodically prepared by brokers and investment advisory companies are designed to aid the investing public in making decisions. An important feature of the established service and agency reports is that they are available to large audiences. Thus, the data that are contained in them could be expected to be incorporated in stock prices just as soon as the information is published. Other information that is not so easily transmitted or interpreted appears in the financial press. Much of this information requires special expertise or training for proper understanding.

The efficient market hypothesis postulates the existence of large, well-behaved securities markets. For a market to generate a unique (no less efficient) price, it must be sufficiently large so that any buyer (or seller) can purchase (sell) whatever quantity he or she wishes without affecting the market price. The securities markets satisfy this requirement for some securities but not for others. In general, stocks traded on an organized exchange will have a broader market than those that are traded over the counter, because exchanges have listing requirements designed to guarantee market depth. Securities trading in the over-the-counter market may be either primary or secondary in nature. The primary market exists for the distribution of new securities. The secondary markets include both listed and OTC securities that are in the hands of the public. New-issue securities (primary market) are sold by investment bankers to investors. The initial public offering (IPO) market has been very popular at times among speculators.

A major element often cited as contributing to the efficiency of the American securities markets is the regulation provided by the U.S. government. After the 1929 stock market crash, a number of laws were passed that were designed to correct some of the abuses that existed in the past. Disclosure requirements were established, and the Securities and Exchange Commission was created to supervise the investments business. One of the excesses of the pre-1933 era was the practice of buying stocks on margin. Use of this device is now regulated by the board of governors of the Federal Reserve System. Another practice that caused problems prior to 1933 was the short sale. This process is still permitted, although the SEC may make rules governing its use.

In order to determine whether one has earned above (or below) market returns, one must have a good measure of the average performance of stocks in general. No perfect indicator exists. The most widely used average is the Dow Jones Industrial Average. The DJIA is primarily a blue-chip measure, although many investors use it for overall market activity. Standard & Poor's computes indices that have a larger base than the DJIA. The S&P measures are also statistically superior to the Dow calculation. There are a number of other indexes and averages that are also computed. Fortunately (at least for measurement purposes), there is a high positive correlation among the price movements of all stocks. Thus, if most stocks are going up (or down), almost any stock price measure will indicate this.

REFERENCES

Adams, C. F., and Adams, H. (1956). *Chapters of Erie*. Ithaca, NY: Great Seal Books.

Allen, F. L. (1931). *Only Yesterday*. New York: Harper and Brothers.

Bernstein, P. L. (1996). *Against the Gods: The Remarkable Story of Risk*. New York: John Wiley & Sons.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25, 2: 383–423.

Fama, E. F. (1996). Multifactor portfolio efficiency and multifactor asset pricing. *Journal of Financial and Quantitative Analysis* 31, 4: 441–446.

Galbraith, J. K. (1955). *The Great Crash*. Boston: Houghton-Mifflin Co.

Glassman, J. K., and Hassett, K. A. (1999). Stock prices are still far too low. *Wall Street Journal*, March 17 (op. ed.).

Graham, B., and Dodd, D. (1934). *Security Analysis*, 1st edition. New York: McGraw-Hill.

Haugen, R. A. (1999). *The New Finance: The Case Against Efficient Markets*, 2nd edition. Englewood Cliffs, NJ: Prentice Hall.

Kennedy, D. M. (1999). *Freedom from Fear*. Oxford: Oxford University Press.

Keynes, J. M. (1936). *The General Theory of Employment, Interest, and Money*. New York: Harcourt, Brace, and World.
Mackay, C. (1869). *Memoirs of Extraordinary Popular Delusions and the Madness of Crowds*. London: George Routledge and Sons.
Malkiel, B. G. (1999). *A Random Walk Down Wall Street: The Best Investment Advice for the New Century*. New York: W. W. Norton & Company.
Smith, E. L. (1924). *Common Stocks as Long Term Investments*. New York: Macmillan.
Thompson, J. R., Williams, E. E., and Findlay, M. C. (2003). *Models for Investors in Real World Markets*. Hoboken, NJ: John Wiley & Sons.
Williams, E. E., and Findlay, M. C. (1974). *Investment Analysis*. Englewood Cliffs, NJ: Prentice Hall.
Williams, J. B. (1938). *The Theory of Investment Value*. Amsterdam: North Holland.

CHAPTER 6

Investment Banking

K. THOMAS LIAW, PhD
Professor of Finance and Chair, St. John's University

Abstract: Investment banks operate in investment banking, principal transactions, and asset management and securities services. Regulatory changes, globalization, and advances in technology have reshaped the industry. Deregulation in many countries has permitted large financial firms to add different services and products while operating on a global basis. Scandals like Enron and WorldCom, however, have prompted regulators to impose stringent requirements, such as the Sarbanes-Oxley Act of 2002, on the conduct of public companies to restore public trust. Wall Street houses no longer can ask their research analysts to push the stocks of their investment banking clients. Research analysts must make independent recommendations about their assessment of the client's business prospects. Furthermore, advances in technology have enabled clients to access financial services offered by investment banks whenever and wherever they choose. Technology also has allowed investment banks to design, price, and trade complex securities.

Keywords: investment banking, full-service investment banks, boutique investment banks, underwriting, mergers and acquisitions, private equity, venture capital, buyouts, merchant banking, trading, financial holding companies, bulge bracket, fairness opinion, restructuring, financial engineering, swaps, credit derivatives, asset management, repurchase agreements, risk management, valuation, prime brokerage, Gramm-Leach-Bliley Act of 1999, Sarbanes-Oxley Act of 2002

Full-service investment banks offer clients a range of services including underwriting, merger and acquisition advice, trading, merchant banking, asset management, and prime brokerage. Goldman Sachs, Morgan Stanley, and Merrill Lynch are examples of such investment banks. Some of the large financial holding companies such as Citigroup, HSBC, Credit Suisse, UBS, JPMorgan Chase, and Bank of America operate full-service investment banking as well. All these large, full-service investment banks are known as the Wall Street "bulge bracket." The so-called "boutique" investment banks specialize in particular segments of the market. This chapter describes the lines of business offered by those institutions, including investment banking, principal transactions, financial engineering, asset management, and securities services. This chapter also discusses market trends and success factors in the investment banking business.

TYPES OF INVESTMENT BANKS

There are two basic types of investment banks: full-service and boutique. *Full-service investment banks* engage in all kind of activities, including underwriting, trading, mergers and acquisitions (M&As), merchant banking, securities services, investment management, and research. In contrast, *boutique houses* focus on particular segments: Some specialize in M&As, some in financial institutions, and some in Silicon Valley business.

Before the *Gramm-Leach-Bliley Act of 1999* (GLB), there used to be large full-service investment banks and smaller boutiques. Section 20 of the Glass-Steagall Act of 1933 prohibited the affiliation of a member bank of the Federal Reserve System with a company that was engaged principally in underwriting or dealing in securities. In 1987, the Federal Reserve board of governors reinterpreted that phrase to allow bank subsidiaries—so-called "Section 20 subsidiaries" or underwriting subsidiaries—to underwrite and deal in securities. The board approved applications by three bank holding companies to underwrite and deal in Tier 1 securities such as commercial paper, municipal revenue bonds, mortgage-backed securities, and securities related to consumer receivables. In 1988, the Board approved applications by five bank holding companies to underwrite and deal in Tier 2 securities (all debt and equity securities).

Initially, a Section 20 subsidiary could not derive more than 5% of its total revenue from activities involving bank-ineligible securities. The board increased the limit to 10% of total revenue in 1989 and raised it to 25% in 1997. Finally, with the passage of the GLB, the limit was effectively eliminated. Under the act, a bank holding company that elects to become a financial holding company may engage in securities underwriting, dealing, or market-making activities through its subsidiaries (called "securities subsidiaries").

The GLB has enabled a financial services firm such as a commercial bank or a securities house to become a one-stop shop that can supply all its customers' financial needs. By allowing banks, insurance companies, and securities firms to affiliate with each other, the act has opened the way for financial services supermarkets that offer a vast array of products and services including savings and checking accounts, credit cards, mortgages, stock and bond underwriting, insurance, M&A advice, commercial loans, derivative securities, and foreign exchange trading.

The GLB has not only opened up new opportunities for banks but has also provided significant protection for investors and consumers while striving to create a level playing field for all financial services firms. It established a new system of functional regulation whereby banking regulators oversee banking activities, state insurance regulators supervise insurance business, and securities regulators supervise securities activities. In this new regulatory environment, commercial banks can engage in formerly forbidden activities such as stock underwriting and dealing. Citigroup, JPMorgan Chase, Bank of America, HSBC, Deutsche, and UBS all operate under this format.

The traditional full-service firms such as Goldman Sachs and Morgan Stanley offer clients a full menu of investment banking services. What they do not have is a bank that can extend large sums of credit to corporate clients. They have established networks, however, and have been successful.

Niche players are smaller in general, but are creative in specializing in a particular type of clients or services. Sandler O'Neill works on the financial institutions segment. Lazard specializes in asset management and M&As. Houlihan Lokey Howard & Zukin focuses on M&A advisory, fairness opinion, and restructuring.

Financial Holding Companies

Large financial holding companies now include investment banking in their menu of services. Under the universal banking scheme, large banks in Europe and Japan have operated in commercial banking and investment banking. In the United States, after the Gramm-Leach-Bliley Act of 1999 took effect, investment banking has become an integral part of their businesses. Furthermore, these global financial holding companies all have operations in most financial centers and are competing on nearly every continent of the world. The advancement of technology has enabled these financial services giants to offer a complete menu of services on a global basis.

Full-Service Investment Banks

Several investment banks are full-service providers and are not part of a financial holding company. Goldman Sachs, Morgan Stanley, and Merrill Lynch are traditionally called the "big three." Lehman Brothers and Bear Stearns both have unique strengths—Lehman in fixed income, Bear Stearns in custodian, prime brokerage, and mortgage-backed securities. Goldman Sachs, Lehman Brothers, and Bear Stearns focus their clientele on institutions and high-net-worth individuals. Merrill Lynch and Morgan Stanley offer services to retail investors as well. Table 6.1 summarizes the business categories offered by those five Wall Street houses. These lines of business by large investment banks are discussed in detail in the next section.

Table 6.1 Business Categories of Full-Service Investment Banks

Firms	Goldman Sachs	Morgan Stanley	Merrill Lynch	Lehman Brothers	Bear Stearns
Business Categories	Investment Banking	Institutional Securities	Capital Markets	Equities	Capital Markets
	Trading and Principal Investments	Individual Investor	Investment Banking and Advisory	Fixed Income	Wealth Management
	Asset Management and Securities Services	Investment Management	Wealth Management	Investment Banking	Global Clearing Services
		Credit Services MSCI	Insurance Banking	Banking	
				Investment Management	

Table 6.2 Specializations of Boutique Investment Banks

Firms	Sandler O'Neill	Greenhill	Lazard	Houlihan Lokey
Specializations	Financial Institutions and Insurance Companies	Advisory Services in M&As and Financial Restructuring	Advisory Services in M&As, Asset Management	Advisory in M&As, Financial Opinion, Restructuring, and Financing

Boutique Investment Banks

Boutique investment banks do not offer a range of services and are not part of a larger financial institution that serves many competing interests. The following provides a brief description of several boutiques with different specializations, as summarized in Table 6.2.

Sandler O'Neill specializes in financial institutions. The company raises capital, provides research coverage, acts as a market maker, advises on M&As, and trades securities. Its services cover mutual-to-stock conversion (from a mutual ownership structure to a public company), loan portfolio restructuring, strategic planning, and balance sheet interest rate risk management. The investment banking team focuses on demutualization, M&A advice, fairness opinion, leveraged and management buyout, and strategic issues. The capital markets group specializes in convertible securities for financial institutions and in pooled trust preferred transactions for banks, thrifts, and insurance companies. Its research covers financial services companies.

Greenhill is a boutique house focused on M&As, financial restructuring, and merchant banking. It does not have research, trading, lending, or related activities. As such, many corporate clients regard it as an independent firm without any conflict of interests. Greenhill's M&A practice covers buy-side, sell-side, merger, special, and cross-border transactions. The merchant banking services are to identify private investment opportunities and partner with strong management teams.

Lazard has two core businesses: one specializing in financial advisory and the other in asset management. Its M&A services include general strategic advice and transaction-specific advice in M&As, divestures, privatizations, takeover defenses, strategic partnerships, and joint ventures. The financial restructuring practice specializes in advising companies in financial distress. The asset management business provides investment management and advisory services to institutions, financial intermediaries, and private clients.

Houlihan Lokey Howard & Zukin (Houlihan Lokey) provides services in M&As, financial opinion, financing, and restructuring. Its focus is middle market transactions. In M&As, Houlihan Lokey groups its bankers by industry sectors to provide in-depth knowledge of the client's industry. The group also works on transactions for distressed companies, both in and out of bankruptcy court. The firm can also arrange financing for a wide range of transactions. In fairness opinion, it performs analysis, assessments of the proposed transaction as well as alternatives, to provide clients its views on the fairness of the proposed transaction.

INVESTMENT BANKING BUSINESS

Investment banks engage in public and private market transactions for corporations, governments, and investors. These transactions include mergers, acquisitions, divestitures, and the issuance of equity or debt securities, or a combination of both. Investment bankers advise and assist clients with specialized industry expertise. The industry or sector groupings generally include Industrial, Consumer, Healthcare, Financial Institutions, Real Estate, Technology, Media and Telecommunications, and others. Investment banks today go beyond securities business to include trading, securitization, financial engineering, merchant banking, investment management, and securities services. For those activities, investment banks earn fees, commissions, and gains from principal transactions.

Investment banking includes capital raising and M&A advisory services. Investment banks help clients raise capital through underwriting in which investment banks purchase the whole block of new securities from the issuer and distribute them to institutional and individual investors. For the service, investment bankers earn an underwriting spread, the difference between the price they receive from investors and the amount they pay to the issuing firm.

Another major line in investment banking is strategic advisory on M&As. Services offered include structuring and executing domestic and international transactions in acquisitions, divestitures, mergers, joint ventures, corporate restructurings, and defenses against unsolicited takeover attempts. Fees are usually negotiable. As transactions grow larger and larger, the M&A advisory fees are generally less than 100 basis points and often much lower. This line of business is attractive because "win, lose, or draw," bankers earn fee income. Another source of fee income is from rendering a *fairness opinion*. A fairness opinion is a professional judgment on the fairness of the financial terms of a transaction.

Full-service investment banks offer a service menu that goes beyond just investment banking. Principal transactions, including proprietary trading and merchant banking, have accounted for a significant portion of total net revenues at major Wall Street houses. In proprietary trading, the investment bank trades on its own capital. Financial engineering has enabled them to design complicated trading strategies. Merchant banking invests the firm's own capital as well as funds raised from outside investors in companies and real estate.

Investment management has become an integral part of the investment banking business. Wall Street houses such as Merrill Lynch, Morgan Stanley, and Goldman Sachs each manages hundreds of billions of dollars for their clients. This is an attractive segment of the financial services industry. The income stream is less volatile than trading or underwriting and, hence, contributes to the stability of earnings.

Another line of revenue-producing business is securities services: prime brokerage, securities lending, and financing. Prime brokerage offers tools and services desired by clients looking to support their operations in trading and portfolio management. In *security lending* services, investment banks find securities for clients to make good delivery so as to cover their short positions. Alternatively, financing services provide funds to finance clients' purchases of securities. In addition, new financial products designed by the financial engineering team often enhance their services to clients.

Investment Banking

The main revenue-producing services are underwriting and financial advisory. Underwriting includes public offerings and private placements of equity and debt securities. Financial advisory covers M&As, fairness opinion, divestures, corporate defense activities, restructuring, and spin-offs.

Underwriting

In the equity underwriting market, initial public offerings (IPOs) are more lucrative than the secondary offerings. In a public offering, the lead manager begins by conducting due diligence research and then coordinates the preparation of the registration statement to be filed with the Securities and Exchange Commission (SEC). The registration process was streamlined in 1992 when the SEC adopted shelf registration, which permits an issuer to register multiple types of securities, both common equity and debt, that it intends to issue over the succeeding two years on a single registration statement. The advantages of the rule include flexibility in the timing of the security issuance, reduced regulatory uncertainty, and lower direct issuance costs.

There are two different types of agreements between the issuing company and the investment bank. The first type is the firm commitment, in which the investment bank agrees to purchase the entire issue and distribute it to both institutional and retail investors. The second type is known as a best efforts agreement. With this type of agreement, the investment bank agrees to sell the securities but does not guarantee the price.

Underwriting of fixed income securities covers Treasury securities and corporate debt. Investment banks, through their primary dealerships, participate in the auction of the Treasury securities. On the corporate side, investment banks underwrite corporate fixed income securities and distribute to institutional investors in a way similar to equity underwriting. One of the key differences is that a larger portion of the debt underwriting is by private placement. Private placements differ structurally from the registered public deals because they are highly negotiated in covenants and pricing, and they do not go through the SEC registration process. A private issue can save substantial amounts of legal and registration expenses against a comparable public issue.

M&A Advisory Services

M&As are one of the major areas of the investment banking business. Advice on M&As ranges from strategic recommendation to clients about which targets are worth pursuing to tactical suggestion about what price to offer and how to best structure the deal. Targets of acquisitions also seek M&A bankers for advice on how to negotiate the best price or how to defend themselves.

M&A transactions generate large sums of fee revenues for investment banks. Wall Street is obsessed with M&As, because win, lose, or draw, they produce fees: fees for advising, fees for lending money, and fees for divesting unwanted assets. Fees are usually negotiable and contingent upon the success of a deal.

After a suitable candidate has been identified, the investment bank conducts valuation of the merger candidate to determine what price to offer. The valuation techniques are used only in determining the price range reference for the target company. Each acquirer uses the technique that fits its objective. Equally important, a risk analysis such as a scenario analysis or sensitivity analysis should be performed. The valuation is not complete until the impact of the acquisition on the acquirer is also carefully examined. The techniques investment banks use to value a target include the following:

- The discounted cash flow (DCF) technique is widely used in evaluating acquisitions. The DCF method determines the value by projecting future cash flows of the target and discounting those projections to the present value. The DCF approach is future oriented, it begins with a projection of sales and operating profit, based on the assessment of historical performance as well as certain assumptions regarding the future. The usefulness of this technique depends on several assumptions including the impact on the company's other areas of business, length of projection period, additional working or fixed capital required, discount rate, and residual value. The value of the DCF should be estimated under different scenarios.
- Comparable transaction analysis is undertaken to analyze transactions involving companies in the target's industry or similar industries over the past several years. Acquisition multiples are calculated for the universe of the comparable transactions. These multiples are then applied to the target's financial results to estimate the value at which the target would likely trade. This technique is effective when data on truly comparable transactions are available.
- The comparable company approach makes an assessment of how the value of the potential acquisition candidate compares with the market prices of publicly traded companies with similar characteristics. This method is similar to the comparable transaction approach that identifies a pricing relationship and applies it to the candidate's earnings or cash flow or book value. A change of control premium should be added to the value

identified by this method to arrive at the estimated valuation range for the target. One weakness of this technique is that it works well only when there are good comparables for the target. Another weakness is that accounting policies can differ substantially from one company to another, which could result in material differences in reported earnings or balance sheet amounts.

- The breakup valuation technique involves analyzing each of the target's business lines and summing these individual values to arrive at a value for the entire company. Breakup analysis is best conducted from the perspective of a raider. The process initially determines the value of the target in his hands. The acquisition cost is estimated in the next step. If value exceeds cost, the raider computes the rate of return. This technique provides a reference under a hostile bid.
- Target stock price history analysis examines the stock trading range of the target over a time period. The target price performance is analyzed against a broad market index and comparable-company performances. The offering price is based on the price index plus some premium. Similar analysis is performed on the acquiring firm if the transaction is a stock-for-stock exchange. The purpose is to determine the exchange ratio. This approach fails to account for future prospects of the company. Nevertheless, it does provide historical information many find useful in framing valuation thoughts.
- The M&A multiples technique analyzes the current and past broad acquisition multiples and the change of control premium. This technique is used when comparable transactions or comparable companies are not available. The limitation is that a broad market average may be inapplicable to a single transaction.
- Leveraged buyout (LBO) analysis is performed when the target is a potential candidate for LBO. The objective is to determine the highest price an LBO group would pay. This is often the floor price for the target. On the other hand, it may set the upper value for the target company if a corporate buyer cannot be identified. The LBO analysis includes cash flow projections, rates of returns to capital providers, and tax effects. The primary difference between the LBO analysis and DCF technique is that LBO approach incorporates financing for the LBO. The availability of financing is dependent on the timing of cash flows, particularly in the first several years after the deal is completed. Clearly, the value derived by the LBO approach can be materially affected by temporary changes in financing conditions.
- Leveraged recapitalization method is aimed at identifying the maximum value that a public company can deliver to its shareholders today. In general, the analysis is performed in the context of a probable or pending hostile offer for the target. The value in a recapitalization is delivered to the shareholders through stock repurchase, cash dividends, and a continuing equity interest in a highly leveraged company. This technique focuses on the target's capital structure, and is largely affected by the availability of debt financing at a particular time.
- Gross revenue multiplier is the so-called price-to-sales ratio. The basic concept is that the value is some multiple of the sales the target generates. The method implicitly assumes that there is some relatively consistent relationship between sales and profits for the business. Obviously, the usefulness of the technique depends on the revenue-profit relationship. In practice, this method may be quite useful when acquiring a private company where gross sales are the only reliable data available.
- The book value approach is an accounting based concept and may not represent the earnings power. Also, the value of intangible assets may not be reflected in the balance sheet. However, it will help provide an initial estimate of goodwill in a transaction.
- The multiple of earnings per share method involves taking the past or future income per share and multiplying that figure by an earnings multiplier, derived from publicly traded companies in the same industry. One difficulty is that the known multipliers do not reflect control premiums, as evidenced by the rise in the multiplier in the event of an acquisition. Another problem is that income does not necessarily represent cash flow from operations.
- Liquidation analysis could be used to establish a floor for valuation. This approach is relevant if a business is being acquired for its underlying assets rather than for its going concern.

Fairness Opinion

Majority of companies involved in M&As also obtain a second opinion, in the form of a fairness opinion, to determine if the transaction is fair from a financial standpoint. Fairness opinions are established on the basis of a valuation report and require an in-depth analysis of the companies involved and the terms and conditions of the transaction. The average fee paid for a fairness opinion is small relative to the overall fees paid to investment banks on M&As. But when the investment bank providing advisory services also offers the fairness opinion, a potential conflict of interest can arise since these banks have an incentive to see the transaction completed in order to receive the success fee. Thus, it is prudent for the board's special committees to use another investment bank for fairness opinion. An independent, unbiased fairness opinion will provide value to executives and boards as an additional form of due diligence, and to shareholders as a mechanism to ensure quality transactions. A fairness opinion is also an insurance policy offering directors a first line of defense against shareholders' lawsuit. This is because a timely, independent analysis may establish for the record that the board has properly exercised their business judgment by having adequately considered the proposed transaction and the potential alternatives. Furthermore, the National Association of Securities Dealers (NASD) has intended to require its members to comply with NASD Rule 2290 to ensure proper disclosure and independence of the fairness opinion. Investment banks that provide fairness opinions are typically registered broker-dealers and NASD members.

Financial Restructuring

Financial restructuring is complex as it often involves transactions for distressed companies, both in and out of bankruptcy. Advisers in this area need to provide structure for a broad array of options for company owners, executives, creditors, and other parties. It is necessary to provide thorough and comprehensive analyses of factors in restructuring transactions and to efficiently implement creative solutions. The purpose is to ensure that the restructuring process is effectively managed to maximize value and minimize delay.

Such financial restructuring involves distressed companies in change of control, asset sales, and other M&A activities. These situations may require asset divesture quickly under extremely distressed circumstances for companies in and out of bankruptcy court. Many of those distressed sale transactions are consummated in chapter 11. Thus, the distressed M&A adviser must be able to articulate to buyers the benefits of purchasing assets from a distressed company and allay the concerns of buyers. Often, the valuation of the company in bankruptcy requires modifications to traditional healthy company valuation and the sale must be done quickly. This is because, as time passes, the value may decline leaving liquidation as the only available alternative.

Restructuring is sometimes more strategic than transactional. In such a circumstance, the pressing job is to create and execute a total solution for the company to grow and to return to profitability. First, the banker thoroughly evaluates the company's finances, industry condition, and capital market environment. Second, the advisor explores and presents to company all strategic alternatives, processes, and the impact on various stakeholders. After such a comprehensive review, a value maximizing strategy is recommended and executed.

Trading and Principal Investments

Trading and principal investments are important revenue producing operations. Trading could be for market making or for the firm's proprietary account. Principal investments are the merchant banking operation in venture capital and buyouts.

Trading

Many investment banks put up large sums of capital for proprietary trading. The first step to successful trading is to ensure survival by making risk management a top priority. Many losers are washed out while trading their way out of a hole. Many of them have difficulties taking a loss and they tend to keep on hanging on to money losing securities. The essential aspect is to understand that a 10% loss requires a gain of more than 11% just to get even, and that a 50% loss will require a gain of 100% to get back in the game. Typically, a trader would place a stop right after he got into a position. The level of stop is chosen in such a way that any loss from a single position will be limited to a small percentage of the account.

Likewise, taking profits is sometimes emotionally hard. When the market moves in the anticipated direction, a trader needs to decide whether to stay put, take profits, or add to position. A successful trader sets certain objective for each position, and once the objective is accomplished, he knows when to close out the account.

There are many markets, many instruments, and many techniques. Each market has its own unique characteristics and its own trading hours. Certain instruments continuously trade in different time zones. Fundamentals in the market where they trade and the events in other markets affect their prices. Most major currencies and the U.S. government securities trade in all major markets, and the economic fundamentals in the United States and the financial market conditions in other countries affect their prices.

There are three basic approaches to trading. The first is fundamental analysis, which bases a security price on corporate and economic fundamentals. The fundamental approach for a security involves the analysis of the economy, industry, and company. This applies to equities and fixed income securities. In commodities, fundamentalists study factors that affect market demand and supply. Currencies are affected by economic fundamentals such as production and inflation, and by political factor as well. In futures, expectations of interest rate and cash market conditions are important. Volatility and expected direction of price movements are key in determining the options valuation.

The second approach is the market efficiency hypothesis, in which securities prices are based on all available information so as to offer an expected rate of return consistent with their level of risk. There are three different degrees of informational efficiency. The least restrictive form is the weak form efficiency, which states that any information contained in the past is already included in the current price and that its future price cannot be predicted by analyzing past prices. This is because many market participants have access to past price information, and hence any free lunches would have been consumed. The second form of informational efficiency, semistrong form efficiency, states that security prices fully reflect all relevant publicly available information. Information available to the public includes past prices, trading volumes, economic reports, brokerage recommendations, advisory newsletters, and other news articles. Finally, the strong form of informational efficiency takes the information set a step further and includes all public and private information. This version implies that even insiders who have access to nonpublic material information cannot make abnormal profits. Most studies support the notion of semistrong form market efficiency, but do not support the strong form version of efficient market hypothesis. In other words, insiders can trade profitable on their knowledge of nonpublic material information. This advantage is unfair and hence insider trading is illegal.

Finally, technical analysis attempts to use information on past price and volume to predict future price movement. It also attempts to time the markets.

Principal Investments

Principal investments represent the bank's merchant banking investments. This involves the commitment of the firm's capital to equity level investments and participation. These include financing guarantees, venture capital, leveraged buyouts, and restructurings. In private equity (venture capital and leveraged buyouts), investment banks are involved, from raising capital for the funds to taking the portfolio company public or selling out to other businesses. An investment bank may simply raise money for external private equity funds such as venture capital or buyout funds. An investment bank, alternatively, can manage the fund itself. Even though many private equity investments turn sour, the successful ones are so profitable that the overall annual returns are often quite attractive. Major Wall Street houses all have private equity operations. *Private equity* is of interest to banks because it has several benefits including management fee, capital gains, and contributing to underwriting and merger and acquisition business.

Data on private companies are limited. Early stage companies generally experience a period of negative cash flows and negative earnings before they produce positive net income. The timing and the amount of future profits are highly uncertain. Thus, valuing private companies is subjective and difficult. Common private equity valuation approaches are comparables, net present value, option valuation, and venture capital methods. Comparables and net present value approaches have been discussed in M&A valuation. Thus, we focus on option valuation and venture capital methods.

The option valuation method assigns a value to the flexibility that the venture capitalist has on making follow-on investments. This right is similar to a call option on a company stock, which is a right, not an obligation, to acquire an asset at a certain price on or before a particular date. Options pricing theory captures this "option" to either invest or not invest in the project at a later date. This valuable option is not accounted for by the DCF approach. The Black-Scholes model was the first widely accepted method to value European options using five variables: (1) exercise price, (2) stock price, (3) time to expiration, (4) standard deviation of stock returns, and (5) time value of money. To value a firm, the five variables used are (1) the present value of expenditures required to undertake the project, (2) the present value of the expected cash flows generated by the project, (3) the length of time that the venture capitalist can defer the investment decision, (4) the riskiness of the underlying assets, and (5) the risk-free rate. The value is then obtained once those input variables have been estimated. This approach is useful because it specifically incorporates the flexibility to wait, to learn more, and then to make the investment decision. The options valuation has its drawbacks, too. Many business people are not aware of this "real option" concept. Furthermore, the real-world problems are often too complicated to be captured in the model.

The venture capital method takes into account negative cash flows and uncertain high future profits. It considers cash flow profile by valuing the target company at a time in the future when it expects to generate positive cash flows and earnings. The terminal value at that projected target date is discounted back to the present value by applying a discount rate, a target rate of return (TRR), instead of cost of capital. The TRR is the rate of return that the venture capitalist requires when making an investment in the portfolio company. The terminal value is generally obtained using price-to-earnings ratio multiplied by the projected net income in the year. The amount of proposed investment is divided by the discounted terminal value to give the required final percentage ownership that the venture capitalist wants to own. The final step is to calculate the current percentage ownership taking into consideration the dilution effects when the portfolio company goes through several rounds of financing. This is done by calculating a retention ratio that factors in the dilutive effects of future rounds of financing on the venture capitalist's ownership. For example, assume that the portfolio company will sell 30% in the second round and then another 25% in the third round before it goes public. The retention ratio is 61.5%, meaning that 1% ownership in the initial investment is diluted to only 0.615% after two rounds of financing. If the venture capitalist invests $10 million and requires a final percentage ownership of 10%, she will require the current ownership percentage of 16.26%. The 16.26% current percentage ownership is necessary for the venture capitalist to realize the target rate of return.

Asset Management and Securities Services

Investment banks operate in asset management and other securities services to better service their clients and to diversify their revenue sources. Those essential services cover financial engineering, prime brokerage, financing, and securities lending.

Asset Management

Asset management provides investment advisory services including mutual funds, separate accounts managed for clients, merchant banking funds, and other alternative assets. Investment management is an important segment of the capital markets and has become an integral part of the investment banking business. Wall Street firms have engaged in investment management because it is one of the most attractive segments of the financial services industry. It expands the menu of products and services that investment banks offer to clients. Furthermore, the income stream is less volatile than trading, underwriting, or M&A activities. The affiliated funds also provide synergy to the bank's underwriting business.

Financial Engineering

Financial engineering is the term used to describe the investment banker's creativity in innovative security design. The rapid pace of financial innovation is driven by the competition among investment bankers in response to increased price volatility, tax and regulatory changes, demand for new funding sources, arbitrage, and yield enhancement. The application of mathematical and statistical modeling, together with advances in computer technology, provides the necessary infrastructure for financial engineering.

Financial engineering helps investment banking professionals to meet the needs of borrowers and investors such as hedging, funding, arbitrage, yield enhancement, and tax purposes. It drives the explosive growth in the structured and derivatives markets. The development of the junk bond and asset-backed markets provides borrowers additional funding sources at lower costs. Structured notes add another dimension in the funding and investment spectrum. Transactions in repurchase agreements provide borrowers lower funding costs and give lenders legal title to the collateral. Through swap contracting borrowers and investors obtain a high degree of flexibility in asset-liability management at better terms. Credit derivatives have widespread applications in hedging credit risks.

Prime Brokerage and Related Services

Prime brokerage is a suite of services providing clients such as hedge funds with custody, clearance, financing, and securities lending. These services make it possible for the hedge fund and other clients to have multiple brokers while maintaining one brokerage account. In prime brokerage, the investment bank acts as the back office for the fund by providing the operational services necessary for the money manager to effectively manage his business. This enables the clients to focus on investment strategies rather than on operational issues. The services a good prime broker provides include centralized custody, clearance, securities lending, competitive financing rates, one debit balance/one credit balance, real time and periodic portfolio accounting, position and balance validation, electronic trade download, wash sale reports, and office facilities in selected markets.

TRENDS AND CHALLENGES

Investment banking is a very fluid and dynamic business. Successful bankers constantly anticipate market trends and opportunities and then align resources to ensure that they serve those opportunities in the best way possible. Thus, investment bankers perform on-going analysis of each client to provide smart solutions so clients achieve superior performance. Several important trends have emerged. First, regulatory compliance and high standards of governance have become an integral part of the business. Second, Europe and Asia now provide a faster growth opportunity. Third, many firms pursue a strategy diversifying and balancing revenue streams to maintain sustained earnings growth.

The Evolving Investment Banking Markets

Investment banks perform several essential functions in the marketplace. At the core of what they do is origination and strategic advisory. As the market evolves, the large full-service Wall Street firms are diversifying and balancing their revenue streams. Thus, most of them have expanded their menu of services and products and allocated resources to pursue higher-growth opportunities in Asia and Europe.

Deregulation and Revenue Diversification

Deregulation in the 1980s and 1990s gradually chipped away some of the barriers between investment banking and commercial banking. By 1999, the main barriers separating the three segments of the financial services industry—banks, securities firms, and insurers—were removed by the Gramm-Leach-Bliley Act of 1999.

Commercial banks are uniting with securities firms to create financial supermarkets that offer one-stop shopping for all financial services. Investment banks are undergoing a similar trend, driven both by the competition from other firms and by their corporate clients' desire to have financial advisors that can address all their needs, regardless of what types of instruments might be required. Thus, investment banks have been aggressively expanding the menu of services they provide, adding money lending, retail and institutional fund management, structured finance, and securities services. Indeed, large investment banks now derive the majority of their revenues from sources other than the traditional investment banking activities. These other sources include trading and principal transactions, commissions, asset management, and securities services.

Globalization

In addition to expanding the products and services they offer at home, large investment banks are also expanding geographically to become financial supermarkets to the world. With rapid advances in information technology and greater cooperation among financial regulators, the international capital markets are now closely linked. Larger sums of money are moving across national borders, and more countries have access to international finance. By going global, investment banks not only can serve their clients better but also can benefit from the higher growth potential of international markets. Regulatory frameworks in Japan, Europe, and developing countries are changing to accommodate and encourage private pension programs, more investments in securities, and greater participation by nonlocal firms.

Big Wall Street houses such as Goldman Sachs, Morgan Stanley, and Merrill Lynch all have a strong global presence and have established leadership positions in core products. Although they are among a select few that have the ability to execute large, complex cross-border transactions, many other Wall Street firms pursue a globalization strategy. As a result, U.S. firms have significantly increased their international securities activities. Major U.S. houses earn a significant portion of their net revenues from international operations. At the same time, foreign financial institutions are expanding their investment banking capabilities in the United States. UBS and Deutsche Bank, for example, have established an investment banking presence through acquisitions.

Internet and Information Technology

The changes in investment banking have been in part aided by the advent of the Internet and advances in information technology. As David Komansky (1999), the former chairman and CEO of Merrill Lynch, observed, together globalization and technology "have collapsed time and distance and opened a floodgate of opportunities for those who embrace them."

The Internet and e-commerce have already changed, and will continue to change, the way that securities are traded and distributed. Many firms now use the Internet for extended trading in markets around the world. Clients now have online access to research, data, and valuation models 24/7. Some have gone one step further to allocate shares of initial public offerings through online auctions. The online auction approach brought out many IPOs for small U.S. businesses. Google's IPO was undertaken via this type of auction process as well. To further enhance their distribution capabilities, major investment banks have established alliance with retail brokerage houses.

In addition to using the Internet, investment banks are developing software and information technology systems that enable them to enhance their service to clients, better manage risks, and improve overall efficiency and control. New software has made it possible for firms to tailor their research and services to each client's particular needs.

Technological advances have also provided investment banks with capabilities to design and price complex contracts and derivatives and to analyze their underlying risks. Every firm has software that enables it to monitor and analyze market and credit risks. Risk management software can not only analyze market risk at the firm, division, and trading desk levels, but can also break down the firm's risk into its underlying exposures. This permits management to evaluate the firm's exposure in the event of changes in interest rates, foreign exchange rates, equity prices, or commodity prices. Without such software, ventures into international markets and complex trading would carry much more risk than without it.

Finally, information technology has been a significant factor in improving the overall efficiency of investment banks (and many other businesses as well). Computerized and electronic trading is both more efficient and more accurate. Management now has real-time information on the firm's operations worldwide. Not only has globalization been made possible, but also better decision making and improvement in the firm's competitive edge have taken place.

Challenges and Opportunities

Since the collapse of WorldCom and Enron (the two largest bankruptcies in North America during 2000–2005), regulators and shareholders pay close attention to transparency, accountability, and corporate governance. The objectives of the *Sarbanes-Oxley Act of 2002* are to ensure corporate accountability and to restore public trust in the financial markets. Corporate responsibility will continue to be the focus. In June 2005, settlements by Citigroup and JPMorgan Chase to pay $2 billion and $2.2 billion, respectively, for their involvement in Enron scandal signify the importance of regulatory compliance. In addition, retention of talent and diversity of workforce will continue to be essential to successful operations. As firms pursue the strategy of sustained earnings growth, it is not optional but necessary to manage risks effectively.

Accountability and Corporate Governance

Public confidence in business principles and practices is the foundation of an efficient financial system. Regulations and corporate policies are evolutionary. Focus has been in the areas of corporate accountability, standards of corporate governance, and conflict of interest. Under the Sarbanes-Oxley Act of 2002, public companies are required to certify that they have established reasonable internal control systems and procedures for completeness of their financial reporting. Such accountability and transparency are expected to continue.

In corporate governance, all public companies now have independent directors and these directors meet without management at least once a year. Most financial institutions have established a sound corporate governance structure.

Another important element for investor confidence is relating to analyst conflict of interest. The Sarbanes-Oxley Act contains provisions to prevent analyst conflict of interest. Investment banks also have adopted policies to ensure compliance of analyst independence. As an example, in May 2005, Merrill Lynch withdrew from underwriting the IPO of Warner Music Group after its media analyst, Jessica Reif Cohen, told Merrill's senior bankers that they were overpricing the shares. Such analyst independence is a striking example of changes that have occurred since Wall Street's settlement of $1.4 billion with the attorney general of New York, Eliot Spitzer. After the settlement, Wall Street research analysts are asked to make separate, independent, and impartial recommendations to the underwriters about their views of the client's business prospects.

Recruitment and Retention of Talents

The most important asset at any investment bank is its people. The success of business is dependent upon the team's ability to provide the most innovative and creative solutions to clients' needs. Investment banks face competition from financial holding companies that offer similar services. These financial holding companies also compete for talent, in addition to clients. To maintain the competitive edge and meet expectations of clients, investment banks must attract, retain, and motivate employees. A performance-based compensation that rewards results is fundamental to the operations of an investment-banking firm. Thus, employee compensation and benefits is the largest item in expenses, reaching almost 50% of total net revenues at many houses. Furthermore, most firms also stress a culture of client focus, integrity, social responsibility, diversity, community service, teamwork, and entrepreneurial spirit.

Risk Management

Effective risk management is of primary importance to the success of an investment bank. The proliferation of products and increasing complexity of regulations has made effective risk management a must. All investment banks establish a comprehensive risk management process to monitor, evaluate, and manage the risks that the firm assumes in conducting its businesses. Important areas include market, credit, liquidity, operational, legal, and reputation exposures. Many believe effective management of those risk exposures will ensure regulatory compliance and maintain sustained earnings growth. More importantly, it protects the reputation and survival of the firm.

Europe and China

The global economy is becoming more and more integrated. In the process, Europe has higher growth opportunities than the Americas and China presents tremendous potentials. One of the big driving forces behind China's growth is the huge sums of money from foreign institutions invested in infrastructure. The incredible economic growth has also fostered an increase in wealth development and accumulation. The surge in consumer purchasing power creates businesses for many corporations. This in turn presents investment banks opportunities in underwriting, financing, and strategic advisory. Wealth management is another promising potential. Morgan Stanley, Goldman Sachs, and Merrill Lynch have all established a presence in China. As the equity culture develops, shareholders will demand accountability and stability of earnings. Therefore, corporate governance and risk management are important subjects.

Success Factors

Securities businesses by their nature are subject to volatility. The volatility comes from changes in industry competition, interest and foreign exchange rates, and global economic and political trends. Each line of business has its ups and downs, and is subject to intense competition. The menu of services is also changing. Investment banks have been forced to change to meet new challenges. In addition, most big houses have expanded overseas operations in all major capital markets and have derived a substantial portion of their revenues from non-U.S. markets. To achieve and maintain a leadership position in investment banking, a firm must have:

- Deep client relationships to obtain a flow of businesses.
- A strong product line to offer the best products and services.
- The ability to provide clients with an integrated solution to help them achieve superior results.
- A strong global presence and local knowledge.
- A strong financial strength to establish the confidence of clients and maintain long-term relationships.
- An effective risk management process to ensure the firm's financial soundness and profitability.
- A solid governance structure to ensure compliance with internal policies and regulations.
- Integrity and professionalism to create trust and provide superior services.
- A compensation system that attracts and retains talents.

SUMMARY

The market for investment banking operations evolves over time. Investment banks are facing increasing competition for talent as well as clients. To compete for clients' businesses and assets, investment banks need to offer integrated solutions so that the clients achieve superior results. To compete for talents, investment banks focus on the compensation system and corporate culture. To maintain stability of earnings, investment banks diversify revenue streams and manage risks.

REFERENCES

Amihud, Y. (2002). *Leveraged Management Buyouts: Causes and Consequences*. Frederick, MD: Beard Books.

Camp, J. J. (2002). *Venture Capital Due Diligence: A Guide to Making Smart Investment Choices and Increasing Your Portfolio Returns*. Hoboken, NJ: John Wiley & Sons.

China Securities Regulatory Commission. (2006). *China's Securities and Futures Markets*. Beijing, China.

Cullinan, G., Le Roux, J. M., and Weddigen, R. M. (2004). When to walk away from a deal. *Harvard Business Review* (April): 1–9.

Dittmar, A., and Thakor, A. (2007). Why do firms issue equity? *Journal of Finance* 62, 1: 1–54.

Fabozzi, F. J. and Mann, S. V. (eds.) (2005). *Securities Lending and Repurchase Agreements*. Hoboken, NJ: John Wiley & Sons.

Frankel, M.E. (2005). *Mergers and Acquisitions Basics: The Key Steps of Acquisitions, Divestures, and Investments*. Hoboken, NJ: John Wiley & Sons.

Kisgen, D. J., Qian, J., and Song, W. (2005). Are fairness opinions fair? The case of mergers and acquisitions. Working paper, Boston College.

Komansky, D. H. (1999). Merrill Lynch: At the threshold of a new world. Speech delivered at the Goldman Sachs Financial Services Conferences on May 12, 1999, in New York.

Lazard. IPO Prospectus 2005.

Liaw, K.T. (2006). *The Business of Investment Banking: A Comprehensive Overview*. Hoboken, NJ: John Wiley & Sons.

Liaw, K.T. (2007). *Investment Banking and Investment Opportunities in China: A Comprehensive Guide for Finance Professionals*. Hoboken, NJ: John Wiley & Sons.

Michaelson, M. (2002). *Restructuring for Growth: Alternative Financial Strategies to Increase Shareholder Value*. New York: McGraw-Hill.

Piskorski, M. J. (2005). Note on Corporate Strategy. Harvard Business School Case 9-705-449.

Securities Regulation Institute. (2005). The Evolving M&A Market. Northwestern University Law School.

CHAPTER 7

Securities Innovation

JOHN D. FINNERTY, PhD
Professor of Finance and Director of the MS in Quantitative Finance Program, Fordham University Graduate School of Business
Managing Principal, Finnerty Economic Consulting, LLC

Abstract: Securities innovation is the process of developing positive net present value financing instruments. Securities innovation improves capital market efficiency by offering more cost-effective means of transferring risks, increasing liquidity, and reducing transaction costs and agency costs. It is a profit-driven response to changes in the economic, tax, and regulatory environment. It involves the design of financial instruments that are better in that they either provide superior, previously unavailable risk-return combinations or furnish the desired future cash-flow profile at lower cost than existing instruments. This often includes combining new derivative products with traditional securities to manage risks more cost effectively. The key to developing better risk-management vehicles is to reallocate risk more cost effectively. Innovations thrive when they provide real value. A new financial instrument is truly innovative only if it makes issuers and investors better off than they were before the new security was developed.

Keywords: securities innovation, risk reallocation, financial engineering, agency costs, clientele effect, asset securitization, debt innovations, preferred stock innovations, convertible securities innovations, common equity innovations, primary capital, mortgage-backed securities, stripped

mortgage–backed securities, mortgage pass-through certificates, collateralized mortgage obligations (CMOs), credit card receivable–backed securities, reserve-fund structure, shifting-interest structure, automobile loan–backed certificates, collateralized debt obligations (CDOs), collateralized bond obligations, collateralized loan obligations, basket default swap, pay-in-kind debentures (PIKs), toggle PIKs, zero-coupon bonds, adjustable-rate notes, floating-rate notes (FRNs), dual-currency bonds, indexed-currency option notes, principal exchange rate–linked securities, reverse principal exchange rate–linked securities, alternative risk transfer (ART), insurance-linked notes (ILNs), catastrophe bonds, global bonds, extendible notes, hybrid capital securities, inverse floaters, interest rate reset notes, credit-sensitive notes, floating-rate, rating-sensitive notes, puttable bonds, increasing-rate notes, euronotes, euro-commercial paper, premium bonds, variable-coupon renewable notes, supermaturity bonds, commodity-linked bonds, covered bonds, equity contract notes, equity commitment notes, structured products, structured notes, structured swaps, hybrid capital security, leveraged inverse FRNs, collared FRN, credit-linked note (CLN), adjustable-rate preferred stock, convertible adjustable preferred stock, auction-rate preferred stock, money market preferred stock, remarketed preferred stock, variable cumulative preferred stock, gold-denominated preferred stock, mandatory convertible preferred stock, preferred equity redemption cumulative stock, dividend enhanced convertible stock, preferred redeemable increased dividend equity security (PRIDES), puttable convertible bonds, zero-coupon convertible debt, convertible-exchangeable preferred stock, adjustable-rate convertible debt, liquid yield option notes, ABC Securities, synthetic convertible debt, conversion price reset notes, convertible interest-rate-reset debentures, contingent convertible bonds, cash-redeemable LYONs, cash-settled convertible notes, Americus trust, SuperShares, unbundled stock units (USUs), callable common stock, puttable common stock, master limited partnerships

Securities innovation is the process of creating positive-net-present-value financial instruments. It has brought about revolutionary changes in the array of available financial instruments. Many factors stimulate this process, the more important of which are interest rate and exchange rate volatility, tax and regulatory changes, globalization of the capital markets, deregulation of the financial services industry, and increased competition within the investment banking industry.

Designing innovative financial instruments to solve financial problems is referred to as *financial engineering* (Finnerty, 1988, 1992). A new financial instrument is innovative when it makes both issuers and investors better off than they were with previously existing securities. An innovative security helps the capital markets operate more efficiently or makes them less incomplete (Van Horne, 1985; Ross, 1989; Merton, 1992). It enables market participants to either accomplish something more efficiently or accomplish something they could not achieve previously. Often, the objective has been more cost-effective hedging vehicles. The challenge for a prospective issuer or investor is to determine whether the new security is truly innovative or just looks different and is intended only to enrich the investment bankers who are promoting it!

Innovations thrive when they provide real value. For example, financial futures have enjoyed ever-expanding growth since their inception in the early 1970s. Other innovations, such as deferred-interest debentures, were issued in large volume for a brief time but have since been issued only infrequently—because changes in tax law eliminated their advantages or more recent innovations superseded them. Still others, such as zero-coupon bonds, were issued in large volume in one form, virtually disappeared because of a change in tax law or regulation, and then reemerged in a new form—liquid yield option notes (LYONs)/zero-coupon convertible debt, which became popular. Extendible notes, medium-term notes, mandatory convertible preferred stock, collateralized mortgage obligations, and fixed-rate capital securities are among the innovations that have thrived.

Numerous securities innovations have been designed to circumvent provisions of the tax code or regulation. Miller (1986) likens the role of regulation in stimulating innovation to that of the grain of sand in the oyster. Since few things in this world are as mutable as the current tax code or a set of investment regulations, securities intended to overcome such obstacles are likely to disappear along with the tax or regulatory quirk that gave rise to them. But just as quickly, new tax or regulatory provisions will spawn a new round of securities innovation.

This chapter provides a survey of innovative corporate securities through July 2007. (For earlier surveys, see Finnerty [1988, 1992] and Finnerty and Emery [2001, 2004].) It identifies the sources of value added of the more significant recent innovations and, for the reader's convenience, retains the descriptions of the most significant past innovations. Innovations are categorized as one of four types of instruments: (1) debt; (2) preferred stock; (3) convertible securities; and (4) common equity. The updated tables describe a total of some 80 distinct new securities we have been able to identify. For each security, the tables provide a brief description of its distinctive features, probable sources of value added, the date of first issue, and an estimate of the number of issues and total new issue volume for each security through July 2007. The value added offers reasons for the "staying power" for enduring securities and the lack of innovation in failed securities.

DEBT INNOVATIONS

Securities innovation can add value in the following ways:

- Reallocate some form of risk from issuers or investors to other market participants who are either less risk averse or else willing to bear them at a lower cost.
- Increase liquidity.
- Reduce *agency costs* arising from conflicts of interest among the firm's stakeholders.
- Reduce issuers' underwriting fees and other transaction costs.
- Reduce the combined taxes of issuer and investors.
- Circumvent regulatory restrictions or other constraints on investors or issuers.

Risk Reallocation

Most *debt innovations* (see Table 7.1) involve some form of risk reallocation as compared to conventional debt instruments. *Risk reallocation,* as mentioned, adds value by transferring risks to others better able to bear them. It may also be beneficial to design a security that better suits the risk-return preferences of a particular class of investors. Investors with a comparative advantage in bearing certain risks will pay more—or, alternatively, have a lower required return—for innovative securities that allow them to specialize in bearing such risks.

For example, suppose an oil producer issued "oil-indexed" debt with interest payments that rise and fall with oil prices. It might have a lower required return for two reasons: (1) the firm's after-interest cash flows will be more stable than if it issued straight, fixed-rate debt, thereby reducing default risk; and (2) some investors may be seeking a "play" on oil prices not otherwise available in the financial markets.

It is clear that investors are willing to pay more for "scarce" securities they value highly. Financial intermediaries have earned considerable profits by simply buying existing securities, repackaging their cash flows into new securities, and selling the new securities (Ross, 1989). The success of stripped U.S. Treasury and municipal securities (created by separating the coupon payments from the principal repayment) illustrates how the sum of the value of the parts can exceed the whole. The benefits were so great that the U.S. Treasury decided to capture for itself the profits that securities dealers were making from it and began issuing registered Treasury STRIPS (separate trading of registered interest and principal of securities). STRIPS permit the coupon and principal payments to be registered and traded separately. In another example, investment banks purchase portfolios of mortgages from originating institutions and place them in trusts or special purpose corporations. The new entities then issue *mortgage pass-through certificates.* The investment bank gets the difference (with an important exception noted later) between the payments the entity gets and those it pays out. Issuers may be able to capture such benefits for themselves by designing new issues of securities appropriately.

Like mortgage pass-through certificates, *credit card receivable–backed securities* and *automobile loan–backed certificates* are undivided ownership interests in portfolios of credit card receivables and consumer automobile loans. Such securities allow the originator to transfer the loan's interest rate risk and default risk (or at least a portion of it) to others. The investors' required return is lower because of the diversification benefit from the pooling.

Managing Reinvestment Risk

Pension funds face reinvestment risk when they reinvest interest payments received on standard debt securities. *Zero-coupon bonds* were designed in part to appeal to such investors because they eliminate reinvestment risk by having no interest payments to reinvest. Instead, interest compounds over the entire life of the security.

Managing Prepayment Risk

Most mortgages are prepayable at par at the option of the mortgagor. Both *collateralized mortgage obligations (CMOs)* and *stripped mortgage–backed securities* address this "prepayment" risk, which investors in mortgage pass-through certificates find troublesome (Fabozzi, 1989, 1995). CMOs repackage the payment stream from a portfolio of mortgages into several series of debt instruments—sometimes more than five dozen—that are ordered by the repayment of principal. In the simplest form of CMO, each series is to be repaid in full before any principal repayment is made to holders of the next series (see Figure 7.1). By so doing, such a CMO effectively shifts most of the mortgage prepayment risk to the lower-ordered classes, and away from the higher-ordered classes.

CMOs are designed to take advantage of the segmentation and incompleteness of the bond market. Prepaid mortgages are not a problem for money market mutual funds and other short-term investors, whereas pension funds and other long-term investors do not want mortgages prepaid. By carving up the payment stream and prioritizing the right to receive payments, the CMO structure creates fast-pay classes that appeal to short-term investors and slow-pay classes that appeal to long-term investors.

Planned amortization class (PAC) bonds and targeted amortization class (TAC) bonds refined mortgage-backed securities to further reduce prepayment risk (Perlman,

Table 7.1 Selected Debt Innovations

■ **Security** □ **Distinguishing Characteristics** ● **Enhanced Liquidity** ○ **Reduction in Transaction Costs**	■ **Year Issued**	■ **No. of Issues**	■ **Aggregate Proceeds ($B)**	□ **Risk Reallocation** ● **Reduction in Agency Costs** ○ Tax and Other Benefits
■ **Adjustable-Rate Notes and Floating-Rate Notes** □ Coupon rate floats with some index, such as the 91-day Treasury bill rate. ● Price remains closer to par than price of fixed-rate note of same maturity.	12/22/70	55,268	14,636.8	□ Issuer exposed to floating-interest-rate risk but initial rate is lower than for fixed-rate issue.
■ **Bonds Linked to Commodity Price or Commodity Price Index** □ Interest and/or principal linked to a specified commodity price or commodity price index.	04/10/80	63	4.9	□ Issuer assumes commodity price risk in return for lower (minimum) coupon. Serves as a hedge if the issuer produces the commodity. ○ Attractive to investors who would like to speculate in commodity options but cannot, for regulatory reasons, purchase them directly.
■ **Catastrophe Bonds** □ The interest payments or the principal payment, or both, are reduced according to a specified formula if the insurance-company issuer suffers insurance losses from certain specified natural disasters, such as a hurricane or an earthquake. ● More liquid than traditional reinsurance contracts.	01/01/97	69	21.2	□ Investors bear a portion of the risk of loss from the specified natural disasters. The bonds securitize reinsurance. Because natural catastrophic risk has very low correlations with financial risks, catastrophe bonds are potentially attractive for diversification purposes. ● Reinsurance involves significant adverse selection and moral hazard risks for investors.
■ **Collateralized Debt Obligations** □ A portfolio of junk bonds or bank loans, or some of each, is placed in a trust or special-purpose entity, which issues multiple classes of debt obligations that are prioritized with respect to their right to receive payments from the debt pool. ● The CDOs are usually more liquid than the underlying debt instruments.	03/24/83	4,167	1,552.3	□ The senior class(es) are investment-grade. Credit risk can be reallocated among investors in a cost-efficient manner.
■ **Collateralized Mortgage Obligations (CMOs) and Real Estate Mortgage Investment Conduits (REMICs)** □ Mortgage payment stream is divided into several classes, which are prioritized in terms of their right to receive principal payments. ● More liquid than individual mortgages.	04/23/81	13,435	9,082.5	□ Reduction in prepayment risk to classes with prepayment priority. Appeals to different classes of investors; sum of parts can exceed whole.
■ **Commercial Real Estate-Backed Bonds** □ Nonrecourse bonds backed by specified piece (or portfolio) of real estate. ● More liquid than individual mortgages.	03/14/84	2,279	1,396.8	□ Reduced yield due to greater liquidity. ○ Appeals to investors who like to lend against real estate properties.
■ **Credit Card Receivable-Backed Securities** □ Investor buys an undivided interest in a pool of credit card receivables. ● More liquid than individual receivables. ○ Investors could not achieve the same diversification as cheaply on their own.	01/16/87	1,729	938.3	□ Supplemental credit support in the form of a letter of credit, surety bond, limited guarantee, over-collateralization, or senior/subordinated structure.
■ **Credit-Sensitive Notes** □ Coupon rate increases (decreases) if the issuer's credit rating falls (improves). Alternatively, another structure provides that the coupon of an investment-grade note would step up by 200 basis points if the note is downgraded to junk status.	05/19/88	15	3.7	□ Protects the investor against deterioration in the issuer's credit quality because coupon increases when rating declines.

Table 7.1 Selected Debt Innovations (*Continued*)

■ **Deferrable Interest Debentures** ☐ Junior subordinated debentures that pay interest monthly or quarterly and that permit the issuer to defer interest payments for up to five years without triggering a default. Interest compounds during the deferral period.	■ 01/06/95 ■ 27 ■ 4.9 ☐ Investors bear more credit risk than a conventional junior subordinated debt issue would involve but the issuer gets some 'equity' credit from the rating agencies.
■ **Deferred-Interest Debentures** ☐ Debentures that accrue – and do not pay in cash – interest for a period.	■ 09/17/82 ■ 33 ■ 9.3 ☐ Reduces bankruptcy risk during the interest-deferral period.
■ **Direct-Issue Demand Notes** ☐ Notes, which are issued directly to investors, that pay interest at a variable money-market rate and that are repayable on demand by the holder. ○ By eliminating the intermediary, the issuer can borrow more cheaply than issuing commercial paper while paying investors a higher interest rate than money market funds.	■ 01/01/97 ■ 12 ■ 1.0
■ **Dollar BILS** ☐ Floating-rate zero-coupon note with effective interest rate determined retrospectively based on the value of a specified corporate bond index.	■ 08/22/88 ■ 3 ■ 0.5 ☐ Issuer assumes reinvestment risk. Useful for hedging and immunization purposes because Dollar BILS have a zero duration when duration is measured with respect to the specified index.
■ **Dual Coupon Bond/Fixed-Floating-Rate Bonds** ☐ Interest is calculated on a fixed-rate basis during the early life of the bond and on an inverse-floating-rate basis for the bond's remaining life.	■ 11/25/85 ■ 603 ■ 41.8 ☐ Issuer exposed to the risk that interest rates may decrease during the inverse-floating-rate period because the coupon will increase if the specified market benchmark interest rate decreases. ○ Useful for hedging and immunization purposes because of the very long duration.
■ **Dual Currency Bonds** ☐ Interest payable in U.S. dollars but principal payable in a currency other than U.S. dollars.	■ 01/21/83 ■ 1,326 ■ 130.9 ☐ Issuer has foreign-currency risk with respect to principal repayment obligation. Currency swap can hedge this risk and lead, in some cases, to yield reduction. ○ Euroyen-dollar dual currency bonds popular with Japanese investors subject to regulatory restrictions and desiring income in dollars without principal risk.
■ **Euronotes and Euro-Commercial Paper** ☐ Euro-commercial paper is similar to US commercial paper. ○ Corporations invest in each other's paper directly rather than through an intermediary.	■ 03/09/82 ■ 282 ■ 47.4 ○ Elimination of intermediary brings savings lender and borrower can share.
■ **Extendible Notes** ☐ Interest rate adjusts every 2-3 years to a new interest rate the issuer establishes, at which time the note holder also has the option to put the notes back to the issuer. ○ Lower transaction costs than issuing 2 or 3-year notes and rolling them over.	■ 03/09/82 ■ 221 ■ 40.2 ☐ Coupon based on 2-3 year put date, not on final maturity. • Investor has a put option, which provides protection against deterioration in credit quality or below-market coupon rate.
■ **Fixed-Rate Capital Securities** ☐ Long-term subordinated debt is issued to a trust or a special-purpose company wholly owned by the parent. This entity issues cash-matching preferred stock. The parent can defer interest payments for up to five years without triggering a default but interest compounds during the deferral period.	■ 10/27/93 ■ 226 ■ 58.2 ☐ Investors bear more credit risk than a conventional subordinated debt issue would involve but the issuer gets some 'equity' credit from the rating agencies. ○ Parent company can deduct the interest payments on the underlying subordinated debt.

(*Continued*)

Table 7.1 Selected Debt Innovations (*Continued*)

■ **Floating-Rate, Rating-Sensitive Notes** ☐ Coupon rate resets quarterly based on a spread over LIBOR. Spread increases if the issuer's debt rating declines.	■ 06/28/88 ■ 9 ■ 1.6 ☐ Issuer exposed to floating-interest-rate risk but the initial rate is lower than for fixed-rate issue.
• Price remains closer to par than the price of a fixed-rate note of the same maturity.	• Investor protected against deterioration in the issuer's credit quality because of increase in coupon rate when rating declines.
■ **Global Bonds** ☐ Debt issue structured so as to qualify for simultaneous issuance and subsequent trading in U.S., European, and Japanese bond markets. • Structure facilitates a relatively large issue. Simultaneous trading in U.S., Europe, and Japan coupled with large size enhance liquidity.	■ 06/14/89 ■ 1,021 ■ 1,225.7
■ **Increasing-Rate Notes** ☐ Coupon rate increases by specified amounts at specified intervals.	■ 01/21/88 ■ 59 ■ 13.2 ☐ Defers portion of interest expense to later years, which increases duration.
	• When issued with bridge financing, step-up in coupon rate compensates investors for the issuer's failure to redeem the notes on schedule.
■ **Indexed-Currency-Option Notes** ☐ Issuer pays reduced principal at maturity if specified foreign currency appreciates sufficiently relative to the U.S. dollar.	■ 10/24/85 ■ 27 ■ 0.7 ☐ Investor assumes foreign-currency risk by effectively selling the issuer a call option denominated in the foreign currency.
	○ For investors who would like to speculate in foreign currencies but cannot purchase currency options directly.
■ **Indexed Sinking-Fund Debentures** ☐ The amount of each sinking-fund payment is indexed to a specified interest-rate index (typically the 10-year constant maturity Treasury yield).	■ 07/12/88 ■ 38 ■ 6.3 ☐ The security's duration and convexity are closer to those of a fixed-rate mortgage than a conventional fixed-rate bond: so it is useful to financial institutions that invest in mortgages for duration-matching purposes.
■ **Interest-Rate-Reset Notes** ☐ Interest rate is reset 3 years after issuance to the greater of (i) the initial rate and (ii) a rate sufficient to give the notes a market value equal to 101% of their face amount.	■ 06/01/83 ■ 71 ■ 13.7 • Reduces the (initial) yield due to the reduction in agency costs.
	○ Investor is compensated for a deterioration in the issuer's credit standing within 3 years of issuance.
■ **Loan-Backed Certificates** ☐ Investor buys an undivided interest in a pool of automobile, manufactured housing, residential second-lien, or other consumer loans.	■ 05/15/85 ■ 4,190 ■ 2,029.4 ☐ Supplemental credit support in the form of letter of credit, limited guarantee, surety bond, over-collateralization, or senior/subordinated structure. Provider of credit support bears residual default risk. Reduced yield due to the benefit to the investor of credit support, diversification, and greater liquidity.
• More liquid than individual loans. ○ Investors could not achieve the same diversification as cheaply on their own.	○ Can be structured as a sale of assets to remove loans from balance sheet.
■ **Make-Whole Call Provision** ☐ The call price is equal to the greater of (1) par value and (2) the present value of the remaining debt service payments discounted at a specified (small) spread to the yield on a Treasury security of like duration.	■ 11/01/91 ■ 3,657 ■ 1,850.2 ☐ Investors face less call risk than a conventional bond redemption schedule involves but the issuer has correspondingly less refunding flexibility.
■ **Medium-Term Notes** ☐ Notes are sold in varying amounts and in varying maturities on an agency basis.	■ 04/17/73 ■ 35,147 ■ 5,711.4 ☐ Issuer bears market price risk during the marketing process.
○ Agents' commissions are lower than underwriting spreads. ■ **Mortgage-Backed Bonds** ☐ Bonds issued by financial institutions (or other borrowers) that are collateralized by a specified pool of mortgages.	■ 12/16/70 ■ 1,083 ■ 958.6 ☐ Collateral provides added security to the investors making possible a lower interest rate than an unsecured issue of like maturity.

Table 7.1 Selected Debt Innovations (*Continued*)

■ **Mortgage Pass-Through Certificates** ☐ Investor buys an undivided interest in a pool of mortgages.	■ 09/21/77 ■ 1,324 ■ 993.1 ☐ Supplemental credit support in the form of a letter of credit, surety bond, limited guarantee, senior/subordinated structure, insurance, or a reserve fund. Provider of credit support bears residual default risk. Reduced yield due to the benefit to the investor of credit support, diversification, and greater liquidity.
• More liquid than individual mortgages. ○ Investors could not achieve the same diversification as cheaply on their own.	○ Can be structured as a sale of assets to remove loans from the balance sheet.
■ **Negotiable Certificates of Deposit** ☐ Certificates of deposit are registered and sold to the public on an agency basis.	■ 07/10/79 ■ 11,675 ■ 836.6 ☐ Issuer bears market price risk during the marketing process.
• More liquid than non-negotiable CDs. ○ Agents' commissions are lower than underwriting spreads.	
■ **Pay-in-Kind Debentures/Variable Duration Notes** ☐ Debentures on which the interest payments can be made in cash or additional debentures, at the option of the issuer. Variable duration notes give the issuer this option throughout the life of the security.	■ 09/18/87 ■ 5 ■ 1.4 ☐ Defers the risk that the issuer will not be able to make timely debt service payments. Reduces bankruptcy risk during the pay-in-kind period.
■ **Principal-Exchange-Rate-Linked Securities** ☐ Principal repayment is linked to a specified foreign exchange rate. Amount of repayment in U.S. dollars increases (decreases) as the specified foreign currency appreciates (depreciates) relative to the dollar.	■ 03/12/87 ■ 24 ■ 1.5 ☐ Investor has effectively purchased a call option on the specified foreign currency and sold a put option on the same currency.
	○ Attractive to investors who would like to speculate in foreign currencies but cannot purchase currency options directly.
■ **Puttable Bonds** ☐ Bond redeemable at holder's option, or in the case of "poison put" bonds, if a certain specified "event" occurs.	■ 08/16/73 ■ 6,917 ■ 1,194.8 ☐ Option to redeem benefits holders if interest rates rise.
	• Put option provides protection against deterioration in the issuer's credit standing.
■ **Real Yield Securities/Inflation-Indexed Bonds** ☐ Coupon rate resets quarterly to the greater of (i) change in consumer price index plus the "Real Yield Spread" (3.0% in the first such issue) and (ii) the Real Yield Spread, in each case on a semi-annual-equivalent basis.	■ 01/20/88 ■ 3 ■ 0.5 ☐ Issuer exposed to inflation risk, which may be hedged in the CPI futures market. Real yield securities have a longer duration than alternative inflation-hedging instruments.
• Real yield securities could become more liquid than CPI futures, which tended to trade in significant volume only around the monthly CPI announcement date. ○ Investors obtain a long-dated inflation-hedging instrument that they could not create as cheaply on their own.	
■ **Remarketed Reset Notes** ☐ Interest rate reset at the end of each interest period to a rate the remarketing agent determines will make the notes worth par. If issuer and remarketing agent can not agree on rate, then the coupon rate is determined by formula, which dictates a higher rate the lower the issuer's credit standing.	■ 12/15/87 ■ 16 ■ 3.8 ☐ Coupon based on length of interest period, not on final maturity.
• Designed to trade closer to par value than a floating-rate note with a fixed interest-rate formula.	• Investors have a put option, which protects against issuer and remarketing agent agreeing to set a below-market coupon rate; flexible interest-rate formula protects investors against deterioration in issuer's credit standing.
○ Intended to have lower transaction costs than auction-rate notes and debentures, which require periodic Dutch auctions.	

(*Continued*)

Table 7.1 Selected Debt Innovations (*Continued*)

■ **Reverse Principal-Exchange-Rate-Linked Securities** ☐ Principal repayment is linked to a specified foreign exchange rate. Amount of repayment in U.S. dollars increases (decreases) as the dollar appreciates (depreciates) relative to the specified foreign currency.	■ 10/03/88	■ 24	■ 1.5	☐ Issuer has effectively purchased a call option on the specified foreign currency and sold a put option on the same currency. ○ For investors who would like to speculate in foreign currencies but cannot purchase currency options directly.
■ **Spread-Adjusted Notes** ☐ The interest-rate spread off a specified Treasury benchmark yield is reset on each interest payment date through a Dutch auction.	■ 05/08/91	■ 1	■ 0.5	☐ Investor protected against credit risk but, unlike conventional auction-rate debt, is still exposed to interest-rate risk. • Interest-rate spread off Treasury benchmark yield will increase if issuer's credit standing deteriorates – whether or not issuer's credit rating changes.
■ **Spread-Protected Debt Securities** ☐ The notes can be redeemed on a specified date (in one case, 2 years after issuance) prior to maturity, at the option of the holders, at a price equal to the present value of the remaining debt service stream calculated on the exercise date by discounting the future debt service payments at a rate equal to a specified Treasury benchmark yield plus a fixed spread.	■ 01/15/87	■ 1	■ 0.1	☐ Investor protected against credit risk up until the put date but is not protected against interest rate risk. • Investor has a put option, which provides protection against deterioration in the issuer's credit standing prior to the put date.
■ **Standard & Poor's 500 Index Notes (SPINs)/Stock Index Growth Notes (SIGNs)/Equity-Indexed Notes** ☐ Zero-coupon note, principal payment on which is linked to appreciation in value of specified share price index above a specified threshold. ○ Cheaper than buying a combination of a zero-coupon note and rolling over a series of shorter-term options.	■ 11/21/85	■ 506	■ 22.6	☐ Equivalent to a package consisting of a zero-coupon bond and a long-dated call option on a specified share price index.
■ **Step-Down Floating-Rate Notes** ☐ Floating-rate notes on which the interest margin over the specified benchmark (e.g., 30-day high-grade commercial paper rate) steps down to a smaller margin on a specified date during the life of the instrument.	■ 07/11/88	■ 28	■ 2.8	☐ Designed to reduce interest rate margin to reflect direct dependence of required margin on remaining maturity of notes.
■ **Step-Up Callable Bonds** ☐ Long-term bonds with an interest rate that steps up if the issue is not called on a specified date, in one case 10 years after issue. Thereafter, the interest rate floats.	■ 07/26/89	■ 1,393	■ 142.2	■ Step up in interest rate at least partially compensates investors if the issuer's credit standing declines and the issuer fails to redeem the bonds.
■ **Stripped Mortgage-Backed Securities** ☐ Mortgage payment stream subdivided into two classes: one with below-market coupon and the other with above-market coupon, or one receiving interest only and the other receiving principal only, from a pool of mortgages.	■ 07/08/86	■ 58	■ 0.0	☐ Securities have unique option characteristics that make them useful for hedging purposes. Designed to appeal to different classes of investors; sum of the parts can exceed the whole.
■ **Super Premium Notes** ☐ Intermediate-term U.S. agency debt instrument (typically maturing in between 1 and 3 years) that carries a coupon rate well above current market rates (and therefore sells at significant premium to its face amount).	■ 11/18/88	■ 11	■ 0.3	☐ Attractive to government bond funds that would like to report very high-coupon debt in their portfolios and do not have to amortize the premium over the life of the instrument (or in some cases, money market mutual funds that do not have to show a capital loss even at redemption). As a result, Super Premium Notes provide a lower cost of funds than conventional U.S. agency notes of like maturity.

Table 7.1 Selected Debt Innovations (*Continued*)

■ **Tobacco Asset Securitization Bonds** □ Political jurisdictions, such as New York City, that are entitled to share in the proceeds of the Master Settlement Agreement (MSA) entered into by four cigarette manufacturers, 46 states, and six other jurisdictions to settle certain smoking-related litigation have issued bonds to securitize their right to receive initial and annual payments under the MSA.	■ 11/05/99 ■ 17 ■ 3.2 □ Investors are exposed to settlement-related credit risk although the senior/subordinated structure and reserve accounts provide strong protection to senior bond holders.
• The issuer can monetize its right to receive future settlement payments in order to use the proceeds to fund capital expenditures or meet its other funding needs currently.	
■ **Variable-Coupon Renewable Notes** □ Coupon rate varies weekly and equals a fixed spread over the 91-day T-bill rate. Each 91 days the maturity extends another 91 days. If put option exercised, spread is reduced.	■ 02/02/88 ■ 9 ■ 5.1 □ Coupon based on 1-year termination date, not on final maturity.
○ Lower transaction costs than issuing 1-year note and rolling it over.	○ Designed to appeal to money market mutual funds, which face tight investment restrictions, and to discourage put to issuer.
■ **Variable-Rate Renewable Notes** □ Coupon rate varies monthly and equals a fixed spread over the 1-month commercial paper rate. Each quarter the maturity automatically extends an additional quarter unless the investor elects to terminate the extension.	■ 02/02/88 ■ 31 ■ 7.5 □ Coupon based on 1-year termination date, not on final maturity.
○ Lower transaction costs than issuing 1-year note and rolling it over.	○ Designed to appeal to money market mutual funds, which face tight investment restrictions.
■ **Yield Curve Notes/Maximum-Rate Notes/Inverse-Floating-Rate Notes** □ Interest rate equals a specified rate minus LIBOR	■ 11/18/85 ■ 82 ■ 5.7 □ Issuer exposed to the risk that interest rates may decrease, which would raise the coupon. Can reduce yield relative to conventional debt when coupled with an interest rate swap against LIBOR.
	○ Useful for hedging and immunization purposes because of long duration.
■ **Zero-Coupon Bonds (sometimes issued in series)** □ Non-interest-bearing. Payment in one lump sum at maturity.	■ 04/22/81 ■ 5,634 ■ 585.6 □ Issuer assumes reinvestment risk. Issues sold in Japan carried below-taxable-market yields reflecting tax advantage over conventional debt.
	□ Straight-line amortization of original issue discount pre-TEFRA. Japanese investors realized significant tax savings.

Source: Reprinted by permission of Thomson Financial.

1989). PAC bonds repay principal according to a given schedule so long as pool prepayments remain within a specified range, for example, 100% PSA to 300% PSA. (The PSA Standard Prepayment Model assumes that mortgages prepay at the annualized rate of 0.2% during the first month, the prepayment rate steps up in increments of 0.2% each month until month 30, and prepayments remain constant at 6% per year for all succeeding months.) Thus, they provide a predictable cash flow over a wide range of interest rate scenarios. It is important to appreciate that prepayment risk is not eliminated; it is reallocated, and other pool classes—companion classes—acting as prepayment "shock absorbers," bear most of the pool's prepayment risk.

TAC bonds evolved from PAC bonds. Figure 7.2 shows how PACs and TACs reallocate prepayment risk. They are targeted to a narrower range of prepayment rates. Both rely on companion classes to function as prepayment shock absorbers. PACs get two shock absorbers and TACs get one.

Stripped mortgage–backed securities or mortgage strips divide the payment stream from a pool of mortgages (or mortgage securities) into two (or in some cases more than two) securities. The IO strips get all the interest and the PO strips get all the principal. When prepayments accelerate, PO strips, which have large positive durations, increase in value because principal is received sooner. IO strips, which have large negative durations, decrease in value because the total interest paid is reduced. Understandably, both have high price volatility.

The introduction of these securities also enhanced market completeness because of their duration and

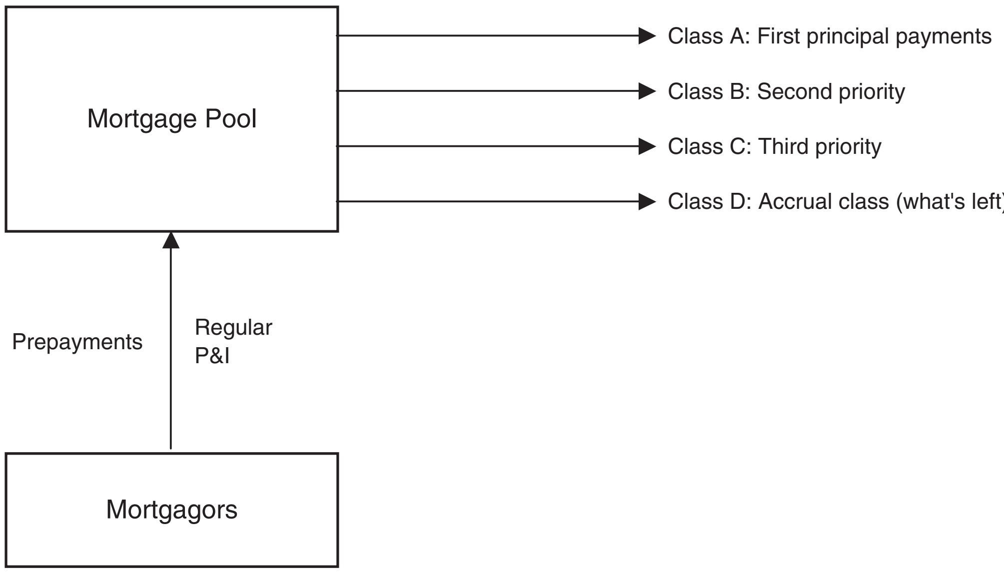


Figure 7.1 How a CMO Works

convexity characteristics. PO strips are recognized as "bullish" investments because a decrease in interest rates, which benefits bond prices generally, tends to boost prepayments and hence PO strip prices. IO strips are recognized as "bearish" investments because their prices move in the opposite direction. Mortgage strips are ideal for speculation on prepayment rates or hedging prepayment risk. For example, the risk that a slower prepayment rate (say, due to an increase in interest rates) will reduce the value of a mortgage portfolio can be hedged by buying IO strips. Those wanting to protect a portfolio of high-coupon mortgages against rising prepayments can purchase high-coupon PO strips.

Another innovation for managing prepayment risk is the *make-whole call provision*, which is an indexed bond call option (Emery, Hoffmeister, and Spahr, 1987). The strike price is indexed to the yield on a comparable duration Treasury bond. The make-whole call provision reduces the security holder's (rather than the issuer's) prepayment risk. Quaker Oats was the first to use this provision in a public debt issue, in October 1995. Since then, it has grown to become virtually standard for investment-grade corporate bonds (Mann and Powers, 2001). For the most part, non-investment-grade corporate bonds that include a call provision have continued to use the fixed-price provision.

Managing Default Risk

Asset-backed securities reduce default risk through diversification. They can also reallocate default risk by prioritizing the right to receive payments from a portfolio of risky securities, such as bonds or bank loans.

An issuer of asset-backed securities, such as a mortgage, automobile, or credit card lender, can retain default risk by providing a limited corporate guarantee, overcollateralizing (by pledging additional assets), or taking a subordinated position. Also, the issuer of the asset-backed security can purchase a guarantee, letter of credit, surety bond, or similar promise of payment from a creditworthy third party.

There are several possible structures for reallocating default risk. With the pass-through structure, the loans are transferred to a special purpose entity (SPE) in

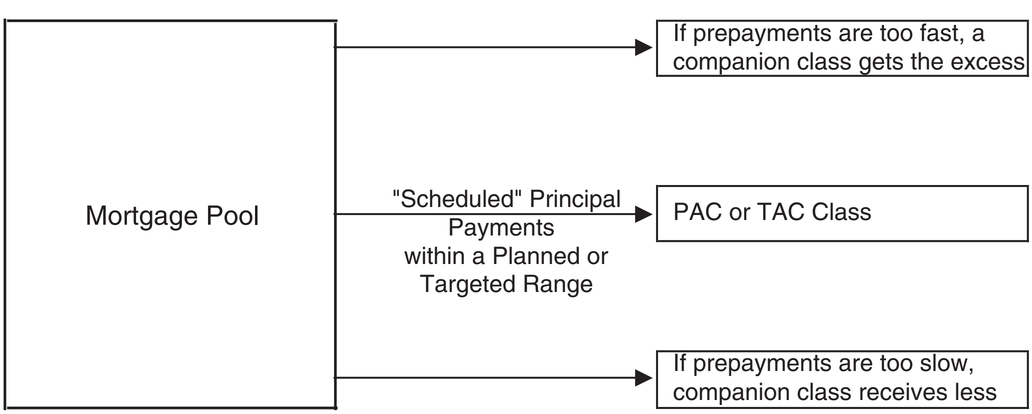


Figure 7.2 How PACs and TACs Reallocate Mortgage Prepayment Risk

exchange for pass-through certificates, which are proportional ownership claims that are sold to investors. The originator/seller can overcollateralize by selling less than 100% of the claims. All payments are collected by the trustee and are passed through to investors on the specified payment dates. If needed, the trustee can draw on credit enhancement mechanisms, such as a reserve fund or a line of credit. The pass-through structure has been used to securitize residential mortgage loans, commercial mortgage loans, automobile installment loans, credit card receivables, recreational vehicle installment loans, equipment leases, boat installment loans, manufactured housing installment loans, and home equity loans.

With the pay-through structure, the loans are transferred to an SPE and provide collateral for SPE-issued notes. The pay-through structure permits the seller/servicer to modify the cash flows received on the underlying collateral. The restructuring is particularly useful when nonmarket, incentive interest rate automobile loans are securitized. Such a structure has been utilized to provide automobile loan–backed and truck loan–backed securities with fixed sinking-fund schedules, which transfers the certificate holder's exposure to prepayment risk to the holders of the SPE's residual interest.

Credit card receivables are handled differently. Because of high payment rates, a credit card receivable–backed issue would mature within about a year if all the payments were passed through. To create an intermediate-term security, credit card receivable–backed certificates typically pay only interest for a specified period, normally 24 to 60 months. Principal payments received during this revolving period are reinvested in newly generated credit card receivables. Typically more receivables are sold to the trust than are sold to investors. The seller/servicer is required to retain a specified minimum ownership of the difference, typically between 5% and 10%.

Senior-Subordinated Structure for Reallocating Default Risk of Asset-Backed Securities With senior/subordinated mortgage-backed securities, senior certificate holders have first claim on the trust's cash receipts. A shortfall in payments to senior certificate holders is prespecified to be handled by one of two possible structures. In the *reserve-fund structure,* the originator establishes a reserve fund either fully at issuance or else by contributing between 0.10% and 0.50% of the original pool balance at issuance and then capturing the excess yield spread until the reserve is fully funded. In the latter case, subordinated certificate holders get no cash until the fund builds up to a prespecified level, typically 1% to 2% of the pool balance for mortgages, and up to 4% to 5% for automobile loans. The reserve fund is then used to cover any senior certificate payment shortfall. Capturing excess spread in a reserve fund is a key form of credit enhancement in agency MBS deals. In the *shifting-interest structure,* shortfalls are met by transferring the amount from the subordinated certificate ownership percentage to the senior certificate ownership percentage. Figure 7.3 compares the two senior-subordinated structures. The shifting-interest structure is usually found only in retail MBS deals. Of course, the reserve-fund structure insures that cash flows to senior certificate holders will be as promised, whereas the shifting-interest structure compensates senior certificate holders through a change in future distributions.

Collateralized Debt Obligations *Collateralized debt obligations (CDOs)* consist of multiple bond classes and an equity class which are backed by a portfolio of debt obligations, such as bonds (in which case they are sometimes referred to as *collateralized bond obligations*) or bank loans (in which case they are referred to as *collateralized loan obligations*). The issuer is a bankruptcy-remote trust. The securities are prioritized with respect to the cash flows from the underlying portfolio in decreasing order: senior tranches,

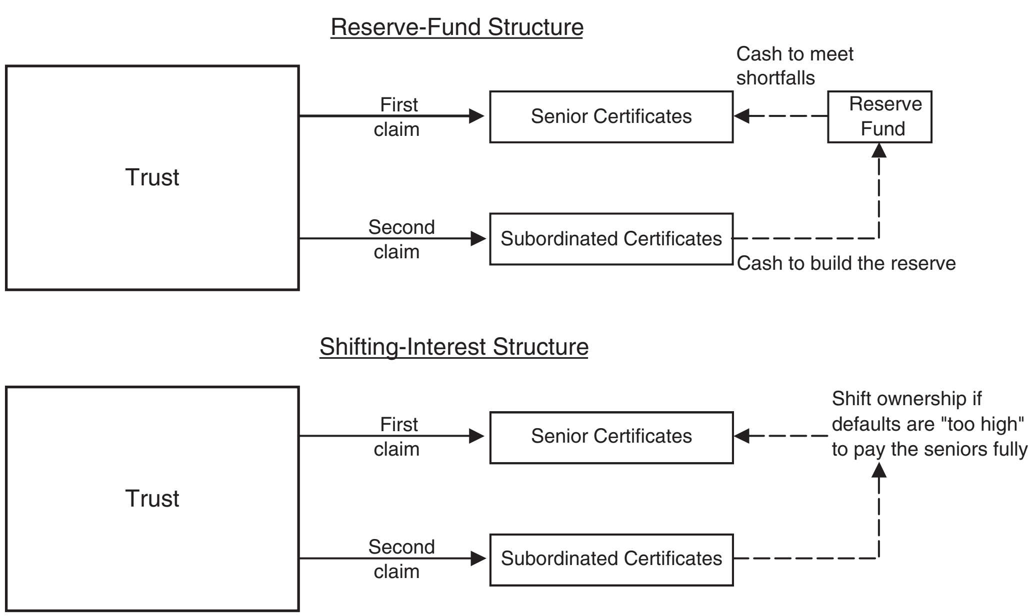


Figure 7.3 Senior/Subordinated Securitization Structures for Managing Default Risk

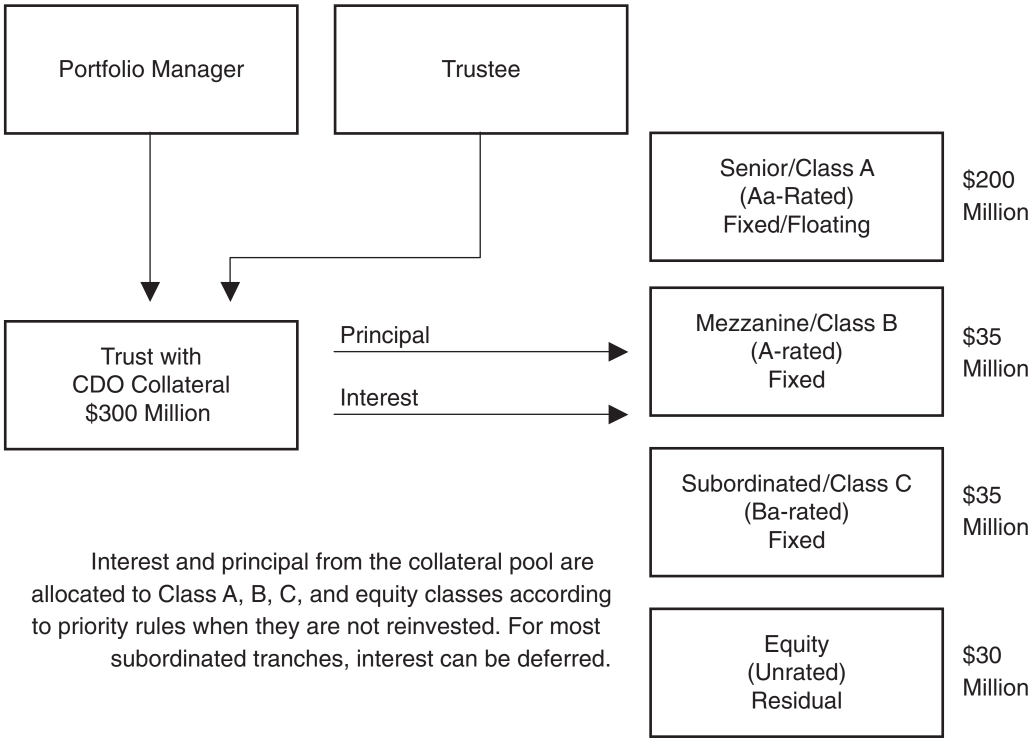


Figure 7.4 Illustration of a Cash Flow CDO Structure

mezzanine tranches, subordinated tranches, and equity tranche, as illustrated in Figure 7.4. The equity tranche takes the first loss and receives all residual cash flows after the payment of all expenses such as administrative and management fees and payments to debt classes in the structure.

CDOs do for default risk what CMOs do for prepayment risk. In both cases, the structure provides a diversification benefit. It also enables investors to choose securities that are most suitable for their specific risk/return preferences.

CDOs are of three basic types, cash flow, market value, and synthetic (*CDO Primer,* 2004; Lucas, Goodman, and Fabozzi, 2006). The cash flow structure is built around a portfolio of regular-cash-paying debt instruments. It can handle a great variety of liability types. The portfolio is usually actively managed so as to eliminate liabilities whose credit is deteriorating. The synthetic structure uses credit default swaps or basket default swaps to replicate a cash flow CDO. A *basket default swap* has a portfolio of assets as the underlying, whereas it is usually a single asset with a credit default swap. The portfolio is either static or else allows, at most, limited substitutions. The synthetic structure is more cost effective than the cash-flow variety when designed to take advantage of favorable pricing in the swap market. The market value structure is designed to take advantage of anticipated security price volatility. It has been used to securitize hedge fund and private equity investments.

Pay-in-Kind Debentures (PIKs)

Pay-in-kind debentures (PIKs) were popular in the 1980s in financing leveraged buyouts. The firm was permitted to issue additional notes (identical to the original notes), rather than pay interest in cash, but could still claim the interest tax deduction. The firm could defer paying cash until it had paid down its more senior debt, thus achieving higher leverage. A tax change in 1986 made it very difficult to get the tax deduction, and PIK issuance dried up. PIKs recently made a comeback in a slightly different form, called *toggle PIKs.* The firm can "toggle" back and forth between paying cash and issuing additional notes, but if it chooses to issue notes, the coupon is higher than if it pays cash. Toggling back to cash at the right time could avoid the tax problem that initially curtailed their issuance. However, investors eventually began to object to the agency problem inherent in the toggle feature: There was nothing to prevent a dicey credit from flipping the note toggle immediately and never switching back.

Managing Interest Rate Risk

Adjustable-rate notes and *floating-rate notes (FRNs)* are among the many innovative debt securities that manage interest rate risk. They reduce the security holder's principal risk by transferring interest rate risk to the issuer. This reallocation can benefit the issuer when the value of its assets is directly correlated with interest rate changes. As a result, banks and credit card companies are frequent issuers of FRNs.

The coupon rate of a typical FRN resets each period at a reference rate, often LIBOR (London Interbank Offer Rate), plus a fixed margin. The margin reflects default risk and other characteristics of the security, including any call or put options. The margin is greater (smaller) (1) the

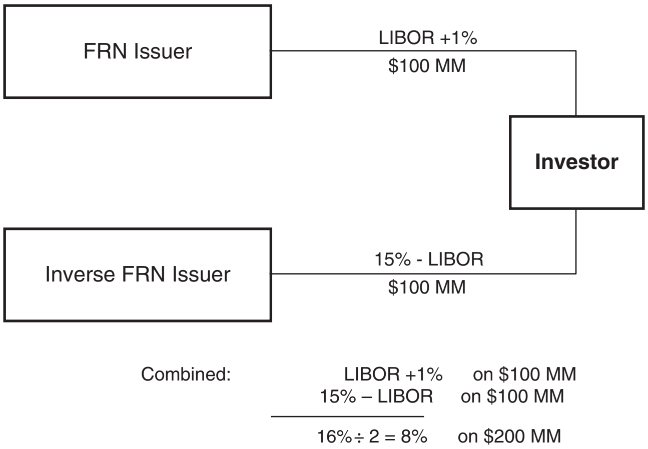


Figure 7.5 Inverse FRNs Can Hedge Interest Rate Risk

greater (smaller) the default risk or (2) the lower (higher) the maximum (minimum) coupon rate. A call option requires a greater margin, a put option a smaller margin.

Inverse FRNs

Inverse FRNs were introduced in 1986 (Ogden, 1987; Smith, 1988). Depending on the sponsoring investment banker, they are variously known as *inverse floaters, reverse floaters, bull floaters, yield curve notes,* or *maximum rate notes.* They have a coupon reset formula in which the London Interbank Offered Rate (LIBOR) is subtracted from a fixed percentage rate, such as 15% − LIBOR. As LIBOR increases (decreases), the coupon payment on an inverse FRN decreases (increases). Inverse FRNs are designed for hedging exposure to floating interest rates (or speculating on interest rate movement), because its payments offset those of the traditional FRN. Therefore, borrowing (or loaning) equal amounts of floating-rate and inverse floating-rate securities produces an approximately fixed rate.

Figure 7.5 illustrates how inverse FRNs can be used to hedge interest rate risk. Combining an inverse FRN paying 15% − LIBOR with an equal principal amount of an FRN paying LIBOR + 1% produces a synthetic fixed-rate note paying 8%.

Managing Price and Exchange Rate Risks

Commodity-linked bonds offer a hedge against volatile prices. The principal repayment and, in some cases, the coupon payments are tied to the price of a particular commodity, such as the price of oil or silver, or a specified commodity price index. A commodity producer's revenues tend to rise and fall with the commodity's price. Therefore, by having debt payments move with the commodity's price, the security reduces the volatility of the firm's (after-interest) cash flow. The security shifts the debt cost from times when the commodity producer is least able to pay to periods when it is most able to do so.

Dual-currency bonds, indexed-currency option notes, principal exchange rate–linked securities (PERLs), and reverse-principal exchange rate–linked securities (reverse PERLs) illustrate different forms of currency risk reallocation. They allow institutions wanting to deal in foreign currencies, but for regulatory or other reasons cannot, to purchase currency options.

Managing Business Risks

Alternative risk transfer (ART), also known as structured insurance, consists of a variety of techniques for transferring business risks that the traditional insurance market cannot handle cost effectively, or may not be able to handle at all. Fabozzi, Drake, and Polimeni (2008) describe these useful techniques, which include issuing *insurance-linked notes (ILNs).*

Catastrophe bonds are insurance-linked notes that are linked to a single peril. Other ILNs are multiperil bonds. Cat bonds reduce the interest or principal payments if the issuer suffers losses from a specified type of natural disaster, such as a hurricane or earthquake. For example, the Japanese owner of Tokyo Disneyland issued a $200 million cat bond in 1999 to insure against earthquake damage to the park. Cat bonds are issued by life and property and casualty insurers as an alternative to traditional reinsurance contracts. Investors bear a portion of the risk of loss from the specified natural disaster. Cat bonds are attractive for diversification purposes because natural catastrophic risks have very low correlations with financial risks. Therefore, the risk investors bear, and the interest rate they demand, is relatively low. Cat bonds have also been issued to cover difficult-to-insure risks. For example, the sponsors of the 2006 World Cup issued a $260 million cat bond to insure against a terror-related cancellation of the tournament.

Enhanced Liquidity

If a firm can securitize an asset or a loan so that it becomes publicly traded, the greater liquidity lowers the required return. Examples of *asset securitization* include residential mortgage-backed securities, commercial mortgage-backed securities, credit card receivable–backed securities, and automobile loan–backed certificates. All are publicly registered securities with yields significantly lower than those on the underlying assets.

Global bonds are structured to qualify for simultaneous issuance and subsequent trading in the U.S., European, and Japanese bond markets. Global bonds have greater liquidity than single-market debt issues. Catastrophe bonds securitize reinsurance contracts, and are more liquid because they are tradable. They are a good example of the trend toward securitizing contracts to appeal to a broader class of investors, including those who are sensitive to liquidity risk.

The tobacco asset securitizations are another good example of how securitization can improve liquidity. In November 1998, four cigarette manufacturers reached a historic settlement of smoking-related litigation with 46 states and six other political jurisdictions. The settlement obligates the manufacturers to pay more than $200 billion over about 25 years. New York City sold $709 million worth of tobacco flexible amortization bonds (TFABs) in November 1999 to monetize part of its share in the settlement and obtain funds for its capital program. It used a senior/subordinated structure and reserve accounts to obtain a single-A rating for the senior bonds.

Reductions in Agency Costs

A new security can add value by reducing agency costs arising out of conflicts of interest among managers, stockholders, and bondholders. For example, debt claim dilution from a leveraged buyout (LBO) may increase shareholder value at the expense of bondholders. *Interest rate reset notes* address this problem by providing bondholders with protection against a drop in the issuer's credit standing. Similarly, *credit-sensitive notes* and *floating-rate, rating-sensitive notes* bear a coupon rate that varies inversely with the issuer's credit standing. All of these securities, however, have a potentially serious flaw. The interest rate adjustment mechanism will tend to increase the issuer's debt service burden just when it can least afford it—when its credit rating has fallen, presumably as a result of diminished operating cash flow.

Puttable bonds provide a series of put options that protect bondholders against deterioration in the issuer's credit standing. Others give the right to put the bonds back to the issuer if there is a change in corporate control or a leverage increase above a specified level. Such poison-put options protect bondholders against event risk.

Increasing-rate notes provide an incentive for the issuer to redeem the notes on schedule. They have been used as bridge financing and replaced by more permanent financing. Unfortunately, the increasing-rate feature can have a perverse effect. In the extreme, the increasing-rate feature can require such a large jump in the interest rate that the issuer is forced into bankruptcy.

Reductions in Transaction Costs

A number of innovative debt securities increase stockholder value by reducing the underwriting commissions and other transaction costs associated with raising capital. *Extendible notes, variable-coupon renewable notes, puttable bonds,* and *euronotes* and *euro-commercial paper* do this with issuer or investor options to extend the maturity. For example, the early extendible notes typically included an interest rate adjustment every two or three years, which avoids the rolling-over refinancing expense. Refinements of this concept, such as certain puttable bonds and remarketed reset notes, give the issuer greater flexibility in resetting the terms of the security.

The more recent generation of extendible notes is designed to qualify as investments for money market mutual funds. Money funds in the United States are restricted to securities with a maturity of 366 days or less. The notes pay interest monthly for added appeal to these funds. However, issuers would prefer longer maturities to reduce their refunding risk. Investors can elect to extend the maturity of the extendible notes by one month on each interest payment date. The notes provide that the coupon will step up periodically. This feature gives investors a strong incentive to exercise the extension option unless the firm's credit deteriorates.

Euronotes and euro-commercial paper extended commercial paper to the Euromarket. Commercial paper reduces transaction costs by allowing corporations to invest directly in one another's securities, without an intermediary.

Reducing Investor Transaction Costs

Innovative securities can reduce investor transaction costs to the extent they offer a lower-cost way of obtaining investor goals, such as diversification.

Reductions in Taxes

Securities that reduce the total amount of taxes paid by the firm and its investors can add value. Such tax arbitrage occurs when a firm issues debt to investors who have a marginal tax rate on interest income that is lower than the firm's marginal income tax rate. For example, *premium bonds,* with above-market interest rates, have been issued in exchange for outstanding high-coupon bonds to create a tax arbitrage (Finnerty, 2001). The exchange is essentially a refunding of high-coupon bonds that preserves debt-service parity. A taxable bondholder can amortize the premium (either new issue or market) over the remaining life of the bond, which conveys tax benefits. The issuer also benefits because the premium paid to retire the old debt is tax deductible.

Tax-Trading Strategies

Flexibility in recognizing capital gains and losses for tax purposes creates valuable tax-timing options. Evidence shows that the value of tax-timing options represented between 3.05% and 11.55% of the prices of bonds issued at close to par, between 1.42% and 4.93% of the prices of bonds issued at a moderate discount, and between 4.23% and 13.45% of the prices of zero-coupon bonds, depending on the investor's tax situation (Sims, 1992; Strnad, 1995). The tax rules for original issue discount (OID) bonds make the tax-timing options of deeper-discount bonds more valuable.

Figure 7.6 illustrates the tax arbitrage opportunity that existed at one time under the U.S. Internal Revenue Code. The firm was permitted to deduct interest faster than it compounded in the early years. It correspondingly deducted it more slowly in the later years. But the front-end loading of the interest deductions was beneficial because of the time value of money. There was no offsetting acceleration in income tax liability when the zero-coupon bonds were sold to pension funds.

Debt Maturity and the Tax-Deductibility of Interest

The Internal Revenue Code requires that a bond have a stated maturity date or else give bondholders the right to force redemption (that is, a put option) on a stated date

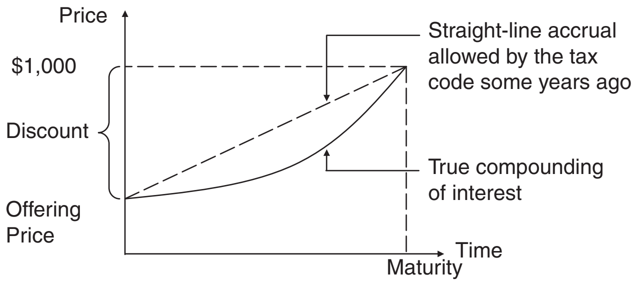


Figure 7.6 Illustration of Zero-Coupon Bonds' Tax Arbitrage

in order for the issuer to deduct the interest payments for tax purposes. Some foreign jurisdictions allow tax-deductibility even for perpetual bonds. In 1995, after U.S. firms began issuing 50-year and 100-year bonds, so-called *supermaturity bonds,* the Treasury Department proposed eliminating the tax deduction for interest on greater-than-40-year bonds but it did not become law. The Treasury Department estimated that eliminating the tax deduction for interest on supermaturity bonds could raise a total of $6.5 billion over 7 years. Understandably, others predicted that firms would simply stop issuing supermaturity bonds. As predicted, with the proposed tax change in 1995, Monsanto Company immediately canceled its previously announced plans to issue $200 million face amount of 100-year bonds.

Circumvention of Regulatory Restrictions or Other Constraints

Bank regulations for debt instruments to qualify as *primary capital* have changed several times since the mid-1980s. Banks have responded predictably with new debt securities designed primarily to meet such regulations. Examples include *equity contract notes,* which subsequently convert into the common stock of the bank (or its holding company); *equity commitment notes,* which the issuer (or its parent) commits to refinance by issuing securities that qualify as capital; and sinking-fund debentures that pay sinking fund amounts in common stock rather than cash.

Variable-coupon renewable notes are a refinement of the extendible note concept discussed earlier. The maturity of the notes automatically extends 91 days at the end of each quarter—unless the holder elects to terminate the automatic extension, in which case the interest rate spread decreases. The reduction in spread can be avoided by selling the notes. These features were designed to meet investment regulations on money market mutual funds at the time.

Commodity-linked bonds allow institutions to speculate in commodity options, or invest in them as an inflation hedge, when regulatory or other reasons disallow them from purchasing commodity options directly. Similarly, bonds with interest or principal payments tied to a foreign exchange rate or denominated in a foreign currency allow institutions to speculate in foreign currencies, or invest in them as a hedge, when they cannot make such investments directly. *Indexed-currency option notes* and many other securities developed in the 1980s and 1990s that contain embedded commodity options or currency options of various forms were motivated by a desire to circumvent regulatory restrictions.

Covered bonds are debt instruments that are issued by European banks and that enjoy special regulatory treatment. They are overcollateralized by sound assets, such as residential mortgage loans or public sector bonds. They are issued under special legal provisions that insulate them from the issuing bank's default on its other debt even though the debt and the assets that collateralize them remain on the bank's balance sheet. Their bankruptcy remoteness enables them to qualify for a triple-A rating even though the bank's other debt is lower rated. Due to their high credit quality, bank regulators assign them a privileged credit risk weighting, and investment funds and insurance companies are allowed to invest a higher percentage of their assets in the covered bonds than in the other bonds of any single issuer. This favored legal and regulatory treatment has led to enormous growth in this asset class in Europe within the last few years.

STRUCTURED PRODUCTS

Interest rate swaps, fixed-rate notes, and floating-rate notes can be structured to take advantage of specific future movements of interest rates (or a currency or commodity price), usually by including one or more interest rate options. These instruments are known as *structured products* (Crabbe and Argilagos, 1994; Brown and Smith, 1996). Generally, they are designed around an interest rate swap (and called a *structured swap*) or around a fixed-rate or floating-rate note (called a *structured note*). Prior to a recent change in the accounting for derivatives, structured swaps had often been preferred because it was easier to keep them off the balance sheet.

What Structured Products Are Designed to Achieve

Market participants often have a "view" on the direction of interest rates, that is, they have an expectation about where interest rates are headed. But even when there is no explicit expectation and no intended bet on future interest rates, there is an implicit view in every decision to issue or invest in debt. Even a simple decision to invest in back-to-back three-month Treasury bills rather than buying six-month Treasury bills has an implicit view—in this case, that the three-month Treasury bill rate on the rollover date is likely to be above the rate the forward curve now predicts.

Structured products are distinguished, however, by the specificity of the view. The view can be tailored in any number of ways, such as a specific interest rate staying within a particular interest rate band, or the difference between specified short-term and long-term interest rates changing in a particular way. One can bet on the direction of interest rates, changes in the shape of the yield curve, shifts in interest rate volatilities, and so on.

Taking a View on Interest Rates

As noted, inverse FRNs can be used for hedging. On the other side of the transaction, an investor who believes interest rates are going to fall can buy inverse FRNs. Their coupon rises while the required yield falls if interest rates decline. An investor can take a more aggressive view by investing in *leveraged inverse FRNs,* which is a refinement of the inverse FRN that provides even greater price sensitivity to interest rate changes.

Leveraged Inverse FRNs

Suppose an inverse FRN pays 15% − LIBOR, and LIBOR is currently 7%. A *leveraged inverse FRN* would typically pay 22% − 2 × LIBOR. The term "leveraged" refers to the multiple applied to LIBOR, in this case 2. The leveraged inverse FRN is more interest rate sensitive because its coupon changes twice as fast. The sensitivity can also be seen in the security's duration. The duration of an inverse FRN is roughly twice the duration of an identical-maturity par-value fixed-rate note. Similarly, the duration of a leveraged inverse FRN is more than twice the duration of a fixed-rate note of like maturity, and an even larger multiple creates an even longer duration.

Collared FRNs

Sometimes investors or issuers want protection against too much interest rate volatility. In such cases, they can purchase another contract that will provide a cap (maximum) or a floor (minimum) on the FRN's rate, or purchase an FRN that includes a cap or floor. A *collared FRN* is an FRN with both a cap and a floor attached. Early issuers of collared FRNs, like those of inverse FRNs, were able to reduce their cost of borrowing (after offsetting the effects of the embedded cap and floor), by offering investors a combination of securities that was cheaper than buying them individually.

Creating Synthetic Instruments through Financial Engineering

An inverse FRN, barring default, is equivalent to purchasing a fixed-rate note, going short a swap (to pay floating and receive fixed), and buying an interest rate cap. Figure 7.7 shows how layering additional short swaps on top of a long position in an FRN creates different *synthetic* (fixed-rate and floating-rate) debt instruments. In panel A, an investor purchases FRNs. In panel B, the investor then sells a swap (to receive 7% fixed and pay LIBOR). The swap's notional amount equals the principal amount of the FRNs. Combining the long position in the FRNs and the short position in the swap produces a synthetic fixed-rate note paying 8%.

In panel C, selling a second swap that is identical to the first (or doubling the notional amount of the first swap) transforms the synthetic fixed-rate note into a synthetic inverse FRN paying 15% – LIBOR. This synthetic instrument combines a long position in FRNs and two receive 7% pay–LIBOR swaps. In panel D, selling a third identical swap converts the synthetic inverse FRN into a synthetic leveraged inverse FRN.

Investors buy inverse FRNs, rather than creating synthetics, because it is simpler and has less credit risk. Issuers of inverse FRNs usually have strong credit, often AAA-rated. The synthetic involves bearing credit risk on the underlying FRN, on the two swaps, and on the cap. Also, some potential investors might be prohibited, by policy or regulation, from buying derivative contracts. Embedding the derivatives in a structured note, such as an inverse FRN, provides a way around such restrictions.

Other investors create synthetic inverse FRNs, rather than buying the security, because they might want to adjust or "unwind" their position. With the synthetic, investors can make adjustments at relatively low cost, and avoid the cost of selling and repurchasing the entire principal. This is another example of using a particular configuration to reduce transaction costs.

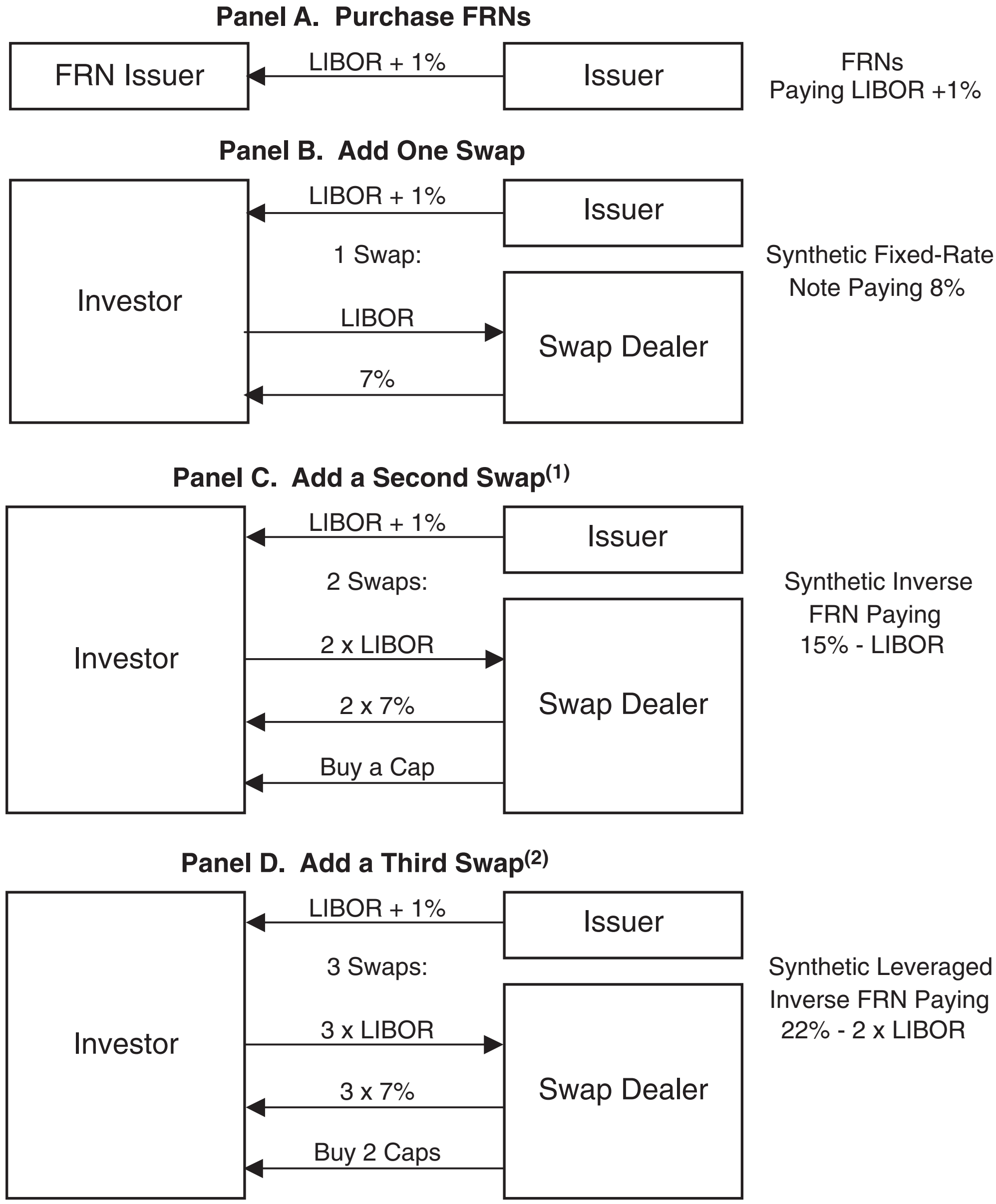
Transferring Default Risk

A *credit-linked note (CLN)* combines a plain vanilla bond and a credit default swap (Fabozzi, Davis, and Choudhry, 2006). The issuer of the CLN is the credit protection buyer, and the investor is the credit protection seller. If no credit event occurs during the life of the CLN, the investor receives a higher coupon rate. But if a credit event occurs, the coupon rate or the principal repayment decreases according to a formula that is tied to the loss resulting from the credit event. For example, banks that issue credit cards and sell unsecured bonds to fund these receivables have issued CLNs that reduce the principal repayment obligation to 85% of face if the credit card default rate exceeds 10%. The credit default swap embedded in the CLN pays 15% (100% – 85%) of face when the 10% default rate credit event occurs, which transfers default risk from the bank to the CLN purchasers.

HYBRID CAPITAL SECURITIES

A popular debt (or debt-like) innovation that has gone through several stages of development is the *hybrid capital security.* Hybrid securities have a variety of different structures, and they seem to be continually evolving. But they all share certain general characteristics. Hybrids are fixed-income instruments that are accorded some degree of equity credit for the issuer by the rating agencies or are accepted as capital by the issuer's main regulator because they combine features of debt and equity.

They provide regular monthly or quarterly income at a stated rate (a percentage of liquidation value), the liquidity of a publicly traded instrument, and an investment-grade credit quality. Monthly distributions are designed to appeal to retail investors and money market funds. Like corporate debt securities, they generally rank senior to common and preferred stock in the issuer's capital structure, have a stated maturity, and the interim payments qualify as interest for corporate income tax purposes.



(1) This result can also be achieved by doubling the notional amount of the first swap.
(2) This result can also be achieved by tripling the notional amount of the first swap.

Figure 7.7 Synthetic Inverse Floaters

However, they also carry certain additional investor risks, including the risks of optional redemption, deferred payments, and extension. They rank junior to corporate unsecured debt. Payments cannot be made on junior subordinated debentures until all required payments have been made on the issuer's "senior" obligations and all other conditions in the indenture have been met.

Trust Preferred Hybrid Capital Securities

The first hybrid capital securities, trust preferred securities, were issued about 10 years ago. Figure 7.8 illustrates their basic structure. A firm sets up a special purpose entity (usually a trust but possibly a partnership) that issues the hybrid capital securities and invests the proceeds in an issue of the parent firm's junior subordinated debentures. Most hybrid capital securities have an initial maturity between 20 and 49 years and have a fixed-price call feature that is activated after 5 or 10 years. Some non-U.S. firms have issued perpetual hybrid capital securities. Although the monthly or quarterly payments are tax deductible, these hybrid capital securities are different from conventional bonds in that they allow deferral of distributions for up to five years. During the deferral period, income continues to accrue for income tax purposes, and the investor is liable for income tax on the deferred income, even though the investor does not receive any cash

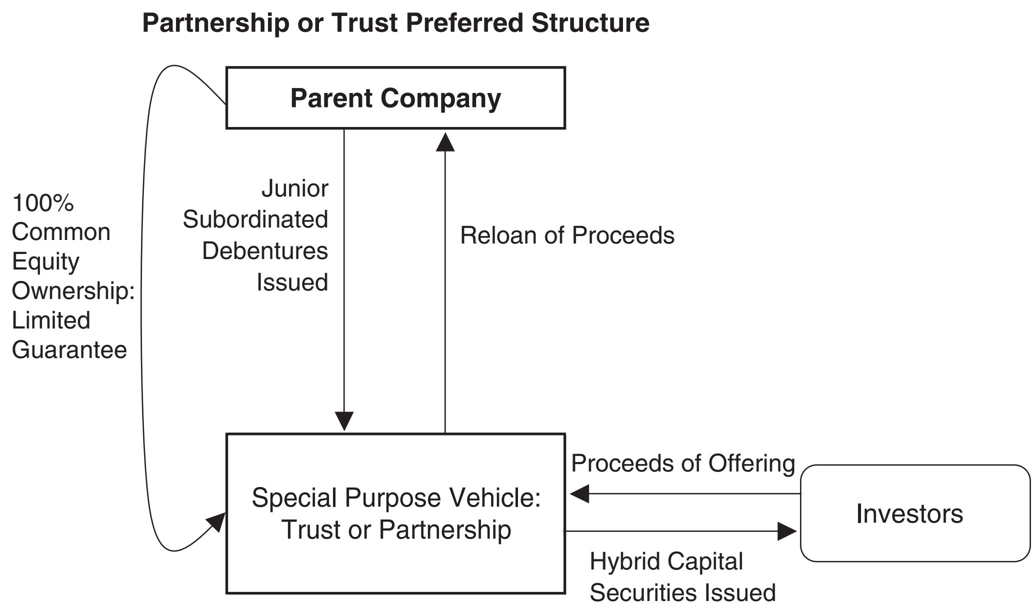


Figure 7.8 Structure of Trust Preferred Hybrid Capital Securities

payments. (To avoid such a tax liability, the securities can be held in a tax-deferred retirement account.) At the end of the deferral period, the issuer must pay all deferred distributions in cash or else be in default. Like regular preferred stock, deferral requires suspension of all cash dividends (common and preferred). In some cases, the issuer also has the option to extend the maturity of the securities, although it cannot be used if payments are being deferred. Neither can the payment-deferral option be used to extend the maturity.

An offering of hybrid capital securities does not affect the parent firm's balance sheet to the same extent as an issue of conventional debt. Rating agencies seem to view hybrid capital securities as similar to preferred stock because they provide long-term capital that permits the issuer to defer payments in case of financial distress. They reflect a common theme in securities innovation, attempts to craft new securities that qualify as debt for income tax purpose but have equity-like features.

Newer Hybrid Structures

The newer hybrids are usually structured either as debt to allow the issuer to deduct the interest payments for income tax purposes or as preferred stock to allow investors who are C corporations to claim the dividends received deduction. Investors expect regular interest or dividend payments and, at least in the United States, the repayment of principal either at a scheduled maturity date or at a synthetic maturity date. Hybrids can be callable, some as early as the fifth anniversary. If there is a step-up in the interest rate on the call date or a conversion from a fixed to a floating interest rate after that date, the securities are typically priced on the assumption that the hybrids will be redeemed on the call date. The step-up or the conversion in interest rate creates a synthetic maturity because investors expect that the issuer will call the hybrid to avoid paying the increased coupon or having the interest rate float.

Hybrid security structures also typically have certain deferral features, either mandatory or optional, which may interrupt the payment stream. The deferral feature, along with other features that may affect scheduled payments, are reviewed by the rating agencies in determining the amount of equity credit to be given the security. Mandatory deferral typically occurs only if the issuer is in significant financial distress. Optional deferral features are not typically utilized by issuers. Often, these securities are issued by investment-grade or highly rated financial institutions. Such issuers are highly unlikely to defer because they need continuous access to the capital markets. Even one deferral would send a signal to the market that the issuer is experiencing financial distress, which would significantly discourage lenders from committing more capital. Even one deferral can prevent the issuer from paying cash dividends, which will in turn prevent it from issuing equity securities to raise capital, at least until the deferral is eliminated. Lastly, issuers are very reluctant to defer payments on hybrids in the form of preferred stock because the terms of such securities often grant affected security holders the right to elect one or more directors in the event that payments have been deferred for a specified period of time. Thus, hybrid investors can be reasonably assured of receiving a steady stream of payments. As a consequence, hybrid securities are particularly attractive to life insurance companies and other investors who desire the combination of high yield and low default risk (Kumar and Shah, 2006).

Newer hybrids differ in some key respects from conventional preferred stock and also from traditional trust preferred hybrids. Table 7.2 summarizes the typical features of four types of hybrid securities that can qualify as capital for bank regulatory purposes: conventional perpetual preferred stock, trust preferred securities, Yankee tier

Table 7.2 Features of Hybrid Securities that Can Qualify as Capital

	Perpetual Preferred Stock	**More Conventional Hybrid Capital Securities**		**Newer Hybrid Securities**
Security:	Perpetual Preferred Stock	Trust Preferred Securities	Yankee Tier 1 Preferred Securities	Hybrid Capital Securities
Structure:	Issued directly by a taxpaying Delaware "C" corporation.	Limited business trust is formed by a company. The trust purchases a junior subordinated debenture of the company, which is the trust's only asset. The corporate issuer makes periodic interest payments on the junior subordinated debentures to the business trust. The trust passes these payments to the trust to pay periodic dividends on the trust preferred securities.	Either issued directly or through a conduit structure (e.g., a partnership purchases a perpetual debt security of the company). In the case of a conduit, the debt is the partnership's only asset. The corporate issuer makes periodic interest payments on the perpetual debenture to the partnership, which passes these payments to a trust to pay periodic dividends on the Yankee preferred securities.	Junior subordinated debentures issued directly or through a trust.
Maturity:	Perpetual security with no fixed maturity date.	At least a 30-year maturity, which in some cases may be extendible at the issuer's option.	Usually no fixed maturity date.	Usually 60-year maturity but can be perpetual.
Optional Redemption:	Typically cannot be put at the option of the holder. Callable at the issuer's option.	Trust preferred securities may be redeemable at the corporate issuer's option. Usually the trust preferred securities are subject to a mandatory redemption upon repayment of the underlying junior subordinated debenture.	Institutionally targeted Yankee preferred securities usually have a synthetic maturity, which is in the form of a rate step-up after the issue becomes callable in the 10th year.	(a) Callable at the option of the issuer after 10 years with a step-up in coupon after the 10th anniversary. (b) Scheduled redemption obligation beginning on the 30th anniversary. If the issuer is unable to redeem in year 30 due to market disruption events, there is a step-up in the coupon. (c) Callable at the option of the issuer after the 5th anniversary with no step-up in the coupon.
Deferral Feature:	Cumulative and can be deferred indefinitely.	Mimics the underlying junior subordinated debenture. It is usually cumulative and deferrable for up to 5 years. If payments are not made at that time, the issue goes into default.	Usually noncumulative.	(a) is cumulative with both mandatory and optional deferral. There is a mandatory alternative payment requirement, which goes into effect immediately when a payment is deferred. (b) and (c) are cumulative with optional deferral only. There is a mandatory alternative payment requirement after a period of time, usually 5 years.
Priority of Claims:	Senior to common stock and pari passu to other preferred stock.	Senior to common stock and preferred stock, pari passu to junior subordinated debt.	Depends on the structure. In all cases, senior to common stock and at least pari passu to preferred stock. In some cases, pari passu to junior subordinated debt and senior to common stock and preferred stock.	Bankruptcy claim is junior subordinated debt. (It is therefore senior to common stock and preferred stock.)

(*Continued*)

Table 7.2 Features of Hybrid Securities that Can Qualify as Capital (*Continued*)

	Perpetual Preferred Stock	**More Conventional Hybrid Capital Securities**		**Newer Hybrid Securities**
Creditors' Rights:	Investor remedy for nonpayment is the right to elect board members after some period (typically 18 months) of deferrals.	Investor has creditors' rights like any bondholder but typically can exercise them only after 5 years of missed payments.	Creditors can accelerate payment and put in a claim for a fixed amount upon a regulatory event or bankruptcy. In addition, for structures done through conduits, investors can also elect board members after 12 months of deferrals.	
Pricing:	Prices like bonds.	Prices like bonds.	Prices like bonds.	(a) prices like a 10-year security (that is, credit spread over 10-year Treasury). (b) prices like a 30-year security (that is, credit spread over 30-year Treasury). (c) prices like a 60-year security (that is, credit spread over 30-year Treasury).
Tax Treatment:	Dividends are not tax-deductible but the dividends are eligible for the dividends received deduction.	The interest is tax-deductible.	The interest is tax-deductible only in the conduit structure.	The interest is tax-deductible.
Accounting Treatment:	Appear on the balance sheet of the issuer as preferred stock.	Issuer's interest on the junior subordinated note is deducted as a business expense. Principal is recorded as a liability on the balance sheet. FDIC generally requires banks to account for investments in trust preferred securities as debt.	Generally, directly issued Yankee preferred securities are accounted for as preference shares. Generally, Yankee preferred securities issued through a conduit are treated like debt.	Liability.
Replacement Language:	None	None	None	States the issuer's intent or included in the form of a covenant.

1 preferred securities, and the newer hybrid capital securities of the type involved in the National Association of Insurance Commissioner's reclassification of hybrid capital securities in 2007. Perpetual preferred stock, a traditional hybrid security, has no fixed maturity, although the security is redeemable at the issuer's option. Dividends are not tax deductible, but they are eligible for the dividends-received deduction when the recipient is a C corporation. The lack of tax deductibility generally makes perpetual preferred stock less attractive to a fully taxable issuer than other hybrids that achieve tax deductibility.

Trust preferred stock and Yankee tier 1 preferred securities have both existed for at least ten years. They are long-dated preferred stock, which can be structured to achieve income tax deductibility for the interest payments when the company issues debt to a conduit that then issues the preferred stock to investors. Yankee tier 1 preferred securities are generally given more equity credit than trust preferreds by the rating agencies as they do not have a maturity date and their dividends are noncumulative.

Newer hybrid capital securities are generally structured as junior subordinated debentures. They are typically investment-grade or highly rated companies and so the opportunity to raise additional capital in a tax-efficient manner while obtaining relatively favorable rating agency treatment is attractive. The usual maturity of newer hybrid securities is 30 or 60 years, although some are perpetual. Some issues have a synthetic maturity as short as 10 years. Typically, the issuer either states its intention or else commits through a covenant to redeem the issue with proceeds from the sale of junior or pari passu securities if the hybrid issue is called, which is referred to as a "replacement" provision. Payment deferrals are cumulative, and some structures require mandatory deferral if certain balance sheet, income statement, or risk-based capital tests are met. Some of these securities have a mandatory alternative payment requirement, which starts immediately in some structures and within five years (20 quarters) in others. This feature requires the issuer to sell junior or pari passu securities in an amount sufficient to pay interest after payments have been deferred for a certain period of time. Failure to make the stipulated payment is an event of default.

Substituting Debt for Equity

Several firms have issued hybrid capital securities and used the funds to retire a portion of their preferred stock. Others have issued junior subordinated debt or conventional debt to replace preferred stock, in some cases offering preferred stockholders the opportunity to exchange their shares for the new debt.

Substituting debt for preferred stock generally elicits a favorable stock market reaction (see Masulis, 1980, 1983). There is both a favorable signaling effect, because interest is a fixed obligation and preferred stock dividends are not, and a favorable tax effect due to the tax-deductibility of interest. However, because of the five-year interest-deferral feature, the signaling effect of issuing hybrid capital securities is less favorable than issuing straight debt, although still positive.

For example, in June 1995, McDonald's Corporation offered to exchange up to $450 million aggregate principal amount of its 8.35% subordinated deferrable interest debentures due 2025 for up to 18 million shares of its 7.72% cumulative preferred stock. The debentures were offered in minimum denominations of $25 and make quarterly payments to match the preferred shares. The debentures allow McDonald's to defer interest payments for up to 20 consecutive quarters; mature in a lump sum; and, like the preferred stock, are redeemable at par beginning in December 1997.

The exchange offer met with only limited success; 26% of the preferred stockholders exchanged their shares. The lower-than-expected response rate was attributed to the higher-than-expected percentage of corporate holders, who could claim the 70% dividends-received deduction, which would be lost with the debt. In general, a debt-for-preferred exchange would not be profitable if the interest rate had to compensate preferred stockholders fully for the loss of the dividends-received deduction.

PREFERRED STOCK INNOVATIONS

U.S. corporations are permitted to deduct from their taxable income 70% of the common and preferred stock dividends they receive from unaffiliated corporations. As a result, the receiving corporation's tax rate on dividend income is effectively 10.5% (30% of the 35% statutory federal corporate tax rate). This offers corporations an incentive to purchase preferred stock rather than commercial paper or other short-term debt instruments, the interest of which is fully taxable. Preferred stock provides a tax arbitrage under current tax law when nontaxable firms issue it to fully taxable corporations. Nontaxable corporate issuers find preferred stock cheaper than debt because corporate investors are willing to pass back part of the value of the tax arbitrage by accepting a lower dividend rate.

Managing Interest Rate Risk with Preferred Stock

Purchasers of long-term, fixed-dividend-rate preferred stock are exposed to interest rate risk, and an interest rate increase could lower the value of the preferred stock by more than the tax saving. A variety of preferred stock instruments (see Table 7.3) have been designed to deal with this problem.

Adjustable-rate preferred stock was designed to reduce interest rate risk by adjusting the dividend rate by a formula specifying a fixed spread over Treasuries. At times, however, the spread investors have required to value the securities at par has differed significantly from the fixed spread specified in the formulas, causing the value of the security to deviate significantly from its face amount.

Convertible adjustable preferred stock (CAPS) was designed to eliminate this deficiency by making the security

Table 7.3 Selected Preferred Stock Innovations

■ **Security** ☐ **Distinguishing Characteristics**	■ **Year Issued** ■ **No. of Issues** ■ **Aggregate Proceeds ($B)** ☐ Risk Reallocation
• **Enhanced Liquidity**	• Reduction in Agency Costs
○ **Reduction in Transaction Costs**	○ Tax and Other Benefits: SEE NOTE
■ **Adjustable-Rate Monthly Income Preferred Securities (Adjustable-Rate MIPS)** ☐ Dividend rate resets each quarter based on a specified fraction of the highest of the 3-month T-bill, 10-year, or 30-year Treasury rates. Dividends paid monthly to appeal to retail investors. Really parent-company subordinated debt repackaged as preferred stock.	■ 03/21/96 ■ 2 ■ 0.5 ☐ Issuer bears more interest-rate risk than a fixed-rate preferred would involve. Investors bear more credit risk than a conventional subordinated debt issue would involve.
• Security is designed to trade near its par value.	○ Issuer can deduct the interest payments on the underlying subordinated debt, giving rise to tax-deductible 'capital' securities.
■ **Adjustable-Rate Preferred Stock** ☐ Dividend rate reset each quarter based on maximum of 3-month T-bill, 10-year, or 20-year Treasury rates plus or minus a specified spread.	■ 05/11/82 ■ 130 ■ 11.8 ☐ Issuer bears more interest-rate risk than a fixed-rate preferred would involve. Lower yield than commercial paper.
• Security is designed to trade near its par value.	
■ **Auction-Rate Preferred Stock (MMP/DARTS/AMPS/STAR)** ☐ Dividend rate reset by Dutch auction every 49 days. Dividend is paid at the end of each dividend period.	■ 08/27/84 ■ 541 ■ 56.2 ☐ Issuer bears more interest-rate risk than a fixed-rate preferred would involve. Lower yield than commercial paper.
• Security is designed to provide greater liquidity than convertible adjustable preferred stock.	• Dividend rate each period is determined in the marketplace, which provides protection against deterioration in issuer's credit standing.
■ **Convertible Adjustable Preferred Stock** ☐ Issue convertible on dividend payment dates into number of issuer's common shares, subject to cap, equal in value to par value of preferred.	■ 09/15/83 ■ 20 ■ 1.9 ☐ Issuer bears more interest-rate risk than a fixed-rate preferred would involve. Lower yield than commercial paper.
• Security is designed to provide greater liquidity than adjustable-rate preferred stock (due to the conversion feature).	
■ **Fixed-Rate/Adjustable-Rate or Auction-Rate Preferred** ☐ Fixed-dividend-rate preferred stock that automatically becomes adjustable-rate or auction-rate preferred after a specified length of time.	■ 11/16/84 ■ 94 ■ 5.8 ☐ Once the adjustment or auction period begins, issuer bears more interest-rate risk than a fixed-rate preferred would involve.
• Security is designed to trade near its par value once the adjustment or auction period begins.	
■ **Indexed-Floating-Rate Preferred Stock** ☐ Dividend rate resets quarterly as a specified percentage of 3-month LIBOR.	■ 10/01/85 ■ 8 ■ 1.8 ☐ Issuer bears more interest-rate risk than a fixed-rate preferred would involve. Lower yield than commercial paper.
• Security is designed to trade closer to its par value than a fixed-dividend-rate preferred.	
■ **Monthly Income Preferred Securities (MIPS)** ☐ Preferred stock issued by a trust or special-purpose company that purchases cash-matching long-term subordinated debt of the parent firm. Really parent-company subordinated debt repackaged as preferred stock.	■ 10/27/93 ■ 395 ■ 87.7 ☐ Investors bear more credit risk than a conventional subordinated debt issue would involve.
	○ Parent company can deduct the interest payments on the underlying subordinated debt, giving rise to tax-deductible 'capital' securities.
■ **Remarketed Preferred Stock (SABRES)** ☐ Perpetual preferred stock with a dividend rate that resets at the end of each dividend period to a rate the remarketing agent determines will make the preferred stock worth par. Dividend periods may be of any length, even 1 day. Different shares of a single issue may have different periods and dividend rates.	■ 06/27/85 ■ 132 ■ 29.9 ☐ Issuer bears more interest-rate risk than a fixed-rate preferred would involve. Lower yield than commercial paper.

Table 7.3 Selected Preferred Stock Innovations (*Continued*)

• Security is designed to trade near its par value.	○ Remarketed preferred stock offers greater flexibility in setting the terms of the issue than auction rate preferred stock, which requires a Dutch auction for potentially the entire issue once every 49 days.
■ **Single-Point Adjustable-Rate Stock** ☐ Dividend rate reset every 49 days as a specified percentage of the high-grade commercial paper rate.	■ 12/13/85 ■ 2 ■ 0.2 ☐ Issuer bears more interest-rate risk than a fixed-rate preferred would involve. Lower yield than commercial paper.
• Security is designed to trade near its par value. ☐ Security is designed to save on recurring transaction costs associated with auction-rate preferred stock.	
■ **Step-Up Preferred Stock** ☐ Long-term preferred stock with a dividend rate that steps up if the issue is not called on a specific date, in one case 10 years after issue.	■ 6/30/97 ■ 3 ■ 0.3 • Step up in dividend rate at least partially compensates investors if the issuer's credit standing falls and the issuer fails to redeem the preferred stock.
■ **Variable Cumulative Preferred Stock** ☐ At start of dividend period issuer can select between auction method and remarketing method to reset dividend rate at beginning of next period.	■ 07/07/88 ■ 24 ■ 1.5 ☐ Issuer bears more interest-rate risk than a fixed-rate preferred would involve. Lower yield than commercial paper.
• Security is designed to trade near its par value.	• The maximum permitted dividend rate increases according to a specified schedule if the preferred stock's credit rating falls.
○ Saves on transaction costs the issuer would otherwise incur if it wanted to change from auction reset to remarketing reset or vice versa.	○ Security is designed to give the issuer the flexibility to alter the method of rate reset.

[1]NOTE: All preferred stock innovations except MIPS are designed to enable short-term corporate investors to take advantage of the 70% dividends-received deduction.
Source: Reprinted by permission of Thomson Financial.

convertible on each dividend payment date into enough shares to make the security worth par. But although CAPS have traded closer to par than adjustable-rate preferred stocks, there have been only 13 CAPS issues. This may reflect the convertibility risk, which could force the issuer to issue common stock or raise considerable cash on short notice.

Auction-rate preferred stock carried the evolutionary process a step further. The dividend rate is reset by Dutch auction every 49 days, which represents just enough weeks to meet the 46-day holding period required to qualify for the 70% dividends-received deduction. A Dutch auction accepts the highest bid, the next highest bid, and so on, until a market-clearing price is found for which all the securities offered for sale will be purchased. All the sales then take place at the market-clearing price. Auction-rate preferred stock has a variety of names (MMP, money market preferred; AMPS, auction market preferred stock; DARTS, Dutch auction rate transferable securities; and STAR, short-term auction rate, among others), depending on the securities firm offering the product.

There have been two attempts to refine the adjustable-rate preferred stock concept further, but only one was successful. Single-point adjustable-rate stock (SPARS) has a dividend rate that adjusts automatically every 49 days to a specified percentage of the 60-day high-grade commercial paper rate. The security is designed to provide the same liquidity as auction-rate preferred stock, but with lower transaction costs since no auction need be held. The problem with SPARS, however, is that the fixed-dividend-rate formula introduces a potential agency cost. Because the dividend formula is fixed, investors will suffer a loss if the issuer's credit standing falls. Primarily for this reason, there have been only two SPARS issues.

Remarketed preferred stock, by contrast, pays a dividend that is reset at the end of each dividend period to a dividend rate that a specified remarketing agent determines will make the preferred stock worth par. Such issues permit the issuer considerable flexibility in selecting the length of the dividend period (it can be as short as one day). Remarketed preferred also offers greater flexibility in selecting the other terms of the issue. In fact, each share of an issue could have a different maturity, dividend rate, or other terms, provided the issuer and holders so agree. Remarketed preferred has not proven as popular with market participants as auction-rate preferred stock. Through July 2007, more than nine times as much auction-rate preferred stock as remarketed preferred stock has been issued.

Variable cumulative preferred stock was born out of the controversy over whether auction-rate preferred stock or remarketed preferred stock results in more equitable pricing. It effectively allows the issuer to decide at the beginning of each dividend period which of the two reset methods will determine the dividend rate at the beginning of the next dividend period.

More Desirable Pattern of Cash Flows

Monthly income preferred securities make monthly dividend distributions, whereas most fixed-rate preferred stock pays dividends quarterly. Monthly dividends appeal to retail investors and other investors looking for a more regular flow of cash from their investments.

Other Preferred Stock Innovations

In view of the similar structures for debt and preferred stock, a variety of preferred stock instruments are understandably adaptations of debt innovations. One of the more interesting examples of this is *gold-denominated preferred stock.* Commodity-linked preferred stock is like commodity-linked bonds. Freeport-McMoRan Copper & Gold Inc. raised $233 million in 1993 by issuing gold-denominated preferred stock. The issue price, quarterly dividend, and redemption value at maturity are each tied to the price of gold. Freeport-McMoRan Copper & Gold used the issue proceeds to help fund the expansion of one of the world's largest gold mines, its Grasberg mine in Irian, Jaya, Indonesia.

Several mutual funds that were unable to own physical gold were among the investors in the new issue, which embodied a 10-year forward contract on gold. The investors purchased the shares for $38.77 each. At maturity, each investor will receive the value of 1/10 of an ounce of gold for each share, whether the price of gold is higher or lower than $387 per ounce. In effect, the investor will trade the stated value of $38.77 per share at maturity for the cash value of 1/10 of an ounce of gold, profiting if the price of gold is above $387 per ounce and losing if it is below.

Freeport-McMoRan Copper & Gold designed the issue as a forward sale of 600,000 ounces of gold, hedging some of its gold price risk exposure. It chose to issue preferred stock rather than arrange a more traditional gold loan because it wanted a longer-dated security that did not require any amortization of principal and did not have the restrictive loan covenants usually found in gold loans.

CONVERTIBLE SECURITIES INNOVATIONS

Securities innovation has resulted in several new convertible securities, some of which have been very successful, such as *mandatory convertible preferred stock, puttable convertible bonds,* and *zero-coupon convertible debt.* These have been popular in large measure because of financial contracting considerations. For example, convertible bonds reduce agency costs arising from possible conflicts of interest between stockholders and other security holders, such as the problems of asset substitution, underinvestment, and claim dilution. Table 7.4 describes several of the innovations in convertible securities.

Reallocation of Investment Risk/More Desirable Pattern of Cash Flows

Preferred equity redemption cumulative stock (PERCS) is a form of *mandatory convertible preferred stock.* Its conversion feature differs from a traditional convertible preferred stock because conversion into common stock is not an American option held by the security holder. In fact, it is not an option at all. At maturity, as the name implies, conversion is mandatory. Furthermore, the issuing firm holds a call option, allowing it to redeem the security for cash prior to maturity, even if the value of the common stock exceeds the redemption price. Therefore, referring to this as a convertible security is somewhat of a misnomer. The contract is in fact more like a futures contract for common stock from the security holder's viewpoint. The holder has essentially purchased the common stock to be delivered at a future date and sold an option to the firm for it to instead pay cash prior the delivery date.

PERCS pay a higher dividend rate than the underlying common stock. However, the firm's call option puts a cap on the payoff. PERCS alter the investor's return distribution, providing greater dividend income in exchange for a cap on capital appreciation. In response to some investor-voiced objections to the cap, securities dealers created *dividend-enhanced convertible stock* (DECS) and *preferred redeemable increased dividend equity securities (PRIDES),* which are similar. Conversion of DECS is also mandatory, but the conversion feature is initially out-of-the-money and the payoff is not capped. Through year-end 2000, the PERCS-type and DECS-type convertible preferred had been issued in nearly equal amounts.

Figure 7.9 compares the payoff patterns for PERCS and DECS. The investors' returns are capped with PERCS. They are not with DECS, which eliminates the initial upside in the underlying common stock but provides full appreciation thereafter.

Reductions in Taxes

Many of the convertible security innovations have a tax motive. *Convertible-exchangeable preferred stock* is attractive to firms that are not currently paying income taxes but

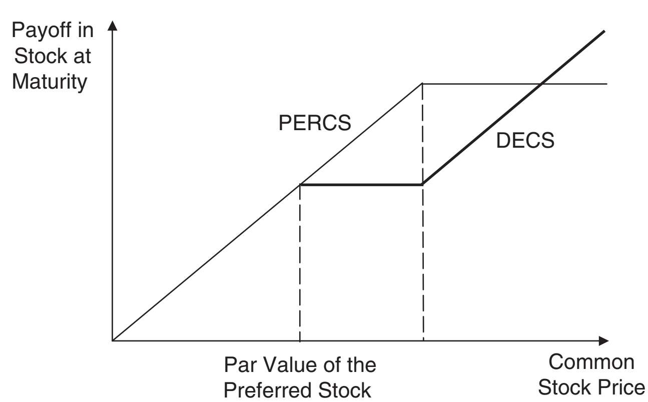


Figure 7.9 PERCS versus DECS Mandatory Convertible Preferred Stock Payoff Patterns

Table 7.4 Selected Convertible Securities Innovations

■ **Security** ☐ **Distinguishing Characteristics**	■ **Year Issued**	■ **No. of Issues**	■ **Aggregate Proceeds ($B)**	☐ Risk Reallocation
• **Enhanced Liquidity**				• **Reduction in Agency Costs**
○ **Reduction in Transaction Costs**				○ Tax and Other Benefits
■ **ABC Securities** ☐ Non-interest-bearing convertible debt issue on which the dividends on the underlying common stock are passed through to bondholders if the common stock price rises by more than a specified percentage (typically around 30%) from the date of issuance.	■ 02/06/91	■ 2	■ 0.4	○ If issue converts, the issuer will have sold, in effect, tax-deductible common equity. If holders convert, entire debt service stream is converted to common equity.
■ **Adjustable-Rate Convertible Debt** ☐ Debt the interest rate on which varies directly with the dividend rate on the underlying common stock. No conversion premium at issuance.	■ 04/18/84	■ 43	■ 6.8	○ Effectively, tax-deductible common equity. Security has since been ruled equity by the IRS. Portion of each bond recorded as equity on the issuer's balance sheet.
■ **Cash-Redeemable LYONs** ☐ Non-interest-bearing convertible debt issue that is redeemable in cash for the value of the underlying common stock, at issuer's option.	■ 06/20/90	■ 21	■ 13.9	○ If issue converts, the issuer will have sold, in effect, tax-deductible common equity. Issuer does not have to have its equity ownership interest diluted through conversion.
■ **Cash-Settled Convertible Notes** ☐ Issuer pays the value of the underlying common stock in cash when the holder tenders shares for exchange. ○ Holders save on the cost of having to sell their common stock when they want cash.	■ 07/13/89	■ 10	■ 3.1	
■ **Conversion Price Reset Notes/Premium Adjustable Notes** ☐ The conversion price adjusts downward on one or more specified dates if the market price of the underlying common stock is below the conversion price. It cannot be adjusted upward but is subject to a floor.	■ 07/13/88	■ 3	■ 0.8	• The reset feature at least partially protects investors if the managers take actions that diminish the firm's value. Cutting the strike price restores (part of) the value of the conversion option.
■ **Convertible-Exchangeable Preferred Stock** ☐ Convertible preferred stock that is exchangeable, at the issuer's option, for convertible debt with identical rate and identical conversion terms. ○ No need to reissue convertible security as debt – just exchange it – when the issuer becomes a taxpayer.	■ 12/15/82	■ 777	■ 110.8	○ Issuer can exchange debt for the preferred when it becomes taxable with interest rate the same as the dividend rate and without any change in conversion features. Appears as equity on the issuer's balance sheet until it is exchanged for convertible debt.
■ **Convertible Interest-Rate-Reset Debentures** ☐ Convertible bond the interest rate on which must be adjusted upward, if necessary, 2 years after issuance by an amount sufficient to give the debentures a market value equal to their face amount.	■ 10/13/83	■ 8	■ 0.6	• Investor is protected against a deterioration in the issuer's credit quality or financial prospects within 2 years of issuance.
■ **Convertible Monthly Income Preferred Securities (MIPS)** ☐ Long-term convertible subordinated debt is issued to a trust or a special-purpose company wholly owned by the parent. This entity issues cash-matching convertible preferred stock that pays dividends monthly. The parent can defer interest payments for up to five years without triggering a default but interest compounds during the deferral period.	■ 03/22/94	■ 33	■ 11.2	☐ Investors bear more credit risk than a conventional subordinated debt issue would involve but the issuer gets some 'equity' credit from the rating agencies. ○ Parent company can deduct the interest payments on the underlying subordinated debt.
■ **Debt with Mandatory Common Stock Purchase Contracts** ☐ Notes with contracts that obligate note purchasers to buy sufficient common stock from the issuer to retire the issue in full by its scheduled maturity date.	■ 04/22/82	■ 4	■ 1.1	○ Notes provide a stream of interest tax shields, which (true) equity does not. Commercial bank holding companies have issued it because it counted as "primary capital" for regulatory purposes

(*Continued*)

Table 7.4 Selected Convertible Securities Innovations (*Continued*)

■ Dividend Enhanced Convertible Stock (DECS)/PRIDES □ Preferred stock that pays a cash dividend significantly above that on the underlying common stock in exchange for a conversion option and a requirement to convert no later than 4 years after issuance.	■ 06/30/93 ■ 41 ■ 12.8 □ Investor trades off a portion of the underlying common stock's capital appreciation potential in return for an enhanced dividend rate. Differs from PERCS because there is no cap on the upside potential.
□ **Equity-Commitment Notes** □ The bank or bank-holding-company issuer commits to refund the notes with securities that qualify as primary capital.	■ N/A ■ N/A ■ N/A □ At one time, the notes were treated as bank capital for regulatory purposes. The bank (or parent) realized tax-deductible capital.
■ **Equity-Contract Notes** □ Notes that obligate investors contractually to convert the notes into the bank's or its parent's common stock.	■ 02/21/96 ■ 4 ■ 0.3 □ The issuer avoids the equity-market risk inherent in waiting to issue common stock.
○ The issuer saves the transaction costs of issuing the common stock in a separate transaction.	○ At one time, the notes were treated as bank capital for regulatory purposes. The bank (or parent) realized tax-deductible capital.
■ **Equity-Index-Linked Notes/Upside Note Securities** □ The interest payments or principal payments, or both, are indexed to a specified equity index or basket of equity securities. The dividend rate can exceed the blended dividend rate of the underlying equities.	■ 10/22/92 ■ 18 ■ 2.0 □ Securities reduce the investors' risk exposure by furnishing part of the return as enhanced dividends, diversifying the equity risk, and providing downside protection in the form of principal repayment.
○ The securities provide some diversification and downside protection more cheaply than retail investors could achieve it on their own.	○ The structured notes enable investors who cannot invest directly in the underlying equity securities to invest indirectly.
■ **Exchangeable Auction-Rate Preferred Stock/Remarketed Preferred Stock** □ Auction-rate preferred stock or remarketed preferred stock that is exchangeable on any dividend payment date, at the option of the issuer, for auction-rate notes, the interest rate on which is reset by Dutch auction every 35 days.	■ 02/12/88 ■ 8 ■ 0.4 □ Issuer bears more interest-rate risk than a fixed-rate instrument would involve.
● Security is designed to trade near its par value. ○ Issuance of auction-rate notes involves no underwriting commissions.	○ Issuer can exchange notes for the preferred when it becomes taxable. Appears as equity on the issuer's balance sheet until it is exchanged for auction-rate notes
■ **Liquid Yield Option Notes (LYONs)/Zero-Coupon Convertible Debt** □ Non-interest-bearing convertible debt issue.	■ 04/12/85 ■ 343 ■ 104.2 ○ If issue converts, the issuer will have sold, in effect, tax-deductible equity. If holders convert, entire debt service stream converts to common equity.
■ **Preferred Equity Redemption Cumulative Stock (PERCS)/Mandatory Conversion Premium Dividend Preferred Stock** □ Preferred stock that pays a cash dividend significantly above that on the underlying common stock in exchange for a conversion option that has a capped share value and requires conversion no later than 3 years after issuance.	■ 06/26/91 ■ 36 ■ 22.1 □ Investor trades off a portion of the underlying common stock's capital appreciation potential in return for an enhanced dividend rate.
■ **Puttable Convertible Bonds/Contingent Convertible Bonds** □ Convertible bond that can be redeemed prior to maturity, at the option of the holder, on certain specified dates at specified prices.	■ 07/21/82 ■ 1,459 ■ 261.2 ● Issuer is exposed to risk that the bonds will be redeemed early if interest rates rise sufficiently or common stock price falls sufficiently. Investor has one or more put options, which provide protection against deterioration in credit quality.
■ **Synthetic Convertible Securities** □ Issuer sells units consisting of subordinated discount notes and warrants that has the same risk-return characteristics as a conventional convertible bond. The synthetic convertible securities may be either coupon-bearing or zero-coupon.	■ N/A ■ N/A ■ N/A □ Investors can sell the debt and warrants separately or they can alter the mix of debt and warrants to tailor the package to suit their individual preferences.

Table 7.4 Selected Convertible Securities Innovations (*Continued*)

	○ The structure enables the issuer to realize larger interest-expense deductions than a traditional convertible because the discount can be amortized for income-tax purposes.
	○ The ability to separate the debt and warrants appealed to a broader market than traditional buyers of convertibles, resulting in better pricing.
■ **Zero-Premium Exchangeable Notes** □ Notes that are exchangeable for another firm's common stock, at the holder's option, without any conversion premium. Interest can be deferred for up to five years without triggering default but interest compounds during the deferral period. Holders can put the notes to the issuer at 95% of the market value of the underlying shares.	■ 09/15/99 ■ 2 ■ 1.5 ○ Under the tax regulations pertaining to contingent-payment debt instruments, the issuer can claim interest deductions based on the (much higher) 'comparable yield' at which it could sell conventional fixed-rate straight debt.
• Put option could improve liquidity.	○ A firm can use a stock investment it does not want to sell immediately to support a long-term debt issue at a relatively low interest rate.

Source: Reprinted by permission of Thomson Financial.

may in the future. It starts out as convertible perpetual preferred stock, but the issuer has the option to exchange for convertible subordinated debt with the same conversion terms and an interest rate equivalent to the original dividend rate. The exchange feature enables the issuer to replace the convertible preferred stock dividends with tax-deductible convertible debt interest payments when the firm begins paying income taxes. Not surprisingly, a large volume of such securities have been issued by firms that were not currently paying federal income taxes.

Contingent convertible bonds are a variant of the basic puttable convertible bond structure plus an added tax and accounting benefit for the issuer. Bond interest payments are contingent on some specified factor, such as the dividend rate on the underlying common stock. Until 2005, provided this contingency was neither remote nor incidental, the contingent interest rules under the Internal Revenue Code allowed the firm to take interest deductions based on the much higher interest rate on its straight debt with like maturity, payment dates, and seniority. In addition, conversion was often contingent on the stock price rising above a specified threshold, such as 30% above the share price at the issue date. This feature allowed the firm to keep the underlying shares out of the fully diluted earnings per share calculation until the share price reached the threshold. The accounting rules were changed in 2005 to treat contingent convertibles just like regular convertibles. Firms began issuing contingent convertible bonds in November 2000. More than $125 billion worth were issued before the accounting and tax rules were changed in 2005.

Adjustable-rate convertible debt was a very thinly disguised attempt to package equity as debt. The coupon rate varied directly with the dividend rate on the underlying common stock, and there was no conversion premium (at the time the debt was issued). After just three issues, the IRS ruled that the security is equity for tax purposes, thereby denying the interest deductions. Not surprisingly, the security has not been issued since that ruling.

Zero-coupon convertible debt, which includes *liquid yield option notes* (LYONs) and *ABC securities*, represents a variation on the same theme. If the issue is converted, both interest and principal are converted to common equity, in which case the issuer will have effectively sold common equity with a tax-deductibility feature. Zero-coupon convertible debt has been very popular among individual investors who have purchased it for their individual retirement accounts. These securities often contain put options that guarantee a minimum holding-period return and reduce agency costs.

Debt and warrants (exercisable into the issuer's common stock) can be combined to create *synthetic convertible debt*, with features that mirror those of conventional convertible debt (Finnerty, 1986). Synthetic convertible debt has a tax advantage over a comparable convertible debt issue because, in effect, the warrant proceeds are deductible for tax purposes over the life of the debt issue.

Reductions in Agency Costs

Puttable convertible bonds reduce agency costs by protecting investors against deterioration in the issuing firm's credit standing by giving them the option to put the bonds back to the issuer. The investors can exercise their put option if credit quality deteriorates or exercise their conversion (call option) if the firm's share price appreciates sufficiently. *Conversion price reset notes* adjust the conversion price downward on one or more specified dates if the market value of the underlying common stock is below the conversion price. This reset feature at least partially protects investors if the firm's managers take actions that diminish the firm's value. Similarly, *convertible interest rate reset debentures* adjust the interest rate upward if the issuing firm's credit standing deteriorates.

Contingent convertible bonds provide downside protection for investors in exchange for reduced periodic cash payments. They are issued at par initially with a zero yield to maturity. Contingent interest becomes payable if the conversion option becomes deep in-the-money. However, investor have a series of put options, exercisable at par, usually beginning as early as one year after the bonds are issued. Contingent convertible bonds differ from traditional convertible bonds in that a strip of in-the-money put options plus a stream of contingent interest payments replace the stream of stated interest payments.

Reductions in Transaction Costs

Cash-redeemable LYONs and *cash-settled convertible notes* pay the value of the underlying common stock in cash, in lieu of conversion. Many investors, such as convertible bond mutual funds, sell the common stock following conversion as soon as they can do it on an orderly basis. For them, cash settlement reduces transaction costs. The security is less attractive to hedge funds because they do not receive conversion shares, but instead have to buy common shares in the open market, to cover their short positions.

Satisfying Regulatory Restrictions

Banks have issued capital notes because they can be substituted for equity (subject to certain restrictions) for regulatory purposes, but the interest payments are tax deductible. For example, prior to the passage of the Financial Institutions Reform Recovery and Enforcement Act of 1989 (FIRREA), banks issued *equity contract notes.* They consisted of interest-deductible debt with a mandatory common stock purchase contract. They qualified as primary capital for bank regulatory purposes because conversion was mandatory.

Example of a Securities Innovation that Solved a Difficult Corporate Finance Problem

One of the innovative securities described earlier, PERCS, was created to help a corporation deal with the otherwise disruptive effects of reducing its dividend payout. The firm had concluded that it had to reduce its dividend per share. At the same time it announced a reduction of its dividend payout, the firm offered stockholders the opportunity to exchange their shares for PERCS, which its investment bankers had designed specifically for that purpose. Since that initial issue, PERCS have been issued in cash transactions presumably as an alternative to a traditional issue of convertible preferred stock.

Dividend Policy

The impact of dividend policy on firm value has been extensively examined, especially since Miller and Modigliani's (1961) demonstration of its irrelevance in a perfect capital market environment. Three broad categories of imperfections can cause dividend policy to affect firm value: asymmetric information, asymmetric taxes, and transactions costs. Among these, asymmetric information—specifically the information content of a dividend announcement—is generally believed to have the potential for causing the largest impact on a firm's market value. This is because the so-called *clientele effect* may be able to mitigate much of the impact of asymmetric taxes and transaction costs.

In equilibrium, investors sort themselves into various clienteles that invest in firms that follow the dividend policy that is most favorable to the clientele in terms of taxes and transaction costs. Of course, when a firm changes its dividend policy, investors in that clientele may be forced to incur transaction costs and tax liabilities as they sell their stock in that firm and purchase new shares in firms that have the dividend policy that is best for them.

Corporate dividend policies are notoriously stable, reflecting management's well-known reluctance to cut the dividend. But what can a firm do if it believes that cutting its cash dividend payout is a shareholder-wealth-maximizing decision? How can a firm communicate such a belief most accurately and at the lowest cost? If the dividend cut is perceived negatively, current shareholders would suffer a loss of wealth if they sell shares for less than their true value during the time it takes the market to see that the dividend cut is not negative after all.

Combining a PERCS-for-common exchange offer with a dividend cut can reduce the disruptive effects of a dividend cut for the following reasons (Emery and Finnerty, 1995):

1. It sends a credible positive signal to market participants about the firm's longer-run prospects.
2. It offers the option of continued capital gains tax deferral to shareholders who might otherwise sell their common shares and trigger a tax liability.
3. It offers the option of lower transaction costs to shareholders who choose to maintain their cash dividend income by exchanging for the PERCS rather than selling their common shares and reinvesting in higher-dividend-paying shares of other firms.

COMMON EQUITY INNOVATIONS

Five of the *common equity innovations* listed in Table 7.5 serve to reallocate risk: the Americus Trust, SuperShares, unbundled stock units, callable common stock, and puttable common stock.

Reallocation of Investment Risk

The first *Americus trust* was offered to owners of AT&T common stock on October 25, 1983. It offered AT&T common stockholders the opportunity to retain their predivestiture shares because AT&T was ordered to divest its regional and local operating companies, which it did by distributing the shares of these companies to its shareholders. The Americus trust received this share distribution. More than two dozen other Americus trusts were later formed. An Americus trust offered the common stockholders of a firm the opportunity to strip each of their common shares into a PRIME component, which carried full dividend and voting rights and limited capital appreciation rights, and a SCORE component, which carried full capital appreciation rights above a threshold price. PRIMEs and SCOREs appeared to expand the range of securities available, thus helping make the capital markets more complete. Unfortunately, a change in tax law made the separation of a share of common stock into a PRIME

Table 7.5 Selected Common Equity Innovations

■ **Security** □ **Distinguishing Characteristics** • **Enhanced Liquidity** ○ **Reduction in Transaction Costs**	■ Year Issued ■ No. of Issues ■ Aggregate Proceeds ($B) □ Risk Reallocation • Reduction in Agency Costs ○ Tax and Other Benefits
■ Americus Trust □ Outstanding shares of a particular company's common stock are contributed to five-year unit investment trust. Units may be separated into PRIME component, which embodies full dividend and voting rights in the underlying share and permits limited capital appreciation, and SCORE component, which provides full capital appreciation above the specified share price.	■ 10/25/83 ■ 27 ■ approx. 3.0[a] □ Stream of annual total returns on a share of stock is separated into a dividend stream (with limited capital appreciation) and a residual capital appreciation stream. ○ PRIME component would appeal to corporate investors who can take advantage of the 70% dividends received deduction. SCORE component would appeal to capital-gain oriented individual investors. ○ PRIME component resembles participating preferred stock if the issuer's common stock dividend rate is stable. SCORE component is a longer-dated call option than the ones customarily traded in the options market.
■ **Callable Common Stock** □ Common stock of a subsidiary sold by the parent subject to a stock purchase option agreement. Exercise prices step up overtime. Callable common stock often issued with warrants to purchase common stock of the parent company.	■ 05/23/91 ■ 2 ■ 0.1 □ Call option causes holders of the callable common stock to forgo capital appreciation in excess of the strike price (unless the callable common stock was sold in units that include warrants to buy the parent company's common stock). • Warrant to purchase parent company's shares enables holders of callable common stock to share in the upside if the common stock is called away. ○ Issuer retains the right to regain 100% ownership of the subsidiary's common stock.
■ **Master Limited Partnership** □ A business is given the legal form of a partnership but is otherwise structured, and is traded publicly, like a corporation.	■ 11/29/82 ■ 309 ■ 34.9 ○ Eliminates a layer of taxation because partnerships are not taxable entities.
■ **Paired Common Stock** □ Common shares of two related companies are paired and trade as a unit. Can be used when a company has a real estate-related business that can be organized as a real estate investment trust (REIT) but wishes to conduct other operations that a REIT is not permitted to engage in.	■ 08/20/86 ■ 8 ■ 1.3 ○ A REIT is not subject to federal income taxation on the income it distributes to its shareholders (except for certain specified classes of income).
■ **Puttable Common Stock** □ Issuer sells a new issue of common stock along with rights to put the stock back to the issuer on a specified date at a specified price.	■ 11/11/81 ■ 3 ■ 0.1 □ Issuer sells investors a put option, which investors will exercise if the company's share price decreases. • The put option reduces agency costs associated with a new share issue that are brought on by informational asymmetries. ○ Equivalent under certain conditions to convertible bonds but can be recorded as equity on the balance sheet so long as the company's payment obligation under the put option can be settled in common stock.

(*Continued*)

Table 7.5 Selected Common Equity Innovations (*Continued*)

■ **SuperShares** ☐ A trust is formed to hold a portfolio of common stocks that comprise the S&P 500 and a portfolio of Treasury bills. The trust sells four hybrid securities: (1) Priority SuperShares paying dividends on the S&P 500 and providing limited capital appreciation (2) Appreciation SuperShares providing appreciation above the Priority SuperShares' appreciation ceiling, (3) Protection SuperShares providing the value of any decline in the S&P 500 below some specified level, and (4) Money Market Income SuperShares paying proceeds from the Treasury bill portfolio after Protection SuperShares have been paid.	■ None Issued ☐ Shareholders can hold the components of total return in any proportions they choose, and Protection SuperShares function like portfolio insurance.	■ N/A	■ N/A
■ **Unbundled Stock Units** ☐ The total return stream from a share of common stock would be divided into three components: (1) a 30-year base yield bond (BYB) paying an interest rate equal to the dividend rate on the underlying common stock at the time the trust was formed plus limited capital appreciation, (2) a 30-year preferred stock instrument paying a dividend rate equal to the excess. If any, of the common dividend rate above the base rate, and (3) a 30-year warrant providing capital appreciation above the BYB's redemption value.	■ None Issued ☐ Shareholders could hold the components of a common share's total return in any proportions they choose.	■ N/A	■ N/A

Source: Reprinted by permission of Thomson Financial.

and a SCORE a taxable event, and no new Americus trusts have been formed since that change (see Francis and Bali, 2000, for more on the saga of PRIMEs and SCOREs).

SuperShares were an attempted extension of PRIMEs and SCOREs. They would have divided the stream of annual total returns on the S&P 500 portfolio into two components that are similar to the two components created by the Americus trust: (1) Priority SuperShares that paid the dividends on the S&P 500 stocks and provided limited capital appreciation and (2) Appreciation SuperShares that provided capital appreciation above the Priority SuperShares' capital appreciation ceiling. The new securities were to be issued by a trust that was to contain a portfolio of common stocks that mirrored the performance of the S&P 500 Index and a portfolio of Treasury bills. The trust also would issue two other classes of securities, one of which (Protection SuperShares) would have functioned like portfolio insurance. Unfortunately, a variety of problems arose and no SuperShares were ever issued.

Unbundled stock units (USUs) also can be thought of as an extension of PRIMEs and SCOREs (Francis and Bali, 2000). They divide the stream of annual total returns on a share of common stock into three components: (1) a 30-year "base yield" bond that would pay interest at a rate equal to the dividend rate on the issuer's common stock (thereby recharacterizing the "base" dividend stream into an interest stream), (2) a preferred share that would pay dividends equal to the excess, if any, of the dividend rate on the issuer's common stock above the "base" dividend rate, and (3) a 30-year warrant that would pay the excess, if any, of the issuer's share price 30 years hence above the redemption value of the base yield bond. Despite extensive marketing efforts, the USU concept failed to gain wide investor support and encountered regulatory obstacles that led to its withdrawal from the marketplace before a single issue could be completed (Finnerty, 1992). Like the Americus trust and SuperShares, USUs were designed to give shareholders more flexibility in choosing among the different components of the total returns from holding common stock; each of these new forms would effectively allow shareholders to tailor the corporation's dividend policy to suit their own preferences.

Callable common stock consists of common stock, typically issued by a subsidiary company, coupled with a call option to repurchase the stock, typically held by the parent company. The call price steps up periodically. The parent company may be required to exercise all the outstanding purchase options if any are exercised. The common stock has capital appreciation potential, however, it is limited by the company's call option.

Reductions in Agency Costs

Puttable common stock consists of common stock coupled with an investor-held put option. This package of securities is comparable to a convertible bond (Chen and Kensinger, 1988). The put option reduces the information "asymmetry" associated with a new share issue by putting a floor under the stock's value. Puttable common stock issues often provide a schedule of increasing put prices in order to ensure a minimum positive holding period rate of return. Essentially, this security offers the downside protection of the put option, often in exchange for not getting the dividends that are paid to other shareholders. The put option may be especially valuable for initial public offerings (IPOs), where investor uncertainty is particularly great (and as a result, IPOs are typically underpriced).

Reductions in Taxes

Publicly traded limited partnerships, often referred to as *master limited partnerships* (MLPs), became popular in the United States in the 1980s. They avoided double taxation of firm income, just like any other partnership, but their units could be traded publicly, just like the common stock of a corporation. The Revenue Act of 1987 eliminated the tax advantage of most MLPs (other than MLPs engaged in the natural resource extraction and oil and gas pipeline industries) by making them taxable as corporations if their units are publicly traded.

SUMMARY

Financial engineering involves the design of innovative securities, which provide superior, previously unavailable risk/return combinations. This process often includes coupling new derivative products with traditional securities to manage risks more cost effectively. The key to developing better risk-management vehicles is to design financial instruments that either provide new and more desirable risk-return combinations or furnish the desired future cash-flow profile at lower cost than existing instruments.

Securities innovation can improve capital market efficiency by offering more cost-effective means of transferring risks; increasing liquidity; and reducing taxes, transaction costs, and agency costs. One of the most important types of securities innovation is asset securitization, which redirects cash flows through an intermediary and offers new risk/return alternatives and improved liquidity to investors. Asset securitization has not only led to new products but it has also attracted new classes of investors to the mortgage and consumer receivables markets. The reallocation of various types of risk, principally prepayment risk, default risk, liquidity risk, and interest rate risk, is one of the main benefits of securitization.

Asset securitization is part of the natural evolution of an increasingly sophisticated global capital market. Legal and regulatory restrictions, tax laws, and institutional impediments have tended to segment the separate domestic markets for asset-backed securities and thereby slowed the process of securitization. The globalization of the capital market will increasingly result in cross-border securitizations. It will also lead to the development of new securities to reallocate currency, political, and other types of risk to achieve greater overall economic efficiency.

Securities innovation is a profit-driven response to changes in the economic, tax, and regulatory environment. If the tax regime remains static, interest rates stabilize, and the regulatory landscape solidifies, diminishing returns to securities innovation are bound to set in. But this is not likely to happen. Securities innovation occurs in response to unexpected changes in these factors, and a steady stream of abrupt shifts can keep the process of securities innovation going indefinitely. It seems more likely that a continuously changing economic and regulatory climate and greater competition within the financial services industry will combine to stimulate further securities innovation for many years to come.

For a corporate treasurer or chief financial officer, the opportunity to issue a new security is tempting. Firms that innovate successfully can increase shareholder wealth. The treasurer and the chief financial officer responsible for it can build reputational capital. But the process is not without risk, as the highly publicized failure of unbundled stock units several years ago so clearly demonstrates. A firm should issue a new security only after determining that it is truly innovative, that there is a market for it, and that issuing it will benefit its shareholders.

Remember that there are two sides to every transaction. Securities innovations that last must benefit both issuers and investors.

REFERENCES

Alderson, M. J., Brown, K. C., and Lummer, S. L. (1987). Dutch auction rate preferred stock. *Financial Management* 16, Summer: 68–73.

Brown, K. C., and Smith, D. J. (1996). Structured Swaps. In J. D. Finnerty and M. S. Fridson (eds.), *The Yearbook of Fixed Income Investing 1995* (pp. 137–156). Chicago: Irwin.

Carow, K. A., Erwin, G. R., and McConnell, J. J. (1999). A survey of U.S. corporate financing innovations: 1970–1997. *Journal of Applied Corporate Finance* 12, Spring: 55–69.

CDO Primer. (2004). New York: The Bond Market Association.

Chen, A. H., and Kensinger, J. W. (1988). Puttable stock: A new innovation in equity financing. *Financial Management* Spring: 27–37.

Crabbe, L. E., and Argilagos, J. D. (1994). The anatomy of the structured note market. *Journal of Applied Corporate Finance* Fall: 73–84.

Emery, D. R., and Finnerty, J. D. (1995). Using a PERCS-for-common exchange offer to reduce the costs of a dividend cut. *Journal of Applied Corporate Finance* 7, Winter: 77–89.

Emery, D. R., Hoffmeister, R., and Spahr, R. (1987). The case for indexing a bond's call price. *Financial Management* 16, Autumn: 57–64.

Fabozzi, F. J. (1989). *Advances & Innovations in the Bond and Mortgage Markets*. Chicago: Probus.

Fabozzi, F. J. (1995). *The Handbook of Mortgage-Backed Securities* (4th ed.). Chicago: Probus.

Fabozzi, F. J., Davis, H., and Choudhry, M. (2006). *Introduction to Structured Finance*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., Drake, P. P., and Polimeni, R. (2008). *The Complete CFO Handbook: From Accounting to Accountability*. Hoboken, NJ: John Wiley & Sons.

Finnerty, J. D. (1986). The case for issuing synthetic convertible bonds. *Midland Corporate Finance Journal* Fall: 73–82.

Finnerty, J. D. (1988). Financial engineering in corporate finance: An overview. *Financial Management* 17, 4: 14–33.

Finnerty, J. D. (1992). An overview of corporate securities innovation. *Continental Bank Journal of Applied Corporate Finance* Winter: 23–39.

Finnerty, J. D. (2001). Premium debt swaps, tax-timing arbitrage, and debt service parity. *Journal of Applied Finance* 1, November: 7–12.

Finnerty, J. D., and Emery, D. R. (2001). *Debt Management*. Cambridge, MA: Harvard Business School Press.

Finnerty, J. D., and Emery, D. R. (2002). Corporate securities innovation: An update. *Journal of Applied Finance* 12, 1 (Spring/Summer): 21–47.

Francis, J. C., and Bali, R. (2000). Innovations in partitioning a share of stock. *Journal of Applied Corporate Finance* 13, Spring: 128–136.

Kumar, A., and Shah, S. (2006). *Insurance hybrid securities: Beyond the current turmoil—a good relative value story—corrected*. New York: J.P. Morgan Securities Inc., June 23.

Lucas, D. J., Goodman, L. S., and Fabozzi, F. J. (2006). *Collateralized Debt Obligations: Structures and Analysis* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Mann, S. V., and Powers, E. A. (2001). Indexing a bond's call price: An analysis of make-whole call provisions. University of South Carolina working paper.

Masulis, R. W. (1980). The effects of capital structure change on security prices. *Journal of Financial Economics* 8, June: 139–178.

Masulis, R. W. (1983). The impact of capital structure change on firm value: Some estimates. *Journal of Finance* 38, March: 107–126.

Merton, R. C. (1992). Financial Innovation and Economic Performance. *Continental Bank Journal of Applied Corporate Finance*, Winter: 12–22.

Miller, M. H. (1986). Financial innovation: The last twenty years and the next. *Journal of Financial and Quantitative Analysis* 21: 459–471.

Miller, M. H., and Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *Journal of Business* 34, 4: 411–433.

Ogden, J. P. (1987). An analysis of yield curve notes. *Journal of Finance* 42: 99–110.

Perlman, S. D. (1989). Collateralized mortgage obligations: The impact of structure on value. In F. J. Fabozzi (ed.), *Advances & Innovations in the Bond and Mortgage Markets* (pp. 417–436). Chicago: Probus.

Ross, S. A. (1989). Institutional markets, financial marketing, and financial innovation. *Journal of Finance* 44: 541–556.

Sims, T. S. (1992). Long-term debt, the term structure of interest and the case for accrual taxation. *Tax Law Review* 47, Winter: 313–326.

Smith, D. J. (1988). The pricing of bull and bear floating rate notes: An application of financial engineering. *Financial Management* 17, 4: 72–81.

Smith, D. J. (1989). The arithmetic of financial engineering. *Journal of Applied Corporate Finance* Winter: 49–58.

Smithson, C. W., Smith, C. W., and Wilford, D. S. (1995). *Managing Financial Risk*. Burr Ridge, IL: Irwin.

Strnad, J. (1995). The taxation of bonds: The tax trading dimension. *Virginia Law Review* 81, February: 47–116.

Tufano, P. (1989). Financial innovation and first-mover advantage. *Journal of Financial Economics* 25: 213–240.

Tufano, P. (1995). Securities innovations: A historical and functional perspective. *Journal of Applied Corporate Finance* 7, Winter: 90–103.

Van Horne, J. C. (1985). Of financial innovations and excesses. *Journal of Finance* 40: 621–631.

Winger, B. J., Chen, C. R., Martin, J. D., Petty, J. W., and Hayden, S. C. (1986). Adjustable rate preferred Stock. *Financial Management* Spring: 48–57.

CHAPTER 8

An Arbitrage Perspective of the Purpose and Structure of Financial Markets

ROBERT DUBIL, PhD
Associate Professor, Lecturer of Finance, David Eccles School of Business, University of Utah

Abstract: From a structural point of view, all financial markets, whether for equities, bonds, currencies, or commodities, are the same. They all have distinct primary issuers and secondary traders. They all have spot, forward, and option vehicles. They all have two types of traders—hedgers and speculators—with two different trading motives. And in all markets, the primary trading strategy driving most of activity is relative risk arbitrage, whose goal is to earn reward for taking on exposure to secondary risk factors while eliminating primary directional risks through static or dynamic hedging.

Keywords: spot markets, forward markets, contingent claims, relative-value arbitrage, pure arbitrage, speculation, hedging, risk sharing, cash-and-carry

Financial markets play a major role in allocating wealth and excess savings to productive ventures in the global economy. This extremely desirable process takes on various forms. Commercial banks solicit depositors' funds in order to lend them out to businesses that invest in manufacturing and services or to home buyers who finance new construction or redevelopment. Investment banks bring to market offerings of equity and debt from newly formed or expanding corporations. Governments issue short- and long-term bonds to finance construction of new roads, schools, and transportation networks. Investors—bank depositors and securities buyers—supply their funds in order to shift their consumption into the future by earning interest, dividends, and capital gains.

The process of transferring savings into investment involves various market participants: individuals, pension and mutual funds, banks, governments, insurance companies, industrial corporations, stock exchanges, over-the-counter dealer networks, and others. All these agents can at different times serve as demanders and suppliers of funds, and as transfer facilitators.

Market participants design optimal securities and institutions to make the process of transferring savings into investment most efficient. "Efficient" means to produce the best outcomes—lowest cost, least disputes, fastest, and so on—from the perspective of security issuers and investors, as well as for the society as a whole. In this chapter we briefly discuss some fundamental questions about today's financial markets. Why do we have things like stocks, bonds, or mortgage-backed securities? Are they outcomes of optimal design or happenstance? Do we really need "greedy" investment bankers, securities dealers, or brokers soliciting us by phone to purchase unit trusts or mutual funds? What role do financial exchanges play in today's economy? Why do developing nations strive to establish stock exchanges even though often they do not have any stocks to trade on them?

Once we have basic answers to these questions, it will not be difficult to see why almost all the financial markets are organically essentially the same. Like automobiles made by Toyota and Volkswagen, which all have an engine, four wheels, a radiator, a steering wheel, and so on, all interacting in a predetermined way, all markets, whether for stocks, bonds, commodities, currencies, or any other claims to purchasing power, are built from the same basic elements.

All markets have two separate segments: original issue and resale. These are characterized by different buyers and sellers and different intermediaries. They perform different timing functions. The first transfers capital from the suppliers of funds (investors) to the demanders of capital (businesses). The second transfers capital from the suppliers of capital (investors) to other suppliers of capital (investors). The two segments are:

1. *Primary markets* (issuer-to-investor transactions with investment banks as intermediaries in the securities markets, and banks, insurance companies and others in the loan markets)
2. *Secondary markets* (investor-to-investor transactions with broker-dealers and exchanges as intermediaries in the securities markets, and mostly banks in the loan markets)

All markets have the originators, or issuers, of the claims traded in them (the original demanders of funds) and two distinctive groups of agents operating as investors, or suppliers of funds. The two groups of funds suppliers have completely divergent motives. The first group aims to eliminate any undesirable risks of the traded assets and earn money on repackaging risks; the other actively seeks to take on those risks in exchange for uncertain compensation. The two groups are:

1. *Hedgers:* Dealers who aim to offset primary risks, be left with short-term or secondary risks, and earn spread from dealing
2. *Speculators:* Investors who hold positions for longer periods without simultaneously holding positions that offset primary risks

The claims traded in all financial markets can be delivered in three ways. The first is an immediate exchange of an asset for cash. The second is an agreement on the price to be paid, with the exchange taking place at a predetermined time in the future. The last is a delivery in the future contingent upon an outcome of a financial event (e.g., level of stock price or interest rate), with a fee paid up front for the right of delivery. The three market segments based on the delivery type are:

1. *Spot* or *cash markets* (immediate delivery)
2. *Forward markets* (mandatory future delivery or settlement)
3. *Options markets* (contingent future delivery or settlement)

We focus on these structural distinctions to bring out the fact that all markets not only transfer funds from suppliers to users, but also risk from users to suppliers. They allow *risk transfer* or *risk sharing* between investors. The majority of the trading activity in today's market is motivated by risk transfer, with the acquirer of risk receiving some form of sure or contingent compensation. The relative price of risk in the market is governed by a web of relatively simple arbitrage relationships that link all the markets. These allow market participants to assess instantaneously the relative attractiveness of various investments within each market segment or across all of them. Understanding these relationships is mandatory for anyone trying to make sense of the vast and complex web of today's markets. Our line of thought here is adapted from Chapter 1 of Dubil (2004).

RISK SHARING

All financial contracts, whether in the form of securities or not, can be viewed as bundles, or packages of unit payoff claims (mini-contracts), each for a specific date in the future and a specific set of outcomes. In financial economics, these are referred to as *state-contingent claims*.

Let us start with the simplest illustration: an insurance contract. A one-year life insurance policy promising to pay \$1 million in the event of the insured's death can be viewed as a package of 365 daily claims (lottery tickets), each paying \$1 million if the holder dies on that day. The value of the policy up front (the premium) is equal to the sum of the values of all the individual tickets. As the holder of the policy goes through the year, he can discard tickets that did not pay off, and the value of the policy to him diminishes until it reaches zero at the end of the coverage period.

Let us apply the concept of state-contingent claims to known securities. Suppose you buy one share of XYZ SA stock currently trading at \$45 per share. You intend to hold the share for two years. To simplify things, we assume that the stock trades in increments of \$0.05 (tick size). The minimum price is \$0 (a limited liability company cannot have a negative value) and the maximum price is \$500. The share of XYZ SA can be viewed as a package of claims. Each claim represents a contingent cash flow from selling the share for a particular price at a particular date and time in the future. We can arrange the potential price levels from \$0 to \$500 in increments of \$0.05 to have overall 10001 price levels. We arrange the dates from today to two years from today (our holding horizon). Overall, we have 730 dates. The stock is equivalent to 10,001 times 730, or 7,300,730 claims. The easiest way to imagine this set of claims is as a rectangular chessboard as shown in Table 8.1, where on the horizontal axis we have time and on the vertical the potential values the stock can take on (states of nature). The price of the stock today is equal to the sum of the values of all the claims, that is, all the squares of the chessboard.

A forward contract on XYZ SA's stock can be viewed as a subset of this rectangle. Suppose we enter into a contract today to purchase the stock one year from today for \$60. We intend to hold the stock for one year after that. The forward can be viewed as 10,001 by 365 rectangles with the first 365 days' worth of claims taken out (that is, we are left with the latter 365 columns of the board; the first 365 are taken out), as in Table 8.2. The cash flow of each claim is equal to the difference between the stock price for that state of nature and the contract price of \$60. A forward carries an obligation on both sides of the contract, so some claims will have a positive value (stock is above \$60) and some negative (stock is below \$60).

Table 8.1 Stock Held for Two Years as a Chessboard of Contingent Claims in Two Dimensions: Time (Days 1 through 730) and Prices ($0 through $500)

500.00	500.00		500.00	500.00	500.00		500.00	500.00	**500.00**	
499.95	499.95		499.95	499.95	499.95		499.95	499.95	**499.95**	
499.90	499.90		499.90	499.90	499.90		499.90	499.90	**499.90**	
499.85	499.85		499.85	499.85	499.85		499.85	499.85	**499.85**	
				...						
60.35	60.35		60.35	60.35	60.35		60.35	60.35	**60.35**	
60.30	60.30		60.30	60.30	60.30		60.30	60.30	**60.30**	
60.25	60.25		60.25	60.25	60.25		60.25	60.25	**60.25**	
60.20	60.20		60.20	60.20	60.20		60.20	60.20	**60.20**	
60.15	60.15		60.15	60.15	60.15		60.15	60.15	**60.15**	
60.10	60.10	...	60.10	60.10	60.10	...	60.10	60.10	**60.10**	
60.05	60.05		60.05	60.05	60.05		60.05	60.05	**60.05**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**60.00**	**Prices**
59.95	59.95		59.95	59.95	59.95		59.95	59.95	**59.95**	
59.90	59.90		59.90	59.90	59.90		59.90	59.90	**59.90**	
59.85	59.85		59.85	59.85	59.85		59.85	59.85	**59.85**	
59.80	59.80		59.80	59.80	59.80		59.80	59.80	**59.80**	
				...						
0.45	0.45		0.45	0.45	0.45		0.45	0.45	**0.45**	
0.40	0.40		0.40	0.40	0.40		0.40	0.40	**0.40**	
0.35	0.35		0.35	0.35	0.35		0.35	0.35	**0.35**	
0.30	0.30		0.30	0.30	0.30		0.30	0.30	**0.30**	
0.25	0.25		0.25	0.25	0.25		0.25	0.25	**0.25**	
0.20	0.20		0.20	0.20	0.20		0.20	0.20	**0.20**	
0.15	0.15		0.15	0.15	0.15		0.15	0.15	**0.15**	
0.10	0.10		0.10	0.10	0.10		0.10	0.10	**0.10**	
0.05	0.05		0.05	0.05	0.05		0.05	0.05	**0.05**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.00**	
1	**2**	...	**364**	**365**	**366**	...	**729**	**730**		
				Days						

Table 8.2 One-Year Forward Buy at $60 of Stock as a Chessboard of Contingent Claims (Payoff in Cells Equal to S-60 for Year 2; No payoff in Year 1)

0.00	0.00		0.00	0.00	440.00		440.00	440.00	**500.00**	
0.00	0.00		0.00	0.00	439.95		439.95	439.95	**499.00**	
0.00	0.00		0.00	0.00	439.90		439.90	439.90	**499.90**	
0.00	0.00		0.00	0.00	439.85		439.85	439.85	**499.85**	
0.00	0.00		0.00	0.00	0.35		0.35	0.35	**60.35**	
0.00	0.00		0.00	0.00	0.30		0.30	0.30	**60.30**	
0.00	0.00		0.00	0.00	0.25		0.25	0.25	**60.25**	
0.00	0.00		0.00	0.00	0.20		0.20	0.20	**60.20**	
0.00	0.00		0.00	0.00	0.15		0.15	0.15	**60.15**	
0.00	0.00	...	0.00	0.00	0.10	...	0.10	0.10	**60.10**	
0.00	0.00		0.00	0.00	0.05		0.05	0.05	**60.05**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**60.00**	**Prices**
0.00	0.00		0.00	0.00	−0.05		−0.05	−0.05	**59.95**	
0.00	0.00		0.00	0.00	−0.10		−0.10	−0.10	**59.90**	
0.00	0.00		0.00	0.00	−0.15		−0.15	−0.15	**59.85**	
0.00	0.00		0.00	0.00	−0.20		−0.20	−0.20	**59.80**	
0.00	0.00		0.00	0.00	−59.55		−59.55	−59.55	**0.45**	
0.00	0.00		0.00	0.00	−59.60		−59.60	−59.60	**0.40**	
0.00	0.00		0.00	0.00	−59.65		−59.65	−59.65	**0.35**	
0.00	0.00		0.00	0.00	−59.70	...	−59.70	−59.70	**0.30**	
0.00	0.00		0.00	0.00	−59.75		−59.75	−59.75	**0.25**	
0.00	0.00		0.00	0.00	−59.80		−59.80	−59.80	**0.20**	
0.00	0.00		0.00	0.00	−59.85		−59.85	−59.85	**0.15**	
0.00	0.00		0.00	0.00	−59.90		−59.90	−59.90	**0.10**	
0.00	0.00		0.00	0.00	−59.95		−59.95	−59.95	**0.05**	
0.00	0.00		0.00	0.00	−60.00		−60.00	−60.00	**0.00**	
1	**2**	...	**364**	**365**	**366**	...	**729**	**730**		
				Days						

Table 8.3 American Call Struck at $60 with an Expiry in Two Years as a Chessboard of Contingent Claims (Payoff in Cells Equal to S – 60 if $S > 60$)

440.00	440.00		440.00	440.00	440.00		440.00	440.00	**500.00**	
439.95	439.95		439.95	439.95	439.95		439.95	439.95	**499.95**	
439.90	439.90		439.90	439.90	439.90		439.90	439.90	**499.90**	
439.85	439.85		439.85	439.85	439.85		439.85	439.85	**499.85**	
0.35	0.35		0.35	0.35	0.35		0.35	0.35	**60.35**	
0.30	0.30		0.30	0.30	0.30		0.30	0.30	**60.30**	
0.25	0.25		0.25	0.25	0.25		0.25	0.25	**60.25**	
0.20	0.20		0.20	0.20	0.20		0.20	0.20	**60.20**	
0.15	0.15		0.15	0.15	0.15		0.15	0.15	**60.15**	
0.10	0.10		0.10	0.10	0.10	...	0.10	0.10	**60.10**	
0.05	0.05		0.05	0.05	0.05		0.05	0.05	**60.05**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**60.00**	**Prices**
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**59.95**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**59.90**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**59.85**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**59.80**	
	...			...						
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.45**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.40**	
0.00	0.00		0.00	0.00	0.00	...	0.00	0.00	**0.35**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.30**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.25**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.20**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.15**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.10**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.05**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.00**	
1	**2**	**...**	**364**	**365**	**366**	**...**	**729**	**730**		
				Days						

An American call option contract to buy XYZ SA's shares for $60 with an expiry two years from today (exercised only if the stock is above $60) can be represented as an 8,800 × 730 subset of our original rectangular 10,001 × 730 chessboard. This time, the squares corresponding to the stock prices of $60 or below are eliminated, because they have no value, as in Table 8.3. The payoff of each claim is equal to the intrinsic (exercise) value of the call. As we will see later, the price of each claim today is equal to at least that.

Spot securities, forwards, and options are discussed in detail in subsequent chapters. Here, we briefly touch on the valuation of securities and state-contingent claims. The fundamental tenet of the valuation is that if we can value each claim (chessboard square) or small sets of claims (entire sections of the chessboard) in the package, then we can value the package as a whole. Conversely, if we can value a package, then often we are be able to value smaller subsets of claims (through a "subtraction"). In addition, we are sometimes able to combine very disparate sets of claims (stocks and bonds) to form complex securities (e.g., convertible bonds). By knowing the value of the combination, we can infer the value of a subset (bullet bond).

In general, the value of a contingent claim does not stay constant over time. If the holder of the life insurance becomes sick during the year and the likelihood of his death increases, then likely the value of all claims increases. In the stock example, as information about the company's earnings prospects reaches the market, the price of the claims changes. Not all the claims in the package have to change in value by the same amount. An improvement in the earnings prospects for the company may be only short term. The policyholder's likelihood of death may increase for all the days immediately following his illness, but not for more distant dates. The prices of the individual claims fluctuate over time, and so does the value of the entire bundle. However, at any given moment in time, given all information available as of that moment, the sum of the values of the claims must be equal to the value of the package, the insurance policy, or the stock. We always restrict the valuation effort to here and now, knowing that we will have to repeat the exercise an instant later.

Let us fix the time to see what assumptions we can make about some of the claims in the package. In the insurance policy example, we may surmise that the value of the claims for far-out dates is greater than that for near dates, given that the patient is alive and well now, and barring an accident, he is relatively more likely to take time to develop a life-threatening condition. In the stock example, we assigned the value of $0 to all claims in states with stock exceeding $500 over the next two years, as the likelihood of reaching these price levels is almost zero. We often assign the value of zero to claims for far dates (e.g., beyond 100 years), since the present value of those payoffs, even if they are large, is close to zero. We reduce a numerically infinite problem to a finite one. We cap

Table 8.4 Stock Plus Short American Call Struck at \$60 as a Chessboard of Contingent Claims (Payoff in Cells Equal to 60 if $S > 60$ and to S if $S < 60$)

60.00	60.00		60.00	60.00	60.00		60.00	60.00	**500.00**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**499.95**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**499.90**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**499.85**	
				...						
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**60.35**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**60.30**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**60.25**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**60.20**	
60.00	60.00	...	60.00	60.00	60.00	...	60.00	60.00	**60.15**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**60.10**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**60.05**	
60.00	60.00		60.00	60.00	60.00		60.00	60.00	**60.00**	**Prices**
59.95	59.95		59.95	59.95	59.95		59.95	59.95	**59.95**	
59.90	59.90		59.90	59.90	59.90		59.90	59.90	**59.90**	
59.85	59.85		59.85	59.85	59.85		59.85	59.85	**59.85**	
59.80	59.80		59.80	59.80	59.80		59.80	59.80	**59.80**	
	...			...			...			
0.45	0.45		0.45	0.45	0.45		0.45	0.45	**0.45**	
0.40	0.40		0.40	0.40	0.40		0.40	0.40	**0.40**	
0.35	0.35		0.35	0.35	0.35		0.35	0.35	**0.35**	
0.30	0.30		0.30	0.30	0.30	...	0.30	0.30	**0.30**	
0.25	0.25		0.25	0.25	0.25		0.25	0.25	**0.25**	
0.20	0.20		0.20	0.20	0.20		0.20	0.20	**0.20**	
0.15	0.15		0.15	0.15	0.15		0.15	0.15	**0.15**	
0.10	0.10		0.10	0.10	0.10		0.10	0.10	**0.10**	
0.05	0.05		0.05	0.05	0.05		0.05	0.05	**0.05**	
0.00	0.00		0.00	0.00	0.00		0.00	0.00	**0.00**	
1	**2**	...	**364**	**365**	**366**	...	**729**	**730**		
				Days						

the potential states under consideration, future dates and times.

A good valuation model has to strive to make the values of the claims in a package independent of each other. In our life insurance policy example, the payoff depends on the person's dying on that day and not on whether the person is dead or alive on a given day. In that setup, only one claim out of the whole set will pay. If we modeled the payoff to depend on being dead and not dying, all the claims after the morbid event date would have positive prices and would be contingent on each other. Sometimes, however, even with the best of efforts, it may be impossible to model the claims in a package as independent. If a payoff at a later date depends on whether the stock reached some level at an earlier date, the later claim's value depends on the prior one. A mortgage bond's payoff at a later date depends on whether the mortgage has not already been prepaid. This is referred to as a *survival* or *path-dependence* problem. Our imaginary two-dimensional chessboards cannot handle path dependence, and we ignore this dimension of risk throughout the book as it adds very little to our discussion.

Let us turn to the definition of risk sharing: *Risk sharing is a sale, explicit or through a side contract, of all or some of the state-contingent claims in the package to another party*.

In real life, risk sharing takes on many forms. The owner of the XYZ share may decide to sell a covered call on the stock (see Chapter 14 of Volume I). If he sells an American-style call struck at \$60 with an expiry date of two years from today, he gives the buyer the right to purchase the share at \$60 from him even if XYZ trades higher in the market (e.g., at \$75). The covered call seller is choosing to cap his potential payoff from the stock at \$60 in exchange for an up-front fee (option premium) he receives. This corresponds to exchanging the squares corresponding to price levels above \$60 (with values between \$60 and \$500) for squares with a flat payoff of \$60, as illustrated in Table 8.4.

Another example of risk sharing can be a hedge of a corporate bond with a risk-free government bond. A *hedge* is a sale of a package of state-contingent claims against a primary position that eliminates all the essential risk of that position. Only a sale of a security that is identical in all aspects to the primary position can eliminate all the risk. A hedge always leaves some risk *unhedged*! Let us examine a very common hedge of a corporate with a government bond. An institutional trader purchases a 10-year 5% coupon bond issued by XYZ Corporation. In an effort to eliminate interest rate risk, the trader simultaneously shorts a 10-year 4.5% coupon government bond. The size of the short is duration-matched to the principal amount of the corporate bond. This guarantees that for small parallel movements in the interest rates, the changes in the values of the two positions are identical but opposite in sign. If interest rates rise, the loss on the corporate bond

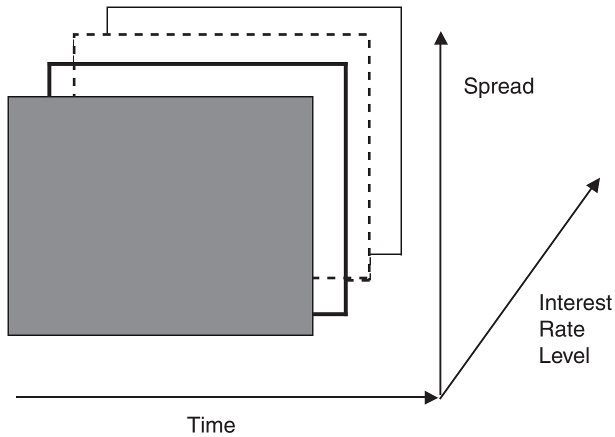


Figure 8.1 Reduction of One Risk Dimension through a Hedge: Corporation Hedged with a Government

holding will be offset by the gain on the shorted government bond. If interest rates decline, the gain on the corporate bond will be offset by the loss on the government bond. The trader, in effect, speculates that the credit spread on the corporate bond will decline. Irrespective of whether interest rates rise or fall, whenever the XYZ credit spread declines, the trader gains since the corporate bond's price goes up more, or goes down less, than that of the government bond. Whenever the credit standing of XYZ worsens and the spread rises, the trader suffers a loss. The corporate bond is exposed over time to two dimensions of risk, interest rates and corporate spread. Our chessboard representing the corporate bond becomes a large rectangular cube, with time, interest rate, and credit spread as dimensions, as illustrated in Figure 8.1. The government bond hedge eliminates all potential payoffs along the interest rate axis, reducing the cube to a plane, with only time and credit spread as dimensions.

Practically any hedge position can be thought of in the context of a multidimensional cube defined by time and risk axes. The hedge eliminates a dimension or a subspace from the cube.

STRUCTURE OF FINANCIAL MARKETS

Most people view financial markets like a Saturday bazaar. Buyers spend their cash to acquire paper claims on future earnings, coupon interest, or insurance payouts. If they buy good claims, their value goes up and they can sell them for more; if they buy bad ones, their value goes down and they lose money.

When probed a little more on how markets are structured, most finance and economics professionals provide a seemingly more complete description, adding detail about who buys and sells what and why in each market. The respondent is likely to inform us that businesses need funds in various forms of equity and debt. They issue stock, lease- and asset-backed bonds, and unsecured debentures; sell short-term commercial paper; or rely on bank loans. Issuers get the needed funds in exchange for a promise to pay interest payments or dividends in the future. The legal claims on business assets are purchased by investors, individual and institutional, who spend cash today to get more cash in the future, that is, they *invest*. Securities are also bought and sold by governments, banks, real estate investment trust, leasing companies, and others. The cash-for-paper exchanges are immediate. Investors who want to leverage themselves can borrow cash to buy more securities, but through that they themselves become issuers of broker margin or bank loans. Both issuers and investors live and die with the markets. When stock prices increase, investors who have bought stocks gain. When stock prices decline, they lose. New investors have to "buy high" when share prices rise, but can "buy low" when share prices decline. The decline benefits past issuers who "sold high". The rise hurts them since they got little money for the sold stock and now have to deliver good earnings. In fixed-income markets, when interest rates fall, investors gain as the value of debt obligations they hold increases. The issuers suffer as the rates they pay on the existing obligations are higher than the going cost of money. When interest rates rise, investors lose as the value of debt obligations they hold decreases. The issuers gain as the rates they pay on the existing obligations are lower than the going cost of money.

In this view of the markets, both sides—the issuers and the investors—*speculate* on the direction of the markets. In a sense, the word *investment* is a euphemism for *speculation*. The direction of the market given the position held determines whether the investment turns out good or bad. Most of the time, current issuers and investors hold opposite positions (long vs. short): when investors gain, issuers lose, and vice versa. Current and new participants may also have opposite interests. When equities rise or interest rates fall, existing investors gain and existing issuers lose, but new investors suffer and new issuers gain.

The investor is exposed to market forces as long as he holds the security. He can enhance or mitigate his exposure, or *risk*, by concentrating or diversifying the types of assets held. An equity investor may hold shares of companies from different industrial sectors. A pension fund may hold some positions in domestic equities and some positions in domestic and foreign bonds to allocate risk exposure to stocks, interest rates and currencies. The risk is "good" or "bad" depending on whether the investor is *long* or *short* an exposure. An investor who has shorted a stock gains when the share price declines. A homeowner with an adjustable mortgage gains when interest rates decline (he is short interest rates) as the rate he pays resets lower, while a homeowner with a fixed mortgage loses as he is "stuck" paying a high rate (he is long interest rates).

While this standard description of the financial markets appears to be very comprehensive, it is rather like a two-dimensional portrait of a multidimensional object. The missing dimension here is the *time of delivery*. The standard view focuses exclusively on spot markets. Investors purchase securities from issuers or other investors and pay for them at the time of the purchase. They modify the risks the purchased investments expose them to by diversifying their portfolios or holding shorts against longs in the same or similar assets. Most tend to be speculators as the universe of hedge securities they face is fairly limited.

Let us introduce the time of delivery into this picture. That is, let us relax the assumption that all trades (that is, exchanges of securities for cash) are immediate. Consider an equity investor who agrees today to buy a stock for a certain market price, but will deliver cash and receive the stock one year from today. The investor is entering into a *forward buy* transaction. His risk profile is drastically different from that of a spot buyer. Like the spot stock buyer, he is exposed to the price of the stock, but his exposure does not start till one year from now. He does not care if the stock drops in value as long as it recovers by the delivery date. He also does not benefit from the temporary appreciation of the stock compared to the spot buyer who could sell the stock immediately. In our time-risk chessboard with time and stock price on the axes, the forward buy looks like a spot buy with a subplane demarcated by today and one year from today taken out. If we ignore the time value of money, the area above the current price line corresponds to "good" risk (that is, a gain), and the area below to "bad" risk (that is, a loss). A forward sell would cover the same subplane, but the good and the bad areas would be reversed.

Market participants can buy and sell not just spot but also forward. For the purpose of our discussion, it does not matter if, at the future delivery time, what takes place is an actual exchange of securities for cash, or just a marked-to-market settlement in cash. If the stock is trading at $75 in the spot market, whether the parties to a prior $60 forward transaction exchange cash ($60) for stock (one share) or simply settle the difference in value with a payment of $15 is quite irrelevant, as long as the stock is liquid enough so that it can be sold for $75 without any loss. Also, for our purposes, futures contracts can be treated as identical to forwards, even though they involve a daily settlement regimen and may never result in the physical delivery of the underlying commodity or stock basket.

Let us now further complicate the standard view of the markets by introducing the concept of *contingent delivery time*. A trade, or an exchange of a security for cash, agreed upon today, is not only delayed into the future, but is also made contingent upon a future event. The simplest example is an insurance contract. The payment of a benefit on a $1 million life insurance policy takes place only upon the death of the insured person. The amount of the benefit is agreed upon and fixed up front between the policyholder and the issuing company. It can be increased only if the policyholder pays additional premium. Hazard insurance (fire, auto, flood) is slightly different from life in that the amount of the benefit depends on the "size" of the future event. The greater the damage is, the greater the payment is. An option contract is very similar to a hazard insurance policy. The amount of the benefit follows a specific formula that depends on the value of the underlying financial variable in the future (see Chapter 14 of Volume I). For example, a put option on the S&P 100 index traded on an exchange in Chicago pays the difference between the selected strike price and the value of the index at some future date and times $100, but only if the index goes down below that strike price level. The buyer thus insures himself against the index's going down, and the more the index goes down, the more benefit he obtains from his put option, just as if he held a fire insurance policy. Another example is a *cap* on an interest rate index that provides the holder with a periodic payment every time the underlying interest rate goes above a certain level. Borrowers use caps to protect themselves against interest rate hikes.

Options are used not only for obtaining protection, which is only one form of risk sharing, but also for risk taking, that is, providing specific risk protection for upfront compensation. A bank borrower relying on a revolving credit line with an interest rate defined as some spread over the U.S. prime rate or the three-month London Interbank Offered Rate (LIBOR) can sell *floors* to offset the cost of the borrowing. When the index rate goes down, he is required to make periodic payments to the floor buyer that depend on the magnitude of the interest rate decline. He willingly accepts that risk because when rates go down and he has to make the floor payments, the interest he is charged on the revolving loan also declines. In effect, he fixes his minimum borrowing rate in exchange for an up-front premium receipt.

Options are not the only packages of contingent claims traded in today's markets. In fact, the feature of contingent delivery is embedded in many commonly traded securities. Buyers of convertible bonds exchange their bonds for shares when interest rates and/or stock prices are high making the postconversion equity value higher than the present value of the remaining interest on the unconverted bond. Issuers call their outstanding callable bonds when interest rates decline below a level at which the value of those bonds is higher than the call price. Adjustable mortgages typically contain periodic caps, which prevent the interest rate and thus the monthly payment charged to the homeowner from changing too rapidly from period to period. Many bonds have credit covenants attached to them, which require the issuing company to maintain certain financial ratios, and noncompliance triggers automatic repayment or default. Car lease agreements give the lessees the right to purchase the automobile at the end of the lease period for a prespecified residual value. Lessees sometimes exercise those rights when the residual value is sufficiently lower than the market price of the vehicle. In many countries, including the United States, the homeowners with fixed-rate mortgages can prepay their loans partially or fully at any time without penalty. This feature allows the homeowners to refinance their loans with new ones when interest rates drop by a significant enough margin. The cash flows from the original fixed-rate loans are thus contingent upon interest rates staying high. Other examples abound.

The key to understanding these types of securities is the ability to break them down into simpler components: spot, forward, and contingent delivery. These components may trade separately in the institutional markets, but they are most likely bundled together for retail customers or original (primary market) acquirers. Not uncommonly, they are unbundled and rebundled several times during their lives.

Proposition. All financial market evolve to have three structural components: the market for spot securities, the market for forwards and futures, and the contingent securities market, which includes options and other derivatives.

All financial markets eventually evolve to have activity in three areas: spot trading for immediate delivery, trading with forward delivery, and trading with contingent forward delivery. Most of the activity of the last two forms is reserved for large institutions, which is why most people are unaware of them. Yet their existence is necessary for the smooth functioning of the spot markets. The trading for forward and contingent forward delivery allows dynamic risk sharing for holders of cash securities who trade in and out of contracts tied to different dates and future uncertain events. This risk-sharing activity, by signaling the constantly changing price of risk, in turn facilitates the flow of the fundamental information that determines the "bundled" value of the spot securities. In a way, the spot securities that we are all familiar with are the most complicated ones from the informational content perspective. Their value reflects all available information about the financial prospects of the entity that issued them and expectations about the broad market, and is equal to the sum of the values of all state-contingent claims that can be viewed as informational units. The value of forwards and option-like contracts is tied to more narrow information subsets. These contracts have an expiry date that is short relative to the underlying security and are tailored to a specific dimension of risk. Their existence allows the unbundling of the information contained in the spot security. This function is extremely desirable to holders of cash assets as it offers them a way to sell off undesirable risks and acquire desirable ones at various points in time. If you own a bond issued by a tobacco company, you may be worried that legal proceedings against the company may adversely affect the credit spread and thus the value of the bond you hold. You could sell the bond spot and repurchase it forward with the contract date set far into the future. You could purchase a spread-related option or a put option on the bond, or you could sell calls on the bond. All of these activities would allow you to share the risks of the bond with another party to tailor the duration of the risk sharing to your needs.

ARBITRAGE: PURE VERSUS RELATIVE

In this section, we introduce the notion of relative-value arbitrage, which drives the trading behavior of financial firms irrespective of the market they are engaged in. Relative arbitrage takes the concept of pure arbitrage beyond its technical definition of riskless profit. In it, all primary market risks are eliminated, but some secondary market exposures are deliberately left unhedged.

Arbitrage is defined in most textbooks as riskless instantaneous profit. It occurs when the *law of one price*, which states that the same item cannot sell at two different prices at the same time, is violated. The same stock cannot trade for one price at one exchange and for a different price at another unless there are fees, taxes, and the like. If it does, traders will buy it on the exchange where it sells for less and sell it on the one where it sells for more. Buying Czech korunas for British pounds cannot be more or less expensive than buying dollars for pounds and using dollars to buy korunas. If one can get more korunas for pounds by buying dollars first, no one will buy korunas for pounds directly. On top of that, anyone with access to both markets will buy korunas through dollars and immediately sell them for pounds to realize an instantaneous and riskless profit. This strategy is a very simple example of pure arbitrage in the spot currency markets. More complicated pure arbitrage involves forward and contingent markets. It can take a *static* form, where the trade is put on at the outset and liquidated once at a future date (e.g., trading forward rate agreements against spot LIBORs for two different terms), or a *dynamic* one, in which the trader commits to a series of steps that eliminate all directional market risks and ensures virtually riskless profit upon completion of these steps. For example, a bond dealer purchases a callable bond from the issuer, buys a swaption from a third party to offset the call risk, and delta-hedges the rate risk by shorting some bullet swaps. He guarantees himself a riskless profit provided that neither the issuer nor the swaption seller defaults.

Pure arbitrage is defined as generating riskless profit today by statically or dynamically matching current and future obligations to exactly offset each other, inclusive of incurring known financing costs. Not surprisingly, opportunities for pure arbitrage in today's ultra-sophisticated markets are limited. Most institutions' money-making activities rely on the principle of relative-value arbitrage. Hedge funds and proprietary trading desks of large financial firms, commonly referred to as *arb desks*, employ extensively relative arbitrage techniques. Relative-value arbitrage consists of a broadly defined hedge in which a close substitute for a particular risk dimension of the primary security is found and the law of one price is applied as if the substitute were a perfect match. Typically, the position in the substitute is opposite to that in the primary security in order to offset the most significant or unwanted risk inherent in the primary security. Other risks are left purposely unhedged, but if the substitute is well chosen, they are controllable (except in highly leveraged positions). Like pure arbitrage, relative arbitrage can be both *static* and *dynamic*. Let us consider examples of static relative arbitrage.

Suppose you buy $100 million of a 30-year U.S. government bond. At the same time you sell (short) $102 million of a 26-year bond. The amounts $100 and $102 are chosen through "duration matching" (see Chapter 14 of Volume III for a discussion of duration), which ensures that when interest rates go up or down by a few basis points the gains on one position exactly offset the losses on the other. The only way the combined position makes or loses money is when interest rates do not change in parallel; that is, the 30-year rates change by more or less than the 26-year rates. The combined position is not risk free. It is speculative, but only in a secondary risk factor. Investors hardly distinguish between 30- and 26-year rates; they worry about the overall level of rates. The two rates tend to move closely together. The relative arbitrageur bets that they will diverge.

The bulk of swap trading in the world relies on static relative arbitrage. An interest rate swap dealer agrees to pay a fixed-coupon stream to a corporate customer, himself an issuer of a fixed-rate bond. The dealer hedges by buying a fixed-coupon government bond. He eliminates any exposure to interest rate movements as coupon receipts from the government bond offset the swap payments, but is left with swap spread risk. If the credit quality of the issuer deteriorates, the swap becomes unfair and the combined position has a negative present value to the dealer.

Dynamic relative arbitrage is slightly more complicated in that the hedge must be rebalanced continuously according to very specific computable rules. A seller of a three-year over-the-counter equity call may hedge by buying three- and six-month calls on the exchange and shorting some of the stock. He then must rebalance the number of shares he is short on a daily basis as the price of those shares fluctuates. This so-called delta hedge (see Chapter 41 of Volume II) eliminates exposure to the price risk. The main unhedged exposure is to the implied volatility differences between the options sold and bought. In the preceding static swap example, the swap dealer may elect not to match the cash flows exactly on each swap he enters into. Instead, he may take positions in a small number of benchmark bonds in order to offset the cash flows in bulk. This shortcut, however, will require him to dynamically rebalance the portfolio of bonds.

Relative-value arbitrage is defined as generating profit today by statically or dynamically matching current and future obligations to nearly offset each other, net of incurring closely estimable financing costs. To an untrained eye, the difference between relative-value arbitrage and speculation is tenuous. To a professional, the two are easily discernible. A popular equity trading strategy called *pairs trading* is a good case in point. The strategy of buying Pfizer (PFE) stock and selling GlaxoSmithKline (GSK) is pure speculation. One can argue that both companies are in pharmaceuticals, both are large, and both have similar research-and-development budgets and new drug pipelines. The specific risks of the two companies, however, are quite different and they cannot be considered close substitutes. Buying Polish zlotys with British pounds and selling Czech korunas for British pounds is also an example of speculation, not of relative-value arbitrage. Polish zlotys and Czech korunas are not close substitutes. An in-between case, but clearly on the speculative side, is called a *basis trade*. An airline needing to lock in the future prices of jet fuel, instead of entering into a long-term contract with a refiner, buys a series of crude oil futures, the idea being that supply shocks that cause oil prices to rise affect jet fuel in the same way. When prices increase, the airline pays higher prices for jet fuel, but profits from oil futures offset those increases, leaving the total cost of acquiring jet fuel unchanged. Buying oil futures is appealing as it allows liquidating the protection scheme when prices decline instead of rising, or getting out halfway through an increase. This trade is not uncommon, but it exposes the airline to the *basis* risk. When the supply shocks take place at the refinery level, not the oil delivery level, spot jet fuel prices may increase more rapidly than crude oil futures.

Most derivatives dealers espouse the relative-value arbitrage principle. They sell options and at the same time buy or sell the underlying stocks, bonds, or mortgages in the right proportions to exactly offset the value changes of the sold option and the position in the underlying financial asset. Their lives are, however, quite complicated in that they have to repeat the exercise every day as long as the options they sold are alive, even if they do not sell additional options. This is because the appropriate proportions of the underlyings they need to buy or sell change every day. These proportions or hedge ratios depend on changing market factors. It is these market factors that are the secondary risks the dealers are exposed to. The dynamic rebalancing of the positions serves to create a close substitute to the options sold, but it does not offset all the risks.

Relative-value arbitrage in most markets relies on a building block of a static or dynamic *cash-and-carry* trade. The static version of the cash-and-carry trade (explained in Chapter 42 of Volume III) consists typically of a spot purchase (for cash) and a forward sell, or the reverse. The dynamic trade, like in the preceding option example, consists of a series of spot purchases or sales at different dates and a contingent payoff at the forward date. The glue that ties the spot and the forward together is the cost of financing, the "carry", of the borrowing to buy spot or lending after a spot sale. Even the most complicated structured derivative transactions are combinations of such building blocks across different markets. When analyzing such trades, focusing on institutional and market infrastructure details in each market can only becloud this basic structure of arbitrage. This book clarifies the essence of such trades by emphasizing common elements. It also explains why most institutions rely on the interaction of dealers on large trading floors to take advantage of inter-market arbitrages. The principle of arbitrage is exploited not only to show what motivates traders to participate *in* each market (program trading of stock index futures vs. stock baskets, fixed coupon stripping in bonds, triangular arbitrage in currencies), but also what drives the risk arbitrage *between* markets (simultaneous trades in currencies in money markets, hedging mortgage servicing contracts with swap options, etc.).

Many readers view *no-arbitrage conditions* found in finance textbooks as strict mathematical constructs. It should be clear from the preceding discussion that they are not mathematical at all. These equations do not represent the will of God like those pertaining to gravity or thermodynamics in physics. They stem from and are continuously ensured by the most basic human characteristic: greed. Dealers tirelessly look to discover pure and relative-value arbitrage (that is, opportunities to buy something at one price and to sell a disguised version of the same thing for another price). By executing trades to take advantage of the temporary deviations from these paramount rules, they eliminate them by moving prices back in line, where riskless money cannot be made and, by extension, the equations are satisfied. Each side of a financial math equation represents the present value of a pure arbitrage strategy. By spotting pure arbitrage and contrasting it with speculation, one is able to identify the in-between case of relative-value arbitrage (sometimes also referred to as *risk arbitrage*). Apart from the ever-shrinking commissions, most traders earn profit from "spread"—a reward for relative-value risk arbitrage. A swap trader who fixes the borrowing rate for a corporate client hedges by selling Treasury bonds. He engages in a relative-value trade (swaps vs. government bonds) that exposes him to swap spread movements. A bank that borrows by opening new checking deposits and lends by issuing mortgages eliminates the risk of parallel interest rate movements (which perhaps affect deposit and mortgage rates to the same degree), but leaves itself exposed to yield curve tilts (nonparallel interest rate movements) or default risk. In all these cases, the largest risks (the exposure to interest rate changes) are hedged out, and the dealer is left exposed to secondary ones (swap spread, default).

Most forms of what is conventionally labeled as *investment* under our definition qualify as *speculation*. A stock investor who does not hedge, or risk-share in some way, is exposed to the primary price risk of his asset. It is expected that in our lives, barring short-term fluctuations, the value of our assets increases over time. The economy in general grows, productivity increases, and our incomes rise as we acquire more experience. We find ourselves having to save for future consumption, family, and retirement. Most of the time, often indirectly through pension and mutual funds, we "invest" in real estate, stocks, and bonds. Knowingly or not, we speculate. Financial institutions, as their assets grow, find themselves in the same position. Recognizing that fact, they put their capital to use in new products and services. They speculate on their success. However, a lot of today's institutional dealers' trading activity is not driven by the desire to bet their institutions' capital on buy-low/sell-high speculative ventures. Institutional traders do not want to take primary risks by speculating on markets rising or falling. Instead, they hedge the primary risks by simultaneously buying and selling or borrowing and lending in spot, forward, and option markets. They leave themselves exposed only to secondary "spread" risks. Well-managed financial institutions are compensated for taking those secondary risks. Even most apt business school students often misunderstand this fine distinction between speculation and relative-value arbitrage. Chief executive officers often do, too. Nearly everyone has heard of the Barings, IG Metallgeselshaft, and Orange County fiascos of the 1990s. The history of finance is filled with examples of financial institutions gone bankrupt as a result of gambling.

Institutional trading floors are designed to best take advantage of relative arbitrage within each market. They are arranged around individual *trading desks*, surrounded by associated marketing and clearing teams, each covering customers within a specific market segment. Trading desks that are likely to buy each other's products are placed next to each other. Special *proprietary desks* (for short called *prop* or *arb desks*) deal with many customer desks of the same firm or other firms and many outside customers in various markets. Their job is to specifically focus on relative-value trades or outright speculation across markets. The distinction between the two types of desks—customer versus proprietary—is in constant flux as some markets expand and some shrink. Trading desks may collaborate in the types of transactions they engage in. For example, a money market desk arranges an issuance of short-term paper whose coupon depends on a stock index. It then arranges a trade between the customer and its swap desk to alter the interest rate exposure profile and between the customer and the equity derivatives desk to eliminate the customer's exposure to equity risk. The customer ends up with a low cost of financing and no equity risks. The dealer firm lays off the swap and equity risk with another institution. Hundreds of such intermarket transactions take place every day in the dealing houses in London, New York, and Hong Kong.

Commercial banks operate on the same principle. They bundle mortgage, car loan, or credit card receipts into securities with multiple risk characteristics and sell the unwanted ones to other banks. They eliminate the prepayment risk in their mortgage portfolios by buying swaptions from swap dealers.

FINANCIAL INSTITUTIONS: ASSET TRANSFORMERS AND BROKER-DEALERS

Financial institutions can be broadly divided into two categories based on their *raison d'être*:

1. Asset transformers
2. Broker-dealers

The easiest way to identify them is by examining their balance sheets. Asset transformers' assets have different legal characteristics from their liabilities. Broker-dealers may have different mixes on the two sides of the balance sheet, but the categories tend to be the same.

An asset transformer is an institution that invests in certain assets, but issues liabilities in a form designed to appeal to a particular group of customers. The best example is a commercial bank. On the asset side, a bank issues consumer (mortgage, auto) and business loans, invests in bonds, and so on. The main form of liability it issues is checking accounts, saving accounts, and certificates of deposit (CDs). Customers specifically desire these vehicles as they facilitate their day-to-day transactions and often offer security of government insurance against the bank's insolvency. For example, in the United States the Federal Deposit Insurance Corporation (FDIC) guarantees all deposits up to $100,000 per customer per bank. The bank's customers do not want to invest directly in the bank's assets. This would be quite inconvenient, as they would have to buy and sell these bulky assets frequently to meet their normal living expenditures. From a retail customer's perspective, the bank's assets often have undesirably long maturity that entails price risk if they are sold quickly, and they are offered only in large denominations. In order to attract funding, the bank repackages its mortgage and business loan assets into liabilities, such as checking accounts and CDs, that have more palatable characteristics—immediate bankomat access, small denomination, short maturity, and deposit insurance. Another example of an asset transformer is a mutual fund (or a unit investment trust). A mutual fund invests in a diversified portfolio of stocks, bonds, or money market instruments, but issues to its customers small-denomination, easily redeemable participation shares (unit trust certificates) and offers a variety of services like daily net asset value calculation, fund redemption and exchange, or a limited check-writing ability. Other large-asset transformers are insurance companies that invest in real estate, stocks, and bonds (assets), but issue policies with payouts tied to life or hardships events (liabilities).

Asset transformers are subject to special regulations and government supervision. Banks require bank charters to operate, are subject to central bank oversight, and must belong to deposit insurance schemes. Mutual funds' regulation is aimed at protecting small investors (e.g., as provided for by the Investment Company Act in the United States). Insurance companies rates are often sanctioned by state insurance boards. The laws in all these cases set specific forms of legal liabilities asset transformers may create and sound investment guidelines they must follow (e.g., percentage of assets in a particular category). Asset transformers are compensated largely for their role in repackaging their assets with undesirable features into liabilities with customer-friendly features. That very activity automatically introduces great risks into their operations. Banks' liabilities have much shorter duration (checking accounts) than their assets (fixed-rate mortgages). If interest rates do not move in parallel, the spread they earn (interest differential between rates charged on loans and rates paid on deposits) fluctuates and can be negative. They pursue relative-value arbitrage in order to reduce this *duration gap*.

Broker-dealers do not change the legal and functional form of the securities they own and owe. They buy stocks, currencies, mortgage bonds, leases, and so on, and they sell the same securities. As dealers, they own them temporarily before they sell them, exposing themselves to temporary market risks. As brokers, they simply match buyers and sellers. Broker-dealers participate in both primary sale and secondary resale transactions. They transfer securities from the original issuers to buyers, as well as from existing owners to new owners. The first is known as *investment banking* or *corporate finance*, the latter as dealing or *trading*. The purest forms of broker-dealers exist in the United States and Japan, where the laws have historically separated them from other forms of banking. Most securities firms in those two countries are pure broker-dealers (investment banking, institutional trading, and retail brokerage) with an addition of asset-transforming businesses of asset management and lending. In most of continental Europe, financial institutions are conglomerates commonly referred to as *universal banks*, as they combine both functions. In recent years, with the repeal of the Glass-Steagall Act in the United States and the wave of consolidations taking place on both sides of the Atlantic, U.S. firms have the possibility to converge more closely to the European model. Broker-dealers tend to be much less regulated than asset transformers, and the focus of the laws tends to be on small-investor protection (securities disclosure, fiduciary responsibilities of advisers, etc.).

Asset transformers and broker-dealers compete for each other's business. Securities firms engage in secured and unsecured lending and offer check writing in their brokerage accounts. They also compete with mutual funds by creating bundled or indexed securities designed to offer the same benefits of diversification. In the United States, the trading on the American Stock Exchange is dominated by ETFs (exchange-traded funds), holding company depositary receipts (HOLDRs), Qubes (so named because of their QQQQ ticker symbol), and the like, all of which are designed to compete with index funds, instead of ordinary shares. Commercial banks securitize their credit card and mortgage loans to trade them out of their balance sheets. The overall trend has been toward *disintermediation* (that is, securitization of previously transformed assets into more standardized, tradeable packages). As burdensome regulations fall and costs of securitization plummet, retail customers are increasingly given access to markets previously reserved for institutions.

PRIMARY AND SECONDARY MARKETS

From the welfare perspective, the primary role of financial markets has always been to transfer funds between suppliers of excess funds and their users. The users include businesses that produce goods and services in the economy, households that demand mortgage and consumer loans, governments that build roads and schools, financial institutions, and many others. All of these economic agents are involved in productive activities that are deemed economically and socially desirable. Throughout most of the history, it was bankers and banks who made that transfer of funds possible by accepting funds from depositors and lending them to kings, commercial ventures, and others. With the transition from feudalism to capitalism came the new vehicles of performing that transfer in the form of shares in limited liability companies and bonds issued by sovereigns and corporations. Stock, bond, and commodity *exchanges* were formed to allow original investors in these securities to efficiently share the risks of these instruments with new investors. This in turn induced many suppliers of funds to become investors in the first place as the risks of holding paper were diminished. Paper could be easily sold and funds recovered. A specialized class of traders emerged who dealt only with trading "paper" on the exchanges or over the counter (OTC). To them, paper was and is faceless. At the same time, the old role of finding new productive ventures in need of capital shifted from bankers to *investment bankers*, who, instead of granting loans, specialized in creating new shares and bonds for sale to investors for the first time. To investment bankers, the paper is far from faceless. Prior to the launch of any issue, the main job of an investment banker or his *corporate finance* staff, like that of a loan banker, is to evaluate the issuing company's business and its financial condition and to prepare a valuation analysis for the offered security.

As we stated before, financial markets for securities are organized into two segments defined by the parties to a securities transaction: primary and secondary markets. This segregation exists only in securities, not in private-party contracts like OTC derivatives. In private contracts, the primary market issuers also tend to be the secondary market traders, and the secondary market operates through assignments and mark-to-market settlements rather than through resale.

In primary markets, the suppliers of funds transfer their excess funds directly to the users of funds through a purchase of securities. An investment banker acts as an intermediary, but the paper-for-cash exchange is between the issuing company and the investor. The shares are sold either publicly, through an initial public offering or a seasoned offering, or privately through a *private placement* with "qualified investors," typically large institutions. Securities laws of the country in which the shares are sold spell out all the steps the investment bank must take in order to bring the issue to market. For example, in the United States, the shares must be registered with the Securities and Exchange Commission (SEC), a prospectus must be presented to new investors prior to a sale, and so on. Private placements follow different rules, the presumption being that large qualified investors need less protection than retail investors. In the United States, they are governed by Rule 144-A, which allows their subsequent secondary trading through a system similar to an exchange.

In secondary markets, securities are bought and sold only by investors without the involvement of the original user of funds. Secondary markets can be organized as exchanges or as OTC networks of dealers connected by phone or computer, or the hybrid of the two. The Deutsche Börse and the New York Stock Exchange (NYSE) are examples of organized exchanges. It is worth noting, however, that exchanges differ greatly from each other. The NYSE gives access to trade flow information to human market makers called *specialists* to ensure the continuity of the market making in a given stock, while the Tokyo Stock Exchange is an electronic market where continuity is not guaranteed, but no dealer can earn monopoly rents from private information about buys and sells. Corporate and government bond trading are the best examples of OTC markets. There, as in swap and currency markets, all participants are dealers who trade one on one for their own account. They maintain contact with each other over a phone and computer network, and jointly police the fair conduct rules through industry associations. For example, in the OTC derivatives markets, the International Swaps and Derivatives Association (ISDA) standardizes the terminology used in quoting the terms and rates and formalizes the documentation used in confirming trades for a variety of swap and credit derivative agreements. The best example of a hybrid between an exchange and an OTC market is the National Association of Securities Dealers Automated Quotation System (NASDAQ) in the United States. The exchange is only virtual, as participants are connected through a computer system. Access is limited to members only, and all members are dealers.

Developing countries strive to create smooth-functioning secondary markets. They often rush to open stock exchanges even though there may only be a handful of companies large enough to have a significant number of dispersed shareholders. In order to improve the liquidity of trading, nascent exchanges limit the number of exchange seats to very few, the operating hours to sometimes only one per day, and so on. All these efforts are aimed at funneling all buyers and sellers into one venue. This parallels the goals of the specialist system on the NYSE. Developing countries' governments strive to establish a well-functioning government bond market. They start by issuing short-term obligations and introduce longer maturities as quickly as the market will have an appetite for them.

The main objective in establishing these secondary trading places is to lower the cost of raising capital in the primary markets by offering the primary market investors a large outlet for risk sharing. Unless investors are convinced that they can easily get in and out of these securities, they will not buy the equities and bonds offered by the issuers (local businesses and governments) in the first place. This "tail wagging the dog" pattern of creating secondary markets first is very typical not only for lesser developed nations, but is quite common in introducing brand new risk classes into the marketplace. In the late 1980s, Michael Milken's success in selling highly speculative high-yield bonds to investors relied primarily on creating a secondary OTC market by assuring active market making by his firm Drexel Burnham Lambert. Similarly, prior to its collapse in 2002, Enron's success in originating energy forwards and contingent contracts was driven by Enron's ability to establish itself as a virtual exchange of energy derivatives (with Enron acting as the monopolist dealer, of course). In both of these cases, the firms behind the creation of these markets failed, but the primary and secondary markets they started remained strong, the high-yield market being one of the booming high performers during the tech stock bubble collapse in 2000 to 2002.

MARKET PLAYERS: HEDGERS VERSUS SPECULATORS

According to a common saying, nothing in life is certain except death and taxes. No investment in the market is riskless, even if it is in some way guaranteed. Let us challenge some seemingly intuitive notions of what is risky and what is safe.

Sparkasse savers in Germany, postal account holders in Japan, and U.S. Treasury bill investors, for most intents and purposes, avoid default risk and are guaranteed a positive nominal return on their savings. T-bill and CD investors lock in the rates until the maturity of the instruments they hold. Are they then risk-free investors and not speculators? They can calculate in advance the exact dollar amount their investment will pay at maturity. After subtraction of the original investment, the computed percentage return will always be positive. Yet, by locking in the cash flows, they are forgoing the chance to make more. If, while they are holding their CD, short-term or rollover rates increase, they will have lost the extra *opportunity* return they could have earned. We are hinting here at the notion of *opportunity cost of capital* common in finance.

Let us consider another example. John Smith uses the $1,000 he got from his uncle to purchase shares in XYZ Corporation. After one year, he sells his shares for $1,100. His annual return is 10%. Adam Jones borrows $1,000 at 5% from his broker to purchase shares in XYZ Corporation. After one year, he sells his shares for $1,100. His annual return is 10% on XYZ shares, but he has to pay 5% or $50 interest on the loan, so his net return is 5%. Should we praise John for earning 10% on his capital and scold Adam for earning only 5%? Obviously not. Adam's cost of capital was 5%. So was John's! His was the nebulous opportunity cost of capital, or a *shadow cost*. He could have earned 5% virtually risk free by lending to the broker instead of investing in risky shares. So his relative return, or excess return, was only 5%. In our T-bill or CD example, one can argue that an investor in a fixed-rate CD is a speculator as he gambles on the rates not increasing prior to the maturity of his CD. The fact that his net receipts from the CD at maturity are guaranteed to be positive is irrelevant. There is nothing special about a 0% threshold for your return objective (especially if one takes into account inflation).

All investors who take a position in an asset, whether by borrowing or using owned funds, and the asset's return over its life is not contractually identical to the investor's cost of capital, can be considered *speculators*. This definition is relative only to some benchmark cost of capital. In this sense both Adam and John speculate by acquiring shares whose rate of return differs from their cost of capital of 5%. An outright CD investment is speculative, as the rate on the CD is not guaranteed to be the same as that obtained by leaving the investment in a variable-rate money market account. A homeowner who takes out a fixed-rate mortgage to finance a house purchase is a speculator even though he fixes his monthly payments for the next 30 years! When he refinances his loan, he cancels a prior bet on interest rates and places a new one. In contrast, an adjustable-rate mortgage borrower pays the fair market rate every period equal to the short-term rate plus a fixed margin.

Most financial market participants can be divided into two categories based on whether their capital is used to place bets on the direction of the market prices or rates or whether it is used to finance holdings of sets of transactions that largely offset each other's primary risks: speculators and hedgers.

Speculators are economic agents who take on explicit market risks in order to earn returns in excess of their cost of capital. The risks they are exposed to through their investments are not offset by simultaneous "hedge" transactions. Hedgers are economic agents who enter into simultaneous transactions designed to have offsetting market risks in such a way that the net returns they earn are over and above their cost of capital. All arbitrageurs, whether pure or relative, are hedgers. They aim to earn nearly risk free returns after paying all their financing costs. A pure arbitrageur's or strict hedger's returns are completely risk free. A relative arbitrageur's returns are not risk free; he is exposed to secondary market risks.

All investors who use their capital to explicitly take on market risks are speculators. Their capital often comes in the form of an outside endowment. Mutual funds obtain fresh funds by shareholders sending them cash. Pension funds get capital from payroll deductions. Insurance companies sell life or hazard policies and invest the premiums in stocks, bonds, and real estate. Individual investors deposit cash into their brokerage accounts in order to buy, sell, or short sell stocks and bonds. In all these cases, the investors use their funds (that is, sacrifice their cost of capital) to bet on the direction of the market they invest in. They "buy" the services of brokers and dealers who facilitate their investment strategies. In order to help these investors improve the precision of the bets they take, broker-dealers, who are hedgers by nature, invent new products, which they "sell" to the investors. These can be new types of bonds, warrants and other derivatives, new classes of shares, new types of trusts, and annuities. Often, the division of the players into speculators and hedgers is replaced by the alternative terms of buy-side participants and sell-side participants.

Buy-side players are investors who do not originate the new investment vehicles. They choose from a menu offered to them by the sell-side players. The sell-siders try to avoid gambling their own capital on the explicit direction of the market. They want to use their capital to finance the hedge, that is, to "manufacture" the new products. As soon as they sell them, they look to enter into a largely offsetting trade with another counterparty or to hedge the risks through a relative arbitrage strategy. Often the sell-sider's hedge strategies are imperfect and take time to arrange. That is when sell-siders act as speculators. The hedger/speculator compartmentalization is not exactly equivalent to sell-/buy-side division. Sell-siders often act as both hedgers and speculators, but their mindset is more like that of the hedger ("to find the other side of the trade"). Buy-siders enter into transactions with sell-siders in order to get exposed to, or alter how they are exposed to, market risks ("to get in on a trade").

We use quotes around the words *buy* and *sell* to emphasize that the sell-sider does not necessarily sell a stock or bond to a buy-sider. He can just as well buy it. But he hedges his transaction while the buy-sider does not.

Geographically, the sell-side resides in global financial centers, like New York or London, and is represented by the largest 50 global financial institutions. The buy-side is very dispersed and includes all medium and smaller banks with mostly commercial business, all mutual and pension funds, some university endowments, all insurance companies, and all finance corporations. The buy/sell and hedger/speculator distinctions have recently become blurred. Larger regional banks in the United States, which have traditionally been buy-side institutions, have started their own institutional trading businesses. They now offer security placement and new derivative product services to smaller banks and thrifts. In the 1990s, some insurance companies established sell-side trading subsidiaries and used their capital strength and credit rating to compete vigorously with broker-dealers. Most of these subsidiaries have the phrase "Financial Products" inserted in their name (e.g., Gen Re Financial Products or AIG FP).

One type of company that can be, by design, on both the buy and sell side is a hedge fund. Hedge funds are capitalized like typical speculators (read: investment companies), similar to mutual funds, but without the regulatory protection of the small investor. Yet almost all hedge fund strategies are some form of relative-value arbitrage; that is, they are hedges. The original capital is used only to acquire leverage and to replicate a hedge strategy as much as possible. Most hedge funds have been traditionally buy-siders. They have tended not to innovate, but to use off-the-shelf contracts from dealers. Sometimes, however, hedge funds grow so large in their market segment that they are able to wrest control of the demand-and-supply information flow from the dealers and are able to sell hedges to the dealers, effectively becoming sellers of innovative strategies. In the late 1990s, funds like Tiger, AIM, and Long-Term Capital Management, sometimes put on very large hedged positions, crowding dealers into speculative choices as the supply of hedges was exhausted by the funds. In the early 2000s, hedge funds retrenched to their traditional buy-side role as the average size of the fund declined. However, by 2006 the number of funds increased dramatically, topping 8,000 with assets over $1.2 trillion and some of the larger funds playing both buy- and sell-side roles.

SUMMARY

From a structural point of view, all financial markets are the same in that they have the same (1) distinct primary issuers and secondary traders, (2) spot, forward, and option vehicles, (3) types of traders with two different trading motives (hedgers and speculators), and (4) primary trading strategy driving most of activity is relative risk arbitrage. The goal of relative risk arbitrage is to earn reward for taking on exposure to secondary risk factors while eliminating primary directional risks through static or dynamic hedging. Taking an arbitrage perspective, in this chapter we explained how risk sharing drives most of the world's trading activity.

REFERENCES

Arrow, K. J. (1951). Alternative approaches to the theory of choice in risk-taking situation. *Econometrica* 19, 4: 404–737.

Baxter, M. and Rennie, A. (1996). *Financial Calculus: An Introduction to Derivative Pricing*, Cambridge: Cambridge University Press.

Cox, J. C. and Rubinstein, M. (1985). *Options Markets*. Englewood Cliffs, NJ: Prentice-Hall.

Debreu, G. (1959). *Theory of Value*. New Haven, CT: Yale University Press.

Dubil, R. (2004). *An Arbitrage Guide to Financial Markets*. Chichester, England: John Wiley & Sons.

Fabozzi, F. J., Davis H. A., and Choudhry, M. (2006). *Introduction to Structured Finance*. Hoboken, NJ: John Wiley & Sons.

Huang, C. and Litzenberger, R. H. (1998). *Foundations for Financial Economics*. New York, NY: North-Holland.

Kolb, R. W. and Overdahl, J. A. (2003). *Financial Derivatives*, 3rd Ed. Hoboken, NJ: John Wiley & Sons.

Merton, R. C. (1990). *Continuous-Time Finance*. Cambridge, MA: Blackwell Publishers.

Neftci, S. N. (2004). *Principles of Financial Engineering*. London, UK: Elsevier Academic Press.

Shiller, R. J. (2003). *The New Financial Order: Risk in the 21st Century*. Princeton, NJ: Princeton University Press.

Strang, G. (1980). *Linear Algebra and Its Applications*. Second Ed. New York, NY: Harcourt Brace Jovanovich.

CHAPTER 9

Complete Markets

LES GULKO, PhD
Principal, Cove Island Ventures

Abstract: A financial market is complete if there exist contracts to insure against all possible eventualities. First, complete markets are desirable because they enable producers, consumers, and investors to allocate scarce resources, invest capital, and share financial risks in a Pareto-efficient way. For example, calls options, put options, and other derivatives are socially beneficial because they enhance completeness. Second, complete markets in the Arrow-Debreu space provide state-of-the-art analysis of capital markets and capital structures. For example, arbitrage-free pricing is feasible only in a complete market; and investor expectations are easy to infer from complete market prices. Finally, the complete market theory offers guidance to financial entrepreneurs regarding new securities, investment strategies and capital market architecture.

Keywords: complete market, Pareto-efficient allocation, optimal allocation of risk, Arrow-Debreu space, Arrow-Debreu prices, risk-neutral probabilities, arbitrage-free preference-free pricing, decoding market prices, missing markets, financial innovation, mean-variance theory

The two main models of securities markets are the mean-variance model, the subject of Chapter 9, and the *complete market* model, the subject of this chapter. A securities market is said to be complete if for every future state of the world there is a *state security* (or a portfolio of securities) that pays \$1 in that state and zero in every other state. Although it may sound abstract, the notion of complete markets is integral to any constructive discussion of asset pricing, hedging, arbitrage and other topics in finance. To be sure, the real markets are not complete, but it is necessary to know the properties of completeness in order to understand what we are missing. In this chapter, I outline the economic origins of market completeness and its use in financial economics.

ECONOMIC ORIGINS OF COMPLETE MARKETS

Hayek (1937, 1945) was the first to point out that competitive market prices carry useful information and, like the invisible hand, coordinate economic activity and allocation of scarce resources. In the absence of the right market structure though, the invisible hand of the price system is not potent enough to ensure efficient resource allocations, for example, *Pareto-efficient allocations* when no one can get better off without making someone else worse off.

The notion of complete markets first appeared in the groundbreaking Arrow-Debreu model of general equilibrium as a simple market structure supporting Pareto efficiency (Arrow, 1951; Debreu, 1951; Arrow and Debreu, 1954). In fact, in complete markets every competitive equilibrium is Pareto-optimal (by the First Welfare Theorem). The original Arrow-Debreu model featured a static setting with production, exchange, and consumption all taking place at once. Without a time dimension, the original model had no room for time uncertainty and financial assets, and market completeness was defined by frictionless exchange of *state goods* rather than state securities.

Apples and oranges are examples of state goods. It would have been inefficient and cumbersome for both producers and consumers if apples and oranges had been sold only as a package and not individually. Yet corporate bonds have always traded as packages consisting of an interest-rate component and a credit component. Only in recent years, with the advent of interest-rate derivatives and credit derivatives, has it become possible to unbundle corporate bonds by separating credit from interest rates. Whether building a portfolio of physical goods or financial securities, a complete set of basic building elements, such as state goods and state securities, helps to achieve better portfolio allocations.

Subsequent versions of the Arrow-Debreu model included a time interval between 0 and 1, time uncertainty and financial assets (Arrow, 1964; Debreu, 1959). At time 0, the agents first assign subjective probabilities to the states of the world at time 1 and next make allocations to time-0 consumption and to financial assets designed to finance time-1 consumption. This economy reaches a Pareto-efficient competitive equilibrium if financial assets form a complete market, that is, if for every future state of the world there is a state security (or a portfolio of securities) that pays $1 in that state and zero otherwise.

Furthermore, Arrow (1964) shows that the agents use financial securities not only to finance future consumption but also to allocate risk associated with different states, just as they allocate physical resources. A state security that pays $1 in an adverse state is, in effect, an insurance contract against the state-specific risk. The agents with different state-time preferences trade state securities as insurance contracts. Those who want to assume state-specific risk can sell state securities to those who loathe this risk. Therefore, financial assets provide hedging and transfer of risk, while complete financial markets offer Pareto-optimal allocation of risk.

COMPLETE MARKETS IN FINANCE

The theory presented in Arrow (1964) is a significant milestone in finance for several reasons, including the introduction of the *state space model* of securities markets. Consider a securities market on a time interval 0 to 1 with m possible states of the world at time 1. In the state space model, a financial security is represented by an $m \times 1$ vector of payoffs in different states at time 1. For example, a risk-free bond is a column-vector of 1s; a stock is a column-vector of state-dependent payoffs, and the state security insuring state i is a column-vector with 1 in position i and zero in all other positions.

Figure 9.1 is the payoff matrix for a market consisting of a risk-free bond and a stock. The state space Ω of the market includes four possible states at time 1. The first column represents a risk-free bond that pays 1 in every state. The second column represents a stock with state-dependent payoffs. The rows are the market payoffs in different states.

The state space model places a securities market in an m-dimensional vector space. It takes m linearly independent vectors to span the entire vector space. Fewer than m vectors span only a subset of the entire space. Similarly,

$$
\begin{array}{l c}
 & \begin{array}{cc} \textit{Bond} & \textit{Stock} \end{array} \\
\begin{array}{l} \textit{State } 1 \\ \textit{State } 2 \\ \textit{State } 3 \\ \textit{State } 4 \end{array} &
\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}
\end{array}
$$

Figure 9.1 An Arrow-Debreu Market for a Stock and Risk-Free Bond; the Columns Indicate the Bond and Stock Payoffs in Different States at Time 1

it takes m linearly independent securities—for example, m state securities—to construct every possible portfolio. Fewer than m securities can replicate only a subset of all possible portfolios.

In the state space model, a securities market is complete if and only if the number of linearly independent securities is equal to the number of states. The stock-bond market in Figure 9.1 is not complete because it has four states and only two securities. The bond and stock are linearly independent securities that span a two-dimensional vector space but not a four-dimensional vector space.

The state space model makes complete-markets theory operational and practical. The state space model is called the *Arrow-Debreu space*, and the state securities are also called the Arrow-Debreu securities and pure state-contingent claims. It is convenient to describe an Arrow-Debreu market by a payoff matrix, as the one describing a stock-bond market in Figure 9.1. The rank of the payoff matrix is a measure of market completeness. A higher rank implies greater completeness. A securities market is complete if and only if the rank equals the number of states m.

Complete financial markets benefit both the economy and society because they facilitate efficient allocation of capital and risk. Therefore, it has been a public policy in recent decades to enhance the completeness of financial markets. In particular, this public policy has fostered a rapid growth of the markets for derivative securities, especially, after Ross (1976a) demonstrated in a simple way that call and put options can be used to complete financial markets.

Let us complete the stock-bond market depicted in Figure 9.1 by introducing two call options written on the stock price S. The options have strike prices $K = 1, 2$, meaning that at time 1 the options pay $\max(0, S - K)$. Matrix $\mathbf{A}$ in Figure 9.2 shows the market payoffs in different states of the economy. The first column on the left is the risk-free bond, the second column is the stock, the third is the call

$$
\begin{array}{l c}
 & \begin{array}{cccc} \textit{Bond} & \textit{Stock} & \textit{Call } 1 & \textit{Call } 2 \end{array} \\
\begin{array}{l} \textit{State } 1 \\ \textit{State } 2 \\ \textit{State } 3 \\ \textit{State } 4 \end{array} &
\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 2 & 1 \end{bmatrix} = \mathbf{A}
\end{array}
$$

Figure 9.2 The Stock-Bond Market Is Completed by Two Stock Options option paying max(0, $S - 1$), and the last is the call option paying max(0, $S - 2$).

Market **A** is complete because is has four states and four linearly independent securities. Matrix **A** has an inverse $\mathbf{A}^{-1}$, a 4×4 matrix such that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$, where **I** is a 4×4 identity matrix with 1's along the diagonal and 0s off the diagonal. Therefore, the identity matrix **I** consists of Arrow-Debreu securities, while matrix $\mathbf{A}^{-1}$ defines the composition of four stock-bond-option portfolios that replicate the Arrow-Debreu securities.

Financial derivatives have grown explosively since their emergence in the 1970s. At present, derivatives contracts are used to reallocate not only traditional risks inherent in equities, fixed income and foreign exchange, but also nontraditional risks such as corporate credit, Fed policy, housing values, commodity prices, weather conditions, political elections and others. The ability to manage and distribute risk by means of financial derivatives is partially responsible for the flexibility and resiliency that the U.S. economy demonstrates in adverse circumstances.

SECURITIES PRICING IN COMPLETE MARKETS

Market completeness simplifies valuation of financial securities. Consider a one-period Arrow-Debreu market described by an $m \times n$ matrix **B**, where m is the number of states and n is the number of securities. There is a risk-free bond in the market whose price P at time 0 is known. Let **p** be a $1 \times n$ vector of securities prices at time 0. Then, by the no-arbitrage principle (Ross, 1976b), there exists a positive $1 \times m$ vector **f** that (in both complete and incomplete markets) justifies the following linear pricing rule:

$$\mathbf{p} = P \cdot \mathbf{f} \cdot \mathbf{B}$$

First, the linear pricing rule implies that securities values are additive, that is, if security a is a linear combination of securities x, y, z, then the price of a is the same linear combination of the prices of x, y, z. Therefore, in a complete market, the Arrow-Debreu securities span both the time-0 price and time-1 payoffs of every security. The prices of Arrow-Debreu securities are called the *state prices* or *Arrow-Debreu prices.*

Second, in complete markets, $P \cdot \mathbf{f}$ is a row-vector of Arrow-Debreu prices. The Arrow-Debreu prices can be inferred from market prices **p** by inverting matrix **B** (that has rank m in a complete market). If any state price is not positive, then there is an arbitrage opportunity in the market. Third, by linear pricing, the bond price P satisfies $P = P \cdot \mathbf{f} \cdot \mathbf{1}$, where **1** is a column-vector of 1s; therefore, $1 = \mathbf{f} \cdot \mathbf{1}$. Since **f** is strictly positive in the absence of arbitrage, it follows that **f** is a probability function defined on the state space of the market, Ω.

The linear pricing rule looks like the risk-free discounting of expected future values, which is consistent only with risk-neutral preferences. Therefore, **f** is called a *risk-neutral probability* function, risk-neutral probability measure and martingale measure. Much of modern valuation theory is devoted to the construction of *risk-neutral probabilities* and Arrow-Debreu prices.

In a complete securities market, a unique risk-neutral probability function obtains by inverting matrix **B**. In fact, market completeness can be characterized by uniqueness of risk-neutral probabilities. By contrast, incomplete markets feature multiple risk-neutral probability functions, leading to ambiguous securities prices. A way to resolve the ambiguity is to introduce utility functions that represent agents' preferences. A major advantage of arbitrage-free pricing in complete markets is that it is independent of agents' preferences.

Most single-period results generalize to multiperiod markets. A multiperiod market is a sequence of single-period markets. The states of the world in a multiperiod market are identified with different price trajectories. A multiperiod market is said to be complete if every state of the world is insurable. A multiperiod market is complete if and only if every intermediate single-period market is complete. See Pliska (1997, section 4.4) for proof and discussion.

The number of states can be very large in discrete-time markets and infinitely large in continuous-time markets. Fortunately, it takes fewer securities than the number of states to complete a multi-period market. Dynamic spanning makes it possible. Consider a multiperiod market for a stock whose price trajectories form a binary tree depicted in Figure 9.3. Every period, the stock price jumps to one of two possible states—up or down. Also, there is a risk-free bond that returns 1 in each state at the end of each period.

Since a single-period market has two states—up and down—and two linearly independent securities—stock and bond, this market is complete and every additional security is redundant; that is, it can be spanned by the stock-bond portfolio and priced by a linear combination of the stock and bond prices. Therefore, it is possible to dynamically complete the multiperiod market by rebalancing the stock-bond portfolio every single period.

The binary tree in Figure 9.3 represents eight trajectories of the stock price over three time periods and eight

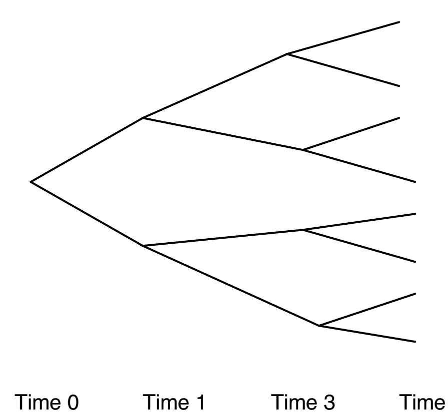


Figure 9.3 The Binary Tree Represents Potential Trajectories of the Stock Price that at Every Node Moves Either Up or Down

final states of the stock price at time 4. These states can be dynamically insured using only a stock and bond. Therefore, it takes fewer securities than the number states to dynamically complete this multiperiod market.

Dynamic rebalancing of the stock-bond portfolio is the essence of the binomial stock option pricing in Cox, Ross, and Rubinstein (1979). Since every period the stock-bond model is complete, a stock option is a redundant security that can be replicated and priced by the stock and bond. Dynamic replication of put options on stock market indices (using cash and stock-index futures) has been commercialized under the name of portfolio insurance. Portfolio insurance failed to fulfill its promise as the stock market crushed in October 1987, revealing that some risks were not insured.

A discrete-time model converges to a continuous-time model as the duration of each single period contracts to zero. In continuous time, a stock and bond are linearly independent every instant while a stock option is redundant. Since the stock, bond, and stock option are linearly dependent, each one of them can be replicated by the other two. For example, the risk-free bond can be replicated by continuous rebalancing of a portfolio that includes the stock and stock option.

This procedure is employed in the Black-Scholes stock option model. Every instant, the stock and stock option are combined in a risk-free portfolio, called the riskless hedge, that eliminates the equity risk. In the absence of arbitrage, the riskless hedge must generate the same rate of return as the risk-free bond. This no-arbitrage condition leads to an equation that is solved for the stock option price in terms of the current stock and bond prices. See Merton (1973) and Black and Scholes (1973).

Arbitrage-free valuation in complete markets is used for pricing both derivative securities and primary securities (such as stocks, bonds, and other securities in positive net supply; by contrast, exchange-traded derivatives are in zero net supply). In particular, the Black-Scholes-Merton options valuation technology has been used extensively for pricing fixed-income securities. Some examples are Vasicek (1977); Brennan and Schwartz (1979); Cox, Ingersoll, and Ross (1985); Ho and Lee (1985); Hull and White (1990); Black, Derman, and Toy (1990); and Heath, Jarrow, and Morton (1992).

Complete markets make arbitrage-free preference-free valuation feasible. The absence of market arbitrage guarantees the existence of risk-neutral probabilities that reduce asset pricing to the risk-free discounting of expected future values. See Cox and Ross (1976) and Harrison and Kreps (1979). However, the absence of arbitrage does not guarantee the existence of unique risk-neutral probabilities. In fact, in incomplete markets, there are many feasible risk-neutral measures, leading to ambiguous pricing. No-arbitrage offers unambiguous preference-free asset pricing only in complete markets.

Financial decision making in complete markets is analogous to solving a system of linear simultaneous equations when the number of unknowns equals the number of (linearly independent) equations. Such systems have unique unambiguous solutions. Whenever the number of unknowns exceeds the number of equations, there are multiple solutions, creating ambiguity and uncertainty. This situation describes incomplete markets and often prevails in the real markets.

Consider, for example, the irrelevance theorems set forth by Modigliani and Miller (1958, 1961): The firm's value is independent of the capital structure policy and dividend policy (in the absence of taxes and bankruptcy costs). In other words, changing leverage and redistributing assets should not affect the firm's value. The proof relies on the absence of arbitrage and linear pricing. As such, this proof can work only in complete markets. In reality, the Modigliani-Miller theorems often fail, and incomplete markets may be one of the reasons.

PRICE INFORMATION IN COMPLETE MARKETS

It has been known at least since Hayek (1937, 1945) that competitive market prices contain consensus expectations and other information of interest to both investors and economists. This information, however, remained inaccessible until Breeden and Litzenberger (1978) and Banz and Miller (1978) showed that risk-neutral probabilities could be inferred from options prices in complete markets. This was a pivotal discovery because it offered a practical way of decoding market prices and because risk-neutral probabilities approximate consensus probability beliefs.

The risk-neutral probability distributions implied in stock option prices have been extensively studied by Rubinstein (1994), Rubinstein and Jeckwerth (1996), Ait-Sahalia and Lo (1998), Buchen and Kelley (1996), and others. The implied risk-neutral distributions vary over time but their general properties remain the same and include a bell shape, realistic dispersions, and fat left-hand tails. Also, implied risk-neutral probabilities resemble plausible consensus beliefs. Risk-neutral probabilities recovered from the prices of fixed-income and foreign-exchange securities have similar properties.

From a technical viewpoint, extracting investor expectations from market prices is similar to identifying weather patterns, mapping out oil reservoirs, decoding noisy radio signals and other problems where a full image is to be constructed on the basis of partial observation. More complete observations generate more accurate images. In particular, greater market completeness means a more complete price set that allows for more accurate inference. In this sense, complete market prices are more informative than incomplete market prices.

Since derivative securities improve completeness, their prices represent a fertile source of recoverable market information. For example, stock index options are used to infer expected stock market volatility. Fed funds futures and options are used to infer the consensus probabilities of the future Federal Reserve policy on interest rates. Credit derivatives are used to infer the expected probabilities of bond defaults.

Table 9.1 Prices of September 1996 Eurodollar Futures Options on July 1, 1996

Strike Price K	93.75	94.00	94.25	94.50	94.75
Call Price	0.52	0.30	0.12	0.03	0.01

Reported in *The Wall Street Journal* on July 2, 1996.

Let's take a sample of Eurodollar futures options and invert their market prices. The expiration price of a Eurodollar futures contract is $F = 100 - y$, where y is the spot rate on a three-month Eurodollar deposit at the time the futures expires. For example, if the spot is $y = 5.5\%$, then $F = 94.5$. There is a liquid market for options on Eurodollar futures. Both the options and underlying futures expire at the same time. At expiration, a call option is worth $\max(0, F - K)$, where K is the contractual strike price. The strike prices are set 25 basis points apart, for example, 95.00, 95.25, 95.50, and so on.

Consider a sample of Eurodollar futures options on a time interval [0, 1]. Time 0 is July 1, 1996 and time 1 is the option expiration date September 16, 1996. A set of call prices at time 0 is listed in Table 9.1, and the price of a Treasury bill, that pays \$1 at maturity at time 1, is $P =$ \$0.98912.

Figure 9.4 describes an Arrow-Debreu market for five call options and the Treasury bill. The state space Ω has six states indexed by possible realizations of the futures price F at time 1. The row-vector $\mathbf{p}$ consists of securities prices at time 0. Matrix $\mathbf{A}$ is the payoff matrix. The first column of $\mathbf{A}$ is the Treasury bill while the other columns are call options that pay $\max(0, F - K)$ at expiration. For example, zeros in the first row mean that all calls expire out of the money in the state $F = 93.75$. This market is complete because it has six states and six linearly independent securities.

The implied risk-neutral distribution $\mathbf{f}$, obtained by inverting the payoff matrix $\mathbf{A}$, is plotted in Figure 9.5. Distribution $\mathbf{f}$ features common properties despite a small market size. First, $\mathbf{f}$ is strictly positive implying no arbitrage

$$
\begin{array}{c c}
 & \begin{array}{cccccc} \text{T-bill} & \multicolumn{5}{c}{\text{Call options with strike prices } K =} \\ & 93.75 & 94.00 & 94.25 & 94.50 & 94.75 \end{array} \\
\begin{array}{c} \text{State Space} \\ \Omega \text{ of } F \\ 93.75 \\ 94.00 \\ 94.25 \\ 94.50 \\ 94.75 \\ 95.00 \end{array} &
\begin{array}{c} \\ \\ \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & .25 & 0 & 0 & 0 & 0 \\ 1 & .50 & .25 & 0 & 0 & 0 \\ 1 & .75 & .50 & .25 & 0 & 0 \\ 1 & 1.00 & .75 & .50 & .25 & 0 \\ 1 & 1.25 & 1.00 & .75 & .50 & .25 \end{array}\right] = \mathbf{A} \end{array}
\end{array}
$$

$$[0.98912 \quad 0.52 \quad 0.30 \quad 0.12 \quad 0.03 \quad 0.01\,] = \mathbf{p}$$

Figure 9.4 An Arrow-Debreu Market for Eurodollar Futures Options

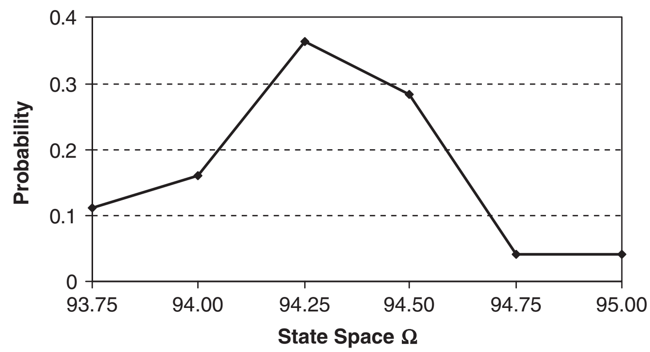


Figure 9.5 Risk-Neutral Probabilities $\mathbf{f}$ Implied in the Prices of Eurodollar Futures Options $\mathbf{f} = \mathbf{p}{\cdot}\mathbf{A}^{-1}/P =$ [0.1103 0.1617 0.3640 0.2831 0.0404 0.0404]

opportunity. Second, $\mathbf{f}$ has a bell shape. Third, the standard deviation of the random futures price under $\mathbf{f}$—about 30 bps—is a reasonable volatility estimate for short-term interest rates over two-and-half months from July 1, 1996, to September 16, 1996. Finally, $\mathbf{f}$ resembles plausible consensus beliefs (if it is drawn as a smooth curve).

Some additional observations are in order. First, persistent empirical properties of the implied probability distributions justify the common use of bell-shaped investor beliefs that are either postulated outright (as the normal distributions in the capital asset pricing model) or parameterized by stochastic price processes (as in many options-pricing models).

Second, while the arbitrage theory treats risk-neutral probabilities as abstract martingale measures, the empirical evidence demonstrates that the implied probability distributions resemble plausible market expectations, shaped as bell curves with realistic volatilities. Therefore, the risk-neutral probabilities are subject to some unknown conditions, yet to be discovered.

Third, it is not necessary to rely exclusively on the prices of options in order to recover the risk-neutral probabilities. For example, in Figure 9.4, the underlying futures contract can replace a call option in the payoff matrix $\mathbf{A}$. Breeden and Litzenberger (1978) find an ingenious method to recover risk-neutral probabilities exclusively from options prices. In the continuous state space, the second derivative of the option premium with respect to the strike price K is proportional to the risk-neutral probability density. This follows from the continuous-space version of the linear pricing rule.

A complete market structure helps us to infer investor expectations from the prices but does not necessarily help investors to form their expectations. In other words, it is possible that the same probability beliefs prevail whether or not the market is complete. But then the same asset prices and the same portfolio allocations may exit irrespective of market completeness. An incomplete market is said to be *effectively complete* if it features the same prices and the same Pareto optimal allocations as a complete market. For example, removing one option from the market in Figure 9.4 leaves the market incomplete but does not necessarily cause any price changes. See Ingersoll (1987) for a discussion of effective completeness.

MISSING MARKETS AND FINANCIAL INNOVATION

Uninsured states create gaps in the Arrow-Debreu space that represent *missing markets* as well as potential opportunities for *financial innovation* aimed at filling in the gaps. The notion of market completeness offers a systematic way to search for missing markets and new business ventures—find uninsured Arrow-Debreu states and create protection against the state-specific risks. Indeed, the design of housing futures, weather derivatives and other securities has followed this formula. Whether or not financial entrepreneurs are aware of the Arrow-Debreu space, their inventions tend to enhance market completeness.

Some of the new financial instruments include inflation-protected securities, mortgage-related securities, asset-backed securities, securitized insurance products, emerging markets, exchange-traded funds as well as a myriad of derivative products. Among recent investment strategies are portfolio indexing, fundamental indexing, fund-of-funds, long/short equities, market-neutral equities, event-driven trading, and other hedge-fund strategies. Electronic trading, decimalization, privatization and globalization of securities exchanges are some of the factors improving trading mechanisms and making execution faster and cheaper.

Since the early 1980s, U.S. investment portfolios have enjoyed improved allocational efficiency and solid returns thanks, in part, to diverse investment choices and risk-management tools that span the domestic Arrow-Debreu space more fully than ever. As the investment universe has extended globally, the Arrow-Debreu space has also expanded to include unfamiliar states of the world. The global state space presents new economic opportunities and risks; for example, in the fledgling markets in Southeast Asia and Eastern Europe, in the countries that ran centrally planned economies as recently as the end of the last century.

As the investment universe expands, the insights of Ross (1976a) remain as useful as ever. First, as we have discussed, Ross (1976a) proves that the (one-period) market for every primary security can be completed by calls and puts written on this security. Hence, there is no need for more complicated options than plain vanilla calls and puts. This is not, however, a cost-efficient way to hedge diversified investment portfolios because of high transaction costs and because an option on a portfolio is cheaper than a set of options on individual securities. It is more cost-efficient to use portfolio options than single-security options.

Second, Ross (1976a) proves a surprising result: given a universe of primary securities, there is a special portfolio, a *Ross portfolio*, such that calls and puts on this portfolio span the same space as all other options on all other portfolios. Hence, plain vanilla calls and puts on a Ross portfolio are at least as powerful as any set of simple or complex options. The linchpin is a Ross portfolio. Despite its appeal, it is unknown how to design this portfolio in practice and whether or not it is unique. As the investment universe changes, so does the Ross portfolio (or portfolios). In reality, therefore, market completeness is always a moving and elusive target.

When Ross (1976a) was published, no portfolio-based products existed; only single-stock options traded in the market. Since then, structured portfolio investments, for example, index funds and exchange-traded funds (ETFs), have gained broad acceptance, while stock index options have grown to dominate single-stock options. Moreover, certain portfolios have become busy hubs of trading activity and liquidity. For example, the S&P 500 stock index is a good proxy for a Ross portfolio. Index derivatives are used extensively even when the equity investments are significantly different from the S&P 500 portfolio.

Similarly, mortgage investors favor the liquidity of Treasury futures to hedge mortgage prepayments despite substantial basis risk (a mismatch between a hedge and the hedged), whereas several mortgage-specific contracts have failed. When investors face a choice between basis risk and liquidity risk, they usually take basis risk in order to avoid unpredictable costs of illiquidity. It is possible to create precise hedges but they will be illiquid and costly. The basis risk is clear evidence of incompleteness. Since poor liquidity causes basis risk, it also inhibits completeness. Conversely, good liquidity helps to reduce basis risk and to complete the market.

In practice, there are popular market benchmarks, for example, the S&P 500 stock index and the 10-year Treasury note, surrounded by large liquidity pools. These benchmarks occupy strategic locations in the market and in the Arrow-Debreu space—they are low-correlated (or near orthogonal) among themselves and high-correlated (or near parallel) to many other investments. As such, they deliver effective spanning and make basis risk manageable in addition to generating pools of liquidity.

THE MEAN-VARIANCE THEORY IN COMPLETE MARKETS

The complete market theory and *mean-variance theory* are two main models of financial markets. In the mean-variance theory, securities are described by the means, variances, and correlations of their returns. This description is so compact that the entire securities market fits in the risk-return plane. The complete market model, by contrast, occupies a multidimensional vector space because of its detailed security format. Every mean-variance security can be reformatted and the risk-return plane can be in embedded in the Arrow-Debreu space, preserving the mathematical mean-variance properties. An economic integration of the two models is more problematic than mathematical integration.

The mean-variance theory has three parts. The first is the Markowitz (1952) portfolio model that selects investments by maximizing the expected portfolio return for a given level of volatility. The solutions, called the

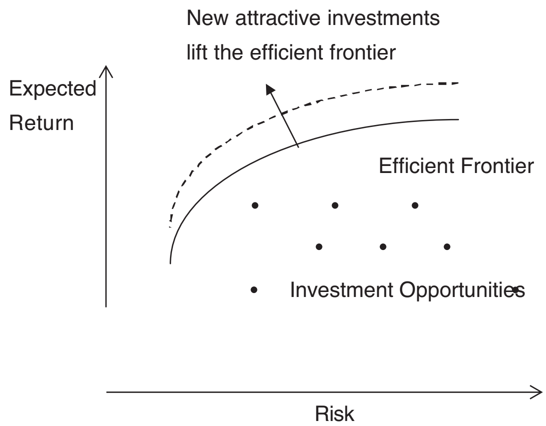


Figure 9.6 Investment Opportunities and the Mean-Variance Efficient Frontier in the Risk-Return Plane; New Attractive Opportunities Lift the Efficient Frontier

mean-variance efficient portfolios, form the efficient frontier depicted in Figure 9.6. The Markowitz portfolio model works in a complete market without restrictions. Moreover, mean-variance efficient portfolios correspond to Pareto optimality—new attractive securities improve allocational efficiency and lift the efficient frontier, as shown by the dotted line in Figure 9.6.

The second part is the two-fund separation: All endowments should be invested in the risk-free asset and market portfolio. The relative allocation between the two may depend on individual risk preferences but there is no need to consider any other investments. Dibvig and Ingersoll (1982) prove that in a mean-variance complete market, the two-fund separation holds only if all investors have quadratic utilities. Quadratic-utility investors can prefer less to more and can create arbitrage opportunities as a result. But arbitrage opportunities preclude market equilibrium that is germane to the mean-variance analysis. Therefore, complete markets generally do not support the two-fund separation.

The third part is the mean-variance pricing based on the capital asset pricing model. Dibvig and Ingersoll (1982) prove that, in complete markets, the mean-variance pricing is valid for primary securities but not for derivative securities. Therefore, mean-variance prices create arbitrage opportunities in a complete market that includes both primary and derivative securities. To sum up, except for the Markowitz portfolio model, the mean-variance theory is not consistent with complete markets. This conclusion may be unsettling to academics but not to professional investors, because, out of the entire mean-variance arsenal, only the Markowitz portfolio model is used extensively in practice.

SUMMARY

Market completeness is a theoretical concept of practical importance. First, complete markets help consumers, producers and investors make efficient allocations of capital and other scarce resources; efficiently manage financial risks; and extract useful information from market prices. Second, complete markets in the Arrow-Debreu space provide the modern framework for analysis of securities markets and corporate finance. For example, arbitrage-free preference-free pricing is valid only in complete markets. Third, the complete market theory lays a road map to financial innovation, including new securities, investment strategies, and capital market architecture. In sum, few economic theories combine abstract models and practical applications as well as the complete market theory.

REFERENCES

Ait-Sahalia, Y., and Lo, A. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance* 53, 2: 499–547.

Arrow, K. J. (1951). An extension of the basic theorems in classical welfare economics. In J. Neyman (ed.), *Proceedings of the Second Berkeley Symposium on Mathematival Statistics and Probability* (pp. 507–532), California: University of California Press.

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk-bearing. *Review of Economic Studies* 31, 2: 91–96. (First appeared in French in 1953.)

Arrow, K. J., and Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica* 22: 265–290.

Black, F., and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 3: 637–654.

Black, F., Derman, E., and Toy, W. (1990). A one factor model of interest rates and its application to Treasury bond options. *Financial Analysts Journal* 46, 1: 33–39.

Breeden, D., and Litzenberger, R. (1978). Prices of state contingent claims implicit in options prices. *Journal of Business* 51, 4: 621–651.

Brennan, M. J., and Schwartz, E. S. (1979). A continuous time approach to the pricing of bonds. *Journal of Banking and Finance* 3: 133–155.

Buchen, P. W., and Kelley, M. (1996). The maximum entropy distribution of an asset inferred from option prices. *Journal of Financial and Quantitative Analysis* 31: 143–159.

Cox, J. C., Ingersoll, J. E. Jr., and Ross. S. A. (1985). A theory of the term structure of interest rates. *Econometrica* 53: 385–407.

Cox, J. C., and Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3: 145–166.

Cox, J. C., Ross, S. A., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics* 7: 145–166.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica* 19: 273–292.

Debreu, G. (1959). *Theory of Value*. New York: John Wiley & Sons.

Dibvig, P. H., and Ingersoll, J. E. Jr. (1982). Mean-variance theory in complete markets. *Journal of Business* 55, 2: 233–251.

Harrison, J. M., and Kreps, D. (1979). Martingales and arbitrage in multiperiod capital markets. *Journal of Economic Theory* 20: 215–260.

Hayek, F. A. (1937). Economics and knowledge. *Economica*, New Series, 4, 3: 33–54.

Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review* 35, 4: 519–530.

Heath, D., Jarrow, R., and Morton, A. (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica* 60, 1: 77–105.

Ho, T. S., and Lee, S. (1985). Interest movements and pricing interest rate contingent claims. *Journal of Finance* 41, 5: 1011–1028.

Hull, J., and White, A. (1990). Pricing interest rate derivative securities. *Review of Financial Studies* 3, 4: 573–592.

Ingersoll, J. E. (1987). *Theory of Financial Decision Making*. Savage, MD: Rowman & Littlewood Publishers.

Jackwerth, J. C., and Rubinstein, M. (1996). Recovering probability distributions from options prices. *Journal of Finance* 51, 5: 1611–1631.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance* 7, 1: 77–91.

Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science* 4: 141–183.

Miller, M., and Modigliani, F. (1961). Dividend policy, growth and the valuation of shares. *Journal of Business* 34: 411–433.

Modigliani, F., and Miller, M. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review* 48, 3: 261–297.

Pliska, S. R. (1997). *Introduction to Mathematical Finance*. Malden, MA: Blackwell Publishers.

Ross, S. A. (1976a). Options and efficiency. *Quarterly Journal of Economics* 90, 1: 75–89.

Ross, S. A. (1976b). Return, risk and arbitrage. In I. Friend and J. Bicksler (eds.), *Risk and Return in Finance* (pp. 189–218), Cambridge, MA: Ballinger.

Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance* 49, 3: 771–818.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics* 5: 177–188.

CHAPTER 10

Introduction to Islamic Finance

MAHMOUD A. EL-GAMAL, PhD
Professor of Economics and Statistics, and Chair of Islamic Economics, Finance, and Management at Rice University

Abstract: Islamic finance began to take shape in the 1970s. It was fueled financially by the flow of petrodollars to Islamic countries in the oil-rich Gulf Cooperation Council (GCC) countries, and fueled ideologically by the nationalist and Islamist movements that took shape during the first half of the twentieth century. This industry has witnessed dramatic growth over the past decade, fueled again by petrodollar flows and resurgence of nationalist and Islamist tendencies. As will become apparent shortly, the demarcation between Islamic and conventional financial practices is almost exclusively a matter of contract form. This makes Islamic finance a branch of structured finance more generally. A third reason for growth in Islamic finance must thus be added to excess liquidity in the GCC and the rise in Islamist and nationalist sentiments, and that is the ready availability of structured-finance methods that were developed during the 1980s and 1990s.

Keywords: Islamic finance, structured finance, riba, gharar, bay` al-`ina, murabaha, salam, takaful, Fiqh, fatawa, tawarruq, securitization, da` wa ta`ajjal, ijara, sukuk al-ijara, sukuk, musharaka, mudaraba, salam, istisna`, `urbun

The purpose of this chapter is to describe Islamic finance briefly. This chapter is not written from the point of view of pious Muslims, or from a rigorous academic perspective. (For discussions of Islamic finance from the former perspective, see El-Gamal [2000]; from the latter perspective, see El-Gamal [2006].) Rather, this chapter is intended for financial practitioners seeking a basic understanding and critical evaluation of the modes of operation in Islamic finance. In my criticism of the industry's modes of operation, I argue that they are costly and unnecessary (that is, inefficient) forms of legal arbitrage.

WHAT IS ISLAMIC FINANCE?

While *Islamic finance* is a form of *structured finance,* there are two distinctive features that distinguish it from other forms of structured finance. The first difference pertains to regulatory constraints. Conventional structured finance aims to improve marketability and reduce costs and tax burdens by adhering to well-defined sets of regulations in various geographical regions. There has been no shortage in the 1980s and 1990s of attempts to standardize the set of regulations that determine whether or not a financial product or service may be marketed as Islamic. However, despite efforts by the Accounting and Auditing Organization for Islamic Financial Institutions (AAOIFI) and the Islamic Financial Services Board (IFSB), housed in Manama, Bahrain, and Kuala Lumpur, Malaysia, respectively, there remains a great deal of heterogeneity in what is deemed Islamic. Providers of Islamic financial products and services and their advisers, therefore, rely on the direct advice of religious scholars who have attained a reputation in Islamic finance.

This is particularly important for new and innovative products, the structure of which is often kept secret to prevent imitation by competitors. One of the major problems in this regard is that different religious scholars may be brought to court to testify against the Islamic nature of products and services approved by other scholars. This has happened, in fact, twice, before English courts. Fortunately for the Islamic financial providers in both cases, English courts decided to disregard Islamic law provisions and to award them the accrued interest, following English law. However, the issue of adherence to multiple legal systems, one of which is not clearly defined, remains a problem for this industry.

The second distinctive feature is that Islamic financial innovation generally tracks innovation in the conventional sector, with potentially substantial lags. On the supply side, providers of Islamic finance have traditionally approached the Islamic market with products that they had already been offering to conventional customers. Conversely, if demand for a new product is presented to conventional financial providers, the cheapest-to-deliver products can normally be found by restructuring conventional products that are already available. The legal and consulting fees required to restructure a conventional product or practice Islamically (e.g., commodity funds, mortgage-backed securities, leveraged buyouts, etc.) can be substantial. However, the lag in bringing those products and services to market are reduced significantly by focusing the discussions of bankers, lawyers, and religious scholars on providing the closest analog to well-defined products and services at minimal added transaction costs.

One of the most intimidating factors for newcomers to Islamic finance is the general use of religious rhetoric and excessive usage of Arabic words. Due to those tendencies, senior administrators at a major University in the United States that had established a program on Islamic finance in the 1990s were disturbed by what they characterized as ecumenical discourse. Many non-Muslim financial professionals also felt apprehensive at first, but then quickly learned to see through the seemingly ecumenical rhetoric. Part of the apprehension, no doubt, was driven by fear of offending the religious-scholar consultants who are needed to certify products and services as Islamic. However much religious rhetoric may tilt the initial power structure in favor of those scholars, the bankers and lawyers ultimately recognize that they control the process.

Three forces give conventional financial and legal professionals this power over the religious scholars: They get to choose whom to hire and publicize as an expert, they get to choose what questions should be asked of the hired scholars, and they get to choose which of the scholars' answers to disseminate. Once this simple first-mover advantage is understood, the next step for a financial professional is to learn how to cut through the religious rhetoric. This barrier to entry is not nearly as difficult as it may appear at first. The number of Arabic words to learn is quite small, and the concepts themselves constitute a rather primitive subset of conventional financial practices. This primer is intended for a financial professional, on the financial or legal side, and it should give them a good first understanding of the bulk of Islamic finance in less than half an hour.

PROHIBITIONS AND THEIR IMPLICATIONS

It is generally accepted in Islamic jurisprudence that all contracts conducted by mutual consent are permissible, unless they contain one of two major prohibitions, known in Arabic as *riba* and *gharar.* There are other prohibitions, of course, that pertain to asymmetry of bargaining power, such as would be the case in monopolistic markets. However, the focus in Islamic finance has been fundamentally focused on contract forms, rather than economic substance or market structure. Therefore, a newcomer to Islamic finance needs only to learn about those two prohibitions. Even then, only the most basic understanding of the general concepts is required, since practitioners often aim to avoid the prohibitions simply by using building-block contracts that have been previously approved, as we shall discuss in the following section.

Riba

The prohibition of *riba* predates Islam, as Islamic scriptures themselves report in discussion of Judeo-Christian prohibitions of *ribit*, a Hebrew word that obviously shares the same etymological root as its cousin Semitic Arabic word. Like the prohibition of usury in Judeo-Christian history, the prohibition of *riba* has been the subject of numerous scholastic disputes over the centuries. Since this primer is for financial professionals, those scholastic debates are largely irrelevant. The rhetoric used by industry practitioners and their pietist customers is quite simple. The forbidden *riba*, according to this rhetoric, is interest. To sell a product in Islamic finance, one must proclaim it to be interest free.

One should not be fooled by this rhetoric, however, even when advocates of Islamic finance announce that Islam does not recognize the time value of money. One needs only to recognize that scholars also accept that the credit price of an asset or commodity may be higher than its spot price. Moreover, Islamic scholars do not place any restrictions on the credit-price markup over the cash price, allowing one to incorporate time value, credit risk, interest rate risk, and other conventional components of financing charges. One must be careful, however, not to jump to hasty conclusions. For instance, a simple two-party sale buy-back would technically satisfy the provisions on using spot and credit sales of a nonmonetary commodity, but it would not be universally accepted as Islamic.

`Ina versus Murabaha

This practice is quite simple: Instead of lending you $1 million at 5% interest, I buy from you a property for $1 million and then sell it back to you at $1,050,000. In principle, the property need not even be worth $1 million on the spot market. Unfortunately, this practice is named in the Islamic Canon as a forbidden form of (or legal ruse for) *riba.* It is called *bay` al-`ina,* or *`ina* for short, meaning multiple trades of the very same property. This contract is in fact used extensively in Malaysia, where the practice is deemed permissible as long as the two sales are not stipulated in the same contract. However, it is not allowed in the vast majority of Islamic countries, including most notably the countries of the oil and cash-rich GCC.

One does not need to add much complexity to the structure to make it permissible, however. For instance, one may simply sell assets or commodities worth $1 million on credit for a deferred price of $1,050,000, even if one knows that the buyer will turn around and sell the assets or commodities for their spot price, thus effectively receiving the loan at 5% interest. Yes, surprising as this may be, and as many hours of consultants' time and religious rhetoric as one may have to endure, Islamic finance is—in the end—just that simple. It helps to know the names of the first contract (*`ina*) and the second (*murabaha*), which we shall discuss in the next section.

As one begins to consider progressively more complex structures, one invariably faces multiple choices of how to characterize a transaction with multiple components. Practitioners in Islamic finance learn to use the correct Arabic names to characterize those components for the purpose of Islamic certification, and to characterize them differently—if necessary—to adhere to regulatory and legal provisions in the relevant jurisdictions. This requires advanced legal-arbitrage skills, which are beyond the scope of this primer, and which the target reader may either possess or have the resources to acquire.

Gharar

The second major prohibition in Islamic financial jurisprudence is even easier to finesse. Like *riba,* the forbidden *gharar* is not definitively defined in the Islamic canon or legal literature. Unlike *riba,* which is definitively forbidden, even if we are not entirely sure what would fall into that category, the forbidden *gharar* is left to the scholars' discretion. Indeed, classical and contemporary Islamic legal scholars have ruled that *gharar,* which is translated variously as uncertainty or risk, cannot be eliminated entirely. What is forbidden, therefore, is excessive and unnecessary *gharar.* In other words, the religious scholar must perform a cost-benefit analysis to determine if the economic benefit from allowing a transaction is sufficient to overcome the potential harm due to uncertainty that is present in the contract language or provisions.

This degree of Islamic legal discretion can eliminate the need for Islamic finance altogether. For instance, while the vast majority of Islamic jurists continue to disallow futures trading based on the prohibition of *gharar,* Malaysian jurists have permitted futures trading based on the argument that modern futures market and clearinghouse structures have eliminated excessive uncertainty. Retaining the prohibition of futures trading, however, may create a legal arbitrage opportunity. Classical Islamic jurisprudence permits credit sales, as we have already seen, as well as a prepaid forward contract known as *salam.* Conservative jurists thus like to say that Islamic law permits trading goods now for money later (credit sale), or money now for goods later (*salam*), but not goods later for money later (forward or future sale).

Forwards and Insurance

It does not take much skill to use spot sales, credit sales, and the *salam* sale to synthesize a forward contract. We have already seen how spot and credit sales can be used to synthesize a loan. All one needs is to use that structure to lend the *salam* buyer the present value of the forward price, and we have a synthetic forward.

Another interesting legal arbitrage opportunity applies to both *gharar* and *riba* prohibitions. Both prohibitions are only observed for commutative (that is, *quid-pro-quo*) financial contracts. Therefore, fixed income securities have been offered for decades in Egypt and Malaysia, under the name "investment certificates," based on guaranteeing only the principal. Interest is paid as a "gift," ostensibly unanticipated, which is based on economic conditions. This structure has not been adopted in other countries, but other noncommutative provisions, for instance, a unilateral promise to sell a property at lease end, have been used to circumvent other prohibitions. Moreover, the notion of noncommutativity has been used extensively to adopt the rhetoric of mutuality in Islamic insurance alternatives—known as *takaful*—even though to the best of my knowledge, there has not been any *takaful* provider that was in fact owned by its policyholders.

NOMINATE CONTRACTS: FROM BASIC BUILDING BLOCKS TO COMPLEX STRUCTURES

We review some of the main nominate contracts that are used as building blocks in contemporary Islamic finance. There are two reasons for the popularity of those nominate contracts. The first is an issue of authority. Contemporary Islamic scholars who are retained as consultants by Islamic financial providers generally lack the authority unilaterally to declare that a transaction is free of forbidden *riba* or *gharar.* This is partially avoided by referring to collective opinions issued by international Islamic jurisprudence (*Fiqh*) academies. Still, the most common method to ensure legitimacy of a juristic pronouncement—in this Islamic common-law tradition—is appeal to precedent, which naturally lends itself to framing the question within the context of nominate contracts that were certified by earlier generations of jurists.

The second reason that nominate contracts have played such a prominent role in Islamic financial practice is that financial professionals often propose the initial structures for new services or products, which are then refined in collaboration with the retained religious-scholar consultants. It is much easier for financial professionals to design the new product or service using the building blocks of other products and services that were approved earlier, often by the same consultants. We will review some of the most commonly used building-block contracts of Islamic financial practice in this section, and then review some of their uses in contemporary Islamic finance in the following section.

Murabaha

The most basic contract, of course, is the sales contract. Classical Islamic jurisprudence distinguished between standard sales and what were called "trust sales." In the latter category, instead of negotiating a final price, the buyers rely on the sellers' truthful revelation of their cost of acquisition of the object of sale. The buyer and seller then negotiate the markup or markdown relative to the seller's invoice. The most common trust sale used in Islamic finance is the trust sale with a markup over invoice, known by the Arabic name *muarabaha.*

Murabaha was transformed from a simple markup-over-invoice sale into a mode of financing in the late 1970s, when it was combined with the permissibility of credit sales—known in Arabic by the name *bay` bi-thaman 'ajil,* although this terminology is generally used only in Malaysia. The first successful Islamic banking product introduced in the late 1970s was based on combining *Murabaha,* credit sale, and a third component: a binding promise by the customer to buy the property on credit from the bank.

Through a series of *fatawa* (religious opinions) by prominent Islamic jurists, the practice of *murabaha* financing was approved as follows. A customer wishes to finance the purchase of some asset or commodity. The Islamic bank may obtain a binding promise that the customer will buy the property on credit, at an agreed-upon markup above the spot price, which markup is characterized as the *murabaha* profit, rather than interest. With the promise in hand, the bank may then buy the property at its spot price, and then promptly sell it to the customer at the agreed-upon credit price. In some applications, the customer may serve as the bank's agent to purchase the property on its behalf and then to sell it to himself, thus reducing transaction costs and time delays.

Murabaha financing was also used by major Islamic banks starting in the late 1970s as a means of extending credit facilities to large corporate customers. It was characterized in this framework as a form of trade financing. The Islamic bank or financial institution would have a standing agreement with the corporate customer to finance the purchase of a certain amount of metal or other commodity with liquid spot markets. Whenever the customer needed credit, they would use the agreement to buy metal or commodities of a specific spot value at a credit price that is equal to the spot price plus a mutually agreed upon implicit interest charge. The customer may indeed need the purchased metal, in which case the transaction ends with this purchase. In other cases, the customer may actually be in need of cash, in which case it can turn around and sell the metal or other commodity at its spot price.

`Ina and Tawarruq

In some cases, especially in retail finance, the transaction costs of ownership transfer and multiple sales may be too large. There may also be basis risk due to movements in the commodity's spot price between the initial credit sale and the ensuing spot sale. In Malaysia, the buyer and seller would simply use *`ina,* described in the previous section, to sell a property on credit and then buy it back at the spot price. In the GCC, a more elaborate three-party alternative called *tawarruq* (literally, turning a commodity into silver, or monetizing it) has become a popular vehicle to extend financial credit.

Tawarruq financing often involves a standing agreement with a metals or other commodity trader. The financier buys the commodity on the spot, sells it to the customer on credit, and then sells it back to the dealer on the customer's behalf at the spot price (less any agreed-upon fees). The standing agreement and speed of transactions—three faxes or other communications sent in quick succession—can eliminate most of the transaction costs and basis risks. Depending on jurisdiction, the jurists may allow only dealing with domestic commodities traders, or add other warehousing restrictions to ensure that the traded commodity actually exists. However, actual physical receipt of the commodities was never made a requirement, allowing this transaction to emulate a credit facility by introducing trades of a real asset or commodity with minimal added cost.

Securitization and Ijara

Simple spot and credit sales can therefore be seen easily to emulate interest-based loans. The result of those transactions, however, is a pure financial debt, which the majority of Islamic jurists deem generally nontradeable. The majority of Islamic legal scholars allow selling the debt back to the debtor at a price below its face value—a practice known in Arabic as *da` wa ta`ajjal,* or discount for prepayment. However, if the debt is sold to a third party, it can be sold (transferred) only at face value, and only with the debtor's consent. The notable exception, again, is Malaysia, where trading debts at market value, known in Arabic as *bay` al-dayn,* is permitted. This allows Malaysian bankers to securitize debts and create secondary markets. However, those securities would not be acceptable to the religious scholars from other parts of the world, most notably from the cash-rich GCC. Consequently, other methods of securitization were required. One of the most popular securitization structures uses leasing, under the Arabic name *ijara.*

Ijara financing, like its *murabaha* sibling, was first approved by Islamic-banking scholars as an implicit mode of secured lending. If a customer wished to finance the purchase of a nonperishable asset (e.g., real estate, automobiles, or equipment), then the bank could buy the property and lease it to the customer. The bank may then give the customer an option to buy the property at an agreed-upon price at the end of the leasing period, thus converting the simple leasing arrangement into a versatile financial tool. The majority of jurists insist that the lessor in this arrangement must retain substantial ownership of the leased property, rendering the transaction an operating rather than a financial lease. This, of course, can have varying tax implications for the implicit interest (financing) charge. From a logistical point of view, lessor obligations for insurance, maintenance, and the like are handled most often through side agreements with special purpose vehicles.

Sukuk

An interesting by-product of the insistence on lessor-ownership of the underlying property has been the ability to securitize lease-generated receivables. The lease is often conducted through a special purpose vehicle (SPV) at any rate—for a variety of tax, regulatory, and bankruptcy-remoteness purposes. Shares in that SPV would entitle the shareholders to the stream of rental payments. More importantly, lessor SPV shares are deemed by the religious scholars to represent ownership of the underlying asset itself, not merely the rent receivables. Thus, the scholars allow trading those shares, often marketed over the past few years under the name *sukuk al-ijara,* or rent certificates. This structure has been the workhorse of the fastest-growing segment of Islamic finance: the issuance of bond alternatives known collectively as *sukuk.*

Partnerships: Musharaka and Mudaraba

Interestingly, since *sukuk* are generally issued as common shares in a special purpose vehicle, they can be advertised as containing an element of partnership, which resonates well with Islamist rhetoric on ideal Islamic finance being based on partnership and profit-and-loss sharing. There are two main partnership models that are discussed at length in Islamic economics and finance. Simple partnership or *musharaka* requires that losses are shared in proportion to capital contributions, but profits can be shared according to any agreed-upon formula. The other form of partnership discussed at length in Islamic economics and finance is silent partnership, known by the Arabic name *mudaraba.* In this structure, the investor or principal provides all the capital and bears all financial losses, but financial gains are shared with the entrepreneur or agent.

Benchmarking to LIBOR

For retail as well as investment banking products, Islamic scholars have allowed the financing rate (characterized as profit in credit sales and rent in leases) to be benchmarked to conventional interest rates. The industry is dominated by English bankers, and therefore the benchmark of choice has been the London Interbank Offered Rate (LIBOR). In sale-based structures, the financing rate is fixed for the duration of the financing facility. In lease-based structures, an added degree of flexibility allows the lessor and lessee to adjust the rent, usually tracking LIBOR by adding a simple spread thereupon.

Short-Term Sukuk and Salam

Sukuk structures that rely on long-term leases are not particularly efficient for short-term debt instruments, since the transaction costs and tax burdens may be substantial. Therefore, commodity-sale based structures were introduced for short-term debt instruments. The contract of choice in this case, especially as used by the Bahrain Monetary Agency to issue Treasury bill-like instruments, has been the prepaid forward contract *salam.* The government would collect the Treasury bill prices by ostensibly selling a commodity, usually aluminum or other metal, forward. Instead of delivering aluminum at the Treasury bill maturity, the government guarantees that it will sell the aluminum on behalf of the bill-holders at a price equal to the initial price that they paid plus the appropriate interest rate on its short-term debt. This structure is shown in Figure 10.1.

Parallel Salam

Other vehicles were also devised, using what is known as parallel-*salam.* Under that structure, a three-month bill can be structured by selling aluminum forward, say, six months, with the price payable now. In three months, the parties may engage in a reverse three-month prepaid forward contract, whereby the obligations to deliver the commodity at the coinciding maturity dates of the two *salam*

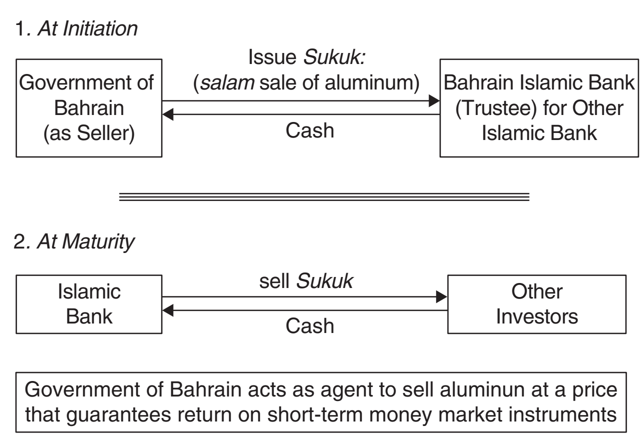


Figure 10.1 *Salam* Certificate Structure for Bahrain Monetary Agency Treasury Bills
Source: Presentation by Sheikh Salman bin Ahmed Al Khalifa, director of Banking Services at the BMA, at the International Islamic Finance Forum in Istanbul, Turkey, September 27–29, 2004.

sales cancel each other out. The price of the second reverse *salam,* of course, would be the price of the first salam plus the appropriate quarterly interest on short-term debt.

Miscellaneous Structures

Other classical nominate contracts are sometimes used, e.g., *istisna`*, or commission to manufacture, is a popular analog to *salam,* where the price is prepaid, possibly in installments, as the object of sale is actually manufactured or constructed. This structure is popular for infrastructure and other construction projects, and often combined with leasing to create a build-operate-transfer (BOT) structure that mimics the financial structures of conventional practice.

Classical conditions of leasing, which allow some types of subleasing, have allowed timeshare agreements to be used under the name *sukuk al-intifa`*, or usufruct securities. Juristic bodies, for example of AAOIFI, have produced lists of permissible contracts and basic structures, both classical and modern (e.g., *salam* itself is a classical contract, *sukuk al-salam* and *parallel salam* as financial tools are modern inventions).

Financial professional who are newcomers to Islamic finance can accomplish much using only simple sales, credit sales, and leases. As they contemplate more complex structures (e.g., shorting assets) or as they aim to reduce transaction costs to gain a competitive advantage in this market, they may consult the lists produced periodically in AAOIFI's *Shari`a Standards* publications to find appropriate building blocks for their structures. As they become more advanced still and need to devise new building blocks, they can get their religious-scholar consultants to approve those structures, and then possibly to add them to the list of AAOIFI standards, as the same scholars who serve as consultants for the banks are the ones who serve on AAOIFI's Shari`ah advisory board.

Derivatives

Derivative securities such as swaps, options, and futures have become essential parts of today's financial world. They are used to hedge various risks, as well as to leverage financial exposure to make it sufficiently attractive to investors. In Islamic finance, call options have been synthesized from a classical contract known as *`urbun*—literally, down payment on a purchase. The down payment is treated as the call premium, and the exercise price is the difference between the original price and that down payment, which the buyer pays if he chooses to exercise the option. We have already seen how a forward contract can be synthesized from the classical prepaid forward contract, *salam,* and a credit facility generated by spot and credit sales. It then follows by put-call parity that we can synthesize a forward contract from the synthetic call and the synthetic forward (long put = short forward + long call). In addition, while the majority of Islamic jurists continue to forbid trading in options, they allow offering the options as unilaterally binding promises that are ostensibly noncommutative. Practitioners in Islamic finance use a combination of synthetic and noncommutative characterizations to include options in hedging mechanisms, as was the case, for instance, in the recent issuance of U.S. $166 million "East Cameron Gas *Sukuk*" in July 2006.

The First U.S. Sukuk

The East Cameron Gas *Sukuk* was rated CCC+ by Standard & Poor's, and it pays 11.25% for the duration of its 13-year tenure. The originator of the *sukuk* is the Houston-based East Cameron Partners, an independent oil and gas exploration and production company. The issuer is an SPV called East Cameron Gas Company, incorporated in the Cayman Islands. The deal was primarily arranged by Bemo Securitization (BSEC), of Beirut, Lebanon, in cooperation with Merrill Lynch. Bemo had earlier arranged for a well-publicized issuance of Islamic lease-based *sukuk* for the rental car fleet of the Saudi company Hanco. The bulk of the East Cameron *sukuk* were initially envisioned for sale in the GCC countries. However, it was reported that a large portion of those *sukuk* were, in fact, bought by conventional hedge funds.

This *sukuk* issuance received significant media coverage because it was the first originated by a U.S. company. However, it must be noted that similar financial structures had been used by Islamic investment banks and private equity firms for leveraged buyouts in the United States. The standard sale-leaseback *sukuk* structure was not viable for the East Cameron Gas issuance because the fixed assets eligible for leasing (mainly rigs) were not of sufficient value to generate the desired Dollar amount. The SPV was thus characterized as a co-owner of the originator's assets (in a *musharaka*). A reserve account was used to cover potential shortage in collected revenues from the originator, thus nearly guaranteeing a fixed rate of return to the issuer SPV, and hence to the *sukuk* holders. Put instruments were also used to protect the *sukuk*-holders from the bulk of ownership risks that would entail sharing in profits and losses. In the meantime, co-ownership of the company's assets allows secondary-market trading of those *sukuk.*

A simplified schematic structure of this issuance is shown in Figure 10.2.

BSEC's director said that they could have achieved an investment-grade rating for the issuance, but it would have taken longer to develop the structure. This is ultimately the trade-off that participants in Islamic finance have to examine: quick but inefficient structures, or more efficient ones that require more time and legal fees. It is no wonder that hedge fund managers were happy to buy the East Cameron Gas *Sukuk* at a yield commensurate with CCC+ rating, when the credit rating of a conventional bond would have been higher. Then, again, there are a host of legal risks associated with those new structures, and the higher yield may—in part—compensate for the resulting uncertainty.

SUMMARY

This short introduction did not provide a comprehensive survey of all products and services in the Islamic

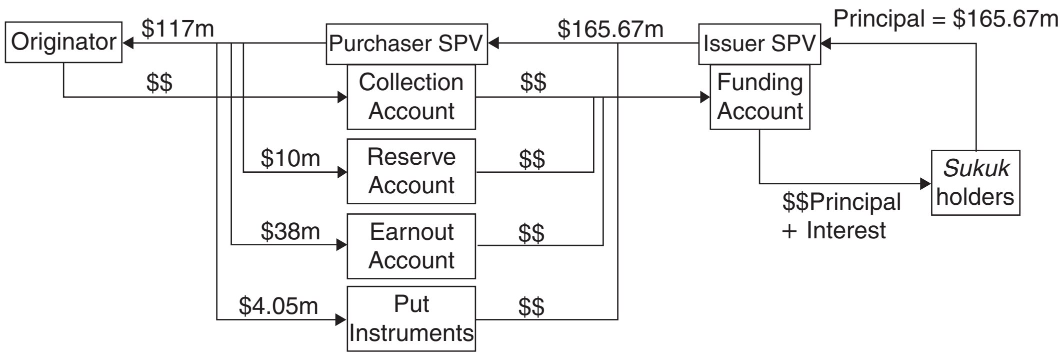


Figure 10.2 Simplified Structure of East Cameron Gas *Sukuk*
Source: Based on *Sukuk Insider,* Issue 01.

finance space. For example, there are a number of screening methodologies, pertaining to lines of business and debt structures, that are used in selecting stocks and companies in which Islamic mutual funds and private equity firms can invest. The tricky financial aspects of setting up investment vehicles, however, pertain to structuring derivatives, credit facilities, and other contract-based structures that are introduced in this chapter. I hope to have shown that the logistical and informational barriers to entry into Islamic finance are actually quite low. It is a theorem that every financial practice can be adapted for the Islamic market as it exists today. Financial providers' considerations in this adaptation are usually restricted to transaction costs of structuring the deal, and the marketability of the product. However, it must be noted that the relative ease of entry into this market segment is a mixed blessing. It means that one can easily bring structured products to market with combinations of simple and well-understood contracts. However, this very ease with which new products can be brought to market masks considerable inefficiencies and poorly understood legal risks, which must be a concern for careful financial providers.

REFERENCES

Accounting and Auditing Organization for Islamic Financial Institutions (2004). Shari`a Standards, 1424–5H/ 2003–4, Manama, Bahrain: AAOIFI.

Al-Zuhayli, W. (2003). *Financial Transactions in Islamic Jurisprudence* (Volumes I and II), Damascus: Dar Al-Fikr (M. El-Gamal, tr.).

El-Gamal, M. (2000). *A Basic Guide to Contemporary Islamic Banking and Finance*. Plainfield, IN: Islamic Society of North America, 2000.

El-Gamal, M. (2006). *Islamic Finance: Law: Economics, and Practice*. New York: Cambridge University Press.

Henry, C., and Wilson, R. (eds.) (2004). *The Politics of Islamic Finance*. Edinburgh: Edinburgh University Press.

Kamali, M. H. (2000). *Islamic Commercial Law: An Analysis of Futures and Options*. Cambridge: Islamic Texts Society.

Kuran, T. (2005). *Islam and Mammon: The Economic Predicaments of Islamism*. Princeton, NJ: Princeton University Press.

Lewis, M., and Algaoud, L. (2001). *Islamic Banking*. Cheltenham, UK: Edward Elgar Publishing.

Maurer, B. (2005). *Mutual Life, Limited: Islamic Banking, Alternative Currencies, Lateral Reason*. Princeton, NJ: Princeton University Press.

Saeed, A. (1996). *Islamic Banking and Interest: A Study of the Prohibition of Riba and Its Contemporary Interpretation*. Leiden: Brill Academic Publishers.

Usmani, M. T. (2002). *An Introduction to Islamic Finance*. Berlin: Springer.

Vogel, F., and Hayes, S. (1998). *Islamic Law and Finance: Religion, Risk, and Return*. Berlin: Springer.

Warde, I. (2000). *Islamic Finance in the Global Economy*. Edinburgh: Edinburgh University Press.

PART 2

Common Stock

Cash Instruments

Equity Derivatives

CHAPTER 11

The U.S. Equity Markets

FRANK J. JONES, PhD
Professor, Accounting and Finance Department, San Jose State University and Chairman of the Investment Committee, Salient Wealth Management, LLC

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: At the beginning of the twenty-first century, the world's stock exchanges were a complex of separate, independent, single-product exchanges. During the first part of this century, however, these exchanges have become interconnected within and across country lines and have become multiproduct exchanges. That is, some of the world's stock exchanges have become international, multiproduct exchanges. The exchanges have and will continue to change rapidly, both diversifying and integrating. In addition to the sanctioned stock exchanges, the off-exchange markets have become much more important in their size and diversity. The U.S. stock market has become a complex of interconnected exchanges and off-exchange markets.

Keywords: market structure, order driven, quote-driven, market orders, limit orders, bid quote, offer quote, pure order-driven market, natural buyers, natural sellers, continuous market, call auction, brokers, agents, dealers, market makers, principals, specialist, national best bid and offer (NBBO), hybrid markets, system orders, floor brokers, commission broker, limit order book, competitive-dealer quote-based system, off-exchange markets, alternative electronic markets, electronic communications networks (ECNs), alternative trading systems, cross networks, dark pools, fragmentation, alternative display facility (ADF), tick size, block, block trade

This chapter offers a snapshot of the current but evolving U.S. stock markets. International exchanges are considered herein only in their relationship to the U.S. stock markets. There are two fundamental differences among U.S. and international exchanges. The first is their method of trading, that is, their *market structure.* The market structures of U.S. and international exchanges have evolved and even changed radically. One cannot appreciate current exchanges without understanding their market structure. In addition, the nature of the exchanges' business organizations have changed considerably, from membership floor–traded organizations to publicly owned electronic trading organizations. This chapter begins with discussions of the market structures and business organizations of the U.S. exchanges. (For a more detailed discussion of some of the topics covered in this chapter, see Schwartz and Francioni [2004].)

EXCHANGE MARKET STRUCTURES

An *exchange* is often defined as a market where intermediaries meet to deliver and execute customer orders. This description, however, also applies to many dealer networks. In the United States, an exchange is an institution that performs this function and is registered with the Securities and Exchange Commission (SEC) as an exchange. There are also some off-exchange markets that perform this function.

There are two overall market models for trading stocks. The first model is *order driven,* in which buy and sell orders of public participants who are the holders of the securities establish the prices at which other public participants can trade. These orders can be either *market orders* or *limit orders.* The second model is *quote-driven,* in which intermediaries, that is market-makers or dealers, quote the prices at which the public participants trade. Market makers provide a *bid quote* (to buy) and an *offer quote* (to sell) and realize revenues from the spread between these two quotes. Thus, market makers derive a profit from the spread and the turnover of their stocks.

Order-Driven Markets

Participants in a *pure order-driven market* are referred to as "naturals" (the natural buyers and sellers). No intermediary participates as a trader in a pure order-driven market. Rather, the investors supply the liquidity themselves. That is, the *natural buyers* are the source of liquidity for the *natural sellers,* and vice versa. The naturals can be either buyers or sellers, each using market or limit orders.

Order-driven markets can be structured in two very different ways: a continuous market and a call auction at a specific point of time. In the *continuous market*, a trade can be made at any moment in continuous time during which a buy order and a sell order meet at a specific time. In this case, trading is a series of bilateral matches. In the *call auction,* orders are batched together for a simultaneous execution in a multilateral trade at a specific point in time. At the time of the call, a market-clearing price is determined; buy orders at this price and higher and sell orders at this price and lower are executed.

Continuous trading is better for customers who need immediacy. However, for markets with very low trading volume, an intraday call may focus liquidity at one (or a few) times of the day and permit the trades to occur. In addition, very large orders—block trades that will be described later—may be advantaged by the feasibility of continuous trading.

Nonintermediated markets involve only naturals; that is, such markets do not require a third party. A market may not, however, have sufficient liquidity to function without the participation of intermediaries, who are third parties in addition to the natural buyers and sellers. This leads to the need for intermediaries and quote-driven markets.

Quote-Driven Markets

Quote-driven markets permit intermediaries to provide liquidity. Intermediaries may be *brokers* (who are *agents* for the naturals); *dealers* or *market makers* (who are *principals* in the trade); or *specialists,* as on the New York Stock Exchange (who act as both agents and principals). Dealers are independent, profit-making participants in the process.

Dealers operate as principals, not agents. Dealers continually provide bid and offer quotes to buy for or sell from their own accounts and profit from the spread between their bid and offer quotes.

Dealers compete with each other in their bids and offers. Obviously, from the customer's perspective, the "best" market is highest bid and lowest offer among the dealers. This highest bid/lowest offer combination is referred to as the "inside market" or the "top of the book." For example, assume that dealers A, B, and C have the bids and offers (also called *asking prices*) for stock Alpha as shown in Figure 11.1.

Stock Alpha

Bids		Offers		
Dealer	**Bid**	**Dealer**	**Offer**	
A	40.50	C	41.00	← Top of the book: 40.50/41.00
B	40.35	B	41.10	
C	40.20	A	41.20	

Figure 11.1 Quote-Driven/Dealer Market

The best (highest) bid is by dealer A of 40.50; the best (lowest) offer is by dealer C of 41.00. Thus, the inside market is 40.50 bid (by A) and 41.00 offer (by C). Note that A's spread is 40.50 bid and 41.20 offer for a spread (or profit margin) of 0.70. A has the highest bid but not the lowest offer. C has the lowest offer but not the highest bid. B has neither the highest bid nor the lowest offer.

For a stock in the U.S. market, the highest bid and lowest offer across all markets is called the *national best bid and offer (NBBO).*

Dealers provide value to the transaction process by providing capital for trading and facilitating order handling. With respect to providing capital for trading, they buy and sell for their own accounts at their bid and offer prices, respectively, thereby providing liquidity. With respect to order handling, they provide value in two ways. First, they assist in the price improvement of customer orders, that is, the order is executed within the bid/offer spread. Second, they facilitate the market timing of customer orders to achieve price discovery. Price discovery is a dynamic process that involves customer orders being translated into trades and transaction prices. Because price discovery is not instantaneous, individual participants have an incentive to "market-time" the placement of their orders. Intermediaries may understand the order flow and may assist the customer in this regard. The intermediary may be a person or an electronic system.

The over-the-counter (OTC) markets are quote-driven markets. The OTC markets began during a time when stocks were bought and sold in banks and the physical certificates were passed over the counter.

A customer may choose to buy or sell to a specific market maker to whom they wish to direct an order. Directing an order to a specific market is referred to as "preferencing."

Order-Driven versus Quote-Driven Markets

Overall, nonintermediated, order-driven markets may be less costly due to the absence of profit-seeking dealers. But the markets for many stocks are not inherently sufficiently liquid to operate in this way. For this reason, intermediated, dealer markets are often necessary for inherently less liquid markets. The dealers provide dealer capital, participate in price discovery and facilitate market timing, as discussed above.

Because of the different advantages of these two approaches, many equity markets are now *hybrid markets.* For example, the NYSE is primarily a continuous auction order-driven system based on customer orders but the specialists enhance the liquidity by their market making to maintain a fair and orderly market. Overall, the NYSE is primarily an auction, order-driven market which has specialists (who often engage in market making), other floor traders, call markets at the open and close, and upstairs dealers who provide proprietary capital to facilitate block transactions. Thus, the NYSE is a hybrid combination of these two models. Another hybrid aspect of the NYSE is that it opens and closes trading with a call auction. The continuous market and call auction market are combined. Thus, the NYSE is a continuous market during the trading day and a call auction market to open and close the market and to reopen after a stop in trading. Thus, the NYSE is a hybrid market.

Nasdaq began as a descendent of the OTC dealer network, and is a dealer quote-driven market. It remains primarily a quote-driven market, but has added some order-driven aspects such as its limit order book, called SuperMontage (discussed below), which made it a hybrid market.

An overview of the nonintermediated, auction, order-driven market and the intermediated, dealer, quote driven markets is provided in Figure 11.2.

Another structural change that has occurred in exchanges is their evolution from membership-owned, floor-traded, organizations to publicly owned electronically traded (that is, no trading floor) organizations. The nature of this evolution (or revolution) is discussed in the next section.

CHANGES IN EXCHANGE OWNERSHIP AND TRADING STRUCTURES

Exchanges have traditionally been organizations built around a physical trading floor.

They have also usually been mutual organizations that are owned and operated on a nonprofit basis for the benefit of their members, those who operate on the trading floor. The ownership by the members is reflected in the memberships or "seats," which provide floor access or trading privileges as well as ownership rights. As the profits derived from these trading privileges increase, the prices of the seats increase and the value of the members' equity in their exchange increases. Thus, a membership organization's goal is to increase the value of the access privileges, which increases the price of a seat. A mutual organization's primary objective, thus, is to increase the income of the individual members not the profit of overall organization, which is a nonprofit organization.

However, membership organizations may not find it beneficial to themselves to adopt some changes which are beneficial to the customers of the exchange, the buyers and sellers of the exchange's products, because such changes may not be beneficial to the owners of the exchange, the members. For example, adopting a new technology may

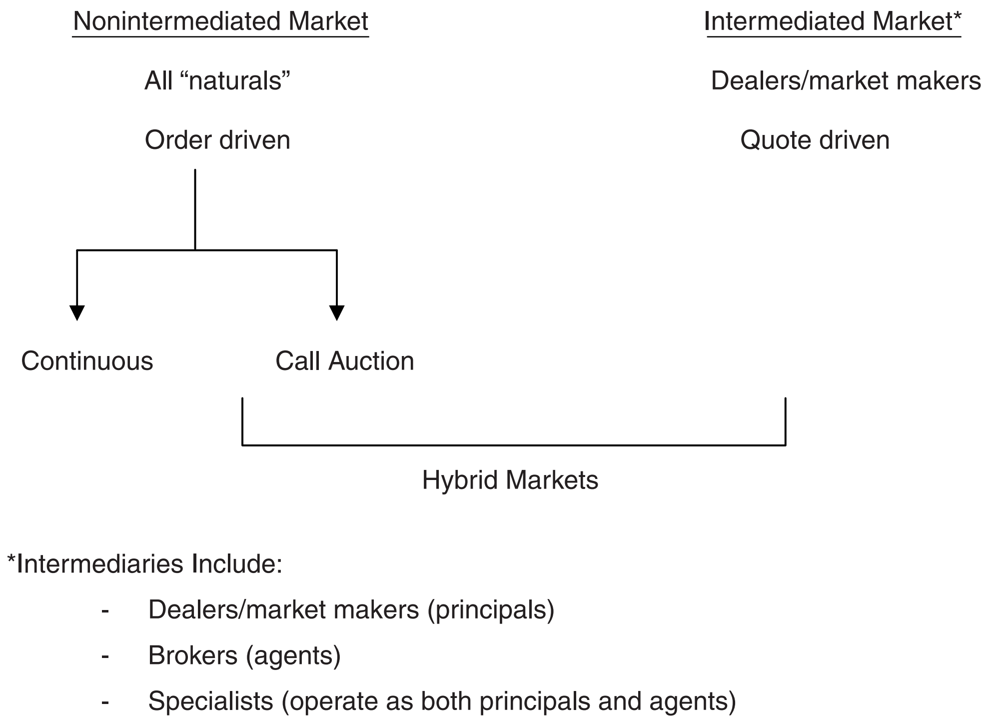


Figure 11.2 Structure of Stock Markets

benefit the customers by reducing the transaction costs but also decrease the value of access privileges and seats of a membership organization. Thus, the members/owners in a membership organization often resist technological innovations which could benefit customers but reduce the value of their own trading income.

In contrast, a publicly owned equity-based organization is a corporation and operated for a profit. And the profit of the overall organization accrues to its shareholders via an increase in the value of its equity shares. Thus, an equity-based organization might adopt the above-mentioned technology if it benefited its customers, increased its profits, and increased its stock price. The equity-based organization is free of the conflicts of a member organization between trader income and organizational profits.

While an equity-based organization may be superior over time in serving its customers, the difficult issue is convincing the members/owners in a mutual organization to agree to a demutualization and public ownership. The demutualization occurs by giving the members shares or equity in the demutualized organization in exchange for their seats in the mutual organization. Thus, the members would receive wealth/equity shares in exchange for income/access privileges. For such reasons, many exchanges have converted from membership organizations to publicly owned equity-based demutualized organizations in recent years.

Such a demutualization will align the interests of the customers of the organization and the owners of the organization. After the demutualization from a mutual company to a stock company, however, one more step is necessary before equity capital can be raised for the exchange and alliances among exchanges can be easily made with stock. Immediately after the demutualization, the stock of the equity company may be privately held and the equity shares do not have a known market value. Knowing this value is necessary if the shares are going to be exchanged, new equity capital is raised, or mergers or acquisitions among such organizations are to be consummated. In order to give the stock a known market value, the newly equity-based organization has to "go public," that is, sell at least some its stock on the public markets via an initial public offering (IPO) and then list its shares on a secondary stock market, such as the NYSE or Nasdaq. Once the IPO is complete, the resulting "corporation" knows its overall value ("market value," "market capitalization," or simply "market cap"), which is its share price multiplied by its number of shares. Corporations can then use their stock to acquire other corporations with their shares. Corporations can also use the value of their corporation as a basis for being acquired by another firm via its stock or cash. Equity or for-profit organizations have the flexibility to raise capital, make acquisitions, and acquire other organizations without resistance from its members, who would be considering their own income.

Overall, before demutualization, the market participants and market owners are the same via memberships, seats or access privileges. This is ideal for a trading floor organization. The members derive their income from trading on the floor. As a result, floor trading organizations tend to be mutual organizations. After demutualization, the market participants and the market owners are not necessarily the same entities and, thus, may have different objectives, the traders motivated by trading income and the shareholders motivated by organizational (that is, corporate) profits. Thus, after the demutualization and

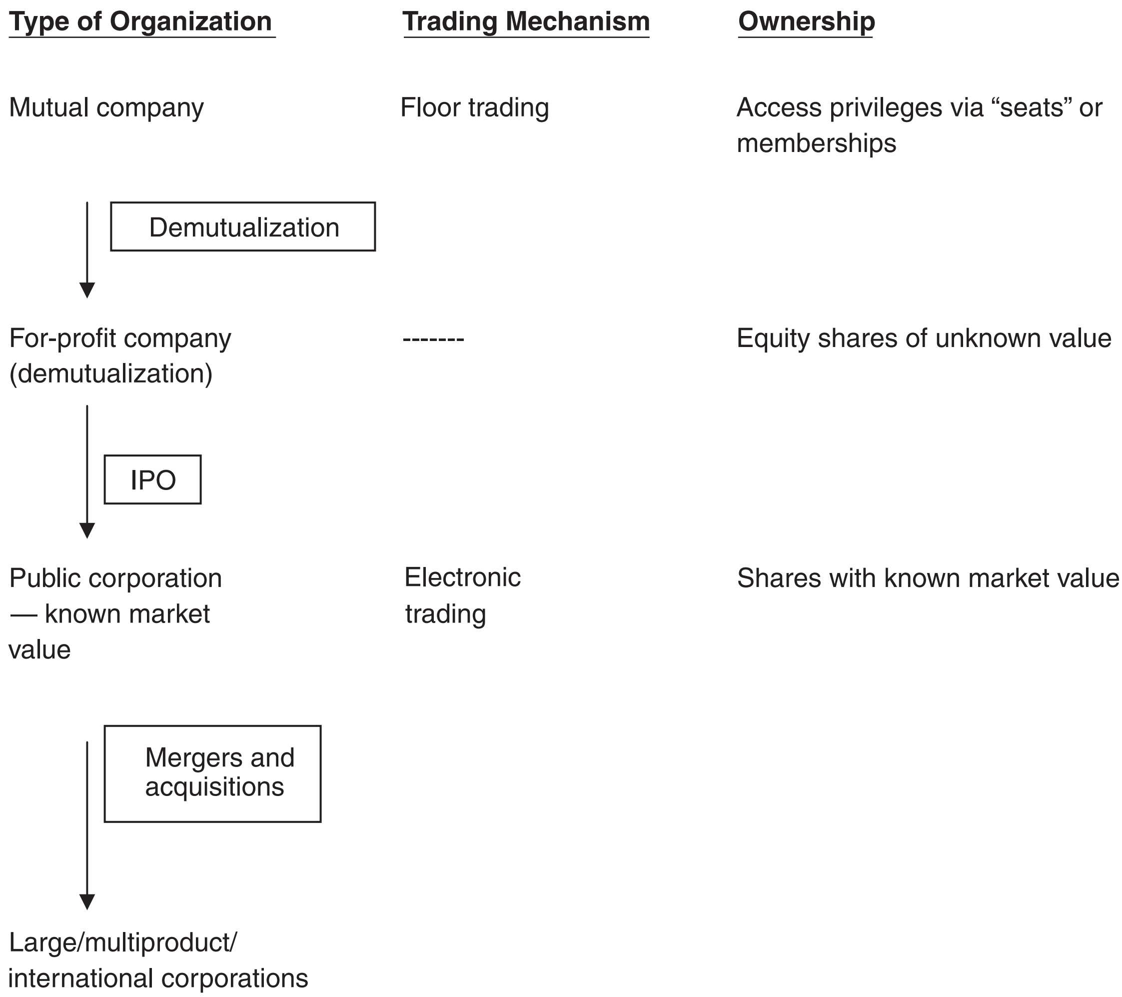


Figure 11.3 Evolution of Exchanges

the subsequent IPO, a common change is that the corporation can take actions which benefit the corporation itself by providing better service to its customers, even though trader profits may be disadvantaged. The degree of electronic trading increases and trading may become exclusively electronic or remain a mix of floor trading and electronic, often called a hybrid. The owners are shareholders, and they do not derive their income from trading on the floor. This sequence of actions is shown in Figure 11.3.

Traditionally, exchanges of most types have been based on floor trading, and the ownership of the exchanges has been with the floor traders, both individuals and firms. Changes in exchange structure and ownership began in the 1980s. Most notable were the changes that related to the "Big Bang" in London during 1986. These reforms related to the London Stock Exchange (LSE) included the abolition of minimum commissions and the introduction of "dual capacity," whereby member firms could be both brokers (agents) and "jobbers" (the British term for dealers who are principals in a transaction). One outcome of these and other changes was that the LSE's trading floor was closed and replaced by "screen trading," which is a dealer OTC market. The LSE became a public company in July, 2001 (ticker symbol: LSE).

Since then, the major futures exchanges, including the Chicago Mercantile Exchange and the Chicago Board of Trade, have become electronic trading corporations and have merged. The International Securities Exchange, an electronic options exchange, began (in 2000) as a membership-owned exchange, subsequently demutualized, and then subsequently did an IPO (in 2005). The NYSE became a publicly owned mainly electronically traded stock exchange. There were other such transformations.

THE U.S. STOCK MARKETS: EXCHANGES AND OTC MARKETS

The view of the U.S. stock market "from 30,000 feet" is that of a large homogeneous market. But while it has been large, it has not been homogeneous since the 1970s. It has become even much more heterogeneous since the 1990s. The U.S. stock market is now composed of the stock exchanges and OTC markets and also, more recently, the off-exchange markets. This section provides the "big picture" of the current stock market—specific parts of the market are examined more closely in subsequent sections.

The U.S. stock market began over two centuries ago and has evolved considerably since then. The U.S. stock market has traditionally been the core of the world capital markets. Over the last few decades, there have been significant changes in the U.S. stock market and also the other international stock markets. But, undoubtedly, during the 1990s the pace and extent of this change has accelerated.

The stock exchanges have been the primary component of the U.S. stock market. Among the types of changes in the U.S. stock exchanges are:

- The market structures of the exchanges.
- The trading mechanisms of the exchanges.
- Consolidation among different types of assets, for example, securities options and futures.
- Growth and diversity of the off-exchange markets.
- Consolidation internationally.

While the exchanges have been the main component of the U.S. stock market, the OTC markets and the off-exchange markets have also become important parts of the U.S. stock market. The off-exchange markets have also grown and become much more diverse since 2005.

This section covers the exchanges and the OTC markets; the next section considers the off-exchange markets. Given the pace and extent of the recent changes, there is a high likelihood that the markets will be much different during the next decade than it is now.

The international stock exchanges have also changed and, in fact, in some cases, have become integrated with the U.S. exchanges. However, the international stock markets are not considered here except for their relationship with U.S. exchanges.

Figure 11.4 provides a general overview, or the "big picture," of the current construct of the U.S. stock markets. The components of the current U.S. stock market are discussed individually in the following sections. This section treats the components of the U.S. stock market, including the national exchanges, the NYSE and the American Stock Exchange (Amex); the regional stock exchanges; Nasdaq, technically an OTC market, not an exchange (until June 2006); other OTC markets; and other stock exchange markets.

I. Stock Exchanges
   A. National Exchanges
      1. New York Stock Exchange Euronext
         a. NYSE Hybrid Market
         b. Archipelago ("Arca")
      2. American Stock Exchange
   B. Regional Exchanges
      1. Chicago Stock Exchange (CHX)
      2. Philadelphia Stock Exchange (PHLX)
      3. Boston Stock Exchange (BeX)
      4. National Stock Exchange (formerly the Cincinnati Stock Exchange) (NSX)
      5. Pacific Stock Exchange (owned by Archipelago; in turn owned by the NYSE)
   C. Nasdaq—the OTC Market (technically became an exchange during June 2006)
      1. Nasdaq National Market (NNM)
      2. Small Cap Market
   D. Other OTC Markets
      1. Bulletin Board ("Bullies")
      2. Pink Sheets
   E. Off Exchange Markets/Alternative Electronic Markets
      1. Electronic Communication Networks (ECNs)
      2. Alternative Trading Systems (ATS)
         a. Crossing networks
         b. Dark pools

Figure 11.4 The "Big Picture" of the U.S. Stock Market

National Exchanges

As of the first quarter of 2008, the U.S. stock markets are dominated by the NYSE and Nasdaq, the two largest exchanges (as discussed below, until June 2006 Nasdaq was not technically an exchange).

New York Stock Exchange

The beginning of the NYSE is identified as May 17, 1792, when the Buttonwood Agreement was signed by 24 brokers outside of 68 Wall Street in New York under a buttonwood tree. The current NYSE building opened at 18 Broad Street on April 22, 1903 (the "main room"). In 1922, a new trading floor (the "garage") was opened at 11 Broad Street. Additional trading floor space was opened in 1969 and in 1988 (the "blue room"). Finally, another trading floor was opened at 30 Broad Street in 2000. Notably, for reasons discussed below, during early 2006, the NYSE closed the trading room at 30 Broad Street with the beginning of the NYSE Hybrid Market and a greater proportion of the trading being executed electronically. The NYSE is referred to as the "Big Board."

The NYSE trading mechanism has been based on the specialist system. This system, as discussed above, is a hybrid of primarily an order-driven market with some quote-driven features. According to this mechanism, each stock is assigned to an individual specialists. Each specialist "specializes" in many stocks but each stock is assigned to only one specialist. Each specialist is located at a "booth" or "post." All orders for a stock are received at this post, and the specialist conducts an auction based on these orders to determine the execution price. The orders arrive at the specialists' posts either physically, delivered via firm brokers, or electronically via the Designated Order Turnaround (DOT) system or its successors. In conducting the auction, typically the specialist is an agent, simply matching orders. At times, however, the specialist becomes a principal and trades for itself in the interest of maintaining an "orderly market."

Limit orders, as opposed to market orders, are kept by the specialist in their "book," originally a physical paper book but now an electronic book. These limit orders are executed by the specialist when the market price moves to the limit. At one time, the book could be seen only by the specialist, which was judged to be a significant advantage for the specialist, but now the book is open to all the traders on the exchange floor. Overall, the NYSE trading mechanism is an auction-based, order-driven market.

This type of mechanism is often judged to provide the best price but, on a time basis, often a less rapid execution. There is, thus, a trade-off between price and speed.

The need for "space" for trading floors for the NYSE derives from its trading mechanism, a floor-based specialist system. The amount of space necessary depends not only on overall trading volume, but also the fraction of this volume that is handled by the specialist.

The NYSE lists stocks throughout the United States (as well as some international stocks) and, thus, is a "national exchange."

Trading Mechanism—The Specialist System Fundamentally, the NYSE is an auction-type market based on orders (order driven). As indicated above, the traditional trading mechanism for the NYSE is the specialist system. However, the volume of trading that has occurred electronically has increased continually. In 2006, with the advent of NYSE Hybrid Market, the degree of electronic trading has increased significantly, as discussed below.

Here, we discuss the traditional NYSE specialist system in more detail. Trading in stocks listed on the NYSE is conducted as a centralized continuous auction market at a designated physical location on the trading floor, called a post, with brokers representing their customers' buy and sell orders. A single specialist is the market maker for each stock. A member firm may be designated as a specialist for the common stock of more than one company; that is, several stocks can trade at the same post. But only one specialist is designated for the common stock of each listed company.

The NYSE began its DOT (Designated Order Turnaround) system during 1976. This system, now called SuperDOT, is an electronic order routing and reporting system that links member firms worldwide electronically directly to the specialist's post on the trading floor of the NYSE.

The NYSE SuperDOT system routes NYSE listed stock orders electronically directly to a specialist on the exchange trading floor, rather than through a broker. The specialist then executes the orders. This system was initially introduced as the DOT system but is now referred to as the SuperDOT system. The SuperDOT system is used for small market orders, limit orders, and basket (or portfolio) trades and program trades.

The SuperDOT system can be used for under 100,000 shares with priority given to orders of 2,100 shares or less. After the order has been executed, the report of the transaction is sent back through the SuperDOT system.

According to the NYSE, as of 2007, over 99% of the orders executed through the NYSE were done through SuperDOT, which meets the continually increasing demand, which stood at 20 million quotes, 50 million orders, and 10 million reports daily.

In addition to the single specialist market maker on the exchange, other firms that are members of an exchange can trade for themselves or on behalf of their customers. NYSE member firms, which are broker-dealer organizations that serve the investing public, are represented on the trading floor by brokers who serve as fiduciaries in the execution of customer orders.

The largest membership category on the NYSE is that of the *commission broker.* A commission broker is an employee of one of the securities houses (stockbrokers or wire houses) devoted to handling business on the exchange. Commission brokers execute orders for their firm on behalf of their customers at agreed commission rates. These houses may deal for their own account as well as on behalf of their clients.

Other transactors on the exchange floor include the following categories. Independent *floor brokers* (nicknamed "$2 brokers") work on the exchange floor and execute orders for other exchange members who have more orders than they can handle alone or who require assistance in carrying out large orders. Floor brokers take a share in the commission received by the firm they are assisting. Another category, registered traders, are individual members who buy and sell for their own account. Alternatively, they may be trustees who maintain memberships for the convenience of dealing and to save fees.

The major type of exchange participant is the specialist.

NYSE Specialist As indicated, specialists are dealers or market makers assigned by the NYSE to conduct the auction process and maintain an orderly market in one or more designated stocks. Specialists may act as both a broker (agent) and a dealer (principal). In their role as a broker or agent, specialists represent customer orders in their assigned stocks, which arrive at their post electronically or are entrusted to them by a floor broker to be executed if and when a stock reaches a price specified by a customer (limit or stop order). As a dealer or principal, specialists buy and sell shares in their assigned stocks for their own account as necessary to maintain an "orderly market." Specialists must always give precedence to public orders over trading for their own account.

In general, public orders for stocks traded on the NYSE, if they are not sent to the specialist's post via SuperDOT, are sent from the member firm's office to its representative on the exchange floor, who attempts to execute the order in the trading crowd. There are certain types of orders where the order will not be executed immediately on the trading floors. These are limit orders and stop orders. If the order is at a limit order or a stop order and the member firm's floor broker cannot transact the order immediately, the floor broker can wait in the trading crowd or give the order to the specialist in the stock, who will enter the order in that specialist's *limit order book* (or simply, the *book*) for later execution based on the relationship between the market price and the price specified in the limit or stop order. The book is the list on which specialists keep the limit and stop orders that are given to them, arranged with size, from near the current market price to farther away from it. Whereas the book used to be an actual physical paper book, it is now electronic. While for many years only the specialist could see the orders in the limit order book, with the NYSE's introduction of OpenBook in January 2002, the book was electronically made available to the traders on the exchange floor.

A significant advantage of the NYSE market is its diversity of participants. At the exchange, public orders meet each other often with minimal dealer intervention, contributing to an efficient mechanism for achieving fair securities prices. The liquidity provided in the NYSE market stems from the active involvement of the following principal groups: the individual investor; the institutional investor; the member firm acting as both agent and dealer; the member-firm broker on the trading floor acting as agent, representing the firm's customer orders; the independent broker on the trading floor acting as agent and handling customer orders on behalf of other member firms; and the specialist, with assigned responsibility in individual securities on the trading floor. Together, these groups provide depth and diversity to the market.

NYSE-assigned specialists have four major roles:

1. As agents, they execute market orders entrusted to them by brokers, as well as orders awaiting a specific market price.
2. As catalysts, they help to bring buyers and sellers together.
3. As dealers, they trade for their own accounts when there is a temporary absence of public buyers or sellers, and only after the public orders in their possession have been satisfied at a specified price.
4. As auctioneers, they quote current bid-ask prices that reflect total supply and demand for each of the stocks assigned to them.

In carrying out their duties, specialists may, as indicated, act as either agents or principals. When acting as an agent, the specialist simply fills customer market orders or limit or stop orders (either new orders or orders from their book) by opposite orders (buy or sell). While acting as a principal, the specialist is charged with the responsibility of maintaining a "fair and orderly market." Specialists are prohibited from engaging in transactions in securities in which they are registered unless such transactions are necessary to maintain a fair and orderly market. Specialists profit only from those trades in which they are involved; that is, they realize no revenue for trades in which they are an agent.

The term "fair and orderly market" means a market in which there is price continuity and reasonable depth. Thus, specialists are required to maintain a reasonable spread between bids and offers and small changes in price between transactions. Specialists are expected to bid and offer for their own account if necessary to promote such a fair and orderly market. They cannot put their own interests ahead of public orders and are obliged to trade on their own accounts against the market trend to help maintain liquidity and continuity as the price of a stock goes up or down. They may purchase stock for their investment account only if such purchases are necessary to create a fair and orderly market.

Specialists are also responsible for balancing buy and sell orders at the opening of the trading day in order to arrange an equitable opening price for the stock. Specialists are expected to participate in the opening of the market only to the extent necessary to balance supply and demand for the security to affect a reasonable opening price. While trading throughout the day is via a continuous auction-based system, the opening is conducted via a single-priced call auction system. The specialists conduct the call and determine the single price.

If there is an imbalance between buy and sell orders either at the opening of or during the trading day and the specialists cannot maintain a fair and orderly market, then they may, under restricted conditions, close the market in that stock (that is, discontinue trading) until they are able to determine a price at which there is a balance of buy and sell orders. Such closes of trading can occur either during the trading day or at the opening, which is more common, and can last for minutes or days. Closings of a day or more may occur when, for example, there is an acquisition of one corporation by another or when there is an extreme announcement by the corporation. For this reason, many announcements are made after the close of trading.

NYSE trading officials oversee the activities of the specialists and trading-floor brokers. Approval from these officials must be sought for a delay in trading at the opening or to halt trading during the trading day when unusual trading situations or price disparities develop.

Because of their critical public role and the necessity of capital in performing their function as a market-maker, capital requirements are imposed by the exchanges for specialists.

American Stock Exchange

The American Stock Exchange (Amex) dates from colonial times when brokers conducted outdoor markets to trade new government securities. Amex began trading at the curbstone on Broad Street near Exchange Place. Until 1929, it was called the New York Curb Exchange. In 1921, the Amex moved inside into the building where it still resides at 86 Trinity Place in New York City. In 1998, Amex merged with the National Association of Securities Dealers (NASD), which then operated Nasdaq, to create the Nasdaq-Amex Market Group wherein Amex was an independent member of the NASD parent. After conflicts between the NASD and Amex members, the Amex members bought Amex from NASD and acquired control in 2004. Amex continued to be owned by its members until its acquisition by the NYSE in early 2008.

Amex, like the NYSE, lists stocks from throughout the United States and also international stocks. Amex is therefore a national exchange. Amex is also an auction-type market based on orders. Its specialist system is similar to that of the NYSE.

Amex developed exchange-traded funds (ETFs). The first ETF, the SPY ETF, based on the S&P 500 index, was listed on the Amex on January 29, 1993. Although ETFs have proved to be a very successful product and most of the listings remain on the Amex, most of the trading volume has migrated to other exchanges, including the NYSE and Nasdaq.

The number of listings on and the trading volume of stocks on the Amex have continued to decline in recent years, and as of early 2008, the Amex is regarded as a minor market in U.S. stocks, although it continues to trade some small to mid-sized stocks.

Amex is the now highly dependent on trading in stock options.

Regional Exchanges

Regional exchanges developed to trade stocks of local firms that listed their shares on the regional exchanges and also to provide alternatives to the national stock exchanges for their listed stocks. Regional stock exchanges now exist in Chicago, Philadelphia, and Boston and have existed in many other U.S. cities. These exchanges have also been specialist-type, auction-based systems. Some of the regional stock exchanges, including Philadelphia and Boston, as well as the Amex, have been driven by trading in stock options and index options rather than stock in recent years.

Chicago Stock Exchange The Chicago Stock Exchange (CHX) was founded on March 21, 1882. In 1949, it merged with the St. Louis, Cleveland, and Minneapolis/St. Paul Stock Exchanges and changed its name to the Midwest Stock Exchange. In 1993, it changed its name back to the Chicago Stock Exchange and is the most active regional exchange.

Philadelphia Stock Exchange The Philadelphia Stock Exchange (PHLX) is the oldest stock exchange in the United States, founded in 1790. In 2005, a number of large financial firms purchased stakes in the PHLX as a hedge against growing consolidation of stock trading by the NYSE and Nasdaq. These firms—Morgan Stanley, Citigroup, Credit Suisse First Boston, UBS AG, Merrill Lynch, and Citadel Investment Group—collectively own about 45% of the PHLX.

During October 2007, PHLX announced that it was for sale by a group of its shareholders. On November 7, 2007 Nasdaq announced a "definitive agreement" to purchase PHLX for $652 million, with the transaction expected to close in early 2008.

The Philadelphia Stock Exchange handles trades for approximately 2,000 stocks, 1,700 equity options, 25 index options, and a number of currency options. As of 2007, it had a 14% U.S. market share in exchange-listed stock options trading.

Boston Stock Exchange The Boston Stock Exchange (BSE) was founded in 1834, the third-oldest stock exchange in the United States. The Boston Options Exchange (BCX), a facility of the BSE, is a fully automated options market. On October 2, 2007, Nasdaq agreed to acquire BSE for $61 million.

National Stock Exchange The National Stock Exchange (NSX), now in Chicago, was founded in 1885 in Cincinnati, Ohio as the Cincinnati Stock Exchange. In 1976, it closed its physical trading floor and became the first all-electronic stock market in the United States. The Cincinnati Stock Exchange moved its headquarters to Chicago in 1995 and changed its name to the National Stock Exchange during November 2003. The NSX handles a significant share, approximately 20%, of all Nasdaq-listed securities.

Pacific Exchange The Pacific Exchange began in 1957 when the San Francisco Stock and Bond Exchange (founded in 1882 with a trading floor in San Francisco) and the Los Angeles Oil Exchange (founded in 1889 with a trading floor in Los Angeles) merged to form the Pacific Coast Stock Exchange (the trading floors were kept in both places). The name was changed to the Pacific Stock Exchange in 1973, and options trading began in 1976. In 1997, its name was changed to the Pacific Exchange. In 1999, the Pacific Exchange (PCX) was the first U.S. stock exchange to demutualize. In 2001, the Los Angeles trading floor was closed and the next year the San Francisco trading floor was closed (the options trading floor still operates in San Francisco.)

On September 27, 2005, the Pacific Exchange was bought by the ECN Archipelago, which was in turn bought by the NYSE in 2006. No business is conducted under the name Pacific Exchange, thus ending its separate identity. All formerly PCX stock and options trading takes place through NYSE Arca.

Overall, as indicated by these brief descriptions of regional exchanges, some of the regional exchanges have diversified into options trading to remain viable. Some have made the transformation from membership-owned, trading floor organizations to publicly owned electronic organizations, and others have remained in their original forms. Finally, the regional exchanges have become attractive acquisition targets for larger exchanges, with some having already been acquired and others remaining potential merger targets.

Nasdaq Stock Market: The OTC Market

A significant change in the U.S. stock market occurred during 1971 when the National Association of Securities Dealers Automated Quotations System, now often referred to as NASDAQ (or Nasdaq, as referred to in this chapter), was founded. When it began trading on February 8, 1971, Nasdaq was the world's first electronic stock market. Nasdaq was founded by the NASD. Fundamentally, Nasdaq is a dealer-type system based on quotes (quote driven).

The NASD divested itself of Nasdaq in a series of sales in 2000 and 2001 to form a publicly traded company, the Nasdaq Stock Market, Inc. The Nasdaq Stock Market is a public corporation, the stock of which was listed on its own stock exchange at its IPO on July 1, 2002 (ticker: NDAQ).

Initially, the Nasdaq was simply a computer bulletin board system that did not connect buyers and sellers. The Nasdaq helped lower the spread (the difference between the bid price and the ask price of the stock) and so was unpopular among brokerage firms because they profited on the spread. Since then, the Nasdaq has become more of a stock market, adding automated trading systems and trade and volume reporting.

The Nasdaq, as an electronic exchange, has no physical trading floor, but makes all its trades through a computer and telecommunications system. Since there is no trading floor where the Nasdaq operates, the stock exchange built a site in New York City's Times Square to create a physical presence. The exchange is a dealers' market, meaning brokers buy and sell stocks through a market maker rather than from each other. A market maker deals in a particular stock and holds a certain number of stocks on its own books so that when a broker wants to purchase shares, the broker can purchase them directly from the market maker.

Nasdaq is a dealer system or OTC system where multiple dealers provide quotes (bids and offers) and make trades. There is no specialist system, and therefore there is no single place where an auction takes place. Nasdaq is essentially a telecommunication network that links thousands of geographically dispersed, market-making participants. Nasdaq is an electronic quotation system that provides price quotations to market participants on Nasdaq listed stocks. Nasdaq is essentially an electronic communications network (ECN) structure that allows multiple market participants to trade through it, increasing competition, as discussed below.

Since Nasdaq dealers provide their quotes independently, the market has been called "fragmented." So while the NYSE market is an auction/agency, order-based market, the Nasdaq is a *competitive dealer quote-based system.*

Until 1987, most trading occurred via telephone. During the October 9, 1987, crash, however, dealers did not respond to telephone calls. As a result, the Nasdaq developed the Small Order Execution System (SOES), which provides an electronic method for dealers to enter their trades. The Nasdaq requires that the market makers honor their trades over the SOES. The purpose of the SOES is to ensure that during turbulent market conditions small market orders are not forgotten but are automatically processed.

Over the years, the Nasdaq became more of a stock market by adding trade and volume reporting and automated trading systems. In October 2002, the Nasdaq started a system, called SuperMontage, which has led to a change in the Nasdaq from a quote-driven market to a market that provides both quote-driven and order-driven aspects; that is, it has become a hybrid market. This system permits dealers to enter quotes and orders at multiple prices and then displays these aggregate submissions at five different prices on both the bid and offer sides of the market. SuperMontage also provides full anonymity, permits dealers to specify a reserve size (that is, they do not have to display their full order), offers price and time priority, allows market makers to internalize orders, and includes preferenced orders. In effect, SuperMontage is the Nasdaq's order display and execution system.

The advent of SuperMontage continues completing Nasdaq's transformation from a quote-driven market to a hybrid market that contains both quote- and order-driven features. The Nasdaq added a third component to the hybrid, which is a call auction that both opens and closes the market. Currently, SuperMontage competes with the alternative display facility (ADF) that is operated by the Financial Industry Regulatory Authority (FINRA). SuperMontage is a key feature in Nasdaq's development.

There are two sections of the Nasdaq stock market: the Nasdaq National Market (NNM) and the Small Cap Market (also known as the Nasdaq Capital Market Issues). For a stock to be listed on the NNM, the company must meet certain strict financial criteria. For example, a company must maintain a stock price of at least $1, and the total value of outstanding stocks must be at least $1.1 million and must meet lower requirements for assets and capital. To qualify for listing on the exchange, a company must be registered with the SEC and have at least three market makers. However, the Nasdaq also has a market for smaller companies unable to meet these and other requirements, called the Nasdaq Small Cap Market. Nasdaq will move companies from one market to the other as their eligibility changes.

During December 2005, Nasdaq acquired Instinet, the largest ECN and a large trader of Nasdaq-listed stocks.

On June 30, 2006, the SEC approved Nasdaq to begin operating as an exchange in Nasdaq-listed securities. Prior to this, as indicated above, Nasdaq had been an OTC stock market but not formally an exchange. This change is more technical then substantive.

Nasdaq was very acquisitive during 2007. During September 2007, Nasdaq agreed to buy the Middle East's Borse Dubai for approximately $4.9 billion. During August 2007 Nasdaq, after failing to acquire the LSE, partnered with Borse Dubai in the Middle East to gain control of Stockholm's OMX, which operates eight Nordic and Baltic exchanges. As part of the deal, Nasdaq sold its 28% position in the LSE to Borse Dubai, which ended up with nearly a 20% stake in Nasdaq. This acquisition made the Middle East's Borse Dubai a minority owner of the combined Nasdaq/OMX. With the purchase of OMX following its agreement with Borse Dubai, Nasdaq captured 47% of the controlling stake in OMX, thereby going closer to taking over the company and becoming a trans-Atlantic exchange.

During October 2007, Nasdaq also announced plans to buy the Boston Stock Exchange (for $61 million).

During November 2007, the Nasdaq Stock Market announced that it would buy the Philadelphia Stock Exchange for approximately $650 million, mainly to trade stock options. This was Nasdaq's first effort in stock options. Nasdaq, a purely electronic exchange, was expected to maintain Philadelphia trading floor.

The NYSE versus Nasdaq

Fundamentally, the NYSE has been an auction-type market based on orders (order driven), while Nasdaq has been a dealer-type market based on quotes (quote driven).

For decades, debates continued about which system—the NYSE or Nasdaq system—was most competitive and efficient. Those who think the Nasdaq OTC market is superior to the specialist-based NYSE often cite the greater competition from numerous dealers and the greater amount of capital they bring to the trading system. They also argue that specialists are conflicted in balancing their obligation to conduct a fair and orderly market and their need to make a profit.

Proponents of the specialist NYSE market structure argue that the commitment of the dealers in the OTC market to provide a market for shares is weaker than the obligation of the specialists on the exchanges. On the NYSE, specialists are obligated to maintain fair and orderly markets. Failure to fulfill this obligation may result in a loss of specialist status. A dealer in the OTC market is under no such obligation to continue its market-making activity during volatile and uncertain market conditions. Supporters of the specialist system also assert that without a single location for an auction, the OTC markets are fragmented and do not achieve the best trade price.

Another difference of opinion comes from traders who say that the specialist system may arrive at the better price, but take a longer period of time, during which the market price may move against the trader, or at least expose the trader to the risk that it will do so. The OTC market may, on the other hand, lead to a faster execution but not arrive at a better, market-clearing price. Professional traders, in this case, often prefer higher speed over better pricing. Retail investors on the other hand may prefer a better price.

While the NYSE has been an auction type/order driven market, it has adopted many dealer-type features. Similarly, while the Nasdaq has been a dealer type/quote-driven market, it has adopted many auction-type features. Thus, while distinct differences continue between these two markets, they have converged considerably and are both currently hybrid markets, although with different mixes of order-driven and quote-driven features.

Exchange Volume Data

This section illustrates the fragmentation of the trading of the stocks listed on an exchange among different trading markets.

During the 1980s, an exchange actually traded all stocks listed by the exchanges and only those stocks. Currently, however, stocks listed on one exchange can be traded by other exchanges, including regional exchanges, by nonexchange markets such as ECN, or via internalization markets, which are discussed below. For example, during the first week of January 2008, based on exchange data, of the 13,222,716 shares of NYSE-listed stocks traded during this week, 41.7% were traded by the NYSE Euronext and 12.3% by NYSE Arca. The remainder were traded by markets not related to the NYSE, including the regional markets, Nasdaq markets, and new stock markets such as the International Securities Exchange and Chicago Board Options Exchange, discussed below.

To generalize this dispersion (or "fragmentation") of trading of an exchange's listed stocks across multiple trading venues on a day in January 2008, according to exchange data, consider that:

- Of the 4,634,118,176 shares of NYSE listed-stocks traded, 39.7% were traded on the NYSE and 12.8% were traded on NYSE Arca, for a total of 52.5% on NYSE affiliated markets.
- Of the 2,573,601,692 shares of Nasdaq-listed stocks, 48.4% were traded on Nasdaq.
- Of the 1,245,043,387 shares of Amex-listed stocks, only 3.4% were traded on Amex.

As indicated in the discussion on Regulation NMS later in this chapter, this increase in the fragmentation of trading among venues is likely to continue or even increase due to Regulation NMS.

Other OTC Markets

The OTC market is often called a market for "unlisted" stocks. As described previously, there are listing requirements for exchanges. And while, technically, the Nasdaq has not been an exchange—it was an OTC market—there are also listing requirements for the Nasdaq National Market and the Small Capitalization OTC markets. Nevertheless, exchange-traded stocks are called "listed," and stocks traded on the OTC markets, including Nasdaq, are called "unlisted."

There are three parts to the OTC market: the two under Nasdaq and a third market for truly unlisted stocks, which are therefore non-Nasdaq OTC markets. The third non-Nasdaq OTC market is composed of two parts: the OTC Bulletin Board (OTCBB) and the Pink Sheets.

Thus, technically, both exchanges and the Nasdaq have listing requirements and only the non-Nasdaq OTC markets are nonlisted. However, in common parlance, the exchanges are often called the "listed market," and Nasdaq, by default, referred to as the "unlisted market." As a result, a more useful and practical categorization of the U.S. stock trading mechanisms is as follows:

1. Exchange-listed stocks
   a. National exchanges
   b. Regional exchanges
2. Nasdaq-listed OTC stocks
   a. Nasdaq National Market
   b. Nasdaq Small Cap Market (capital market issues)
3. Non-Nasdaq OTC stocks—unlisted
   a. OTC Bulletin Board
   b. Pink Sheets

The OTCBB, also called simply the Bulletin Board or Bulletin (often just the "Bullies"), is a regulated electronic quotation service that displays real-time quotes, last sale prices, and volume information in the OTC equity securities. These equity securities are generally securities that are not listed or traded on the Nasdaq or the national stock exchanges. The OTCBB is not part of or related to the Nasdaq Stock Market.

The OTCBB provides access to more than 3,300 securities and includes more than 230 participating market makers. The traded companies do not have any filing or reporting requirements with Nasdaq or FINRA, which is discussed later in the chapter. However, issues of all securities quoted on the OTCBB are subject to periodic filing requirements with the SEC or other regulatory authorities. Companies quoted on the OTCBB must be fully reporting (that is, current with all required SEC fillings) but have no market capitalization, minimum share price, corporate governance, or other requirements. Companies that have been "delisted" from stock exchanges for falling below minimum capitalization, minimum share price, or other requirements often end up being quoted on the OTCBB.

The Pink Sheets is an electronic quotation system that displays quotes from broker-dealers for many OTC securities. Market markers and other brokers who buy and sell OTC securities can use the Pink Sheets to publish their bid and ask quotation prices. The name "Pink Sheets" comes from the color of paper on which the quotes were historically printed prior to the electronic system. They are currently published today by Pink Sheets LLC, a privately owned company. Pink Sheets LLC is neither a NASD broker-dealer nor registered with the SEC; it is also not a stock exchange.

To be quoted in the Pink Sheets, companies do not need to fulfill any requirements (e.g., filing statements with the SEC). With the exception of a few foreign issuers (mostly represented by American Depositary Receipts, or ADRs), the companies quoted in the Pink Sheets tend to be closely held, extremely small, and/or thinly traded. Most do not meet the minimum listing requirements for trading on a stock exchange such as the NYSE. Many of these companies do not file periodic reports or audited financial statements with the SEC, making it very difficult for investors to find reliable, unbiased information about those companies.

For these reasons, the SEC views companies listed on Pink Sheets as "among the most risky investments" and advises potential investors to heavily research the companies in which they plan to invest. Buying Pink Sheets stocks is intended to be difficult. Broker-dealers are enjoined to weed out unsophisticated investors who may get an e-mail or word-of-mouth tip about a small stock.

Most OTCBB companies are dually quoted, meaning they are quoted on both the OTCBB and Pink Sheets. Stocks traded on the OTCBB or Pink Sheets are usually thinly traded microcap or penny stocks and are avoided by many investors due to a well-founded fear that share prices are easily manipulated. The SEC issues a stern warning to investors to beware of common fraud and manipulation schemes.

Options Markets

In general, options trading is composed of two components: (1) options on individual stocks (stock options) and options an indexes (index options).

Options exchanges are a combination of exchanges of two different origins. The first group began as options exchanges (and, as discussed elsewhere, diversified into stock exchanges). They are the Chicago Board Options Exchange (CBOE) and the International Securities Exchange (ISE). The second group consists of stock exchanges that diversified into options exchanges. They are the American Stock Exchange, the Philadelphia Stock Exchange, the Boston Options Exchange (a subsidiary of the Boston Stock Exchange), and the New York Stock Exchange through its Archipelago holding, which had bought the Pacific Stock Exchange, which had added options to its original stock business.

With the acquisition of the Philadelphia Stock Exchange, Nasdaq will join the NYSE as a newcomer in the options market.

A significant difference between stock and options trading is that stock trading is predominantly institutional, but stock options trading has a larger retail component, as shown below:

	Institutional	Retail
Stock	85%–90%	10%–15%
Options	50%	50%

Other Stock Exchange Markets

Since options exchanges are registered with the SEC, they, too, can initiate and operate stock exchanges. During 2007, the ISE and CBOE began stock exchanges, called the ISE Stock Exchange and the Chicago Board Options Stock Exchange, respectively.

The ISE Stock Market has two components, the first of which began during September 2006. The first component, called the MidPoint Match, is a nondisplayed market or "dark pool" (discussed below), where users can trade in a continuous, anonymous pool in which trades are executed at the midpoint of the national best bid and offer. The second component is a fully displayed continuous and anonymous electronic market wherein quotes are integrated in an auction market. Thus, investors can benefit from the interaction between a nondisplayed dark pool, the MidPoint Match, and the displayed liquidity pool. In this combination of systems, orders will have the opportunity for price improvement from the MidPoint Match system, or be executed or be displayed on the market's order book; or be routed out to other exchanges as required by the SEC's Regulation NMS.

Interestingly, the CBOE and ISE were options-only exchanges that subsequently developed stock exchanges. Some of the regional stock exchanges—Philadelphia, Boston, and Pacific—later developed options exchanges. In addition, the NYSE is in the stock options business through its purchase of Archipelago, which had previously bought the Pacific Stock Exchange. And Nasdaq entered the stock options business through its purchase of the Philadelphia Stock Exchange.

OFF-EXCHANGE MARKETS/ ALTERNATIVE ELECTRONIC MARKETS

As explained earlier, the national and regional exchanges have continued to evolve and, in particular, have become much more electronically oriented. As of early 2008, however, a large volume of U.S. stock trading is done off any of the regulated stock exchanges. There has been significant growth and innovation in this sector of the U.S. stock markets in recent years. The *off-exchange markets* (also called *alternative electronic markets*) have continued to grow rapidly and become much more diverse.

Innovation in nonexchange (or off-exchange) trading began even before Nasdaq began. For example, Instinet began trading in 1969 and was essentially the first electronic communications network (although, as discussed below, it was not called an ECN until the late 1990s, when the SEC introduced the term as part of the development of its order-handling rules).

In general, these off-exchange markets are divided into two categories: electronic communications networks and alternative trading systems.

Electronic Communications Networks

Electronic communications networks (ECNs) are essentially off-exchange exchanges. They are direct descendants of (and part of) Nasdaq, not the NYSE. ECNs are privately owned broker-dealers that operate as market participants, initially within the Nasdaq system. They display bids and offers; that is, they provide an open display. They provide institutions and market makers with an anonymous way to enter orders. Essentially, an ECN is a limit order book that is widely disseminated and open for continuous trading to subscribers, who may enter and access orders displayed on the ECN. ECNs offer transparency, anonymity, automated service, and reduced costs, and are therefore effective for handling small orders. ECNs may also be linked into the Nasdaq marketplace via a quotation representing the ECN's best buy and sell quote. In general, ECNs use the Internet to link buyers and sellers, bypassing brokers and trading floors. ECNs are informationally linked, even though they are distinct businesses. ECNs are subject to some best execution responsibilities including the SEC's Regulation NMS, which is discussed later.

Consider the background of ECNs. Instinet, the first ECN, began operating in 1969 before Nasdaq was founded in 1971. Instinet was designed to be a trading system for institutional investors (hence its name, which stands for "institutional network"). Instinet was viewed as an alternative to and competitor of the traditional Nasdaq dealer market. Instinet was intended to be a trading system for institutional investors, which allowed them to meet in an anonymous, disintermediated market.

Instinet seemed very similar to an exchange but was registered with the SEC, not as an exchange, but initially as a broker-dealer and subsequently as an ECN. Instinet took the position that they were just a broker-dealer that operated in the off-exchange ("upstairs") market as does any other broker-dealer that puts trades together for large customers. The only difference, according to Instinet, was that it operated electronically. This view emphasized the difficulty of distinguishing an exchange from a broker-dealer in a technological environment. The SEC acknowledged this difficulty by using a new category to apply to Instinet, that is, electronic communications Network, as discussed below when we explain order handling rules.

The number of ECNs increased considerably after the SEC imposed the order handling rules in 1997, as discussed below. As a result, ECNs significantly affected Nasdaq during the late 1990s after the SEC adopted its new order-handling rules in 1997. ECNs such as Archipelago, Brut, Island, and Instinet captured a majority of Nasdaq volume in about two years. Instinet acquired Island in September 2002.

Archipelago, which began operating in 1997, handles both institutional and retail order flow. Another ECN, Island, was primarily retail. Prior to these developments, all the off-exchange systems were designed for institutional customers.

As many as a dozen ECNs existed by early 2000. Then a wave of consolidations and acquisitions began that within only two years whittled that number down to a handful.

Some of the large ECNs were acquired—Instinet by Nasdaq during December 2005, and Archipelago by the NYSE during March 2006. Prior to its acquisition, Archipelago, an ECN at the time, acquired the Pacific Stock Exchange, to form a fully electronic stock exchange.

1969	Instinet (Institutional Network)—first ECN, formed in 1969, before Nasdaq • Electronic block-trading system for institutional investors
1997	OHRs (order-handling rules) approved • ECNs grew quickly in number
1997	Archipelago formed in December 1996; began trading in 1997; granted exchange status by SEC in October, 2001
1997	Island included in the Nasdaq Montage in January, 1997
1999	NYSE Rule 390 eliminated
2000	Archipelago, an ECN, bought the Pacific Stock Exchange, to form the first fully electronic stock exchange
2001 (May 18)	Instinet went public—IPO was a success
2002	Instinet acquired Island
2005	Instinet acquired by Nasdaq
2006 (Jan.)	BATS ECN initiated
2006	Archipelago acquired by NYSE
2008	Only a few ECNs remain, including BATS, Direct Edge, and LavaFlow

Figure 11.5 ECN Highlights

As of early 2008, there were a few ECNs operating; the largest is BATS, which provides trades to Nasdaq, the NYSE, the ISE, and some regional exchanges. BATS began in January 2006 and applied to the SEC to become a fully licensed securities exchange during 2007. Among the others are Direct Edge, Bloomberg, LavaFlow, and Track Data.

Prior to 2000, ECNs could not penetrate the NYSE-listed stocks as they did the Nasdaq market. The main reason was the impediment imposed by the NYSE's Rule 390, also called the order consolidation or order concentration rule. According to Rule 390, dealers who traded NYSE-listed stocks in the OTC could not be members of the NYSE. For this reason, only a few dealers actively participated in the OTC market for NYSE-listed stocks, and NYSE-listed stocks were traded mainly on the NYSE.

All central markets have incentives to impose order consolidation rules on their members. However, the SEC, to open up the NYSE market, pressured the NYSE to eliminate Rule 390. The NYSE eliminated Rule 390 in December 1999. This elimination exposed the NYSE to the same type of fragmentation to which Nasdaq had been exposed. But in the years immediately after the elimination of Rule 390, the NYSE continued to conduct most of the trading in its stocks; that is, the NYSE market did not experience nearly as much fragmentation as the Nasdaq markets. Subsequently, however, the NYSE has lost considerable market share to ECNs and other exchanges.

Some of the key events for ECNs are summarized in Figure 11.5.

Alternative Trading Systems

In addition to ECNs, other alternative trading systems (ATSs) developed as alternatives to exchanges.

It is not necessary for two natural parties conducting a transaction to use an intermediary. That is, the services of a broker or a dealer are not required to execute a trade. The direct trading of stocks between two customers without the use of a broker or an exchange is called an ATS.

A number of proprietary ATSs have been developed. These ATSs are for-profit "broker's brokers" that match investor orders and report trading activity to the marketplace via the Nasdaq or the NYSE. More recently, such trades have been reported through Trade Reporting Facilities, as discussed later. In a sense, ATSs are similar to exchanges because they are designed to allow two participants to meet directly on the system and are maintained by a third party, who also serves a limited regulatory function by imposing requirements on each subscriber.

Broadly, there are two types of ATSs: crossing networks, which have functioned since the 1980s; and dark pools, which are much more recent.

Crossing Networks

Crossing networks are electronic venues that do not display quotes but anonymously match large orders. Crossing networks are systems developed to allow institutional investors to cross trades—that is, match buyers and sellers directly—typically via computer. These networks are batch processors that aggregate orders for execution at prespecified times. Crossing networks provide anonymity and reduce cost, and are specifically designed to minimize market impact trading costs. They vary considerably in their approach, including the type of order information that can be entered by the subscriber and the amount of pretrade transparency that is available to participants.

A crossing network matches buy and sell orders in a multinational trade at a price that is set elsewhere. The price used at the cross can be the midpoint of a bid-ask spread (such as the national bid and offer, as discussed below) or the last transaction price at a major market (such as the NYSE or Nasdaq) or linkage of markets. Thus, no price discovery results from a crossing network.

The major drawbacks of the crossing networks are (1) that their execution rates tend to be low and (2) that if they draw too much order flow away from the main market, they can, to their own detriment, undermine the quality of the very prices on which they are basing their trades. These limitations can be overcome in a call auction environment that includes price discovery.

ATS began developing during October 1987 when Investment Technologies Group's (ITG) Posit began. Posit is a crossing network that matches customer buy and sell orders that meet or cross each other in price (this is the way crossing networks were named) at a price established by the NYSE or the Nasdaq markets or the overall national market.

Another crossing network, LiquidNet, started operation in 2001. LiquidNet is an ATS that enables institutional customers to meet anonymously, negotiate a price, and trade in large sizes (average trade size is nearly 50,000 shares). Part of LiquidNet's ability to attract order flow is attributable to its customers' being able to negotiate their trades with reference to quotes prevailing in the major market centers. In other words, LiquidNet's customers do not have to participate in significant price discovery. Further, LiquidNet customers' anonymity and knowledge that counterparties in the system also wish to trade in size offers them some assurance that their orders will not have undue market impact. A key feature of the LiquidNet system is that customer matches are found electronically, and negotiations are also conducted electronically by the natural buyer and seller. LiquidNet has also developed in Europe.

Instinet, in addition to its continuous ECN, also developed an after-hours crossing, the Instinet Crossing Network. Instinet's after-hours cross was the first crossing network.

The Burlington Capital Markets, Burlington Large Order Cross (BLOX) also provides crossing systems. These systems enable institutions to trade with no price impact in a batched environment; the crosses are made at prices set in other stock market places. In addition, Harborside, which started operations in 2002, provides crossing services. These systems assist institutional customers to meet anonymously and negotiate their trades in an anonymous manner in an electronic environment that uses current quotes from external stock markets as benchmarks.

These crossing system are designed exclusively for institutional order flow. Among the major current crossing networks and their area of specializations are:

- **LiquidNet:** for the buy-side to buy-side only.
- **Pipeline:** for buy-side to buy-side block business only.
- **ITG Posit:** provides timed crossings 5 to 10 times per day for buy-side to buy-side only.
- **BIDS:** unlike the first three is an agency broker, that is it does not engage in proprietary trading and, thus, compete with its customers; launched in spring 2007.

Crossing networks have provided attractive alternatives to institutions to trade without their orders having any impact on the prices. However, due to lack of liquidity, their execution rates tend to be low, and if they draw too much order flow from the established markets, they could undermine the quality of the prices, which are the bases for the trades. In effect, crossing networks that use prices from the central stock markets to price their crosses are "free riding" on the price discovery of the central markets. These limitations could be resolved in a call-auction environment, which does provide price discovery.

In a call auction, sometimes called a period call, orders from customers are batched together for a simultaneous trade at a specific point in time. At the time of the call (in a "timed call") a market clearing price is determined—that is, there is a price discovery—and buy orders at this price and higher and sell orders at this price and lower are executed.

But the two systems based on call auction methods have not developed liquidity. The two ATSs based on call auction principles were the Arizona Stock Exchange (which started operations in 1991 and has been inactive since 2001) and Optimark (which started in 1999 and has been inactive since 2000). Neither of these systems succeeded in attracting critical mass order flow. Their experiences point up the difficulty of implementing an innovative new trading system that has to compete with an established market center, especially when the new system provides independent price discovery. These call auction systems provided price discovery and, thus, competed with established market centers and had difficulty attracting order flow.

Crossing markets are offered by some of the major broker-dealers who may also use such systems to "internalize" their order flow, that is match or cross bids and offers "upstairs," that is in their own organization. These orders may both be customer orders, or one may be a customer order that they cross with their own proprietary orders. A selection of the firms involved in internalization is Citigroup, Credit Suisse, Goldman Sachs, Lehman Bros., Merrill Lynch, Morgan Stanley, and UBS.

Dark Pools

Another step in the evolution of nonexchange trading is the use of *dark pools.* Dark pools fulfill the need for a neutral gathering place and fulfill the traditional role of an exchange in the new paradigm. Dark pools are private crossing networks in which participants submit orders to cross trades at externally specified prices and, thus, provide anonymous sources of liquidity (hence the name "dark"). No quotes are involved—only orders at the externally determined price—and, thus, there is no price discovery.

Dark pools are electronic execution systems that do not display quotes but provide transactions at externally provided prices. Both the buyer and seller must submit a willingness to transact at this externally provided price—often the midpoint of the NBBO—to complete a trade. Dark pools are designed to prevent information leakage and offer access to undisclosed liquidity. Unlike open or displayed quotes, dark pools are anonymous and leave no "footprints." The advent of pricing in pennies led to less transparent markets and was, thus, instrumental in the initiation of dark pools.

Dark pools, as well as crossing networks, are creating very fragmented markets for large trades and block trades. Customers are also using algorithmic trading (discussed later) to respond to such hidden liquidity.

Among the advantages of dark pools are:

- Nondisplayed liquidity.
- Prevent information leakage (anonymous trading).
- Volume discovery.
- Reduced market impact.

Among the disadvantages are:

- Less or no visibility.
- Difficulty to interact with order flow.
- No price discovery.

The sponsors of dark pools can be:

- Exchanges (e.g., NYSE Euronext, the Nasdaq stock market, and the International Securities Exchange).
- Broker-dealers (e.g., Credit Suisse, Lehman Brothers, Morgan Stanley, Goldman Sachs, Merrill Lynch, and others; can be used for brokerage internalization).
- Independent organizations (e.g., Instinet, Liquidnet, Pipeline Trading Strategies and Investment Technologies Group (ITG) Posit).
- Consortia of other organizations.

I. No Intermediary (All "Naturals")
  A. Continuous
  B. Call auction
    1. Provides price discovery (that is, priced at an internally determined price)
II. Via Intermediaries
  A. Exchanges
    1. National
      a. NYSE Hybrid—auction (specialist) and electronic (Direct +)
      b. NYSE—Arca
      c. Nasdaq
      d. Amex
    2. Regional
      a. CHX (Chicago)
      b. PHLX (Philadelphia)
      c. BSE (Boston)
      d. National (NSX)
      e. Pacific Exchange
  B. Off-Exchange Systems/Alternative Electronic Markets
    1. ECNs
    2. Alternate Trading Systems
      a. Crossing networks
      b. Dark pools
      c. Internalization

Figure 11.6 Structure of Equity Transactions

Summary of the Structure of the Mechanisms for Equity Transactions

In this section, several mechanisms for transacting equities, both conceptual and actual, have been discussed. Figure 11.6 summarizes the structure of the mechanisms for equity transactions. While the exchange markets continue to evolve, the off-exchange markets also continue to grow, diversify, and add advantages to the overall market. ECNs are used for either institutional or retail order flow, while crossing networks and dark pools are used mainly for institutional order flow.

Another difference between sectors of the off-exchange market is that ECNs are quote driven, while crossing networks and dark pools are order driven. One disadvantage of these off-exchange markets is that they tend to fragment the overall market, that is, remove quotes and trades from the central exchange markets.

Some analysts observe that there is a hierarchy in the quality of execution, with ATS being the highest, ECN being next, and exchanges being the lowest. Some analysts say that the exchanges are getting the so-called "exhaust" of executions.

THE CURRENT NYSE STOCK MARKET

This section discusses recent changes in the NYSE. The section focuses on the NYSE because of the significant recent changes in its market structure, advances in technology, ownership, and acquisition of other market venues.

Background

For years, even decades, the NYSE gradually increased the use of electronic trading to supplement it specialist-based, human-touch trading. Then, during 2006 and early 2007, this evolution turned into a revolution.

During the previous years, the functionality and capacity of DOT and SuperDOT, discussed earlier, continually increased, reducing the need for specialist intervention for a larger number of trades. A significant indication of the rapid change in the need for specialist-oriented NYSE floor space was that a new NYSE trading room that was opened at 30 Broad Street during October 2000 was closed in February 2007.

There were several components of this revolution. The two major components were as follows. First, in December 2005, the NYSE initiated its NYSE Hybrid Market, which gave customers the choice of the traditional auction-based specialists system or new electronic trading. The basis for the electronic component of the NYSE Hybrid Market was NYSE Direct+, an automatic execution service, the pilot of which was launched in October 2000 and which was expanded in August 2004.

Second, the crescendo of this revolution occurred in March 2006. On March 7, the NYSE merged with Archipelago Holding Inc. (commonly called "Arca") and, as a result, became a for-profit, publicly owned company. On the following day, the shares of the newly formed NYSE Group began trading on an exchange—the NYSE, of course. As a result, there were no more NYSE memberships or "seats" (which reached a high of $4 million in December 2005). These seats were replaced by NYX shares as a measure of value and "access" rights for floor trading privileges became available separately.

The period 2006–2007 was exceptionally active for the NYSE. Specifically, the major events during this period were as follows:

- On March 3, 2006, the NYSE bought Archipelago Holdings, a publicly owned, for-profit exchange (Arca was granted exchange status by the SEC on October 20, 2001). Archipelago bought the Pacific Stock Exchange during January 2005. NYSE Group Inc., a public company, was formed out of the merger of NYSE and Archipelago.
- On the next day, NYSE Group Inc. conducted its IPO and began trading (ticker symbol: NYX; initial price: $67). Thus, the IPO was arranged in conjunction with the acquisition of Archipelago.
- Over the period from October 2006 to January 2007, the NYSE introduced the NYSE Hybrid Market, a blend of an auction and an electronic market. Archipelago remained a distinct electronic market.
- On April 4, 2007, the NYSE Group completed a merger with Euronext NV, a Paris-based European stock exchange, making the NYSE the first trans-Atlantic exchange group. Thus, the NYSE became a global company by buying Euronext. The NYSE went public later than many other exchanges but became an international company before many others.

Thus, during a brief period from 2006 to 2007, the NYSE went public, initiated a hybrid market, and became global.

While the NYSE Hybrid was introduced in the period from October 2006 to January 2007, it was based on systems initiated and developed previously. Some of these are described below:

- **Designated Order Turnaround Systems (DOT and SuperDOT).** This system allows brokers to route orders, usually retail orders, directly to the specialist posts or the trading floor for execution. The original DOT system was initiated during 1976 and has been continually expanded and improved.
- **Direct+.** In October 2000, the NYSE introduced this system, which is an automatic execution service on limit orders up to 1,099 shares at the published NYSE quote. The only option for market orders was the standard method. NYSE Direct+ was subsequently expanded and became the foundation on which the electronic component of the NYSE Hybrid Market was built.
- **OpenBook.** In January 2002, the NYSE introduced this system, which provides limit order book information to traders on the exchange floor. This was the first step in opening the previously closed specialist's order book.

A summary of the key dates and activities is summarized in Table 11.1.

Table 11.1 Summary of Key NYSE Events

1976	Designated Order Turnaround (DOT) initiated
October 2000	Direct+ initiated
January 2002	OpenBook initiated
March 7, 2006	NYSE buys Archipelago; becomes public company named "NYSE Group Inc."
March 8, 2006	NYSE conducts IPO; listed on the NYSE with ticker symbol NYX
October 2006–January 2007	NYSE Hybrid Market initiated on January 24, 2007 for all NYSE stocks (except for a few high-period stocks).
April 4, 2007	NYSE completes merger with Euronext; now named "NYSE Euronext."

The Prelude to the NYSE Hybrid Market

The remainder of this section provides additional information on these elements of the NYSE Hybrid Market and the overall Hybrid market.

Designated Order Turnaround Systems (DOT or SuperDOT)

Traditionally, the NYSE has conducted its execution via the specialist system wherein the specialists execute orders presented by floor brokers. However, as explained earlier, the use of electronic trading has continually increased. The key mechanism for electronic trading has been the DOT or SuperDOT.

DOT is an electronic system that increases order efficiency by routing orders for listed securities directly to a specialist on the trading floor, instead of through a broker to the specialist. The DOT system is used by the NYSE for small orders, limit orders, basket trades, and program trades.

Electronic trading on the NYSE began in 1976 with the DOT for market orders of 100 shares. DOT's capabilities were expanded over time to accommodate limit orders and larger sizes. DOT was renamed SuperDOT in 1984. The system can be used for orders up to 100,000 shares, and it has the capacity to handle in excess of 2 billion shares daily. Orders that come in through SuperDOT are referred to as system orders. Floor brokers provide the main alternative delivery method.

To some extent, however, separate order handling may be unavoidable. Namely, whereas small retail orders are typically sent electronically by SuperDOT to specialist posts, large orders are commonly worked by floor brokers on a not-held basis or are negotiated in the upstairs institutional market. These procedures, however, are coordinated and the exchange's order flow is reasonably well integrated. This coordination is of critical importance because price discovery should reflect the desires of the broad market to hold shares of stock.

Data on the participation of specialists, floor brokers, and the entire floor (specialists plus floor broker) by Hendershott and Moulton (2007) found that since 1999 there has been a continual decline in the floor volume, both floor broker and specialist. The data indicate that floor trading volume began to decline significantly during 2002, the time of the introduction of OpenBook, which provides limit order book data to traders off the exchange floor. By June 2006, floor trading had declined to slightly more than 20% of the total volume from in excess of 50% in January 1999. The specialist share was less than 10% in May 2006. The data by Hendershott and Moulton also show that these shares continued to decline during June 2006 through May 2007 with the total flow volume approximately 10% and the specialists' share approximately 5% during May 2007.

According to the NYSE, during 2007, 99% of all orders pass through SuperDOT.

Anonymous SuperDOT (ADOT) enables institutional investors sponsored by a member firm to submit orders directly to the NYSE without the exchange, member firm, specialist or floor brokers knowing their identity. Institutions that have been sponsored by member firms can use ADOT to send orders directly and anonymously to the point of sale on the NYSE trading floor. With ADOT, the institution will receive transaction reports as they occur, and the member firm will receive a copy only at the end of the day or after an agreed-upon time has elapsed, for clearing purposes.

Direct+

Automatic execution was introduced on the NYSE in October 2000 when Direct+ was launched. Initially, Direct+ provided an automatic execution service only on limit orders of up to 1,099 shares at the published NYSE quote. There was also a 30-second rule for repeat executions belonging to the same beneficial owner. The only option for market orders was the standard auction method.

In 2003, the NYSE began automatically updating the best bid and offer quotes that reflected changes in the limit order book. Prior to this, the best bid and offer quotes were updated manually by the specialist. Also during 2003, the NYSE launched LiquidityQuote, which disseminated executable, sizable quotes outside the best bid and offer so that users could find greater market size and depth.

After filing during August 2004, the NYSE expanded the NYSE Direct+ automatic trading system, by eliminating the limits on the size, trading, and types of orders that could be submitted via Direct+, thus significantly increasing electronic trading at the NYSE.

NYSE Direct+ is the foundation upon which the electronic component of the Hybrid Market is built. NYSE Direct+ enables users to opt for an immediate execution at the best bid or offer, without a fee and with anonymity and speed. The average execution time is 0.36 seconds. This trade volume represents approximately 17% of NYSE's average daily volume. Direct+ has had two main restrictions: (1) a maximum order size of 1,099 shares; and (2) a minimum reload time of 30 seconds. Within the Hybrid Market, these two restrictions have been removed, allowing customers automatic execution up to 1 million shares, sweep the best bid/offer, and to trade without time restrictions.

An important distinction between SuperDOT and Direct+ is that SuperDOT routes the quote or order directly to the specialist and then the specialist actually provides the execution whereas with Direct+ the system, not the specialist, provides the execution automatically.

OpenBook

In January 2002, the NYSE introduced OpenBook, which provides limit order book information to traders off the exchange floor. Immediately after its introduction, OpenBook information was released every 10 seconds. As of May 1, 2006, this information was released continuously (as fast as the NYSE system would permit). OpenBook does not include floor trader interest, and there are lags in floor executions.

The NYSE Hybrid Market

The changes identified above, and others, led to the launch of the NYSE Hybrid Market during late 2006; actually, the NYSE Hybrid Market was introduced across the various stocks listed on the NYSE from October 2006 through January 2007. As of January 2007, all NYSE stocks were traded via its electronic Hybrid Market (except for a small group of very high-priced stocks). Customers could send orders for immediate electronic execution, or route orders to the floor for trade in the auction market. In excess of 50% of all order flow is now delivered to the floor electronically. The NYSE Hybrid Market offers customers the choice between the auction system with the opportunity for price improvement provided by the specialist system and very fast automated trade execution provided by the electronic system. Thus, the NYSE Hybrid Market provides investors with a blend of market models, from completely automated to the value-added auction market.

Among the reasons for the NYSE to launch the Hybrid Market were that (1) it wanted to give users the choice of the existing auction mechanism to attain better prices or the new electronic mechanism for faster executions;

(2) higher trading volumes could be executed more efficiently; and (3) the SEC Regulation NMS order protection rule affords protection for better-priced quotes from being traded only in markets that offer automatic protection at the posted quotes, defined as "fast" markets. After the Hybrid Market, the NYSE market qualified as fast a very large fraction of the time.

The Hybrid Market involved several changes in the NYSE's systems and rules which enhanced the efficiency of the trading process and increased automatic execution and speed. The intent and effect of the Hybrid Market was to expand automatic execution. The specific changes included: (1) orders were no longer limited to 1,099 shares—the new limit is 1 million shares; (2) the 30-second frequency restriction was eliminated; (3) market orders as well as limit orders are eligible for automatic execution; and (4) market and marketable limit orders are automatically executed by default, rather than requiring a special code.

The Hybrid Market automates much of the specialists' activity, helping them become much more efficient at the point of sale. In order to react to customers, specialists interface with the market through proprietary algorithms that interface with the NYSE Display Book via an NYSE-provided application programming interface (API).

On the trading floor, the NYSE still trades in a continuous auction format. Here, the human interaction and expert judgment with respect to order execution differentiates the NYSE from a fully electronic market. There is still one specific location on the trading floor where each listed stock trades as before. The specialists do on occasion (approximately 10% of the time) facilitate the trades by committing their own capital and, as a matter of course, disseminate information to the crowd that helps to bring buyers and sellers together.

The NYSE Hybrid Market is a NYSE-developed electronic trading system. Archipelago is an NYSE-acquired electronic trading system. The NYSE could have used Archipelago as its only electronic trading system but did not because it believed that its NYSE Hybrid system was better suited to support the diversity of trade-execution choices and information flows that comprise the NYSE Hybrid Market. The NYSE's platform includes a range of technology tools that help specialists and brokers provide value to customers. The NYSE Group's plan has been to retain the distinct value of both the NYSE and NYSE Arca marketplaces. Eventually, the infrastructure supporting the trading platform will be consolidated to provide operating efficiencies while retaining the value-added functionalities of each marketplace.

The NYSE Hybrid Market changes apply only during the continuous trading day, not during the opening and closing call auctions, which are handled manually by the specialist.

A new regulatory requirement, which is part of the SEC's Regulation NMS, is the order-protection rule, which requires market centers to satisfy better-priced quotes in other fast markets before filling orders in their own markets. In order to be a fast market, subsecond automatic execution must be offered, as well as the immediate or cancel capability. The Hybrid Market transformed the NYSE from an auction market with an average trade-execution time of 9.0 seconds to a fast market with sub-second turnaround times.

One expectation of electronic trading is that it will provide faster executions and higher executions costs than the floor trading. This expectation is supported by Hendershott and Moulton (2007, p. 23) who state that: "The Hybrid experience suggests that while investors who favor faster execution benefit, investors who are more concerned with lower execution costs or low volatility than speed, may be worse off in a world without floor trading."

In fact, the NYSE Hybrid Market has considerably increased electronic trading at the expense of floor trading. The Hybrid Market expansion of automatic execution has reduced the potential for specialists and floor brokers to participate manually in executions.

In November 2007, NYSE Euronext announced that it would overhaul its Hybrid Market system during 2008, possibly permitting market making firms to act as specialists.

Impact of the NYSE Hybrid Market

As of February 2001, there were 18 specialist firms on the NYSE. Traditionally, NYSE specialist firms were small, private family-owned businesses. However, recently, NYSE specialist firms have consolidated, been acquired by public companies, or have gone public themselves. The increased capital requirements for specialists have been one factor that has led to a consolidation of the specialist firms. These specialists have continued to face increasing competition from other sources such as ECNs. The increase in electronic trading has reduced the volume of trading through the specialists and, as a result, the number of specialists.

In 2007, the number of specialist firms continued to decline. In November 2007, two specialists firms (Van Der Moolen Specialist USA and Susquehanna International Group LLP) stopped trading on the NYSE floor. These departures left only five specialists firms. These five are units at Bank of America Corp., Goldman Sachs Group Inc., Labranche & Co., Kellog Specialists Group, and Bear Wagner Securities (a unit of Bear Stearns Cos.). Some analysts regarded these departures as a natural move toward more sophisticated electronic trading systems. There had been over 40 specialist firms in the early 1990s.

In addition, the fraction of shares that required human/specialist intervention decreased as the share executed electronically increased. During 2007, on normal trading days, only 15% of trading required human intervention. On high-volume, volatile days, however, this share increases to 25% because some traders continue to want the service of a specialist on these days (Lucchetti, 2007).

U.S. stock exchanges no longer enjoy a monopoly on trading in their own listed stock. They have lost share to competing exchanges and alternative electronic systems. For example, over the period from 2005 to October 2007, the NYSE lost market share, which declined to approximately 40% of the trading on its listed shares versus over 80% in earlier years (Lucchetti, 2007).

Thus, overall, the NYSE has lost market share in its listed stocks to alternative electronic systems such as ECNs, crossing networks, and dark pools. In addition, the market share it has kept has gone primarily electronic via the NYSE Hybrid Market, which has lessened the need for NYSE's traditional specialist franchises. As a result, some question whether an NYSE listing is worth as much if a large share of the listed companies shares trade away from the NYSE and even the NYSE-traded segment is traded electronically, which is more common across the markets, rather than electronically, which is unique to the NYSE.

Euronext NV

Euronext NV was a pan-European stock exchange based in Paris with subsidiaries in Belgium, France, Netherlands, Portugal, and the United Kingdom. In addition to equities and derivatives markets, the Euronext group provides clearing and information services. As of January 31, 2006, markets run by Euronext had a market capitalization of US$2.9 trillion, making it the fifth-largest exchange in the world before it merged with the NYSE Group to form NYSE Euronext, the first global stock exchange.

Euronext was formed on September 22, 2000, in a merger of the Amsterdam Stock Exchange, Brussels Stock Exchange, and Paris Bourse, in order to take advantage of the harmonization of the European Union financial markets.

In December 2001, Euronext acquired the shares of the London International Financial Futures and Options Exchange (LIFFE) and Euronext.liffe was formed in January 2002.

The derivatives activities of the other constituent exchanges of Euronext (Amsterdam, Brussels, Lisbon, and Paris) were merged into Euronext.liffe. Euronext.liffe continues to operate under its own governance. Trading is done electronically through the LIFFE CONNECT platform. Euronext.liffe offers a wide range of futures and option products on short-term interest rates, bonds, swaps, equities, and commodities.

NYSE Group offered €8 billion (roughly $10.2 billion) in cash and shares for Euronext on May 22, 2006. The merger was completed on April 4, 2007, forming the NYSE Euronext. NYSE Euronext then owned the NYSE Group, Euronext, and Archipelago.

American Stock Exchange

During January 2008, the NYSE announced an agreement to buy the American Stock Exchange. The purchase price was expected to be approximately $260 million (plus the proceeds from the sale of Amex's classic lower Manhattan building, expected to be $50 million to $100 million), considerably less than the Nasdaq paid for the Philadelphia Stock Exchange during November 2007. The Amex will relocate its traders to NYSE's trading floor, which has extra space, as discussed.

The purchase of the Amex will (1) increase the NYSE's relatively new stock options business, (2) increase its Arca stock exchange listings of emerging company shares by several hundred companies, and (3) strengthen NYSE's already large position in the ETF market.

The Amex had been losing market share in almost everything it traded, including stocks, options, and ETFs, a product it had started in 1993. Its business of listing new companies peaked decades ago. For example, on a trading day in early 2008 when both the NYSE and Nasdaq traded between 40% and 50% of the shares in their own listed stocks, the Amex traded less than 5% of the shares it listed (Lucchetti, 2008a).

Many of the key dates referred to previously are summarized in the appendix to this chapter.

EVOLVING STOCK MARKET PRACTICES

In this section we describe evolving stock market practices, including:

- Order-handling rules
- Smart order routers
- SEC Regulation NMS
- Internalization
- Alternative display facility
- Trade-reporting facility
- Direct market access
- Algorithmic trading

Order-Handling Rules

During the 1990s, the SEC continued its emphasis on greater quote and price transparency. In this regard, the SEC instituted new order-handling rules in 1997. First, any market maker who held a customer order had to display that order in their quote. Second, a market maker could place a more aggressive quote in an ECN, if the ECN displayed the top of its book in the Nasdaq quote montage. Third, if the ECN's own best quote was not shown in the quote montage, then the market maker had to update its own quote in Nasdaq to match the ECN quote.

While these rules may seem narrow and technical, their effect was significant. They were the basis for ECNs to become major participants in the stock market. Before these rules, Instinet was the only ECN. By 1999, there were nine ECNs.

All a new ECN needed to capture order flow was to be a gateway that attracted some customers to place limit orders on its electronic book. Connectivity with other markets (either directly or through one of Nasdaq's systems) would allow market orders from the customers of other firms to reach its books and become traders. ECNs did not have an impact on NYSE trades until Rule 390 was eliminated in 1999, as discussed earlier.

The Archipelago ECN began in December 1996. Archipelago stated on its web site:

> In January 1997, the U.S. Securities and Exchange Commission (SEC) implemented new Order Handling Rules that revolutionized trading in Nasdaq securities. The new rules created the opportunity for Electronic Communications Networks (ECNs), such as the Archipelago ECN, to interact directly with the Nasdaq National Market System. The Archipelago ECN was formed in December 1996 in response to these rules.

The SEC-enforced consolidation, transparency, and accessibility of price information caused by these SEC changes quickly caused the flow of limit orders to fragment onto multiple books and the ECNs' cheap, fast, and anonymous trading forced the Nasdaq to alter its trading systems and organizational structure.

Then, with the passage of time, consolidation started taking place among the ECNs: Instinet acquired Island, and Archipelago acquired the Pacific Stock Exchange (Archipelago and Instinet/Island accounted for most of the ECN volume). In May 2004, the Nasdaq acquired the Brut ECN, previously owned by Sungard Data Systems.

Smart Order Routers

In concept, it might be expected that most of the trading volume of stocks listed on an exchange would be traded directly with the exchange. This has not, however, been the case. The construct of the Nasdaq as a dealer system has made it easy for ECNs and others to conduct the trades directly and report them to the exchange. And since NYSE Rule 390 was removed in 1999, it has been easier to trade NYSE listed stocks off the exchange. This trading of stock listed on an exchange off the exchange is called *fragmentation*. While the fragmentation of the NYSE has increased, the Nasdaq remains much more fragmented than the NSYE.

One of the many results of this fragmentation has been the need by customers for some new systems to provide order management, handling, and routing services. These services select a market often on the basis of recent trading activity, resulting in its user's receiving the best executions across the markets by "consolidating" the information across the markets.

One outcome of these smart order routers (also called consolidators) is that customer order flow is "less sticky"; that is, the order flow will switch from one execution service to another quickly based on short-term, quantitative information.

SEC Regulation NMS

The Securities Act amendments of 1975 mandated a U.S. national market system (NMS). The core of this national market was the Intermarket Trading System (ITS), which began operating in 1978. The ITS electronically linked eight markets (NYSE, Amex, Boston, Cincinnati, Chicago, Pacific, and Philadelphia Stock Exchanges, and the NASD OTC market) via ITS computers. The ITS permitted traders at any of these exchanges to go to the best available price at the other exchanges on which the security could trade. The NMS also included a consolidated electronic tape, as discussed elsewhere, which combined the last sale prices from all the markets onto a single continuous tape. The use of the ITS, however, was voluntary.

Even though ITS evolved, by 2007 it was based on obsolete technology. During 2007, the ITS system was replaced by the new NMS. Regulation NMS was designed by the SEC for electronic exchanges. Regulation NMS requires that orders be executed at the market (exchanges or other execution venues) that offers the best price for the customers. Thus, exchanges must compete with each other on a level playing field. Regulation NMS's impact is attributable to two of its component rules.

The first rule is the Order Protection Rule (Trade-Through Rule). This rule requires that trades be executed at the best displayed prices provided by an electronic trading system and accessible in under one second. This means that markets will have to route out their orders to other markets if the other markets have better prices (bids or offers). That is, a market cannot "trade through" a better price from another market and trade on its own market. As a result of NMS, each exchange has to send its orders to other exchanges if the other exchange has a better price.

The Trade-Through Rule provides price protection to top-of-book orders (best bids or best offers) placed on exchanges that are electronically accessible. Reserve and hidden orders are not protected. Only electronic quotes are protected. All exchanges are required to have capabilities to route orders to the market with the best bid or offer if they are not able to match the price to execute an order on their own exchange. Only the BBO (or top-of-book) is displayed. In this environment, competitive pricing and low-latency systems are essential in attracting order flow.

The second rule is the Access Rule, which requires the use of private linkages among exchanges to facilitate access to quotes and sets a limit on the access fees by the markets. These private linkages replaced the ITS.

Exchanges had to go to electronic trading to offer protection under Regulation NMS. ECNs and broker-dealers are also covered by Regulation NMS.

A common view is that Regulation NMS will fundamentally change U.S. stock trading by creating a virtual centralized market in which all exchanges will be automated and interconnected. The NMS is expected to divert orders from the two major exchanges, NYSE Euronext and Nasdaq, to the regional exchanges and ECNs. Some regional exchanges view Regulation NMS as a great equalizer. The reason for this position is the goal that all investors get the best price when transacting stocks, regardless of which exchange posts these prices.

Internalization

Internalization refers to off the exchange ("upstairs") trades, mainly of retail trades. As opposed to block trading, internalization involves keeping retail orders within the firm ("internalized") with the broker-dealer buying from its sell orders and selling from its buy orders, generally at the published best bid/offer or a penny better. This practice results in proprietary trading revenue for the broker-dealer. A broker-dealer with a large number of customer orders thereby has a trading opportunity to make a "dealer spread" (buying at the bid and selling at the offer) without interference and will lay off any unwanted positions in the primary market. Brokerage firms internalize through proprietary ATSs. Some believe that this practice reduces transparency, impairs price discovery, and harms investors.

The equity markets permit broker-dealers to internalize retail order flow upstairs. The trades are reported on the trade reporting facility (TRF), as discussed later. While the SEC has approved such internalization in the stock markets, in the stock options markets the orders must be shown to the public market before internalizing them upstairs. Internalization is thus treated differently in the stock and stock options markets.

Off-exchange (internalized) trades as of the end of 2007 account for about 30% of volume in Nasdaq's listed stocks and 16% of the consolidated volume in listed NYSE stocks. Brokers that internalize are reporting 500,000 trades a day in the Nasdaq world and about 350,000 trades a day in NYSE-listed stocks. This volume comes from broker-dealers interacting with their customers' order flow in their upstairs environment and then printing those trades on Nasdaq, as discussed next (Schmerken, 2007).

The NYSE is opposed to internalization. It believes that orders are best represented when they interact with the broader marketplace.

Alternative Display Facility

An *alternative display facility (ADF)* is an entity independent of a registered securities exchange that collects and disseminates securities quotes and trades. It is a display-only facility.

The ADF is an alternative to exchanges for publishing quotations and for comparing and reporting trades. This differs from a trading facility with execution capabilities (a stock exchange) in that the exchange would simply send back to the owner of the displayed order a notice of execution. The NASD has operated an ADF since July 2002. It is now operated by FINRA, as discussed later.

The ADF provides members with a facility for the display of quotations, the reporting of trades, and the comparison of trades. As of March 2007, Consolidated Tape Association (CTA)-listed securities (NYSE, Amex, and the regional exchanges), as well as Nasdaq-listed securities, are eligible for posting quotations through the ADF. ADF best bid and offer and trade reports are included in the consolidated data stream for CTA and Nasdaq-listed securities.

The ADF competes with Nasdaq's SuperMontage system.

These organizations exist to capture some of the values of this information (which has historically been captured by exchanges) for the ADF's information suppliers, usually ECNs.

Trade-Reporting Facility

As indicated above, in the stock market trades can be internalized or arranged upstairs, that is, not traded on an exchange. But if they are traded via internalization they are not reported on the exchange on which the stock is listed. Traditionally, NASD has had a trade-reporting facility (TRF) where these off-exchange trades are printed on NASD's ADF. So, for example, if Merrill Lynch internalized a trade, it would be printed on NASDs TRF through its ADF.

Until recently, broker-dealers that internalized trades—as well as crossing networks and ECNs that matched trades among their subscribers—had to report these trades to their regulator, the NASD via the Nasdaq's Automated Confirmation Transaction System, known as ACT.

The SEC issued an order on June 30, 2006 approving Nasdaq to begin operating as an exchange in Nasdaq-listed securities. The order included approval of the TRF, a new limited liability company operated by Nasdaq and subject to NASD's oversight. In the wake of Nasdaq's receiving SEC approval to become a national stock exchange, NASD and Nasdaq separated, and the SEC allowed Nasdaq to keep its operated TRF, but also opened this up to competition, thereby ending Nasdaq's and NASD's monopoly. Nasdaq's TRF went live August 1, 2006, the same day that it began operating as an exchange in Nasdaq-listed stocks. With the consolidation of the regulatory functions of NASD and NYSE, FINRA has taken over the reporting function. However, FINRA keeps the tape revenue. For this reason, exchanges can set up their own TRFs, report these trades, and keep some of the tape revenues.

The National Exchange (NSX) (the former Cincinnati Stock Exchange) has set up a TRF, so currently such trades can be reported through this exchange as well as the Nasdaq. The NYSE is also considering such an arrangement; if this happens, these trades would be reported by NYSE's TRF as ADF.NYSE.

The ADF NYSE has entered into discussions with the NASD to create a TRF, serving customers reporting off-exchange trades in all listed NMS stocks. If this materializes, stock brokers that internalize trades in NYSE-listed stocks would be able to report those trades to a jointly operated NYSE/NASD TRF. The TRF for NYSE-listed stocks began in 2007.

While the NYSE remains opposed to the internalization of trades, the exchange is developing a TRF for competitive reasons. The NYSE believes that retail orders for individual investors are best represented when they interact with the broader marketplace (Schmerken, 2007).

Regional exchanges, including the Philadelphia Stock Exchange (PHLX), the Boston Equity Exchange (BeX), and the Chicago Stock Exchange (CHX), are also moving forward with plans to create TRFs.

Direct Market Access

In general, buy-side firms continued to take more control over the way their transactions were executed. *Direct market access (DMA)* refers to the use of electronic systems to access various liquidity pools and execution venues directly, without the intervention of a sell-side firm trading desk or broker. There are several advantages of DMA to a buy-side firm:

- It is faster, allowing traders to benefit from short-term market opportunities.
- It has lower transactions costs.
- It provides anonymous transactions.
- It is not handled by brokers, so there is less chance for error.

With respect to cost, it is estimated that DMA commissions are about one cent per share; program trades are two cents per share; and block trades cost four to five cents per share. As of 2004, 33% of buy-side equity shares were transmitted by DMA, and it is estimated that by 2008 this share will be 38%. Hedge funds are aggressive users of DMA (Schmerken, 2005).

Initially, the providers of DMA electronic services were independent firms. But increasingly, traditional sell-side firms have either acquired the independent firms or developed their own DMA systems to provide DMA services to sell-side firms. Among the major providers of this type are Goldman Sachs, Morgan Stanley, CSFB, Citigroup (which acquired Lava), and Bank of New York.

Initially, DMA was used only for U.S. equities, but this focus has expanded to U.S. fixed income and derivatives and into the international markets, Europe, and Asia/Pacific.

DMA has become commoditized and its providers now often provide a comprehensive set of services including program trading, block trading, and also the more sophisticated technique, algorithmic trading, which is discussed in the next section.

Algorithmic Trading

Traditionally, orders for stock executions have been conducted by traders who execute the trades on a trading desk for a portfolio manager or whomever determines what trades should be executed. Traders are judged to have "market information and savvy," which permits them to conduct the trades themselves at a lower cost and with less market impact than the portfolio manager conducting the trades on a less formal basis.

The effectiveness of these traders is often measured by execution evaluation services, and traders are often compensated partially on the basis of their effectiveness. But some observers believe that the traders, in the interest of maximizing their compensation, may have different incentives than the portfolio managers and do not optimize the portfolio manager's objectives—this is referred to as an agency effect. In addition, some think that trades could be conducted more efficiently by electronic systems than by human traders.

As a result, due to improved technology and quantitative techniques, and also regulatory changes, electronic trading systems have been developed to supplement or replace human traders and their trading desks. Such trading is called *algorithmic trading*, or "algo."

Algorithmic trading is a relatively recent type of trading technique whereby an overall trade (either buy or sell) is conducted electronically in a series of small transactions instead of one large transaction. Such trades are conducted via computers, which make the decision to trade or not trade depending on whether recent price movements indicate whether the market will be receptive to the intended trade at the moment or, on the other hand, will cause the price to move significantly against the intended price. Algorithmic trading also permits the traders to hide their intentions. Trading may involve small trades on a continuous basis rather than a large trade at a point in time. The algo is often said to leave no "footprint" and is a "soft touch" way of trading.

Algos, like dark pools, provide anonymity, which the "visible markets," like exchanges and ECNs, do not. Algos are often described as "hiding in plain view."

The advent and wide use of algos is due primarily to both technology and regulation. The technology element is based on faster and cheaper technological systems to execute via improved quantitative methods. The regulatory element is the adoption of pennies and the approval of the order-handling rules, which provided for the growth of ECNs by the SEC. The adoption of "pennies," which provides for smaller pricing increments, and technological advancements, which provide for low-latency trading, have made algorithmic trading more necessary and feasible (low latency refers to a short period of time to execute an instruction, that is, high speed).

One important result of pennies in conjunction with algorithmic trading has been that the average trade size has been decreased significantly, again increasing the requirements for reporting and systems. There has been a significant reduction of the average trade size at the NYSE, beginning with 2001. Order size on the NYSE has declined significantly from over 2,000 shares in 1998 to slightly over 330 in 2007.

The use of algorithmic trading is significant by large traders such as hedge funds and mutual funds. Some traders maintain their own algorithmic trading facilities and others use the systems provided by another organization. Overall, algorithmic trading has the advantages of being scalable, anonymous, transparent, and very fast.

BASIC FUNCTIONING OF STOCK MARKETS

In this section we describe the basic functioning of stock markets which includes

- Price reporting
- Regulation
- Clearance and settlement
- Tick size
- Short-selling rules
- Block trades
- Commissions

Price Reporting

Price reporting in the U.S. stock markets is conducted by the Consolidated Tape Association (CTA). The CTA oversees the dissemination of real-time trade and quote information (market data) from the NYSE and Amex-listed securities (stocks and bonds). The CTA is an independent, industry-wide organization. The CTA manages two systems to govern the collection, processing, and dissemination of trade and quote data. The two systems are: the Consolidated Tape System (CTS), which governs trades, and the Consolidate Quotation System (CQS), which governs quotes. Since the late 1970, all SEC-registered exchanges and market centers that trade NYSE or Amex-listed securities send their trades and quotes to a central consolidator where the CTS and CQS data streams are produced and distributed worldwide.

The data collected by the CTA are provided on two networks, Network A (or Tape A) for NYSE-listed securities which is administered by the NYSE and Network B (or Tape B) for Amex and regional exchange-listed securities which is administered by the Amex. Nasdaq operates a similar tape for its listed securities, which is called Network C (or Tape C).

CTS is the electronic service that provides last sale and trade data for issues listed on the NYSE, Amex, and U.S. regional stock exchanges and was introduced in April 1976. CTS is the basis for the trade reports from the consolidated tape that run across television screens on financial news programs or on the Internet sites. The "consolidated tape" is a high-speed, electronic system that constantly reports the latest price and volume data on sales of exchange-listed stocks.

CQS is the electronic service that provides quotation information for issues listed on the NYSE, Amex, and U.S. regional stock exchanges.

For every quote message received from a market center, CQS calculates an NBBO based on a price, size, and time priority schema. If the quote is a Nasdaq market-maker quote, CQS also calculates a Nasdaq BBO. CQS disseminates the market center's root quote with an appendage that includes the National and Nasdaq BBOs.

In general, Tapes A and B are referred to as the "CTS Tapes" and Tape C as the "Nasdaq Tapes."

Regulation

The basis for the federal government regulation of the stock market resides with the SEC. The SEC's authority is primarily based on two important pieces of federal legislation. The first is the Securities Act of 1933 (the "Securities Act"), which covers the primary markets, that is, the new issues of securities. The second is the Securities Act of 1934 (the "Exchange Act"), which covers the secondary markets. The SEC was created by the Exchange Act.

In addition to the SEC regulations, the exchanges also play a role in their own regulation through self-regulating organizations (SROs). The SRO of the NYSE has been responsible for the member regulation, enforcement, and arbitration functions of the NYSE. In addition, the NASD has had the SEC authority to set standards for its member firms and standards of conduct for issuing securities and selling securities to the public. The NASD has also monitored the Nasdaq stock market. There have, however, been some overlapping responsibilities of these two SROs and, thus, some competition between them.

As a result, these two SROs merged and in July 2007 were replaced by the single organization, FINRA, which consolidated the NASD and the member regulation, enforcement, and arbitration functions of NYSE. This consolidation resulted in all firms dealing with only one rule book, one set of examiners, and one enforcement staff, thereby reducing costs and inconsistencies. Thus, FINRA is the single remaining SRO.

Clearance and Settlement

After a stock trade is completed, the delivery of the shares by the seller and the payment of cash by the buyer must occur quickly and efficiently. The efficiency of the trade settlement affects the total speed and the overall cost of the transaction. In the United States, there are several execution mechanisms (exchanges and other) for stocks. There is, however, only a single clearance and settlement mechanism for securities, the Depository Trust and Clearing Corporation (DTTC). In stock options, similarly, there is only one clearing mechanism, the Options Clearing Corporation (OCC).

In the futures markets, however, there are several clearing organizations, each typically associated with the related exchange. This control over clearing by futures exchanges makes it easier for them to preserve their monopolies in trading and also gives them a significant source of profitability. During February 2008, however, the U.S. Department of Justice questioned whether futures exchanges should be allowed to own or control clearing businesses that process their trades. The issue is whether futures clearing will become centralized like stock and options clearing (Lucchetti, 2008c).

All clearance and settlement services for U.S. equities market (as well as corporate bonds, municipal bonds, exchange-traded funds, and unit investment trusts trades) are provided by the National Securities Clearance Corporation (NSCC). NSCC is a wholly owned subsidiary of the DTCC. NSCC generally clears and settles trades on a T +3 basis (that is, three business days after the trade date). DTCC is essentially a utility organization for the exchanges.

Other subsidiaries of DTCC provide clearance and settlement for other products and also trust services.

Tick Size

The minimum price variation for a security is referred to as its *tick size.* The U.S. stock market historically had a tick size of one-eighth of 1 point. The SEC wanted to reduce the bid-offer spread to increase the competition and lower costs. As a result, the NYSE and Nasdaq reduced the tick size first to one-sixteenth and then in 2001 to pennies (1 cent).

This reduction in the tick size narrowed the bid-offer spread considerably, which reduced the costs to customers and the profits of the market makers. In addition, as an unintended consequence, it negatively affected the liquidity of the market. With pennies, there are 100 pricing points per dollar, while with eighths there are only eight. So with pennies, there is less liquidity at each pricing point and therefore less depth at the inside market (the best bid and best offer or the "top of the book"). Since only the top of the book is displayed, the advent of pennies reduced transparency and was one of the reasons for the development of dark pools, as discussed above.

In addition, with pennies, quotes (bids or offers) are changed more frequently, and so the technology must have lower latency. Low latency, that is, a small amount of time necessary to complete an instruction, means high speed. Latency below one millisecond is now common. As indicated above, the advent of pennies is one reason for the development and growth of algorithmic trading.

Subpenny pricing is prohibited except for stocks that trade for less than a dollar.

Short-Selling Rules

During and after the 1929 stock market crash, part of the blame for the crash was attributed to "short selling." In short selling, an investor borrows stock and sells the borrowed stock to profit from an expected subsequent decline in the price when the investor buys the stock back at a lower price to repay the borrowed stock. In the short run at least, short selling the stock causes a decline in the market. In the longer run, the stock must be bought back to "cover the short" and causes the stock price to increase. As a result of the stock market crash, Rule 10A-1 was adopted a decade after the 1929 stock market crash to prevent short sellers from adding to the downward pressure on a stock whose price is already declining. This short-selling rule permitted short sales only when the last sales price was higher than the previous price (an "uptick trade") or if there was no change in the last sales price but the previous sales price was higher than the sales price that proceeded it (a "zero uptick trade"). Short sales were not permitted on a downtick.

In June 2007, the SEC eliminated the short-selling rule. The SEC stated that this rule was obsolete due to decimalization, changes in trading strategies, and increased market transparency. These strategies included standard short selling and long-short strategies such as 130/30 strategies which are practiced by hedge funds, mutual funds, and other investors. In addition, the rapid growth of ETFs which were not subjected to the short-sale rule, provided a convenient way of shorting portfolios of stocks, although not individual stocks.

Block Trades

A *block* is a large holding or transaction of stock, generally 10,000 or more shares or any amount over $200,000. In a *block trade* (a "block facilitation trade") a broker-dealer commits capital to accommodate a large trade for an institutional customer. These trades are conducted "upstairs" (off the exchange) and "shown to" the market, for potential price improvement. Block trades are reported through the standard price reporting systems. With the growth in algorithmic trading, however, blocks which have traditionally been accomplished "upstairs" via internalization are now accomplished via algorithmic trading.

Commissions

Before 1975, stock exchanges were allowed to set minimum commissions on transactions. The fixed commission structure did not allow the commission rate to decline as the number of shares in the order increased, thereby ignoring the economies of scale in executing transactions.

Pressure from institutional investors, who transacted large trades, led the SEC to eliminate fixed commission rates in 1975. Since May 1, 1975, popularly referred to as Black Thursday, commissions have been fully negotiable between investors and their brokers. Black Thursday began a period of severe price competition among brokers, with many firms failing, and a consolidation of firms taking place in the brokerage industry. This liberalization of commissions was the U.S. version of the "Big Bang," which was subsequently followed in the United Kingdom in 1986 and Japan in 1996.

Since the introduction of negotiated commissions, the opportunity has arisen for the development of discount brokers. These brokers charge commissions at rates much less than those charged by other brokers, but offer little or no advice or any other service apart from the execution of the transaction. Discount brokers have been particularly effective in inducing retail investors to participate in the market for individual stocks.

SUMMARY

During the 1980s and 1990s competition among equity transactions providers has changed considerably and became much more intense. In the 1980s, the competition for equity transactions was mainly among the NYSE, the Amex, the regional exchanges, and Nasdaq. Gradually, the Amex and the regional exchanges lost ground and the competition became mainly between the two largest exchanges, the NYSE versus Nasdaq. This competition was significantly based on very different market structures: the order-driven, specialist, floor-based mechanism of the NYSE on the one hand, and the quote-driven, dealer-based, electronic mechanism of Nasdaq on the other. Debates on the merits of these competing market structures flourished. The advantage of the specialist system is the opportunity for price improvement and the advantage of the electronic system is speed of execution. The common view is that retail investors prefer better prices and professional traders prefer greater speed. This argument could be summarized, in the potential words of an advertising specialist, as "better pricing versus higher speed."

Gradually, however, these two market structures have converged significantly toward hybrid markets. The NYSE continued to improve its SuperDOT system, developed Direct+, and finally initiated its NYSE Hybrid Market wherein orders could be executed either via a specialist or electronically. In addition, the NYSE purchased Archipelago, an electronic stock exchange which the NYSE operates in parallel with its Hybrid Market.

Similarly, Nasdaq developed SuperMontage, an electronic order display and execution system, which allows market participants to enter quotes and orders at multiple prices and display aggregate interest. Nasdaq has also added fixed-time call auctions to open and close the market. These changes altered Nasdaq from a pure quote-driven market to a hybrid market which contains both quote-driven and order-driven features. In June 2006, the SEC approved Nasdaq to begin operating as an exchange in Nasdaq-listed securities.

Although both the NYSE and Nasdaq have converged toward hybrid markets from very different positions, their differences have remained significant. But as they became more similar in market structure and in other ways, the competition from other market venues became more intense. After the SEC approved new order handling rules in 1997, ECNs took larger shares first from Nasdaq and then, after 1999 when NYSE's Rule 390 was eliminated, from the NYSE.

The NYSE and Nasdaq have also grown by acquiring other market venues, initially ECNs (Instinet by Nasdaq and Archipelago by the NYSE). However, the two major exchanges have begun acquiring other exchanges, the Philadelphia Stock Exchange and the Boston Stock Exchange by Nasdaq, and the Pacific Stock Exchange (through Archipelago) and Amex by the NYSE. These acquisitions reduced the number of regional exchanges. New stock exchanges, however, have been developed by the previous options-only exchanges, ISE and CBOE.

In addition, both the NYSE and Nasdaq have developed global relationships, Euronext and Borse Dubai for the NYSE and Nasdaq, respectively. And the ISE has been acquired by the large German stock exchange, Deutsche Borse, and its derivatives subsidiary, Eurex.

Meanwhile, some other forces have made it very difficult for either exchange or the ECNs to compete profitably for equity executions. Institutions, intermediaries, and, increasingly, hedge funds have become demanding with respect to three characteristics of execution services. The first is cost. Transactions costs in a very competitive, high-turnover environment have become increasingly important. The second is speed. The volatility of the markets and the competition among market users have made "latency" a common concern. The third is anonymity. Hedge funds pursuing proprietary strategies and mutual funds whose buy and sell programs may endure over a long period require anonymous transactions.

As a result, alternative trading systems (ATSs) have developed to satisfy these institutional market participant needs. Thus, while ECNs have been increasingly successful at competing with exchanges directly, crossing networks, dark pools, and internalization have competed with exchanges in very different ways based on their costs, speed, and anonymity. These mechanisms, however, fragment the central markets and depend on the central markets for price discovery. Thus, there are limits to their potential total activity or size.

Aggressive clients often consider the execution process in three levels. First, they try to cross orders internally where there would be no execution costs. If they do not have an internal match, they will then send their orders to a crossing network which does not display quotes to avoid adverse price reactions. Finally, they send their orders to open markets or exchange only as a last resort (Clary, 2007).

Some observers describe the orders that go to exchanges as "exhaust." In this regard, the data referred to herein indicate that the exchanges are trading a small share of their own listings.

In addition, Regulation NMS with its Trade-Through Rules will level the playing field among exchanges and ECNs. A common expectation is that this regulation will work to the advantage of the regional exchanges versus the NYSE and Nasdaq. There may, however, be few or no regional exchanges remaining.

But the competition is becoming more complex than simply the different execution venues competing with each other from different structures. Rather, they have begun competing against each other from similar structures. ECNs have and will become exchanges. While ATS were the originators of dark pools and crosses, exchanges have also experimented with them (with the exception of buy side-to-buy side crosses). In addition, the importance of block trades diminishes as algorithmic trading reduces the trade size.

The upshot of these changes is that these execution venues all compete with each other and that the categories that formerly distinguished them are disappearing. The exchanges, ECNs, and ATSs will compete with each other for many of the same customers on the basis of the same functions and fees. Such competition may cause commoditization and a resulting low profitability for the providers of these services. Could equity transaction services follow the path of the U.S. inland water canals and the providers of some Internet services?

There may, however, be a limiting issue in such a development. The visible markets, the exchanges and ECNs, provide price discovery. Some of the other markets, crosses, dark pools, and internalization, do not provide price discovery, but rather have a free ride on the price discovery of the central and visible markets. But if the executions shift predominantly to the nonprice discovery markets, how will prices be determined? This issue is similar to the issue in investments regarding active and passive investing and price discovery. That is, if all or most of the investors use a passive investment approach, there will be no active investors to provide the price discovery of the stocks. Opinions are not unanimous in either arena.

Overall, the nature of competition in U.S. stock executions has changed considerably in the 1980s and 1990s among the various market venues: NYSE and Nasdaq; Amex and the regional exchanges; ECNs; ATSs, and international participants.

The only certainty is that the equilibrium of the U.S. stock transaction environment has not arrived. And the ultimate equilibrium, if there will be one, will be much more complex than the original equilibrium with some major and minor exchanges. Continued change is inevitable given the magnitude of equity executions and the level of expertise involved. The future will be beneficial to the users of these execution venues and offers opportunities for innovations by equity execution providers.

APPENDIX: KEY DATES

Dates	Events
1969	Instinet formed
1971	Nasdaq formed, the world's first electronic stock market; began trading on 2/8/1971
1978	Nasdaq and Amex merged to form the Nasdaq-Amex group, Inc.
1987	Instinet acquired by Reuters
12/9/1996	Archipelago founded
1997	Order-handling rules approved by SEC
1997–2000	Proliferation of ECNs
5/18/2001	Instinet conducts an IPO
2001	Consolidation of ECNs
7/01	LSE IPO (ticker symbol: LSE)
10/20/2001	Archipelago granted exchange status by SEC
2002	Advent of Algorithmic Trading
07/01/2002	Nasdaq IPO
9/20/2002	Island acquired by Instinet
12/06/2002	CME IPO (ticker symbol: CME)
1/2005	Archipelago announced plan to buy the Pacific Stock Exchange
10/19/05	CBOT Holdings IPO (ticker symbol: CBOT)
12/2005	Nasdaq buys Instinet Group (and its INET trading platform)
12/15/2005	NYSE launches its Hybrid Market, creating a blend of auction and electronic trading.
12/30/2005	The NYSE, in anticipation of its transformation into a publicly held company, ends member seat sales which were replaced by the sale of annual trading license.
3/7/2006	NYSE buys electronic bourse operator Archipelago Holding (a for-profit, publicly owned company) NYSE Group Inc. formed out of merger of the New York Stock Exchange and Archipelago holdings, Inc.
3/8/2006	NYSE Group IPO (ticker symbol: NYX)
4/4/2007	NYSE Group completed merger with Euronext NV, creating the first trans-Atlantic exchange group, called NYSE Euronext
7/2007	CME merged with CBOT
7/2007	Regulation NMS implemented

ACKNOWLEDGMENTS

Discussions with Andrew Brenner and Steve Wunsch of the International Securities Exchange were constructive in developing this chapter.

REFERENCES

Anderson, A. M., and Dyl, E. A. (2007). Trading volume: NASDAQ and the NYSE. *Financial Analysts Journal* 63, 6: 79–86.

Byrne, J. A. (2007). Hooked on speed. *Alpha Magazine*, January: 53–56.

Clary, I. (2007). Wall Street strikes back. *Pensions & Investments*, October 1: 41.

Hendershott, T., and Moulton, P. C. (2007). The shrinking New York Stock Exchange floor and the Hybrid Market. Working paper.

Leinweber, D. (2007). Algo vs algo. *Alpha Magazine*, February: 45–51.

Lucchetti, A. (2007). "Niederauer's first challenge: NYSE floor traders' future. *Wall Street Journal*, November 21: C1.

Lucchetti, A. (2008a). NYSE Euronext is in talks to buy Amex as its old rivals market share erodes. *Wall Street Journal*, January 10: 2008: C3.

Lucchetti, A. (2008b). NYSE, Amex finally marrying. *Wall Street Journal*, January 18: C3.

Lucchetti, A. (2008c). CME fires back on clearing proposal. *Wall Street Journal*, February 7: C3.

McGee, S. (2006). Revenge of the regionals. *Institutional Investor*, July: 41–48.

Moulton, P. (2006). Who trades with whom, and when? Working paper.

Safarik, D. (2005). Direct market access: The next frontier. *Wall Street & Technology*, February 28.

Schmerken, I. (2001). Exchange consolidation war is expected to continue in 2006. *Advanced Trading*, October 31.

Schmerken, I. (2005). Direct-market-access trading. *Wall Street & Technology*, February 4.

Schmerken I. (2006). NYSE explores point trade reporting facility with NASD. *Wall Street & Technology*, October 17.

Schmerken I. (2007). Stock exchange create trade reporting facilities to earn market data fees from internalized trades. *Wall Street and Technology*, January 5.

Schwartz, R. A., and Francioni, R. (2004). *Equity Markets in Action*. Hoboken, NJ: John Wiley & Sons.

CHAPTER 12

The Information Content of Short Sales

STEVEN L. JONES, PhD
Associate Professor of Finance, Indiana University, Kelley School of Business–Indianapolis

GLEN LARSEN, Jr., PhD, CFA
Professor of Finance, Indiana University, Kelley School of Business–Indianapolis

Abstract: Short interest in a stock is the aggregate number of shares that have been sold short and not yet covered. There has been a long-running debate over whether short interest contains valuable information about a stock's future performance. Weak-form market efficiency suggests that competitive trading should erode any information content in the signal. However, Wall Street analysts have traditionally viewed high short interest as a bullish technical indicator since covering short positions creates upward price pressure and recalls and short squeezes may force premature coverage of short positions. Alternatively, academic studies find that short-sale constraints clearly result in overpricing and that high short interest predicts negative future returns, consistent with the theories developed by Edward Miller in 1977 and Douglas Diamond and Robert Verrecchia in 1987, respectively. Miller's theory of divergent opinions predicts that short-sale constraints lead to overpricing, while Diamond and Verrecchia's theory predict that an unexpected increase in short interest is bad news since it indicates a higher proportion of past sales, than previously realized, came from the presumably more-informed short sellers. Thus, the analysts' traditional bullish view relies on a reversion in prices back, *up*, to the mean, while the academics' bearish view is that short sellers' profits come from taking advantage of the reversion of prices back down to the mean.

Keywords: short interest, relative short interest, determinants of short interest, costs of short selling, short-sale constraints, short squeeze, recall, rebate rate, overvaluation, market efficiency

Transactions data on short sales are not publicly available in the United States; however, the New York Stock Exchange (NYSE), American Exchange (AMEX) and Nasdaq report *short interest* figures for individual stocks on a monthly basis. The short interest in a stock is the aggregate number of shares that have been sold short and not yet covered. Whether these short interest figures contain valuable information about future performance has been a long-running controversy. Wall Street technicians, on the one hand, have traditionally viewed high short interest as a bullish technical indicator. On the other hand, most academic studies find that high short interest predicts negative future returns and therefore signals bearish sentiment.

At first glance, it may seem surprising to suggest that short interest can reliably predict anything about future performance since competitive trading should erode the information content of a technical indicator. Trading is what impounds information into prices and competitive trading should result in an "efficient capital market." Even in a weak-form efficient market, by Fama's definition, it is not possible to reliably predict future performance from technical indicators or trading rules that are based on public market data. Fama recognizes, however, that traders will impound information into prices only to within the cost of attaining and trading on the information. It follows that the high cost of short selling, relative to regular sell or buy orders, constrains the trading necessary to fully impound bad news into security prices, and as a result, some academic studies hypothesize that overpricing may exist in stocks that are costly to short sell.

Academic studies also suggest that the high costs of short selling imply that those who are willing to short sell, despite these costs, are likely to be trading based on superior information, in which case increases in short interest may signal that informed traders have become more bearish about a stock; hence, the price should drop. On the other hand, the technician's view of short interest, as a bullish indicator, is based on the idea that short interest represents latent demand because short positions must eventually be covered by repurchasing the stock; thus, the price should increase in the future. Implicit in the technician's view is the risk of a so-called "short squeeze," in which prices move up very quickly as short sellers are forced to cover.

In this chapter, we analyze the theory and evidence on the information content of short interest in individual stocks. The very limited evidence on short-sale transactions is also considered. We start with brief explanations of how short interest is reported and the constraints on short selling. We then consider the theoretical academic work on short-sale constraints and contrast its predictions for short interest to the traditional technical analyst's view of short interest. Most of the remainder of the chapter synthesizes the empirical evidence. This begins with a review of the early work on predicting short-term returns with short interest and proceeds to the motives for short selling, as well as the use of options and their implications for the information content of short interest. We also investigate whether long-term returns are predictable from short interest and identify the determinants of short interest. Then the costs of short selling are considered as limits to arbitrage. Finally, we conclude and offer some implications for investors.

SHORT SALES: REPORTING, FREQUENCY, AND CONSTRAINTS

The Securities and Exchange Commission (SEC) requires that a short-sale order be marked as such, while a regular sell order, in which the person placing the order owns the shares, is marked as long sale. The NYSE, AMEX, and Nasdaq compile the short interest in individual stocks from member firms' reports as of settlement on the 15th of each month, or the prior business day if the 15th is not a business day. The NYSE and AMEX release the data within four business days, while the Nasdaq takes eight business days.

A popular indicator for the intensity of short interest is the *short interest* ratio (SIR). This is the aggregate short interest in a stock as a percent of its average daily trading volume over some preceding period, usually four weeks. The denominator is sometimes modified to account for seasonality in volume, or measured over longer intervals to smooth out the effects of unanticipated changes in trading activity. In addition, many academic studies focus on the relative SIR (RSI), the aggregate short interest in a stock as a percent of the firm's total shares outstanding.

Although short selling is fairly common, most stocks have relatively little short interest. Arnold, Butler, Crack, and Zhang (2005) report that about 5,000 Nasdaq and about 3,000 NYSE stocks had short interest at sometime between 1995 and 1999, but the RSI was less than 0.5% for the typical stock, and 3% to 4% was average for the quintile of stocks with the highest RSI.

Constraints on short sales include: (1) the direct monetary costs of borrowing shares, (2) the difficulty (or impossibility) of establishing a short position, (3) the risk that the short position cannot be maintained, and (4) the legal and institutional restrictions on short selling. Items 1, 2, and 3 are normally referred to as the costs of short selling. While the nouns, *constraint* and *restriction* have subtly different meanings in this context, we will use their verb forms interchangeably. The most widely known constraints are the "up-tick" and "zero-plus-tick" rules, which prohibit short selling in a stock except at a price higher than the price of the last trade, or at a price equal to that of the last trade if the previous price change was positive. As of November 2003, the SEC was considering a proposal to no longer apply the up-tick and zero-plus-tick rules to widely traded stocks. The motivation is to reduce the incentive to use put options.

While these rules restrict short selling in the near term, there are several other constraints that make short selling much more costly, or may prevent it all together. For example, short sellers must: (1) maintain a margin requirement of 50% (per the Federal Reserve Board's Regulation T), (2) locate the shares to borrow, (3) leave the proceeds of the sale as collateral with the lender of the borrowed shares, and (4) pay the amount of any dividend to the lender and possibly interest (that is, incur a negative rebate rate) if the borrowed shares are in high demand. The borrowed shares are usually located with the assistance of a broker, but this may be difficult if the shares are in high demand. In addition, to have any reasonable expectation of success, short sellers must be able to maintain the position (that is, avoid having the shares recalled by the lender) long enough to give their contrary view a chance of being realized in the market price. Finally, many institutions are restricted from short selling all together.

ACADEMIC THEORY VERSUS THE TECHNICAL ANALYST'S VIEW

Edward Miller was one of the first to recognize the implications of costly short-sale constraints for capital market efficiency. Miller argues that stocks with a wide divergence of opinion, as to intrinsic value, are likely to become overpriced if the more optimistic investors can absorb the shares and short sales are constrained such that the less optimistic investors cannot fully participate in setting the price. We refer to this as Miller's overpricing hypothesis. Miller does not, however, offer suggestions for how one might take advantage of this potential overpricing. Should one short the stocks that are already under the most intense pressure from short sellers, or might high short interest indicate that the price has already bottomed out?

Diamond and Verrecchia (1987) assume that investors glean information from trading activity with the knowledge that short-selling is costly. In other words, investors form expectations rationally (as efficient markets theory assumes). For example, higher costs prevent short sellers from trading as frequently on private information; thus, a prolonged period of trading inactivity portends that the next trade is likely to reflect bad news, rather than good news. The overpricing predicted by Miller cannot survive the assumption of rational expectations; however, two relevant pricing effects still result from short-sale constraints. First, for stocks under heavy short-selling pressure, the distribution of returns is skewed heavily to the left (that is, toward negative returns), such that incremental price changes are likely to be larger on the down side. Rational market makers will respond to this by widening their bid-ask spreads. Second, the reduction in informed trading lowers the speed of price adjustment, especially to bad news.

Diamond and Verrecchia recognize that the high costs may drive out uninformed liquidity-based short sellers, and they consider whether this might actually improve informational efficiency, as an unintended consequence. They dismiss this, however, as highly unlikely on the grounds that few short sales are motivated by liquidity. Regardless, as long as the high costs of short selling are more likely to prevent uninformed trades, as opposed to driving out informed traders, the resulting pool of short sales will reflect proportionally more informed trades than the combined pool of all short sell and long sell orders, in which case, their model predicts that an unexpected increase in short interest is bad news since it indicates a higher proportion of past sales, than previously realized, came from short sellers, who should be more informed than long sellers.

It is worth emphasizing that Diamond and Verrecchia do not require the systematic overpricing of Miller to generate information content from unexpected changes in short interest. In fact, their argument relies on unexpected changes, not absolute short interest; thus, it is consistent with weak-form market efficiency. They do, however, predict slower price adjustment to bad news, and this suggests that the opportunity to profit from unexpected changes in short interest (or any other signal of bad news) may persist for longer than we might otherwise expect. In an efficient capital market, stock prices fully reflect available information in equilibrium. Once information is released, prices adjust to new equilibrium levels. As the market searches for a new equilibrium, it is said to be in "disequilibrium." The faster is this adjustment process, the greater the informational efficiency of the market. Hence, Diamond and Verrecchia imply that short-sale constraints reduce the general informational efficiency of the market; however, the weak-form version of market efficiency is not violated because it is a description of prices in equilibrium.

Diamond and Verrecchia also consider the traditional technical analysts' view that increased short interest in a stock foreshadows positive returns due to latent buying pressure from short sellers as they cover. They dismiss this view, however, on the grounds that it necessitates relatively uninformed short sellers. Technical analysts, however, do not think so highly of short sellers.

The traditional technical analysts' view is that relatively high short interest indicates a buy signal. This view is based largely on two points: (1) that short sales represent latent future demand to cover and (2) the proposition that high short interest results from speculative excess in the form of increased short selling into lengthy price declines that tend to eventually reverse. The first point reflects not only the fact that all short positions must eventually be covered, but also the risk that a short seller may be forced to cover early. This can happen when a short seller's broker recalls the borrowed shares at the request of the lender, with no other shares available to lend, or if the price of the shorted asset increases until the short seller receives a margin call.

The risk of being forced to cover may be at its highest during a so-called *short squeeze*, where one or more buyers intentionally drive the price of an asset up until the shorts are forced to cover at a loss. Hence, high short interest can attract buyers and make a short position extremely risky. The second point, that short selling tends to increase after sustained price declines, reflects the possibility of short sellers creating downward price pressure in which the last short sellers are more likely to be the least informed, especially if short interest was high to begin with. Thus, the price may have been pushed too low and a rebound is inevitable. This, of course, is simply the analogue of the view that the least informed investors usually wait and jump on the bandwagon just before the market peaks.

It is apparent that the traditional technical analysts' view of short interest is not nearly so naïve as Diamond and Verrecchia suggest. In fact, although less impressive in terms of formal rigor, one could argue that its logic is at least as compelling. It does ignore the higher costs of short selling that are the key in Diamond and Verrecchia, but then they fail to recognize that a short seller's information may depend on whether he or she short sold early on or late, as short interest was accumulating.

THE EMPIRICAL EVIDENCE

In this section, we synthesize the evidence on the information content of short interest in individual stocks and relate it back to the theories. The theories serve as a useful framework for following the progression of the investigation and for understanding why at least some information content appears to survive. We start with a review of the early work on predicting short-term returns and proceed to the motives for short selling as well as the use of options and their implications for the information content in short interest. We also investigate whether long-term returns are predictable from short interest and identify the determinants of short interest. Then the costs of short selling are considered as possible limits to arbitrage. Finally, we consider the information content of transaction level short-sales data and whether it should be made publicly available on a timely basis.

Predicting Short-Term Returns with Short Interest: The Early Evidence

Hurtado-Sanchez (1978) set out to test the technical analyst's traditional view of high short interest as a bullish indicator, but his results apply to the academic models as well. He wondered if the inclusion of hedging and arbitrage-motivated trades in short-interest data obscures the information content of speculative short sales. Rather than test directly for the prevalence of these trades, he considers whether short interest predicts future returns using a sample of stocks from the Standard & Poor's 425 Industrials of 1966 and 1967. He fails to detect any evidence that levels or changes in absolute short interest, the SIR, or RSI, can predict future performance in individual stocks. He does find, however, that stocks with high (low) return performance experience increases (decreases) in short interest in the following month. His conclusion is that short-interest data contain no information about future returns, but short sales help stabilize the market by adding to selling pressure after prices have risen.

Figlewski (1981) was one of the first to consider the implications of Miller's overpricing hypothesis. Figlewski assumes that observed levels of short interest proxy for the amount of unfavorable information excluded from market prices as a result of the constraints on short sales. In other words, a relatively high level of short interest in a stock indicates that short interest would have been even higher yet, if unconstrained. He also refines Miller's overpricing hypothesis by pointing out that rational investors, with knowledge of the effects of short-sales constraints, would not overprice some stocks without underpricing others. Thus, he hypothesizes that high (low) levels of short interest predict overpricing (underpricing) in individual stocks. Figlewski's appeal to rational expectations is somewhat of a precursor to Diamond and Verrecchia except Figlewski allows for informational inefficiency at the firm level. That is, in his model investors have yet to learn that short interest proxies for the amount of unfavorable information excluded from market prices. Diamond and Verrecchia get around the assumption of firm-level inefficiency by focusing on the information content of unexpected changes in short interest.

He finds mixed support for this hypothesis in a sample of Standard & Poor's 500 Index stocks from the years 1973 to 1979. Specifically, a short position in the stocks ranking the highest on RSI outperforms a long position in the lowest RSI stocks by a statistically significant amount, but only if the short seller captures the interest on the proceeds from the short sale. Of course, most small traders receive no interest on short-sale proceeds, and even large traders must pay a loan free as compensation to the lender. Excluding the interest on proceeds, the mean return to stocks ranking highest in terms of RSI is actually positive in the post-ranking month, although insignificant.

The inability of both Hurtado-Sanchez and Figlewski to detect compelling evidence of return predictability from short interest suggests that hedging and arbitrage-motivated trades may be obscuring any information content in the data. Examples of such trades include the arbitrage of going long convertibles or warrants and short the converting common stock, the arbitrage of mergers (that is, going long the targets stock while shorting the acquirer's stock), and general "pairs trading." Pairs trading is a general term used to describe strategies that involve buying a stock that is thought to be underpriced, for any of a number of reasons, and shorting a statistically paired stock to neutralize risk and possibly to further enhance return. The need to understand the motivations of short sellers took on added importance with the introduction of Diamond and Verrecchia's previously mentioned work. They, of course, indicate that large unexpected increases in short interest predict negative future returns because short sellers are better informed. They also claim that the information content in short interest is obscured for stocks that have traded options.

Predicting Short-Term Returns with and without Hedging and Traded Options

Brent, Morse, and Stice (1990) considered the motivations of short sellers using random samples of 200 NYSE stocks from the years 1981 to 1984. Their tests confirm the results of Hurtado-Sanchez in that changes in RSI fail to predict future returns, but stocks with high returns subsequently experience large increases in RSI. The latter finding is in direct opposition to one of the key assumptions behind the technical analysts' bullish view of short interest: that short selling supposedly increases in down markets. Thus, it appears that short sellers are attempting to anticipate mean reversion in returns. They also observe that stocks with high short interest tend to have high betas, traded options, and listed convertible securities. They therefore conclude that hedging and arbitrage, as opposed to speculation, motivates a material amount of short selling.

Another hedging strategy that may obscure information in short interest is "shorting against the box" (that is, selling short a stock already held long) at the end of the year to delay capital gains to the following year. Using NYSE and Nasdaq short interest data from 1995 to 1999, Arnold, Butler, Crack, and Zhang demonstrate the popularity of this strategy prior to the Tax Payer Relief Act of 1997. The Act disallowed this practice as a means to delaying taxes, and they find that year-end short interest declined significantly with the introduction of the Act. They also show that the Act had the effect of strengthening the negative relation between changes in a stock's RSI and its return in the following month. This clearly indicates that short interest announcements contain information about subsequent returns, in the manner of Diamond and Verrecchia, as long as information-motivated trades make up an adequate proportion of the short interest.

Senchack and Starks (1993) tested the predictive power of short interest with an event study on a sample 2,419 stocks selected so as to be less susceptible to the problem of obscured information content. They begin with all NYSE and AMEX stocks, whose short interest was published in the *Wall Street Journal* from 1980 through 1986. The sample is then purged of stocks reported to be the subject of arbitrage activities, although this does not account for pairs trading and shorting against the box. They also eliminate all observations in which the reported increase in short interest, from the previous month, is less than 100%. This is done to better reflect the model of Diamond and Verrecchia, which applies only to large, unexpected increases in short interest. Finally, the sample differentiates between stocks that have traded options and those that do not.

Senchack and Starks point out that buying puts and writing calls is a low cost alternative to short selling, and this means that any unfavorable private information about a stock may be observable from option premiums and volumes, well before the short interest announcement. Note that the short interest figures may be relatively unaffected if put writers hedge by selling short. They find that stocks without traded options have a small but statistically significant negative price reaction to the announcement of large percent increases in short interest. The cumulative negative returns over both five- and nine-day event windows are slightly less than one-half of 1%. In addition, the larger the percent increase in short interest, the more negative is the price reaction to the announcement. Stocks with traded options, on the other hand, display no significant reaction to announcements of large percent increases in short interest. These results support both Diamond and Verrecchia's prediction that large, unexpected increases in short interest are bearish signals, as well as the claim that traded options obscure the information content in short interest announcements.

Figlewski and Webb (1993) take a somewhat different approach in their study of the effect of options on short-sale constraints. They recognize that options decrease the costs of effectively going short and suspect that this improves informational efficiency by making constraints on short sales irrelevant. Note that the combination of reduced trading costs and increased informational efficiency should weaken, if not eliminate, the ability of short interest to predict future returns.

Using samples of Standard & Poor's 500 stocks from the 1970s and 1980s, they establish that the options market is actively used as a complement to short selling. Stocks with traded options have significantly higher RSI levels than stocks without traded options, and the introduction of traded options in a stock tends to increase the stock's RSI. They also find that option premiums tend to be higher in puts than in calls for stocks with high levels of RSI. These results suggest that option trading enables more negative information to enter the market, and impact stock prices, than would have otherwise. The impact on stock prices occurs as a result of put writers selling short to hedge, as well as from the arbitrage when the puts become expensive relative to the calls. This arbitrage involves writing the put, buying the call, and shorting the stock.

For stocks with high levels of RSI, Figlewski and Webb find that those without traded options earn negative returns, in the month after the announcement, but these negative returns are not significantly less than the returns to the stocks with traded options. Senchack and Starks, of course, find this difference to be significant, as is expected if options actually improve informational efficiency. The discrepancy is likely due to the cleaner sample used by Senchack and Stark, as well as the concentrated focus of their five- and nine-day event windows. In addition, Senchack and Starks analyze only large percent changes in short interest, while Figlewski and Webb analyze levels of RSI.

A study by Choie, Huang, and James (1994) supports the view that large percent changes in short interest signal more about short-term returns than do high levels or large increases in short interest. They find that a short position in the stocks with the largest percent increases in short interest, as reported by the *Wall Street Journal* in the years 1988 to 1991, earned a mean return of more than 1% in excess of the S&P 500 Index in the month following publication. This is about double the excess return from shorting the stocks with the highest short interest levels or the largest SIRs. In addition, the stocks with the largest simple increases in short interest actually outperformed the S&P 500 Index, on average, in the month following publication. This suggests that percent changes are more difficult to predict and therefore are unexpected in the manner of Diamond and Verrecchia.

Most of the work we have reviewed, up until now, finds that large changes and, to a lesser extent, high levels of short interest predict small negative returns in the month or days after the announcement. However, these returns are statistically significant in only a few cases, and their economic significance is even less certain. Probably the most compelling evidence comes from Senchack and Starks, who focus on large percent increases in short interest and find support for the predictions of Diamond and Verrecchia.

Focusing on the short-term price reaction to large percent increases in short interest is an appropriate test of Diamond and Verrecchia, but it is not clear that any of the above work provides a fair test of Miller's overpricing hypothesis because it results from short-sale constraints. Thus, it will not be eliminated by a short-interest announcement, whether the focus is on short interest levels or changes. The price adjustment process may be much slower, and therefore, detectable only over longer horizons. This implies that short interest may need to accumulate for some unspecified time before any correction occurs.

Predicting Long-Term Returns with Short Interest

Asquith and Meulbroek (1995) investigate the long-term returns to NYSE and AMEX stocks with very high RSI at some point from 1976 to 1993. While the previously mentioned work relies on short interest data reported in the financial press, Asquith and Meulbroek construct their own comprehensive data set. This is done because the financial press reports this data only for stocks with high levels or large changes in short interest. In August 1995, for example, the *Wall Street Journal* reported short interest only for those stocks with positions larger than 300,000 shares or changes of more than 50,000 shares from the previous month. Asquith and Meulbroek, on the other hand, wish to analyze RSI, not large levels or changes in short interest. This is because RSI reflects the supply of shares outstanding in the denominator, and they believe that supply together with demand (the numerator in RSI) will dictate the longer-term return. (Note that relying on the *Wall Street Journal* might preclude some stocks with high RSI if they do not also satisfy the reporting cutoffs.)

Asquith and Meulbroek focus on the excess returns to stocks that attain relatively high levels of RSI for as long as the high levels persist and for up to two years afterwards. In this way, they avoid the timing problem of earlier studies that requires precise alignment of the price reaction with the short interest announcement. They also point out several reasons why traded options may not obscure the information content in short interest. First, interviews with practitioners, including hedge fund managers, reveal that establishing large short positions with put options on hard-to-borrow stocks is more expensive and offers less liquidity than direct short selling. In addition, although one may be forced to cover a short sale early, there is no definite expiration date as with options, and this can be a serious disadvantage when speculating on a possible downturn in a stock. Finally, very few stocks under heavy selling pressure have listed put options. For stocks with RSI at or above the 95th percentile, less than 2% have listed put options traded.

Slightly under 24% of the stocks in the sample reach the 95th percentile of RSI at some point from 1976 to 1993; the RSI at this percentile is roughly 2.5%, on average, over the period. Stocks that attain this 95th percentile, or above, earned mean size-adjusted returns of −18% while remaining at or above this level, plus an additional −23% in the two years subsequent to falling below this level. The excess returns to stocks at the 99th percentile of RSI are even more stunning, but only about 7.5% of the stocks ever reached this level, and it is probably safe to assume that it is almost impossible to borrow these stocks. Note also that these returns do not include the rebate interest that institutional short sellers may receive. The statistically significant negative excess returns persist over the entire 18-year period, and they are even more negative for firms that are heavily shorted for more than one month.

Although it may be difficult to borrow stocks with RSI at or above the 95th percentiie, these returns would still appear to be of economic significance. Even if these stocks cannot be sold short, a high RSI should still serve as a sell signal to those who are long the stock, and at a minimum, these results would seem to relegate to myth status the traditional technical analysts' view that high short interest is a bullish indicator. In addition, the slow reaction of stock prices, that takes months if not years, is strong support for the overpricing hypothesized by Miller, as well as Figlewski.

Desai, Ramesh, Thiagarajan, and Balachandran (2002) extend the work of Asquith and Meulbroek to Nasdaq market stocks with comprehensive monthly short interest data obtained directly from the Nasdaq for the years 1988 to 1994. Based on improved methods from the performance measurement literature, they measure long-term excess returns by controlling for market-to-book ratios and momentum, as well as size and beta. Their results suggest that short sellers target highly liquid stocks whose prices have recently improved relative to fundamentals.

Stocks with RSI of 2.5% or more earn mean excess returns of −6.6% within one year and −8.8% within two years of attaining this level. Upon falling back below this 2.5% level, they continue to earn negative excess returns, on average, of −7.3% within one year and −11.2% within two years. These negative returns increase with higher RSI levels. They also find that the heavily shorted stocks are liquidated or forced to delist with a higher frequency than their size, book-to-market, and momentum-matched control firms. Farinella, Graham, and McDonald (2001) verify these results independently. Thus, Asquith and Meulbroek's conclusion that high short interest signals bearish sentiment about future returns applies to the Nasdaq market as well as the NYSE and AMEX.

Although these studies detect highly negative long-term returns without removing the stocks with traded options, it would be a mistake to assume that traded options have little or no effect on overpricing. Danielson and Sorescu (2001) of options introductions between 1981 and 1995 clearly shows that options improve informational efficiency by reducing the cost of short selling. They find that prices decline and short interest increases for stocks just after their options are first listed. The increase in short interest appears to be due to the purchase of puts by previously constrained short sellers whose intent is then transferred into short sales by the hedging activities of the put writers. As long as the marginal put writer is a market professional, with transactions cost advantages at short selling, the put contracts will represent a reduction in the cost of constructing an effective short position.

Diamond and Verrecchia predict that the lower costs of options will obscure the information content of short interest, but Danielson and Sorescu's price declines are unique to the overpricing hypothesized by Figlewski and Miller. Also consistent with the overpricing hypothesis, Danielson and Sorescu find that the price declines are larger in stocks with higher betas and greater dispersion of investor opinions, as proxied for by volume, return volatility, and analysts' forecasts. They suggest, however, that these predictable price declines are not exploitable because of the high cost of short selling these stocks prior to the listing of their options.

The magnitude of these negative returns, reported by Asquith and Meulbroek as well as Desai, Ramesh, Thiagarajan, and Balachandran, raises an important question. That is, beyond the point that high short interest predicts negative future returns, what factors determine the level of short interest in a stock? The fact that excess returns remain negative for up to two years suggests that accumulated short selling does, eventually, move prices in the direction of fundamentals. Understanding the determinants of short interest may offer some insights into identifying short-sale candidates early, before short interest increases until costs are prohibitive or borrowing becomes impossible. Of course, acting early is less costly, but there is also the added risk of acting too soon. The negative returns may take longer, or they may not materialize at all.

Determinants of Short Interest: Strategies, Profitability, and Information Content

It is well known that stocks with relatively low fundamental-to-price ratios experience systematically lower returns in the future. Using data on NYSE and AMEX stocks from 1976 to 1993, Dechow, Hutton, and Meulbroek (2001) document that short sellers target stocks that rank low based on ratios of cash-flow-to-price, earnings-to-price, book-to-market, and value-to-market. A stock is considered "targeted" if its RSI is 0.5% or higher. Short positions in these stocks earn positive excess returns in the year after they are targeted, as prices fall, and the ratios mean-revert. Further more, short sellers refine this strategy in three ways by avoiding stocks (1) that are expensive to short, such as small stocks with low institutional ownership and high dividends, (2) with low book-to-market ratios that appear justifiable due to high growth potential, and (3) with justifiably low fundamentals. These motives are confirmed by a telephone survey of major global hedge fund managers whose responses indicate that they short sell to profit from overpriced stocks.

Gintschel investigated the determinants of short interest in all the Nasdaq stocks eligible for margin trading between 1995 and 1998. Proxies for the float (that is, the supply of shares available to borrow), such as market capitalization and turnover, explain almost 60% of the cross-sectional variation in RSI. The significant time-series determinants of short interest are firm size, turnover, put option volume, as well as variables relating to technical and fundamental strategies, including future operating performance. He finds that short interest is equally sensitive to both positive and negative innovations in value and operating performance, suggesting it is motivated by hedging, while the short interest attributable to past returns is motivated by overpricing.

From an expectations model based on these findings, Gintschel computes unexpected changes in RSI and finds a significantly negative mean return of about 0.5% in the 15 days after the announcement of unexpectedly high RSI. He also detected a negative mean return of about 1% from the time short interest data are collected until the actual announcement, which indicates considerable leakage. In addition, he suggests that the negative long-term returns reported by Asquith and Meulbroek and Desai, Ramnesh, Thiagarajan, and Balachandran may be due to very high market capitalizations and low book-to-market ratios, rather than overpricing.

Boehme, Danielson, and Sorescu argue that tests of overpricing should use a two-dimensional framework based on Miller (1977). Recall that Miller indicates that binding short-sale constraints and high dispersion of investor beliefs are both necessary conditions for overpricing. Using RSI as a proxy for short-sale constraints, and return variance as well as share turnover as proxies for dispersion of beliefs, Boehme, Danielson, and Sorescu find that controlling for both yields low returns in constrained, high-dispersion Nasdaq and NYSE stocks between 1988 and 1999. Specifically, these stocks have a mean raw return of zero and a mean excess return of −20% over a one-year horizon, although this underperformance is less severe in stocks with traded options. (Considering either short interest or dispersion of beliefs separately does not yield significant excess returns.) Boehme, Danielson, and Sorescu suspect, however, that much of this underperformance cannot be arbitraged due to the high costs of short selling and the difficulty in borrowing these shares.

Pownall and Simko (2003) examine the fundamentals of stocks that are targeted by short sellers in "short spikes" (that is, abnormally large increases in short interest), as announced in the *Wall Street Journal* during the years 1989 through 1998. They also consider the price response to the announcement of a spike in short interest as well as whether the short sellers are profitable. The stocks targeted by short sellers are not materially different, in terms of fundamentals, from the population of NYSE firms during the period immediately prior to the spike. However, in the year subsequent to the short spike, the targeted stocks experience significant declines in key earnings-based fundamentals, such as earnings-to-price and earnings growth.

Their sample-wide mean excess return over the five-day intervals beginning with the announcement of the short spike is negative but small. For individual stocks, excess returns are more negative the larger the price run-up in the months prior to the spike. The profitability of short selling is measured by computing excess returns from the date the spike is announced until short interest returns to normal levels. The mean return for stocks that revert to normal levels of short interest within four months is −1% and significant, with all of this return coming in the month the reversion occurs. The sample-wide mean cumulalive excess return is −5% and significant; however, most of this profit is attributable to the one-third of the sample that takes more than nine months to revert to normal levels of short interest. (Over 75% of the sample stocks revert to normal levels within less than a year.)

These cumulative excess returns are significantly larger for stocks without traded options, for stocks with RSI greater than 2.5%, and for spikes that occur prior to 1994 (when hedge fund trading began in earnest). This last finding is of particular importance since the large post-announcement returns reported by Asquith and Muelbroek and Desai, Ramesh, Thiagarajan, and Balachandran were observed from samples that end in 1993 and 1994, respectively. The implication is that hedge fund managers are either exploiting (through speculation) or obscuring (through hedging) the information content of short interest such that it no longer persists for long periods, post announcement.

Pownall and Simko conclude that the profits to trading on short spikes are small, except in extended positions, which may be difficult to maintain and thus are more risky. This is similar to Boehme, Danielson, and Sorescu's conclusion, as well as that of Gintschel. Although it would appear that the emergence of hedge funds has eroded much of the highly negative pre-1994 returns, it may be slightly premature to dismiss the post-1994 returns as unexploitable. Instead, it would be better to more carefully consider the various costs of short selling.

The Costs of Short Selling as Limits to Arbitrage

In an earlier section, we briefly described the constraints on short sales: (1) the direct monetary costs of borrowing shares, (2) the difficulty (or impossibility) of establishing a short position, (3) the risk that the short position cannot be maintained, and (4) the legal and institutional restrictions on short selling. Now we wish to more carefully consider items 1, 2, and 3 since these are costs that limit the arbitrage of information contained in short interest data. Legal and institutional restrictions, in item 4, constrain short selling, but they do not represent a cost that an individual short seller actually faces.

The direct monetary cost of short selling is reflected in the rebate rate the lender of the stock pays to the borrower. Recall that the borrower sells the stock and the lender then has the use of the short-sale proceeds. Thus, the rebate rate represents the stock lender's cost of accessing funds less a compensating loan fee for lending the stock. Although rebate rates are usually positive, they can be negative if a stock is in such high demand (to borrow) that the loan fee is greater than the cost of funds. Rebate rates apply almost exclusively to institutional investors. Individual investors usually receive no interest on the proceeds from their short sales.

There is no centralized market for lending shares in the United States, and rebate rates are not publicly available. However, the activities of a large institutional lending intermediary during 2000 and 2001 are revealed in a study by D'Avolio (2002). He finds that fewer than 10% of the stocks this institution had available to loan are so-called *specials*, which have loan fees above 1%. The value-weighted loan fee across the entire available supply of shares is 0.25%. The average loan fee for specials is 4.3%, but fewer than 10% of these specials (less than 1% of all available stocks) are in such high demand that their rebate rates are negative.

For the stocks in the highest decile of short interest, D'Avolio reports an average loan fee of just under 1.8%, while about 33% of these stocks are specials. Stocks in the second highest short interest decile have an average loan fee of 0.8% and about 15% of these stocks are specials. Unfortunately, we do not know if the specials with high short interest experienced lower future returns than the general population of high-short-interest stocks. We do know, however, from Jones and Lamont (2002) that stocks with low or negative rebate rates have high market-to-book ratios and low subsequent returns, consistent with overpricing. Their results are based on a centralized market for lending stocks that was operated on the floor of the NYSE from 1919 to 1933. When stocks were newly listed on this lending market, they were overpriced by more than can be explained by the direct monetary costs of short selling. Jones and Lamont suggest that some other constraint on short selling must be limiting the arbitrage of this apparent opportunity.

The most obvious candidate is difficulty in borrowing the shares. However, Geczy, Musto, and Reed (2002) find that at least some of the profits to a number of popular shorting strategies are available to a hypothetical small investor who cannot short specials nor receive rebate interest. Their data are from a major institutional equity lender for 1998 and 1999. Unfortunately, they do not consider strategies based on short interest. Study Chen, Hong, and Stein suggest that overpricing survives because most institutional investors are restricted from short selling, and the rest of the market simply cannot absorb the opportunities. If this is true, it may bode well for the exploitation of carefully constructed short interest strategies that consider the accumulation of short interest over time. However, D'Avolio points out that loan fees are sticky in these decentralized lending markets; if so, stocks under increasing demand may be rationed prior to becoming specials, and this too could explain the Geczy, Musto, and Reed's results.

If short sellers worry that the risks of an early recall are high, or about being caught in a short squeeze, then they will require a premium for risky arbitrage. D'Avolio reports that the unconditional probability of a recall is low, with only 2% of the stocks on loan recalled in a typical month of his sample, but he also notes that recalls often occur when lenders receive negative information about a stock, which causes them to recall the shares, either to sell them or to reprice the loan. The possibility that negative information, possibly in the form of a rumor, could result in a recall is potentially unnerving for a short seller, and this introduces noise trader risk as an additional limitation to risky arbitrage. That is, a lender may rationally recall shares based on how less than fully rational investors may react to news, rather than based on fundamentals. Some short sellers request the identity of a potential lender to minimize the possibility of such a recall.

It is clear that constraints on short selling result in overpricing. It is also apparent from the studies by Gintschel, Boehme, Danielson, and Sorescu, and Pownall and Sirnko that even the post-1994 short interest data contain some information about future returns. Although there is no direct evidence, it would appear from D'Avolio as well as Geczy, Musto, and Reed that the monetary costs of short selling are probably not large enough to render short interest data unexploitable, at least not totally. It may, however, be difficult to borrow shares with high short interest, and possibly even more difficult to maintain the short position for long enough to realize a profit. In addition, D'Avolio points out that there is considerable risk associated with the early recall of a short position. It follows that these results may be viewed as consistent with market efficiency, at least to the extent that arbitrage opportunities are pursued to the limits of the costs and risks.

It is worth emphasizing that the existence of overpricing does not necessarily imply that short interest data contain information. Persistent overpricing relies on Miller's claim that the high costs of short selling constrain the less optimistic investors from trading based on their information, so that the market clearing price is determined by the overly optimistic investors. High short interest is a proxy for high costs only to the extent that short interest would have been proportionally that much higher, if unconstrained. Clearly, some stocks have low short interest precisely because short selling them is relatively costly.

The other academic justification for analyzing short interest comes from Diamond and Verrecchia's rational expectations model, which relies on short sellers with superior information. In their model, overpricing occurs only when the current level of short selling is higher than anticipated, and the entire correction comes with the short interest announcement that follows. It follows from Diamond and Verrecchia that higher frequency reporting of short interest, or transparency in short-sales transactions, should improve the informational efficiency of the U.S. stock markets. Next, we consider whether improvements are likely to actually result from any such changes.

Short-Sales Transactions and the Implications of More Frequent Reporting

Aitken, Frino, McCorry, and Swan (1998) were the first to provide evidence of the information content in short-sales transactions. Their data are from the Australian Stock Exchange for the years 1994 to 1996. This exchange reports transactions-level data, including short-sales information, to brokers and institutions on-line in real time. They report that short sales cause a rapid reassessment of price, with a mean of −0.2% within 15 minutes or 20 trades. There is less of a reaction to short sales associated with hedging activities, just as Diamond and Verrecchia would predict.

Aitken, Frino, McCorry, and Swan interpret their results as evidence that transparent short sales convey information as suggested by Diamond and Verrecchia. Note that this is claiming more than just short sellers have superior information. This is claiming that the execution of a short sale in this transparent market must immediately be recognized as an informed trade by other market observers who then, in turn, quickly sell long (or possibly short), and the price then, moves accordingly. In other words, the price moves directly as a result of other traders reacting to the short, seller's perceived information, rather than as a result of the short seller's actual information. Of course, the short seller does have to be informed if market efficiency is to improve as a result of transparency.

Angel, Christophe, and Ferri (2002) use daily transactions data from late 2000 to show that short sellers in Nasdaq-listed stocks have the ability to predict the direction of future earnings surprises as well as stock returns. But does this mean that the U.S. stock markets should become more transparent and issue more frequent and detailed reports about short sales?

The problem is that the very price adjustment process that should make a transparent market more efficient, that of Diamond and Verrecchia, is also a process that is ripe for manipulation and abuse. For example, almost daily we hear of short sellers being accused of "ganging up" on some stock in the hopes up driving its price down and then exiting at the opportune moment. Imagine how much easier this type of manipulation would be in a market with transparent short sales. This might result in the marginal short seller being a noise trader rather than an informed trader. In which case, the market would be less efficient than before. Finally, greater transparency can only address temporary mispricing that is consistent with rational expectations, as in Diamond and Verrecchia's model. Greater transparency does not reduce the costly constraints on short selling that drive the persistent overpricing Miller's model predicts. Thus, transparency may be of little benefit given that there is considerable support for Miller's overpricing hypothesis.

SOME PRACTICAL IMPLICATIONS

Some practical implications are listed below.

- Large percent increases in short interest are a weak signal of negative short-term returns. Other measures of short interest are weaker yet.
- Accumulating and sustaining levels of RSI are strong signals of negative returns in the long-term, although this relation is somewhat weaker post 1994. In addition, optimal entry and exit may be tricky with the accumulating short interest strategy. "Short spikes," especially those that have been sustained, represent an attractive point of entry.
- Traded put options in a stock may obscure the information content of the stock's short interest figure.
- Arbitrage and hedging activities in a stock may obscure the information content of the stock's short interest figure.
- The short interest data reported in the print media are incomplete and includes only stocks with very large levels or changes in aggregate short interest.
- Rebate rates are usually not available to individual investors.
- For stocks in high demand to borrow, rebate rates may be negative: meaning that the short seller must pay interest to the equity lender because the loan fee exceeds the cost of funds.
- It may be difficult to borrow stocks in high demand, especially if their loan fee is "sticky" low, and the risk of recall is higher in this situation.
- Identifying stocks before they are in high demand to borrow insures the ability to borrow at a modest loan fee. This may be done by studying the determinants of short interest. Recall that stocks with high valuations attract short sellers. Unfortunately, an early recall is more likely

if the stock larer becomes popular to borrow but your loan fee is low.

- Watch out for short squeezes! Avoiding them, as well as recalls, appears to be the logic behind the traditional technical analysis' view of high short interest. An example of a possible short squeeze set off by high short interest is that of Martha Stewart Living Omnimedia stock in January 2004. Investors scorned the stock through much of 2003 because in June 2002, Stewart had been tied to an insider-trading scandal at ImClone Systems. She was also charged with illegally trying to prop up the stock of her own company and deceive its shareholders. Although Stewart stepped down as CEO and chairwoman of the company after being indicted, Martha Stewart Living continued to struggle with slumping sales and earnings. But from mid-December 2003 to the end of January 2004, shares of Martha Stewart Living climbed from just over $9 to $13.39—its highest level in 19 months. Those bullish on the stock stated that the rally was a result of investors believing that closure would soon come with the end of the case and that, regardless of the outcome, the company would thrive once its executives got back to focusing on the business, rather than the trial. Technician's, however, claimed the rise was due in part to a short squeeze resulting from high short interest and the associated increase in demand to cover. More than 50% of the shares available for trading had been shorted during the December 2003 through January 2004 period.
- The only reason to buy or hold a stock with high short interest is if you have reason to believe that a short squeeze may soon come into play.
- Higher frequency reporting of short interest or greater transparency of short-sale transactions may actually reduce the informational efficiency of a market.

SUMMARY

Large percent increases in short interest predict negative future returns over short horizons, of a month or several days, although the relation is weak. It is clear, however, that short sellers tend to target stocks that have recently increased in price, or that have historically optimistic fundamentals, such as low book-to-market ratios. This indicates that short sellers attempt to profit from mean reversion, and since it is well known that mean reversion in stock prices is a long-horizon process, it should not be surprising that we observe that short sellers earn larger profits over long horizons, of up to two years. This, however, implies that short interest must accumulate, over time, before it contains any material information about future returns. Considering this accumulative process in their tests was thus the key insight of Asquith and Muelbroek who detect a very strong negative relation between accumulating RSI and long-term future returns.

More recent (post-1994) evidence, however, suggests that the emergence of hedge funds has weakened this signal, either as a result of their speculation on short interest or their hedging activities, both of which would obscure the information content of short interest. The post-1994 returns, to trading on short interest, appear large enough to survive the direct monetary costs of short selling. Whether they represent excessive compensation, however, is not so clear given the potential difficulties in borrowing shares and the risks of an early recall or a short squeeze. Thus, on the one hand, these results may be interpreted as consistent with Fama (1991) who defines an efficient capital market as one in which traders reflect information in prices only to within the cost of attaining and trading on the information. On the other hand, if noise traders impact the risks of a recall or a short squeeze, and they certainly may, then market efficiency exists only in the sense of the limits to arbitrage argument of Shleifer and Vishny (1997).

Most of the evidence presented here is consistent with the academic theories of either Miller or Diamond and Verrecchia. Short-sale constraints clearly result in overpricing, and there definitely is information content in short interest data, although it may be difficult to exploit. Short sellers' profits come from taking advantage of the reversion of prices back, down, to the mean. There is no evidence to support the traditional technical analysts' bullish view of high short interest, which actually relies on a reversion in prices back, up, to the mean. This bullish view of short interest appears to be rooted more in a fear of recalls and short squeezes than anything else.

REFERENCES

Aitken, M. J., Frino, A., McCorry, M. S., and Swan, P. L. (1998). Short sales are almost instantaneously bad news: Evidence from the Australian Stock Exchange. *Journal of Finance* 53, 6: 2205–2223.

Angel, J. J., Christophe, S. E., and Ferri, M. G. (2003). A close look at short selling on NASDAQ. *Financial Analysts Journal* 59, 6: 66–74.

Arnold, T., Butler, A., Crack, T., and Zhang, Y. (2005). The information content of short interest: A natural experiment. *Journal of Business* 78, 4: 1307–1336.

Asquith, P., and Meulbroek, L. K. (1995). An empirical investigation of short interest. Unpublished working paper, Harvard University.

Boehme, R. D., Danielson, B. R., and Sorescu, S. M. (2006). Short-sale constraints, differences of opinion, and overvaluation. *Journal of Financial and Quantitative Analysis* 41, 2: 455–508.

Brent, A., Morse, D., and Stice, E. K. (1990). Short interest: Explanations and tests. *Journal of Financial and Quantitative Analysis* 25, 2: 273–288.

Chen, J., Hong, H., and Stein, J. C. (2002). Breadth of ownership and stock returns. *Journal of Financial Economics* 66, 2–3: 171–205.

Choie, K. S., Huang, S., and James, S. (1994). Profitability of short-selling and exploitability of short information. *Journal of Portfolio Management* 20, 2: 33–38.

Christophe, S. E., Ferri, M. G., and Angel, J. J. (2002). Short-selling prior to earnings announcements. *Journal of Finance* 59, 4: 1845–1875.

D'Avolio, G. (2002). The market for borrowing stock. *Journal of Financial Economics* 66, 2–3: 271–306.

Danielson, B. R., and Sorescu, S. M. (2001). Why do option introductions depress stock prices? A study of diminishing short-sale constraints. *Journal of Financial and Quantitative Analysis* 36, 4: 451–484.

Dechow, P., Hutton, A., Meulbroek, L., and Sloan, R. (2001). Short-sellers, fundamental analysis, and stock returns. *Journal of Financial Economic* 61, 1: 77–106.

DeLong, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. (1990). Noise trader risk in financial markets. *Journal of Political Economy* 98, 4: 703–738.

Desai, H., Ramesh, K., Thiagarajan, R., and Balachandran, B. (2002). An investigation of the informational role of short interest in the NASDAQ market. *Journal of Finance* 57, 5: 2263–2287.

Diamond, D. W., and Verrecchia, R. E. (1987). Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics* 18; 2: 277–311.

Fama, E. F. (1991). Efficient capital markets: II. *Journal of Finance* 46, 5: 1575–1617.

Farinella, J. A., Graham, J. E., and McDonald, C. G. (2001). Does high short interest lead underperformance? *Journal of Investing* 10, 2: 45–52.

Figlewski, S. (1981). The informational effects of restrictions on short sales: Some empirical evidence. *Journal of Financial and Quantitative Analysis* 16, 4: 463–476.

Figlewski, S., and Webb, G. P. (1993). Options, short sales, and market completeness. *Journal of Finance* 48, 2: 761–777.

Fosback, N. G. (1976). *Stock Market Logic: A Sophisticated Approach to Profits on Wall Street*. Fort Lauderdale, FL: The Institute for Econometric Research.

Geczy, C., Musto, D., and Reed, A. V. (2002). Stocks are special too: An analysis of the equity lending market. *Journal of Financial Economics* 66, 2–3: 241–269.

Gintschel, A. (2001). Short interest on NASDAQ. Working paper, Emory University.

Hurtado-Sanchez, L. (1978). Short interest: Its influence as a stabilizer of stock returns. *Journal of Financial and Quantitative Analysis* 13, 5: 965–985.

Jones, C. M., and Lamont, O. A. (2002). Short-sale constraints and stock returns. *Journal of Financial Economic* 66, 2–3: 207–239.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *Journal of Finance* 32, 4: 1151–1168.

Pownall, G., and Simko, P. (2003). The information intermediary role of short sellers. *The Accounting Review* 80, 3: 941–966.

Senchack, A. J., and Starks, L. T. (1993). Short-sale restrictions and market reaction to short-interest announcements. *Journal of Financial and Quantitative Analysis* 28, 2: 177–194.

Shleifer, A., and Vishny, R. (1997). The limits to arbitrage. *Journal of Finance* 52, 1: 35–55.

CHAPTER 13

Emerging Stock Market Investment

LARRY SPEIDELL, CFA
General Partner, Ondine Asset Management, LLC

JARROD W. WILCOX, PhD, CFA
President, Wilcox Investment Inc.

Abstract: The emerging market phenomenon tends to come to our attention most after extremely good or bad returns, but it is a pervasive long-term process of better returns and useful diversification stretching over many decades. It can be understood in terms of innovation diffusion and the several factors that influence it: awareness of, resistance to, and ability to implement new ideas through entrepreneurial activity, combined with increasing transparency, liquidity and low barriers to capital movement that allow globalization of portfolios. There are strong theoretical reasons for expecting relatively high risk-adjusted returns to the fully diversified long-term global investor. Diversification opportunities are best within the frontier markets, both within this group and relative to the developed country investor's home market. Remaining emerging market inefficiencies appear to offer relatively more scope for active management than do large capitalization markets in the United States and other developed markets, but they require old-fashioned attention to detail for their full exploitation.

Keywords: emerging markets, diversification, innovation diffusion, globalization, portfolio construction, demographic, industrialization, population, transparency, corruption, bureaucracy, World Bank, International Finance Corporation (IFC), Morgan Stanley Capital International (MSCI), frontier markets

Investments in the emerging stock markets have become an important part of many portfolios in the more developed world. The likelihood of continued rapid growth in countries such as China, India, Russia, Brazil, and many smaller, less developed markets has become widely recognized. How should we as investors think about these markets? To clarify this issue, we first identify the members of this group, their long-term transitional economic process and why it may be expected to continue. Then we discuss their investment characteristics. We fill out this framework with discussions of why progress in individual countries may stall, some practical details of investment, and the variety of investment strategies commonly employed. Our treatment is from the viewpoint of the practitioner, incorporating broad nonfinancial qualitative description and analysis, but references to quantitative academic work are also provided.

ECONOMIES IN TRANSITION

Many less developed economies in countries with collectively large populations are becoming industrialized, increasing their rate of growth in gross domestic product (GDP) per capita, and becoming stable and open enough to attract global investors from developed countries. This process, once fairly begun, appears to be self-reinforcing.

Emerging Market Countries

The dawn of the modern age of investing in *emerging markets* came in 1987 with the introduction of the emerging stock market indexes of Capital International Perspective (now Morgan Stanley Capital International [MSCI]). At about the same time, the *International Finance Corporation (IFC),* part of the *World Bank,* began introducing country funds, hired managers to create specialized funds targeted at Malaysia, Thailand, Brazil, and other countries in Asia and Latin America. The IFC pioneered in the development of many new capital markets around the world. These efforts have born greater fruit than most observers would have predicted.

In the mid-1980s, it would have been very difficult to predict the rapid demise of the Soviet Union and the new freedoms that came to Eastern Europe. Equally surprising was the statement by Deng Xiaoping that "it doesn't matter if the cat is black or white, so long as it catches mice." That remark made capitalism legitimate in Communist China and led to an unparalleled economic revolution that has raised living standards for over 1 billion people.

For practical investing purposes, the emerging markets may be considered in three categories: the MSCI emerging markets countries, the frontier markets (those countries outside the MSCI EM Index that have stock markets), and the less developed countries that do not have stock markets.

The *MSCI Emerging Markets Index* is a free float-adjusted market capitalization index that is designed to measure equity market performance in the global emerging markets. In 1987, the Index consisted of eight countries. The list has expanded greatly as more countries have created viable stock markets. As of June 2006, the index included the following 25 countries: Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Hungary, India, Indonesia, Israel, Jordan, Korea, Malaysia, Mexico, Morocco, Pakistan, Peru, Philippines, Poland, Russia, South Africa, Taiwan, Thailand, and Turkey.

Despite the name "emerging," only one country has so far emerged to MSCI developed status—Portugal. Another, Malaysia, was anointed with developed world status from 1993 to 1998, but lost it when its stock market was shuttered to foreign investors in response to the Asian currency crisis. Table 13.1, however, leaves many active stock markets out of the emerging markets "club," including Sri Lanka and Venezuela, which were expelled, as well as Middle East markets such as those of Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates—though these are also now included in the Morgan Stanley Capital International Frontier Market Index.

In addition, over 60 other countries have some kind of stock market of interest to global investors. Many of these are followed in the S&P/IFC database of *frontier markets*—Bangladesh, Botswana, Bulgaria, Cote d'Ivoire, Croatia, Ecuador, Estonia, Ghana, Jamaica, Kenya, Latvia, Lebanon, Lithuania, Mauritius, Namibia, Romania, Slovak Republic, Slovenia, Trinidad & Tobago, Tunisia, Ukraine, and Vietnam. Other frontier markets include Costa Rica, Mongolia, Malawi, Nigeria, Zambia, and many more. Many additional countries yet have no stock markets. Not all countries with stock markets are freely open to foreign investors, but many countries that are outside the mainstream MSCI Emerging Markets Index may nevertheless deserve global investor attention.

Table 13.1 MSCI Emerging Markets Index Constituent Countries

Country	Date Added	Changes
Argentina	1988	
Brazil	1988	
Chile	1988	
Jordan	1988	
Malaysia	1988	Developed 1993–1997
Philippines	1988	
Thailand	1988	
Mexico	1988	
Greece	1989	
S. Korea	1989	
Portugal	1989	Developed since 1997
Taiwan	1989	
Indonesia	1990	
Turkey	1990	
Colombia	1993	
India	1993	
Pakistan	1993	
Peru	1993	
Sri Lanka	1993	Removed in 2001
Venezuela	1993	Removed in 2006
China	1995	
Israel	1995	
Poland	1995	
S. Africa	1995	
Russia	1996	
Czech Rep	1996	
Hungary	1996	
Egypt	2001	
Morocco	2001	

Source: MSCI.

The Process of Emergence

At the end of the 1980s, the Iron Curtain and the Soviet Union collapsed, ushering in a period of euphoria over emerging markets. Communism had been discredited as an economic engine, politicians and businessmen seemed committed to market-based reforms, and there was a virtuous circle of rising capital flows, falling capital costs, rising liquidity, falling volatility, and strong economic growth.

However, the optimism of many investors proved to be naïve. Beginning with the Mexican peso crisis in 1994, the world of emerging markets was rocked by the Asian crisis, the Russian Crisis, and several crises elsewhere, including Argentina, Brazil, Venezuela, and Turkey. Some blamed "crony capitalism." Others blamed inexperienced investment bankers who peddled deals that were bound to fail. Inadequate regulations and ineffective currency policies also played a role.

Yet economic growth rates for emerging countries still generally exceeded those of the developed world. According to the World Bank (2005), in the period from 2000 to 2005, the United States had average real GDP growth of 2.8%, EAFE countries had 2.6%, Emerging Markets (in the MSCI EM Index) had 4.4%, and Frontier Countries had 4.3%. Some countries have done much better than the average, with China having an annual growth rate of 9.3%, Kazakhstan 10.2%, and Armenia 11.0%. Among others showing excellent growth were Latvia (8.1%), Ukraine (7.4%), Vietnam (7.4%), and Tanzania (6.6%).

The Internet and other improved communications technologies provide education, business, and job opportunities in emerging countries that never had them before. Possibly most at variance with prior expectations is the outsourcing from more developed countries of jobs in fields such as technology, software, customer service, and medical diagnostics. Also supporting this development process has been the mobilization of labor from less productive rural areas to more productive urban and industrialized environments. Just as the past 150 years saw massive migration from farms to factories in the developed world, this migration is happening now in the emerging world. In China alone, nearly 20 million people a year are reported moving from rural villages to urban centers. A somewhat slower but still comparable movement is occurring in India.

On the other hand, countries must compete for capital, and some emerging countries will not quickly succeed. Global investors can be fickle and alter capital flows promptly at signs of trouble. Some countries have difficulty grasping this concept. After the fall of the Iron Curtain and the breakup of Czechoslovakia, one of the authors was part of a delegation that met with senior government ministers in Bratislava, capital of Slovakia. They wanted to attract foreign capital and described their companies as "sleeping beauties, awaiting the kiss of foreign investors to awaken." Sadly, the factories they spoke about were moribund Soviet-era tank factories, old industrial plants, and inefficient steel mills.

Many investors still harbor unfavorable images of emerging markets—images of dangerous working conditions, children picking through garbage to survive, inefficient factories belching pollution, and corrupt governments siphoning off their country's wealth into personal Swiss bank accounts. These problems do exist, but they are not the whole story.

Some past success stories of countries in transition are illustrated in Figure 13.1. According to the World Bank, over the 40 years through 2004, Japan's per capita GDP rose in U.S. dollars (constant year 2000) from $10,615 to $38,609 versus a rise for the United States from $16,416 to $36,655. Even more dramatically in percentage growth terms, Singapore's per capital GDP rose from $2,675 to $24,164. Over the same period, China has outpaced India: from $100 to $1,323 versus $187 to $538. Finally, one can see the trials that Russia has endured on its way from Communism to today's free-for-all capitalism: GDP per capita dropped from $2,693 in 1989 to $1,564 in 1996 before recovering to $2,286 in 2004.

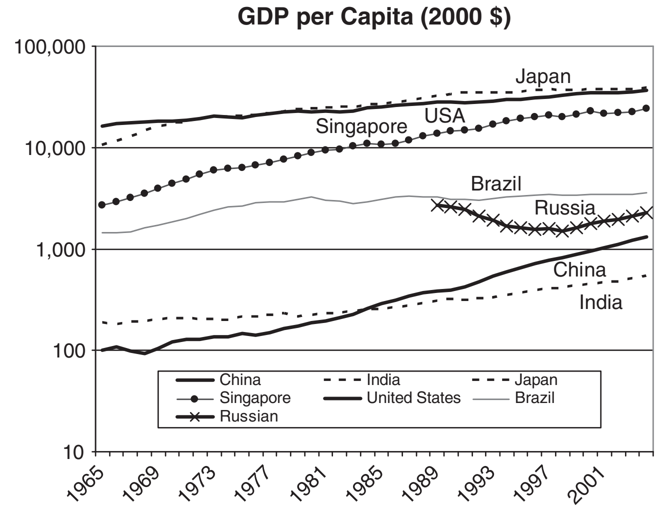


Figure 13.1 Growth in GDP per Capita in 2000 Dollars
Source: World Bank, 2006.

Why the Process Will Continue

One way to understand the emerging market process is to think of it in terms of a "contagion" model, producing the familiar S-shaped curve of *innovation diffusion.* Ideally, the transition of people to advanced market-based economies from traditional and command economies is a function of their exposure to the perceived relative attractiveness of the standard of living available through free-market capitalism and its supportive legal and political structure. In the real world, this picture is complicated by barriers to idea communication and receptivity, by lack of capacity to initiate implementation of different ideas, and by lack of resources, including time, to build up the physical and social infrastructure necessary to carry them out fully. Barriers to trade instituted by developed countries threatened by new competitors can also be a factor. But the process will continue at least as long as the gap in standard of living is perceived and the barriers of poor communication, disbelief, incapacity to initiate, and lack of resources keep falling.

How do the reservoirs of *population, industrialization,* and stock market value stand?

According to the World Bank, in 2006 the world stood very unequally divided. Emerging countries had 82% of the world's population, 77% of the land mass, only 32% of GDP (though 52% adjusted to purchasing power parity), and yet only 8% of world stock market capitalization. It is clear that the reservoir of the undeveloped economy is still much larger than that of the developed economy, and the higher population growth rate in less developed countries suggests that situation will endure for many years. The potential for transition remains enormous.

If one were to project current high relative growth rates into the future, as, for example, Jeremy Siegel (2005) does, it would imply that 77% of world GDP (purchasing power parity basis) will come from emerging countries in 2050, and they will account for 67% of world stock market capitalization. This does not imply that the transition will be smooth; in a process involving both rapid growth and high uncertainty, cycles of boom and bust along the way are to be expected.

To further justify confidence in high economic growth rates from "catch up" in many of today's emerging markets, we may examine the factors that govern the rate of "contagion."

Appeal

The appeal offered by market-oriented societies of both a much higher material standard of living and greater opportunity for personal advancement is intact, despite widespread resentment of what is often perceived to be cultural imperialism. The proof is that whenever barriers to communication and emigration are removed, significant portions of the population have voted with their feet.

Communication

Widespread global interest in motion pictures, television programs, and popular music originating in the advanced economies has occurred wherever the facilities for its expression are made available. A technological revolution in the means of communication based on mobile phones and the Internet still has far to go in many countries before reaching saturation points, but is moving steadily forward. As means of communication reach rural areas, farmers and small business people are empowered. Fisherman in India now use mobile phones so they can deliver their catch to the village paying the highest prices rather than seeing it spoil where there is a glut. The potential for global education through projects such as "One Laptop per Child" is enormous and largely untapped.

Acceptance

The fall of the Berlin Wall was a milestone in the recognition of the futility of holding back transition to market-based economies by the command governments in Eastern Europe and in the former Soviet Union. However, the real power behind the movement may have been the appeal of demonstrations by other countries in transition. In Europe, the success of the Common Market and its successors and the economic resurgence of West Germany were models. In Asia, the economic revival of Japan, followed by South Korea, Taiwan, Hong Kong, and Singapore, demonstrated what could be done. Others observed and learned.

In many countries today, nationalistic resentment, religious reaction, and understandable dissatisfaction with capitalism's excesses if poorly governed, might make it seem that *globalization* of marketplace values is stalled. This is apparent in parts of the Middle East, Africa, and South America. Nevertheless, the large populations of China and India appear fully committed to transition, and it will be hard to argue with their demonstrated success in the future.

Capacity to Innovate

It does no good to learn of technological advances unless there is a capacity to innovate. To accelerate innovation, markets must be allowed to make decisions long made by governments or by an elite. Those in high status positions can no longer be as protected by the status quo. At the most basic level, there must be escape routes from abject poverty for small businesspeople and entrepreneurs.

High birth rates, even if not supported by social strictures, keep women from participating fully in the trading economy; high birth rates are characteristic of low income levels. Lack of even primary education and means of communication and transportation closes channels for specialization in production where the output could be traded. The rule of man rather than that of law reduces the incentive to develop valuable economic assets. All these barriers are falling, as evidenced by United Nations and World Bank data, not everywhere, but in many places. The success of micro financing demonstrates the potential here.

Population changes are glacial but powerful. Many countries in the developing world, and almost all in the developed world, have falling birth rates and increasing life spans. For developed countries, the portion of workers over age 65 is already high, and projections by the United Nations Population Division show these percentages reaching even higher levels, especially in Japan and Europe over the next 40 years (nearly one dependent per worker in Japan). Meanwhile, by mid-century, populations will likely be actually shrinking in Japan, Russia, Europe, and China. The opposite problem exists in most emerging markets today—too many very young people. Birth rates, though they are falling, are still very high. In Pakistan, for example, 41% of the population is under age 14. It seems likely that this *demographic* is an aggravating factor, along with inequalities in income heightened by economic transition, underlying the instability in political and economic conditions in these countries. With economic growth and natural maturation, these populations of youth are likely to assume a more productive and stable economic role over the next two decades. They may, in fact, be needed to do much of the economic heavy lifting for the aging populations in the developed world.

From GDP to Market Value

For many years, the International Finance Corporation has produced a chart that shows that the size of a country's stock market versus its level of GDP per capita. The inference of causality indicates that as economies make real progress measured by incomes, they tend to develop more sophisticated financial markets, including larger stock markets. For example, a change in per capita GDP from $1,000 to $5,000 would imply a change in the stock market value from 28% of GDP to 66%. Thus, a five times increase in GDP per capita suggests that the stock market capitalization might grow by over 10 times.

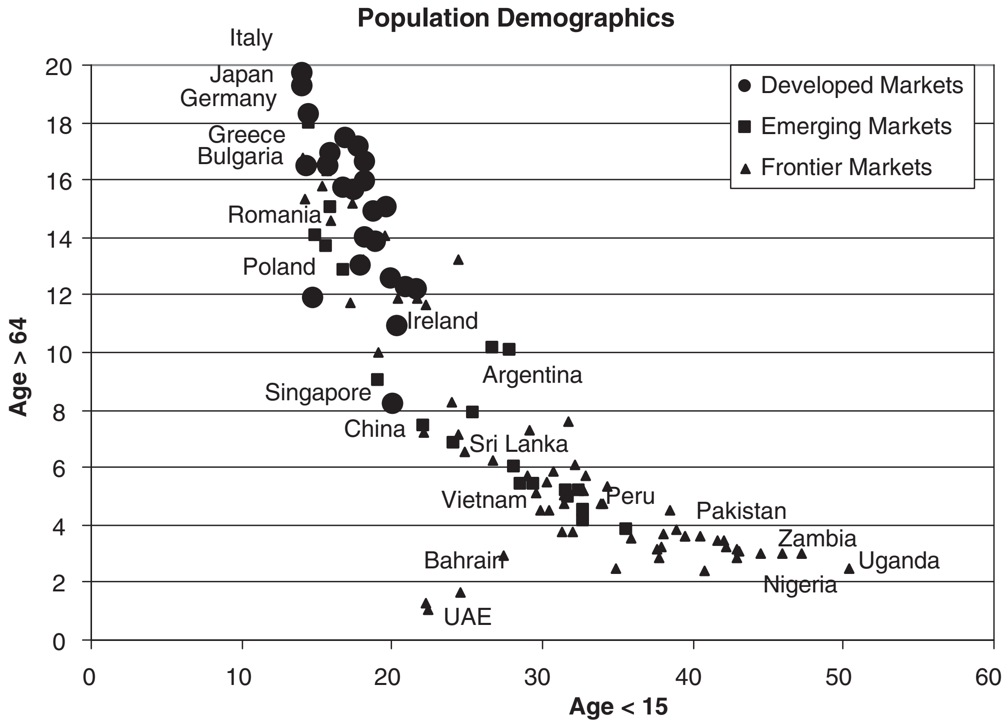


Figure 13.2 Demographics: Old versus Young
Source: World Development Indicators, 2006.

Why Progress Is Uneven

The logic we use for understanding why the emerging market process will continue may be equally valuable in understanding what causes its progress to move in fits and starts, and in some locations far more rapidly than in others.

Communication and Receptiveness

Market-oriented ideas can be extremely disruptive of the social fabric. If capitalism is unrestrained by government, it can produce staggering income inequalities as well as periodic mass unemployment. At the other end of the social scale, the required openness to the rise of new entrepreneurs and increased competition may prove disruptive to the power and privileges of existing elites. Apparently, this extends to political and religious spheres, and even to family relationships. Social disruption naturally causes resistance by those who lose thereby.

Economic Policies Affect Entrepreneurship

The 2007 Economic Freedom Score published by the Heritage Foundation ranks each country in terms factors such as regulations, size of government, labor, trade, fiscal and monetary policies, property rights, and corruption. The average score for 75 developed, emerging, and frontier countries has improved from 60.2 in 1995 to 63.7 in 2007, with gains in all categories, as shown in Table 13.2. There have been glaring exceptions, such as Zimbabwe and Venezuela, but most countries are moving toward conditions allowing entrepreneurship greater scope. But the investors might have profited if they could have predicted backsliding in this respect by the countries' shown in Table 13.3.

Table 13.2 Economic Freedom Ranks

	Heritage Freedom Score 1995	Heritage Freedom Score 2007
Developed countries	72.3	75.9
Emerging markets	60.1	61.5
Frontier markets	55.6	60.2

Source: The Heritage Foundation.

Table 13.3 Changes in Economic Freedom Score

Country	Heritage Rank Change (2000–2007)
Peru	−5.3
Morocco	−5.6
Paraguay	−5.7
El Salvador	−5.9
Turkey	−6.8
Venezuela	−8.1
Zimbabwe	−8.6
Bolivia	−10.4
Bahrain	−10.8
UAE	−14.5
Argentina	−16.7

Source: The Heritage Foundation, 2007.

Table 13.4 Corruption Perceptions Index, Transparency International

	2005 CPI Score	Chg 2000–2005
Developed countries	8.22	3.4%
Emerging countries	3.90	−0.1%
Frontier countries	3.71	2.6%

Corruption Affects Entrepreneurship

Corruption is a block to growth, a deterrent to foreign investors and a parasite that threatens to destroy its host. Those apologists who claim that corruption is cultural need only look to Singapore, surrounded by Asian neighbors where corruption is high. Singapore has insisted on integrity . . . and has enjoyed one of the world's highest economic growth rates.

Transparency International uses its own surveys as well as other sources to compile the Corruption Perceptions Index. It gives highest scores to the best countries, such as Iceland and Finland, and the lowest to those where corruption is the worst, like Haiti. Clearly, emerging and frontier countries have a long way to go to catch up with the developed world, as shown in Table 13.4. Still, the picture is not as bleak as it may seem, because the performance of emerging countries was clouded by declines in a few countries: Venezuela, Argentina, and Poland.

Bureaucracy Affects Entrepreneurship

The World Bank measures several dimensions of the costs of doing business, such as the number of procedures required to start a new business and the fees required relative to average incomes, but the most dramatic is the number of days required before one can open a business legally in different countries. It is an astonishing 152 days in Brazil versus two days in Australia!

Violence affects the entire process of economic transition. The daily news seems as filled as ever with news of violence, but the work done by the Human Security Center in conjunction with Uppsala University's Conflict Data Program and the International Peace Research Institute, Oslo (PRIO) indicates that violent global conflicts have declined from a peak over 50 in 1992 to 30 in 2002. (See Harborn, Hogbladh, and Wallensteen, 2006.) Despite the recent trend downward, anywhere violent conflict breaks out, normal economic progress is severely disrupted.

INVESTMENT CHARACTERISTICS

To fully understand the appeal of emerging markets to developed country investors, we need to look beyond the good financial performance of recent years. We can examine the factors that have tended to increase returns and may do so over a longer period, as well as those that have tended to reduce risks when viewed in a total portfolio context. How have these markets performed?

As we can see in Figure 13.3, from 1987 to 1994, emerging market returns were strong, as the Iron Curtain fell and the emerging market "asset class" became recognized. From 1994 to 2002, the markets moved sideways under the burden of persistent financial crises (the peso, the ruble, the baht, etc.). From 2002 to 2007, both the Emerging Markets Index and the Frontier Markets Composite have done well. Over long periods, emerging markets have provided superior returns as compared to both the S&P 500 Index and the MSCI EAFE index of developed country stock markets, but at the cost of extended periods of retreat. This has led to disappointment during downturns, and viewed in isolation, some investors continue to find the risk/reward ratio unappealing.

Factors Favoring Higher Returns

The simple story is one of the impact of globalization. Economically advanced practices spread from places where they benefit only a few people to places where they benefit the great majority of the world's population. Being less developed, emerging market countries can catch up simply by copying developed nations; being poor in the age of television, their people can see what the rich nations have and become motivated to achieve it; and being home to 84% of the world's population, emerging nations can gain support from developed nations to avert the specter of global violence in a nuclear age. We know that such growth stories can become the stuff of overvaluation and even speculative bubbles. There will be setbacks. However, the longer-term trend is clear.

Markets can be rather efficient in incorporating information into prices. Why then, should we believe that current prices of emerging market stocks have not already reflected this prospect? The following two factors should be considered.

Market Inefficiency in Pricing Long-Term Changes

Suppose one believes that markets are more efficient in pricing information regarding near-term events and less efficient regarding longer-term trends. This might be because no knowledge exists regarding the longer-term future. Alternatively, it may be because such knowledge of longer-term trends is not as widely shared, or because some decision makers do not care to anticipate events stretching beyond their lifetime as active investors. As with climate change, running out of cheap oil and other foreseeable long-term events, most investors appear to wait until major changes become self-evident before reacting.

Emerging markets have some of the flavor of this kind of phenomenon. The economic catch-up period for an economy once it has been opened up to market forces is quite long. For example, it has been two decades since the fall of the Berlin Wall, yet the former East Germany is still considerably behind West Germany in economic development. Moreover, that has been in a favorable circumstance of common language, a common national government, and heavy capital inflows. Where cultural gaps are larger, we may with reason expect much longer periods of transition. It would be requiring a great deal for markets to be sufficiently farsighted to adequately discount even a single generation's accelerated economic progress.

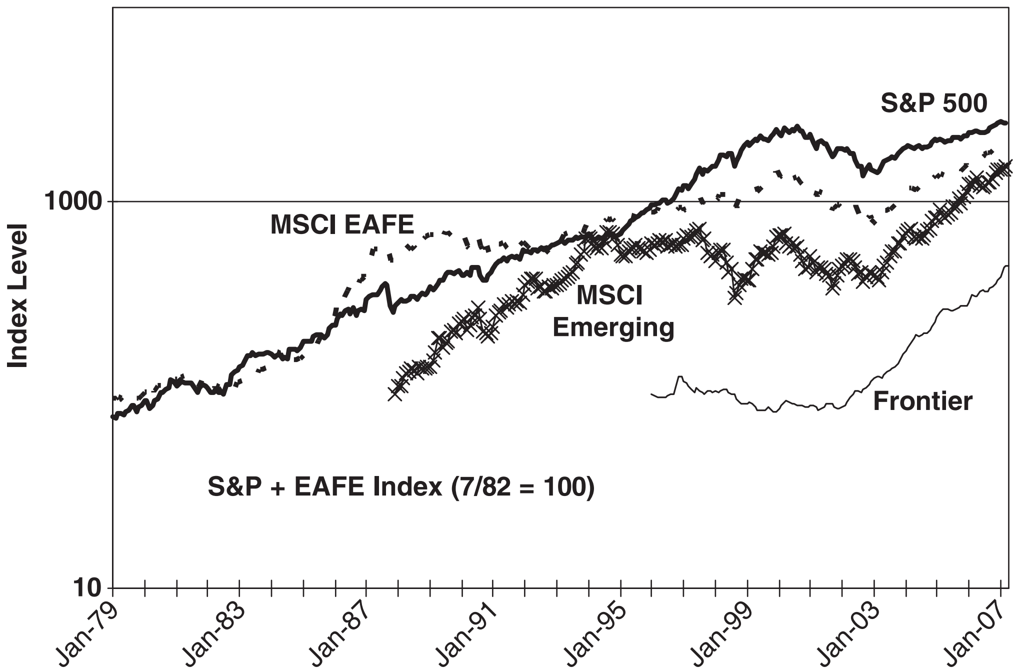


Figure 13.3 Returns of Global Market Indexes

Segmentation

The second factor is that there remain barriers to investment, not only for the global investor trying to put funds to work in the emerging market, but more importantly, for local investors to diversify abroad. Can we expect investors in Malaysia or Turkey or Brazil to be efficiently diversified globally? Lack of knowledge, language barriers, and government intervention through capital controls are common. In other words, emerging markets retain country and regional segmentation. For the local investor with less effective *diversification*, higher risks require compensation in the form of higher returns (Merton, 1987). The global investor with superior access to financial intermediaries and information can more effectively diversify and consequently take advantage of these higher returns with less sacrifice in risk.

It is possible that lower apparent prices for existing assets also reflect inflated accounting. However, the higher return requirements caused by segmentation, higher interest rates reflecting increased risk even after diversification, and the behavioral biases of global investors all suggest there is still room for further convergence for existing emerging markets. In addition, there are more countries waiting in the wings.

The Issue of Valuation

An interesting phenomenon in recent years has been the convergence of valuations around the world, illustrated by the chart of price-to-book ratios in Figure 13.4.

The convergence in price-to-book ratios since the end of 2002 appears to have been driven by an improvement in the quality of markets and companies themselves and is reflected in a narrowing of returns on equity in the different market sectors. For example, in March 2007, the return on equity (ROE) of MSCI Europe and the United States was identical, at 16.6%, while emerging markets were close, at 16.4%, and frontier markets were at 13.4% (Japan was still low at 10.4%, but much improved from zero in 2003).

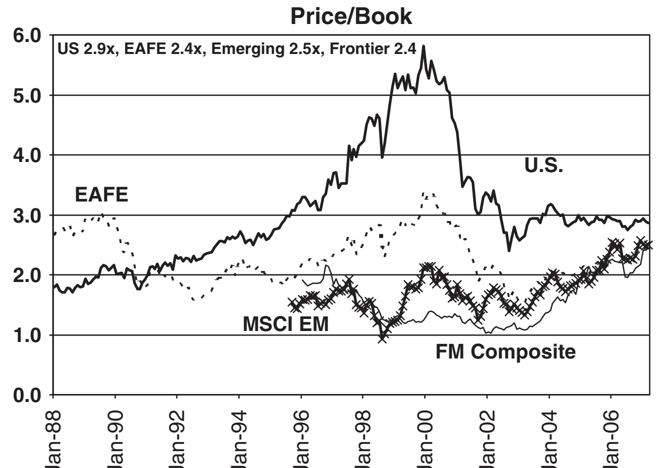


Figure 13.4 Price-to-Book Value
Source: MSCI, S&P.

Factors Favoring Lower Portfolio Risks

When we look at risk in isolated security or even country return characteristics, most emerging markets appear very risky to the developed country investor. Not only

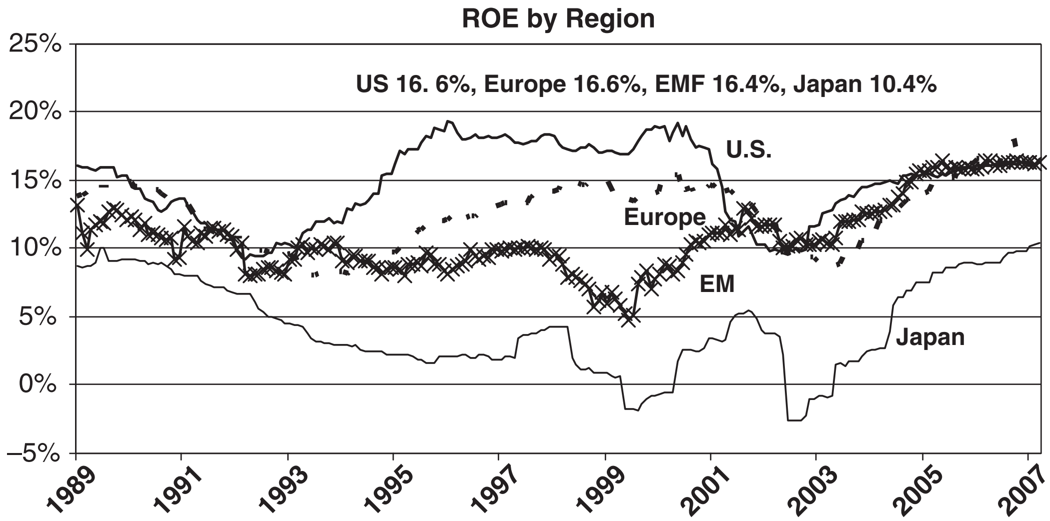


Figure 13.5 ROE by Region
Source: MSCI, S&P.

past return volatility but also future uncertainties as to government policy and the possibility of political disruption can create a daunting prospect. However, the risk picture is considerably alleviated when we consider a diversified basket of different emerging markets, and it is even more favorable when we consider the diversification potential of such a basket relative to developed country investments.

Risk Evolution
As individual country markets become more mature and fully integrated into global capital markets, volatility tends to decline, but simultaneously the benefit of low return correlation is also reduced as return correlations increase. For some investors this has meant the investments have become more mainstream and acceptable; others find that they must continually seek out the least developed markets, perhaps today's frontier markets, to get the full diversification and long-term return benefits.

Volatility
Countering the tendency for correlations with developed country correlations to increase in the countries where globalization has overcome preexisting segmentation, there is the benefit that the volatility of baskets of emerging markets has fallen. Consider Figure 13.6. Although volatility was high during the 1990s, the 36-month standard deviation of monthly returns for the MSCI emerging-market basket has recently fallen from over 30% to 17.7%.

Paradoxically, although frontier markets have high volatilities individually, when we look at the S&P/IFC Composite of 22 frontier markets, their return volatility as a group is only 11.1%, lower than MSCI emerging markets and close to that of the EAFE Index. This surprising result is due to the low return correlations among the frontier markets themselves. Botswana is not influenced by Bangladesh, and neither market cares much about what is happening in Bulgaria.

Correlations
Return correlations of most stock markets with one another have been rising over the past 20 years, due to the globalization of economies and the globalization of investing. As Figure 13.7 shows, the correlations of both the EAFE Index and the Emerging Markets Index with the S&P 500 have risen from 1986 to 2006, from roughly 0.50 in the 1980s to 0.75 and 0.72, respectively, recently. This appears to be a consequence of the rising popularity of global investing.

The earlier R-squared (the correlation squared or coefficient of determination) of 0.25 for emerging markets implied that only 25% of the movement of emerging markets in the 1980s was related to movement in the S&P. Since about 2000, R-squared has risen to about 0.5, implying that there is still significant diversification, but it is not what it used to be. Still, the return correlation of the MSCI Emerging market index with the S&P 500 is lower than that for the developed countries, as represented by the EAFE Index. If we look at the markets of countries very early in the transition to market development, the "frontier markets," the correlations remain very much lower. For the Frontier Markets Composite, the correlation is quite low, at 0.27, giving an R-squared of only 0.07, although both could be expected to rise as these markets become more popular.

Implementation Obstacles

In this section, we discuss some of the practical obstacles facing the global investor. Some of these are macro challenges—the destruction of shareholder value during the prosecution of Russian Mikael Kordokovsky

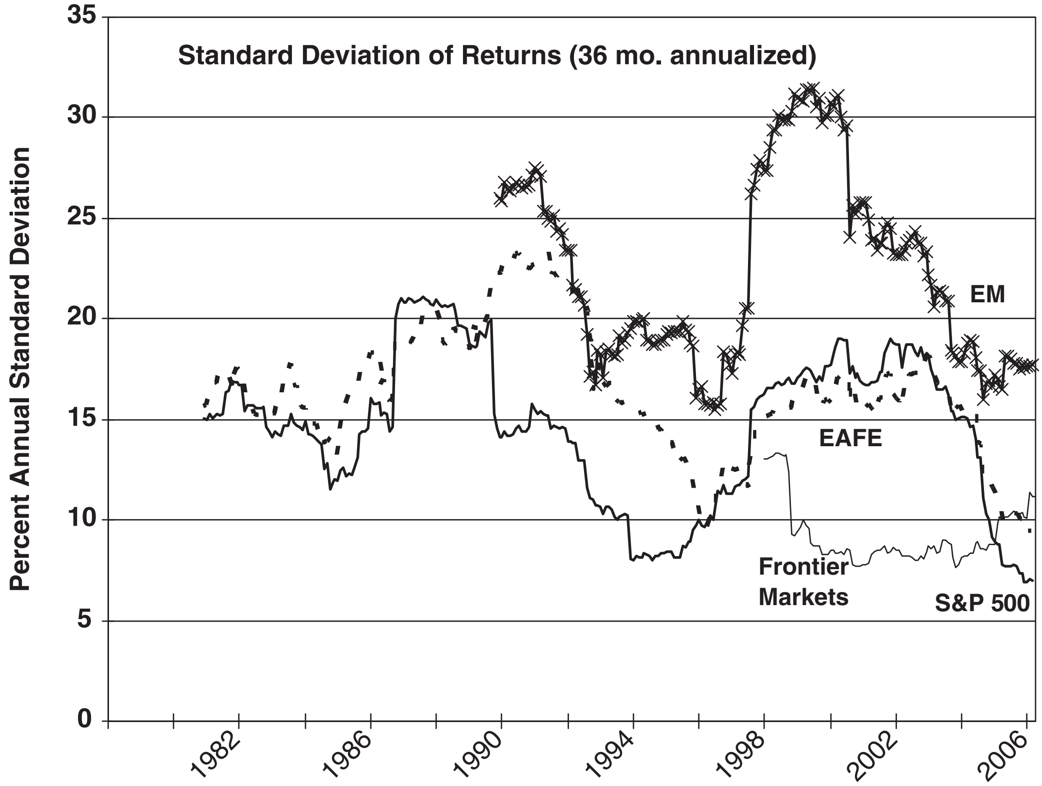


Figure 13.6 Market Volatilities
Source: MSCI, S&P.

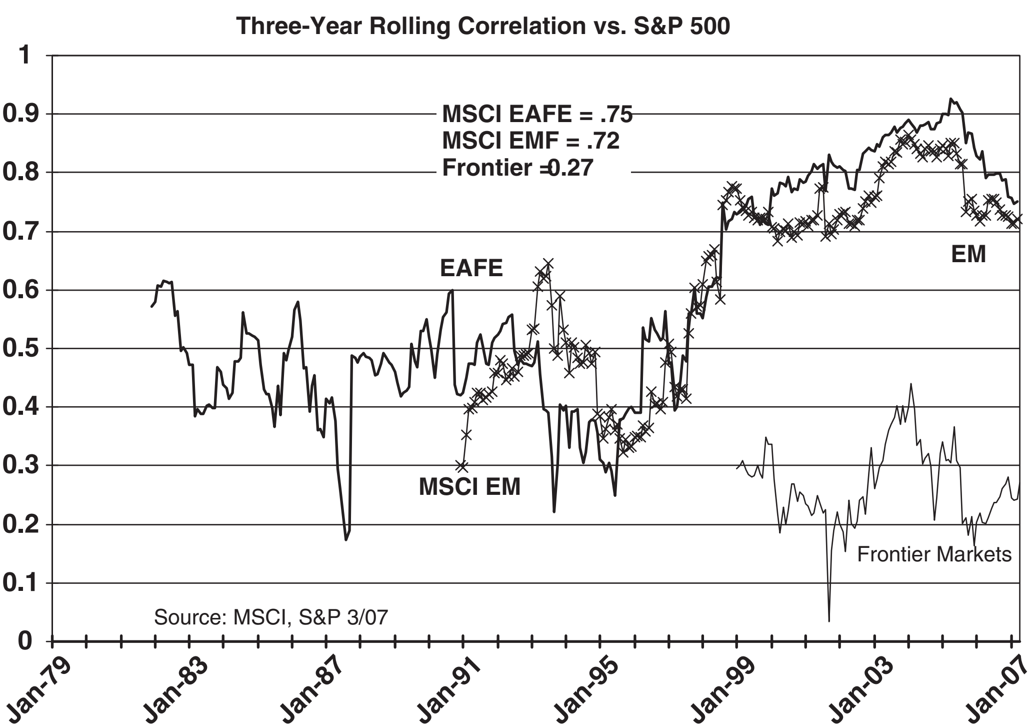


Figure 13.7 Correlations with the U.S. Market
Source: MSCI, S&P.

and the de facto nationalization of his oil company Yukos, and the economic ruin brought on by "big men" rulers in Zimbabwe and Kazakhstan are not isolated examples.

However, many obstacles are niggling details that add up in total. In frontier countries, simply getting annual reports can be difficult. Once published, often many months after year-end, they may not be posted on the company's web site (if there is one) for a year, and companies rarely respond to emailed requests for them. Once annual reports are obtained, they are often short and lacking in candid comments about the business outlook. The financial statements can be hard to interpret, with obscure accounting treatments and local terms such as "headline earnings" or "embedded value." In the case of initial public offerings, the planning and execution can be quite different from procedures in more developed markets. For example, a recent road show for an African bank failed to provide the prospectus, then commitment funds were tied up for two months before allocations were made and, finally foreign investors were allocated only $4,000 each. Local company visits can be very difficult, given the challenges of primitive local infrastructure and logistical challenges. Conference calls can be useful, but phone service to many frontier countries is poor and intermittent. Moreover, language barriers can exist even when both parties believe they are speaking the same English language!

Research reports can be hard to find as well, and their quality varies widely. Although the number of analysts who have passed the Chartered Financial Analyst examinations is growing rapidly in emerging countries, many reports on local companies are too often simply just reports: brief and factual rather than analytical and insightful. Financial forecasts are frequently lacking—in Bangladesh, they are even forbidden. Ask a local broker about his or her market, and the first stocks you will hear about are the biggest, because they would like you to do a large trade. Ask about their favorite stocks, and you will hear about the three stocks that went up the most in the past year. Ask again and you may hear about an obscure name that is cheap but rarely trades. These challenges mean that market prices may be extremely inefficient, but it takes diligence and patience to find the most attractive buy candidates.

Then there are issues of market access. Opening a local account can require coping with sluggish *bureaucracy* and Byzantine regulations. There can be local limits on foreign purchases: For example, until recently only one stock in Tanzania was below the 60% foreign ownership limit. Placing orders presents challenges too, with close monitoring needed to prevent them from being "lost," executed in the wrong stock or executed with unexpectedly high currency conversion charges. Initial public offerings may favor local investors, sometimes practically excluding foreigners who nevertheless have their commitment funds tied up for months. Finally, commission costs can be as high as 200 basis points in Bulgaria, Cote d'Ivoire, and Romania, 250 basis points in Zambia, 275 basis points in Nigeria, and 300 basis points in Ghana and Montenegro.

INVESTMENT APPROACH

Participation in emerging stock markets through a broad-based index fund or exchange-traded fund (ETF) is the most sensible approach for many investors. Language differences, lack of information, and high fees and transaction costs make passive investments in emerging markets through index funds, ETFs, and a basket of country funds the most widely used approach. On the other hand, because emerging markets remain quite inefficient in terms of security pricing, active management can be relatively more rewarding than in more efficient developed markets. The challenges are greater in assembling a basket of securities in frontier markets, but for large investors they can be approached through investment limited partnerships and special funds licensed for foreigners. For active investors, not only country funds, but for some of the larger emerging market stocks traded in the United States or London, often in American Depositary Receipt (ADR) or Global Depositary Receipt (GDR) form, or with the help of a local broker, may be practical vehicles.

Investment Strategies

Active management in emerging countries may take advantage of inefficiencies missed by a passive approach. Both top-down country analysis and bottom-up stock picking can be rewarding, but investors need to be very aware of high transaction costs. Given the volatility of individual countries, local prices can overreact in either direction to anticipated events. In 2006, in Brazil with Lula and in Russia with Putin, political events interpreted negatively by the markets were in fact major buying opportunities. Country volatility presents an opportunity for disciplined rebalancing of emerging countries within a basket, in a strategy that often beats a capitalization-weighted benchmark.

Style Investing

There may be a natural inclination for global investors to invest in the most successful companies in emerging markets, giving their portfolios a tilt toward growth styles. However, there is adequate information available to make at least broad cuts toward value and smaller capitalization investing. Figure 13.8 illustrates at a country-picking level the trade-off between price, expressed as price-to-book ratio, and profitability, expressed as return on equity.

Stock Picking

Understanding companies requires understanding the global industries in which they compete as well as the sometimes fierce competitors located in emerging countries. Investing in these emerging competitors need not mean stepping down in quality. The Indian software company Infosys, for example, does work at the frontier of technology and sets a world standard in terms of its financial disclosure. Its web site, www.infosys.com, is replete with full geographical and product line details that are

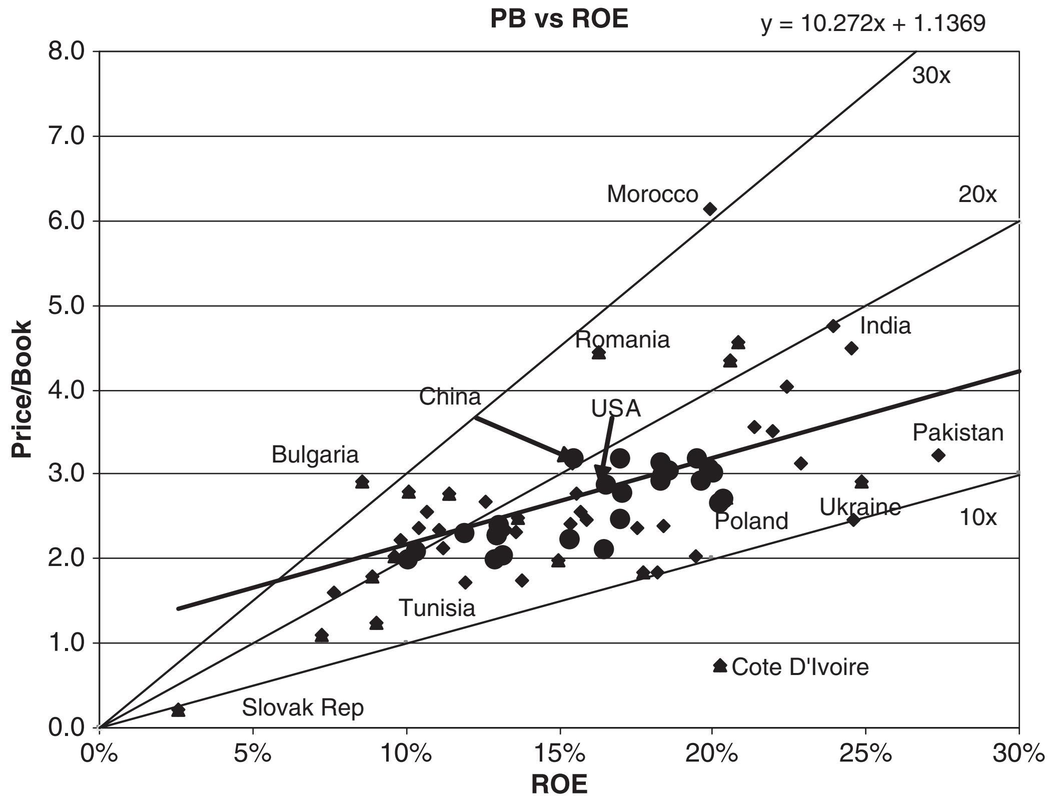


Figure 13.8 Relative Country Valuations Using PB-ROE
Source: MSCI, S&P.

too lacking from most companies in the developed world. The company loses nothing with this degree of openness, since competitors already find out what they need to know from insiders and former employees. However, the information turns security analysts into friends of the company rather than adversaries. Though it starts from a low base, this kind of openness is increasing in emerging markets because of demonstrated increases in value when global investors feel more comfortable (Gelos and Wei, 2005).

RELATED RESEARCH

As earlier noted, this chapter is written from a practitioner's perspective. However, there is much to be gained from academic research. The following studies are starting points from which many other references may be found.

In an impressive thread of continuing study, Harvey (1995); Erb, Harvey, and Viskanta (1996); Bekaert, Erb, Harvey, and Viskanta (1998); Bekaert and Harvey (2000); Bekaert, Harvey, and Lumsdaine (2002); and Bekaert, Harvey, Lundblad, and Siegel (2007) document the growth of understanding of emerging market investment characteristics. Important insights have also been contributed by Gelos and Wei (2005) on the impact of greater *transparency,* and by Dvorak (2005) on the information advantages of local investors.

Investigating the basis for skepticism, we have Errunza, Hogan, and Hung (1999) on whether diversification can be achieved without trading abroad; Conover, Jensen, and Johnson (2002) on the need for timing investing in emerging markets; Stulz (2005) on the limits of globalization; and Chen, Bennett, and Zhang (2006) responding to assertions that globalization has made country analysis less interesting than analysis of industry sectors.

SUMMARY

The history of development in today's emerging markets reveals a pattern of innovation diffusion that clearly has much further to go. During a long transition period, obstacles to open and market-based economic decision making, such as lack of information, resistance by special interest groups, and lack of resources for entrepreneurs, will continue to decline. Large populations in less developed countries will develop economies that are more productive and more open and reliable markets will leverage this productivity to create high asset values.

Many developed country investors still perceive emerging markets as being overpriced or, on the contrary, too ravaged by wars, disease, famine, and authoritarian governments to merit investment. We believe both these views represent deceptive stereotypes. Many emerging market countries have sound macroeconomic fundamentals.

Frequently, real per capita GDP is rising, inflation is low, currency exchange rates are becoming more stable, and corporate profits and return on investment are relatively high. From an economic standpoint, they are clearly emerging rather than stagnating. Looking ahead, they are likely to move on to become part of the developed world, and the stage is set to bring along a new set of emerging candidates from today's frontier markets. While available information is often sparse, local regulations are complex, and research coverage by analysts and brokerage firms is limited, these were also the characteristics of now successful emerging markets 20 years ago.

The remaining lacks of transparency and liquidity mean that security analysis and *portfolio construction* must be based on more care, diligence, and a longer-term perspective than investments in the developed markets. This opens opportunities for active investing styles that may be diminishing in more efficient developed capital markets. There are many risks in emerging markets, and setbacks are very apparent, but they can be reduced substantially through diversification.

As emerging markets develop further, the still enormous gap versus developed economies in terms of security value, GDP, and human capital will likely continue to shrink. Of course, as their equity markets become more efficient and populated by global investors, their return correlations with other markets are likely to continue to rise and the low valuations in overlooked segments will be harder to find. Consequently, the most attractive period for investment is in countries undergoing transition in expectations.

REFERENCES

Bekaert, G., Erb, C., Harvey, C., and Viskanta, T. (1998). Distributional characteristics of emerging market returns and asset allocation. *Journal of Portfolio Management* 24, 3: 102–116.

Bekaert, G., and Harvey, C. (2000). Foreign speculators and emerging equity markets. *Journal of Finance* 55, 2: 565–613.

Bekaert, G., Harvey, C., and Lumsdaine, R. (2002). The dynamics of emerging market equity flows. *Journal of International Money and Finance* 21, 5: 295–350.

Bekaert, G., Harvey, C., Lundblad, C., and Siegel, S. (2007). Global growth opportunities and market integration. *Journal of Finance* 62, 3: 1081–1137.

Chen, J., Bennett, A., and Zheng, T. (2006). Sector effects in developed vs. emerging markets. *Financial Analysts Journal* 62, 6: 40–51.

Conover, C., Jensen, G., and Johnson, R. (2002). Emerging markets: When are they worth it? *Financial Analysts Journal* 58, 2: 86–95.

Darroch, F. (2005). Case study: Lesotho puts international business in the dock. *Global Corruption Report 2005*. Transparency International.

de Soto, H. (2000). *The Mystery of Capital*, New York: Basic Books.

Dvorak, T. (2005). Do domestic investors have an information advantage? Evidence from Indonesia. *Journal of Finance* 60, 2: 817–839.

Erb, C., Harvey, C., and Viskanta, T. (1996). Expected returns and volatility in 135 countries. *Journal of Portfolio Management* 22, 3: 46–58.

Errunza, V., Hogan, K., and Hung, M. (1999). Can the gains from international diversification be achieved without trading abroad? *Journal of Finance* 54, 6: 2075–2107.

Friedman, T. (2005). *The World Is Flat: A Brief History of the Twenty-First Century*. New York: Farrar, Straus and Giroux.

Gelos, R., and Wei, S. (2005). Transparency and international portfolio holdings. *Journal of Finance* 60, 6: 2987–3020.

Harborn, L., Hogbladh S., and Wallensteen, P. (2006). Armed conflict and peace agreements. *Journal of Peace Research* 43, 5: 617–631.

Harvey, C. (1995). Predictable risks and returns in emerging markets. *Review of Financial Studies* 8: 773–816.

The Human Security Report (2005). Human Security Centre, University of British Colombia.

Merton, R. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42, 3: 483–510.

Siegel, J. (2005). *The Future for Investors: Why the Tried and True Triumph over the Bold and New*. New York: Crown Publishing Group.

Siegel, J. (2007). *Stocks for the Long Run, 4th Edition*. New York: McGraw-Hill.

Stulz, R. (2005). The limits of financial globalization. *Journal of Finance* 60, 4: 1595–1638.

World Population Prospects. (2004). Population Database, United Nations Population Division.

Wilcox, J. (1984). The P/B-ROE valuation model. *Financial Analysts Journal* 40, 1: 58–66.

World Development Indicators Online. (2005). The World Bank.

CHAPTER 14

Listed Equity Options and Futures

BRUCE COLLINS, PhD
Professor of Finance, Western Connecticut State University

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: Listed equity derivative contracts, options and futures, provide investors and market makers with an important tool for managing risk. These instruments serve as means to manage an equity investment strategy for portfolio managers and as a hedging device for dealers making markets in over-the-counter (OTC) derivatives. Both options and futures contracts are available on individual stocks and equity indexes. Like other listed derivatives, these instruments are contractual agreements with an exchange. Listed equity options and futures are standardized contracts that offer liquidity and leverage to the investor. In response to the explosive growth of the OTC market, options exchanges have developed longer-term options (LEAPS) and options with flexible terms (FLEX options). In the case of futures, the agreement is an obligation to deliver or receive an asset based on terms written in the contract. Futures contracts are marked to the market daily in that the gain or loss is realized every day in a futures account. Futures serve as a low cost substitute for a transaction in the underlying stock or stock index. Listed equity options are contracts that are paid for in full up front, the option premium or price, and treated as the right to purchase or sell a stock or stock index based on the terms of the contract.

Keywords: FLEX options, stock options, index options, LEAPS, multiplier, single-stock futures, cash settlement contracts, maintenance margin, variation margin

Listed equity option and futures contracts are simply exchange-traded equity derivatives where the exchange serves as the counterparty on every contract traded. The exchange establishes the contract specifications when the contract is created and sets the rules for trading. Because these contracts are exchange listed they have several benefits. The first and foremost is the reduction of default (counterparty) risk because the exchange takes the other side of each contract. A second important benefit is liquidity leading to price discovery. Liquidity enables the investor, the hedger, or speculator a quick and easy way to close out a position with minimum cost. Moreover, listed equity derivatives can help investors manage risk, enhance returns, manage costs more effectively or avoid regulatory hurdles. Investors wanting to change the risk characteristics of their portfolio without transacting in the cash market or wanting easy exposure to foreign markets can find listed equity derivatives a key tool in the investment management process. The purpose of this chapter is to provide an overview of these important instruments; listed equity options and futures.

LISTED EQUITY OPTIONS

An option is a contract in which the option seller grants the option buyer the right to enter into a transaction with the seller to either buy or sell an underlying asset at a specified price on or before a specified date. The specified price is called the strike price or exercise price and the specified date is called the *expiration date*. The option seller grants this right in exchange for a certain amount of money called the option premium or option price.

The option seller is also known as the option writer, while the option buyer is the option holder. The asset that is the subject of the option is called the underlying. The underlying can be an individual stock, a stock index, or another derivative instrument such as a futures contract. The option writer can grant the option holder one of two rights. If the right is to purchase the underlying, the option is a call option. If the right is to sell the underlying, the option is a put option.

An option can also be categorized according to when it may be exercised by the buyer. This is referred to as the exercise style. A European option can only be exercised at the expiration date of the contract. An American option, in contrast, can be exercised any time on or before the expiration date.

The terms of exchange are represented by the contract unit, which is typically 100 shares for an individual stock and a multiple times an index value for a stock index. The terms of exchange are standard for most contracts. In 1993, however, the CBOE introduced a FLexible EXchange Option™ (FLEX™) in response to growing investor demand for customized terms. *FLEX options* were originally introduced on index options, but have since been expanded to include listed equity options. These equity FLEX options or E-FLEX allow the investor to customize equity options in the same way as FLEX options to better manage an investment strategy and to structure risk exposure. The contract terms can be customized along four dimensions: underlying, strike price, expiration date, and settlement style. These options are discussed further below.

The option holder enters into the contract with an opening transaction. Subsequently, the option holder then has the choice to exercise or to sell the option. The sale of an existing option by the holder is a closing sale.

Basic Features of Listed Options

The basic features of listed options are summarized in Table 14.1. The table is grouped into four categories with each option category presented in terms of its basic features. These include the type of option, the underlying, strike price, settlement information, expiration cycle, exercise style, and some trading rules.

Stock Options

Stock options refer to listed options on individual stocks or American Depositary Receipts (ADRs). The underlying is 100 shares of the designated stock. All listed stock options in the United States may be exercised any time before the expiration date; that is, they are American-style options.

Index Options

Index options are options where the underlying is a stock index rather than an individual stock. An index call option gives the option buyer the right to buy the underlying stock index, while a put option gives the option buyer the right to sell the underlying stock index. Unlike stock options where a stock can be delivered if the option is exercised by the option holder, it would be extremely complicated to settle an index option by delivering all the stocks that constitute the index. Instead, index options are *cash settlement contracts*. This means that if the option is exercised by the option holder, the option writer pays cash to the option buyer. There is no delivery of any stocks.

Index options include industry options, sector options, and style options. The most liquid index options are those on the S&P 100 index (OEX) and the S&P 500 index (SPX). Both trade on the Chicago Board Options Exchange (CBOE). Index options can be American or European style. The S&P 500 index option contract is European, while the OEX is American. Both index option contracts have specific standardized features and contract terms. Moreover, both have short expiration cycles.

The dollar value of the stock index underlying an index option is equal to the current cash index value multiplied by the contract's multiple. That is,

$$\text{Dollar value of the underlying index} \\ = \text{Cash index value} \times \text{Multiple}$$

For example, suppose the cash index value for the S&P 500 is 1,410. Since the contract multiple is \$100, the dollar value of the SPX is \$141,000 (= 1,410 × \$100).

For a stock option, the price at which the buyer of the option can buy or sell the stock is the strike price. For an index option, the strike index is the index value at which the buyer of the option can buy or sell the underlying stock index. The strike index is converted into a dollar value by multiplying the strike index by the multiple for the contract. For example, if the strike index is 1,400, the dollar value is \$140,000 (= 1,400 × \$100). If an investor purchases a call option on the SPX with a strike index of 1,400, and exercises the option when the index value is 1,410, then the investor has the right to purchase the index for \$140,000 when the market value of the index is \$141,000. The buyer of the call option would then receive \$1,000 from the option writer.

LEAP and FLEX Options

LEAPS and FLEX options essentially modify an existing feature of either a stock option, an index option, or both. For example, stock option and index option contracts have short expiration cycles. *Long-Term Equity Anticipation Securities* (*LEAPS*™) are designed to offer options with longer maturities. These contracts are available on individual stocks and some indexes. Stock option LEAPS are comparable to standard stock options except the maturities can range up to 39 months from the origination date. Index options LEAPS differ in size compared with standard index options having a multiplier of 10 rather than 100.

FLEX options allow users to specify the terms of the option contract for either a stock option or an index option. The process for entering into a FLEX option agreement is well documented by the CBOE. The value of FLEX options

Table 14.1 Basic Features of Listed Equity Options

Stock Options	
Option Type	Call or Put
Option Category	Equity
Underlying Security	Individual stock or ADR
Contract Value	Equity: 100 shares of common stock or ADRs
Strike Price	$2\frac{1}{2}$ points when the strike price is between \$5 and \$25, 10 points when the strike price is over \$200. Strikes are adjusted for splits, recapitalizations, etc.
Settlement and Delivery	100 shares of stock
Exercise Style	American
Expiration Cycle	Two near-term months plus two additional months from the January, February or March quarterly cycles.
Transaction Costs	\$1–\$3 commissions and 1/8 market impact
Position and Size Limits	Large capitalization stocks have an option position limit of 25,000 contracts (with adjustments for splits, recapitalizations, etc.) on the same side of the market; smaller capitalization stocks have an option position limit of 20,000, 10,500, 7,500 or 4,500 contracts (with adjustments for splits, recapitalizations, etc.) on the same side of the market.
Index Options	
Option Type	Call or put
Option Category	Indexes
Underlying Security	Stock index
Contract Value	Multiplier × index price
Strike Price	Five points. 10-point intervals in the far-term month.
Settlement and Delivery	Cash
Exercise Style	American
Expiration Cycle	Four near-term months.
Transaction Costs	\$1–\$3 commissions and 1/8 market impact
Position and Size Limits	150,000 contracts on the same side of the market with no more than 100,000 of such contracts in the near-term series.
LEAP Options	
Option Type	Call or Put
Option Category	LEAP
Underlying Security	Individual stock or stock index
Contract Value	Equity: 100 shares of common stock or ADRs Index: full or partial value of stock index
Strike Price	Equity: same as equity option Index: Based on full or partial value of index. 1/5 value translates into 1/5 strike price
Settlement and Delivery	Equity: 100 shares of stock or ADR Index: Cash
Exercise Style	American or European
Expiration Cycle	May be up to 39 months from the date of initial listing, January expiration only.
Transaction Costs	\$1–\$3 commissions and 1 /8 market impact
Position and Size Limits	Same as equity options and index options
FLEX Options	
Option Type	Call, put, or cap
Option Category	Equity: E-FLEX option Index: FLEX option.
Underlying Security	Individual stock or index
Contract Value	Equity: 100 shares of common stock or ADRs Index: multiplier × index value
Strike Price	Equity: Calls, same as standard calls Puts, any dollar value or percentage Index: Any index value, percentage, or deviation from index value
Settlement and Delivery	Equity: 100 shares of stock Index: Cash
Exercise Style	Equity: American of European Index: American, European, or Cap
Expiration Cycle	Equity: 1 day to 3 years Index: Up to 5 years
Transaction Costs	\$l–\$3 commissions and 1/8 market impact.
Position and Size Limits	Equity: minimum of 250 contracts to create FLEX Index: \$10 million minimum to create FLEX No size or position limits

is the ability to customize the terms of the contract along four dimensions: underlying, strike price, expiration date, and settlement style. Moreover, the exchange provides a secondary market to offset or alter positions and an independent daily marking of prices. The development of the FLEX option is a response to the growing OTC market. The exchanges seek to make the FLEX option attractive by providing price discovery through a competitive auction market, an active secondary market, daily price valuations, and the virtual elimination of counterparty risk. The FLEX option represents a link between listed options and OTC products.

EQUITY FUTURES CONTRACTS

A futures contract is an agreement between two parties, a buyer and a seller, where the parties agree to transact with respect to the underlying at a predetermined price at a specified date. Both parties are obligated to perform over the life of the contract, and neither party charges a fee. Once the two parties have consummated the trade, the exchange becomes the counterparty to the trade, thereby severing the relationship between the initial parties. The terms of futures contracts are standardized, which means the contracts are known in advance of a transaction and it makes the contracts interchangeable or fungible. In addition, standardized contracts traded on an exchange created a viable secondary market for futures contracts.

Each futures contract is accompanied by an exact description of the terms of the contract, including a description of the underlying, the contract size, settlement cycles, trading specifications, and position limits. The fact is that in the case of futures contracts, delivery is not the objective of either party because the contracts are used primarily to manage risk or costs.

The nature of the futures contract specifies a buyer and a seller who agree to buy or sell a standard quantity of the underlying at a designated future date. However, when we speak of buyers and sellers, we are simply adopting the language of the futures market, which refers to parties of the contract in terms of the future obligation they are committing themselves to. The buyer of a futures contract agrees to take delivery of the underlying and is said to be long futures. Long futures positions benefit when the price of the underlying rises. Since futures can be considered a substitute for a subsequent transaction in the cash market, a long futures position is comparable to holding the underlying without the financial cost of purchasing the underlying or the income that comes from holding the underlying. The seller, on the other hand, is said to be short futures and benefits when the price of the underlying declines.

The designated price at which the parties agree to transact is called the futures price. The designated date at which the parties must transact is the settlement date or delivery date. Unlike options, no money changes hands between buyer and seller at the contract's inception. However, the futures broker and the futures exchange require initial margin as a "good faith" deposit. In addition, a minimum amount of funds referred to as *maintenance margin* is required to be maintained in the corresponding futures account. The initial margin and the maintenance margin can be held in the form of short-term credit instruments.

Futures are marked-to-the-market on a daily basis. This means that daily gains or losses in the investor's position are accounted for immediately and reflected in his or her account. The daily cash flow from a futures position is called *variation margin* and essentially means that the futures contract is settled daily. Thus, the buyer of the futures contract pays when the price of the underlying falls and the seller pays when the price of the underlying rises. Variation margin differs from other forms of margin because outflows must be met with cash.

In Table 14.2 we trace the cash flows of a daily margin account for 100 March 2007 S&P 500 futures contracts. We assume a long position is initiated by a speculator on February 20 and held through March 12. The initial margin requirement at the time was \$17,500 per contract and the maintenance margin was \$14,000. Thus, the total margin requirements for 100 contracts is \$1,750,000, which is reflected in the margin account. Whenever the margin

Table 14.2 Daily Margin Account Cash Flows: March 2007 Contract: February 20–March 12

Date	Futures Price	Value of Position	Daily Profit or Loss	Margin Cash Flows	Margin Account
2/20/2007	1461.8	\$36,545,000			\$1,750,000
2/21/2007	1460.4	36,510,000	−\$35,000		1,715,000
2/22/2007	1459.2	36,480,000	−30,000		1,685,000
2/23/2007	1453.8	36,345,000	−135,000		1,550,000
2/26/2007	1452.6	36,315,000	−30,000		1,520,000
2/27/2007	1395.3	34,882,500	−1,432,500	\$1,312,500	1,400,000
2/28/2007	1408.9	35,222,500	340,000		1,740,000
3/1/2007	1417.4	35,435,000	212,500		1,952,500
3/2/2007	1398.1	34,952,500	−482,500		1,470,000
3/5/200.7	1384.3	34,607,500	−345,000	275,000	1,400,000
3/6/2007	1407.8	35,195,000	587,500		1,987,500
3/7/2007	1405.5	35,137,500	−57,500		1,930,000
3/8/2007	1417.2	35,430,000	292,500		2,222,500
3/9/2007	1417.4	35,435,000	5,000		2,227,500
3/12/2007	1419.5	35,487,500	52,500		2,280,000

account falls below \$1,400,000, the investor will be required to restore the account to the maintenance margin level. Thus, the margin account is not allowed to fall below that amount.

The value of the position is the product of the futures price, the S&P 500 futures contract multiplier of \$250, and the number of contracts. The daily profits and losses reflect the daily marking to the market of the futures contracts. The margin cash flows are the amount the investor must raise to satisfy the maintenance margin requirement when the price moves against the position. For example, on February 27, the market experienced a significant decline of 3.95%, which resulted in a daily loss of \$1,432,500, a margin cash flow of \$1,312,500 was necessary.

Over the course of the futures contract, the net sum of the variation margin adjusted for financing costs and interest income ought to approximate the difference between the initial futures price and the spot price at the settlement date, when the futures price converges to the spot price. The difference between the initial futures price and the spot price at final settlement would already have been tallied across the life of the contract.

The vast majority of equity futures contracts use a stock index as the underlying. However, there are several exchanges outside the United States that list futures contracts on individual stocks.

Unlike options, both parties to a futures contract are exposed to counterparty risk. That is, there is bilateral counterparty risk. The clearinghouse for the futures exchange becomes the counterparty to the trade once a futures transaction is consummated by the initial transacting parties.

Futures contracts have a settlement cycle and there may be several contracts trading simultaneously. The contract with the closest settlement is called the nearby futures contract and is usually the most liquid. The next futures contract is the one that settles just after the near contract. The contract with the furthest away settlement is called the most distant futures contract.

Stock Index Futures Contracts

The underlying for a stock index futures contract can be a broad-based stock market index or a narrow-based index. Examples of broad-based stock market indexes that are the underlying for a futures contract are the S&P 500, S&P Midcap 400, Dow Jones Industrial Average, Nasdaq 100 Index, NYSE Composite Index, Value Line Index, and the Russell 2000 Index.

A narrow-based stock index futures contract is one based on a subsector or components of a broad-based stock index containing groups of stocks or a specialized sector developed by a bank. For example, Dow Jones MicroSector Indexes℠ are traded on OneChicago. There are 15 sectors in the index.

The dollar value of a stock index futures contract is the product of the futures price and a "multiple" that is specified for the futures contract. That is,

$$\text{Dollar value of a stock index futures contract} = \text{Futures price} \times \text{Multiple}$$

For example, suppose that the futures price for the S&P 500 is 1,410. The multiple for this contract is \$250. (The multiple for the S&P 500 futures contract is \$250.) Therefore, the dollar value of the S&P 500 futures contract would be \$352,500 (= 1,410 × \$250).

If an investor buys an S&P 500 futures contract at 1,410 and sells it at 1,430, the investor realizes a profit of 20 times \$250, or \$5,000. If the futures contract is sold instead for 1,360, the investor will realize a loss of 50 times \$250, or \$12,500.

Stock index futures contracts are *cash settlement contracts.* This means that at the settlement date, cash will be exchanged to settle the contract. For example, if an investor buys an S&P 500 futures contract at 1,410 and the futures settlement price is 1,430, settlement would be as follows. The investor has agreed to buy the S&P 500 for 1,410 times \$250, or \$352,500. The S&P 500 value at the settlement date is 1430 times \$250, or \$357,500. The seller of this futures contract must pay the investor \$5,000 (\$357,500 – \$352,500). Had the futures price at the settlement date been 1360 instead of 1,430, the dollar value of the S&P 500 futures contract would be \$340,000. In this case, the investor must pay the seller of the contract \$12,500(\$352,500 – \$340,000). (Of course, in practice, the parties would be realizing any gains or losses at the end of each trading day as their positions are marked to the market.)

Clearly, an investor who wants to short the entire market or a sector will use stock index futures contracts. The costs of a transaction are small relative to shorting the individuals stocks comprising the stock index or attempting to construct a portfolio that replicates the stock index with minimal tracking error.

Single-Stock Futures

Single-stock futures are one of the latest additions to listed equity futures contracts providing the same benefits as equity index futures contracts. Single-stock futures are equity futures in which the underlying is the stock of an individual company. The contracts are for 100 share of the underlying stock. At the settlement date, physical delivery of the stock is required. As of March 2007, single-stock futures are traded in the United States electronically on a market known as OneChicago, which is a joint venture of Chicago-based exchanges and Nasdaq London International Financial Futures and Options Exchange (LIFFE) Markets.

Single-stock futures of only actively traded New York Stock Exchange and Nasdaq stocks are traded. Consequently, an investor interested in short selling using single-stock futures is limited to those traded on both the exchanges. There are three advantages of using single-stock futures rather than borrowing stock in the cash market (via a stock lending transaction) if an investor seeking to short a stock has the choice.

The first advantage is the transactional efficiency that it permits. In a stock-lending program, the short seller may find it difficult or impossible to borrow the stock. Moreover, an opportunity can be missed as the stock loan department seeks to locate the stock to borrow. After a short position is established, single-stock futures offer a second advantage by eliminating recall risk, the risk of the stock lender recalling the stock prior to the investor wanting to close out the short position.

A third potential advantage is the cost savings by implementing a short sale via single-stock futures rather than a stock-lending transaction. The financing of the short-sale position in a stock-lending transaction is arranged by the broker through a bank. The interest rate that the bank will charge the broker is called the broker loan rate or the call money rate. That rate with a markup is charged to the investor. However, if the short seller receives the proceeds to invest, this will reduce the cost of borrowing the stock.

In addition to the advantage over short selling, single stock futures contracts provide investors with all the benefits of index futures contracts and another tool to manage their investment management process. Single stock futures can be used, for example, in hedging an existing cash stock position, for spread trading or pairs trading, as a low-cost alternative to transactions in the cash market and a means to managing nonsystematic risk.

Also, OneChicago has added ETF futures contracts to its list of products. These contracts have similar features to single stock futures contracts except the underlying the fund and not the actual stock.

SUMMARY

Listed equity options and futures contracts provide the foundation for the proliferation of derivative products that have developed over the last two decades. Listed equity derivatives are exchange-traded standardized contracts where the exchange takes the other side of every contract. Listed equity futures are contractual agreements between a party and the exchange where the short party agrees to make or take delivery of the underlying index, stock or fund. Index futures contracts are used to manage risk exposure and are not designed to take delivery of the underlying asset, but are based on cash delivery. Single-stock futures, in contrast, are physical delivery. Whereas futures contracts are obligations, listed equity options are contractual agreements where the buyer purchases the right to buy or sell an asset at an agreed upon price on or before a specific date. The advent of the modern OTC market prompted options exchanges to develop long-dated options (LEAPS) and options with more flexible structures (FLEX and E-FLEX). Whether the investor is hedging an existing position or wants to speculate on a pair of stocks, listed equity derivatives can serve as a useful tool. These contracts can play a pivotal role in risk management, returns enhancement, cost management, and regulatory management for institutional and individual investors alike.

REFERENCES

Angel, J. J., Gastineau, G. L, and Weber, C. J. (1999). *Equity FLEX Options*. New York: John Wiley & Sons.

Black, F., and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 3: 637–654.

Chance, D. M., and Brooks, R. (2007). *Introduction to Derivatives and Risk Management*: 7th edition. Mason, OH: Thomson South-Western.

Collins, B., and Fabozzi, F. J. (1999). *Derivatives and Equity Portfolio Management*. New York: John Wiley & Sons.

Cox, J. C., and Rubinstein, M. (1985). *Option Markets*. Englewood Cliffs, NJ: Prentice Hall.

Hull, J. (1997). *Introduction to Futures and Options Markets: Third Edition*. Englewood Cliffs, NJ: Prentice Hall.

Hull, J. (2006). *Options, Futures, and Other Derivative Securities*. Upper Saddle River, NJ: Prentice Hall.

Kolb, R. W., and Overdahl, J. A. (2006). *Understanding Futures Markets*. New York: Blackwell.

CHAPTER 15

OTC Equity Derivatives

BRUCE COLLINS, PhD
Professor of Finance, Western Connecticut State University

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: The development of the OTC equity derivatives market offers investors investment opportunities that are simply not available in the listed market or cash market. The international banking community has created a marketplace that has improved the efficiency of investing. Serious investors can no longer ignore the value of using OTC derivative products as a major part of their overall investment strategy. It is incumbent on plan sponsors, money managers, insurance companies, mutual funds, and corporations to consider the use of equity derivatives in achieving their investment objectives. The OTC equity derivatives market can be divided into three main components: OTC options and warrants, equity-linked debt instruments, and equity swaps. OTC equity options are customized option contracts that can be applied to any equity index, basket of stocks, or an individual stock. Equity derivatives have a variety of applications to investment management. Among the considerations for equity portfolio management are ways to take advantage of promising returns through the reallocation of funds within the portfolio. This portfolio rebalancing might include sector rotation, international diversification, style rotation, or return enhancement. These strategies mostly focus on stock selection in some way.

Keywords: OTC options, barrier options, compound options, knock-in options, knock-out options, equity swap, equity-linked debt, rainbow option, overperformance option, lookback option, quanto option, chooser option, Asian option, average option, basket option, binary option

An equity derivative can be delivered on a stand-alone basis or as part of a structured product. Structured products involve packaging standard or exotic options, equity swaps, or equity-linked debt into a single product in any combination to meet the risk/return objectives of the investor and may represent an alternative to the cash market even when cash instruments are available. The purpose of this chapter is to provide an overview of over-the-counter (OTC) equity derivatives.

PRODUCT FUNDAMENTALS AND APPLICATIONS

The three basic components of OTC equity derivatives are *OTC options* and OTC warrants, *equity swaps*, and equity-linked debt. These components offer an array of product structures that can assist investors in developing and implementing investment strategies that respond to a changing financial world. The rapidly changing investment

Table 15.1 OTC Equity Derivative Applications

Derivative Structure	Application	Benefit

OTC options	Risk management
	Return enhancement
	Equity investment
	  Single stock
	  Stock portfolio
	Sector rotation
	Traditional option strategies
	Currency hedged investment
		Customization
		Cost reduction
		Leverage
		Accessibility

Equity swap	Asset allocation Diversification Accessing foreign markets Index fund alternative Currency hedged investment	Cost reduction Leverage Customization Simplicity of deal

Equity-linked debt	Risk management
	Accessing foreign markets
	Equity investment
	  Single stock
	  Stock portfolio
		Customization
		Leverage
		Debt instrument

climate has fundamentally changed investor attitudes toward the use of derivative products. It is no longer a question of what can an investor gain from the use of OTC derivatives, but how much is sacrificed by avoiding this marketplace. OTC derivatives can assist the investor with cost minimization, diversification, hedging, asset allocation, and risk management. In this chapter we examine the product fundamentals across each category of OTC equity derivatives.

Table 15.1 summarizes various OTC equity derivative structures, their use, and their benefits. The benefits of using derivatives range from cost reduction to market access. There are several applications within each derivative category listed in the table. OTC structures can be devised to aid almost any style of equity management. OTC options can be used to buy or sell securities with lower market impact costs. All three OTC equity derivative structures provide a means for risk management. The benefits are not restricted to one investor group. Money managers and pension funds can utilize these derivatives as an integral part of their strategic and tactical investment plans. Every investment strategy has a derivatives application.

Let's take a quick look at the landscape of the equity investment world in order to make the connection between OTC equity derivative products and equity investments. There are two interrelated issues that all investment managers must address—risk and return. Applications of OTC derivatives can emphasize return enhancement or risk management issues. For example, a zero-cost collar structure in the OTC market can address the issue of hedging market risk by selling off a piece of the upside potential of the investment. A barrier option can isolate the precise conditions that the investor believes are most likely to occur without the need to buy all possible outcomes. In both cases, the manager is focusing on risk management.

However, an OTC structure such as an equity swap may be designed to take advantage of higher expected returns in a foreign market. In either application, one cannot separate return from risk, but we can separate the choice of an asset within an asset class from managing the market risk associated with the asset class itself. Thus, we can view equity investment as part of a strategic asset allocation strategy and utilize, when necessary, OTC structures to manage the asset allocation exposure to equities.

Another way to explore equity investments is to separate passive from active management. Passive management via indexing involves the construction of a portfolio of securities designed to exactly replicate the returns and risk profile of an established index. No attempt is made to time the market in order to enhance returns. In contrast, active management is based on the premise that managers with superior knowledge can add value to realized risk-adjusted returns above a corresponding passive strategy. Active management takes on many forms and in some sense can be linked through the methods of implementation. For example, an active strategy based on value involves the purchase of an equity portfolio of stocks that meet a certain fundamental criteria such as a high earnings-to-price ratio. The OTC market offers value managers a means of risk management and strategy implementation that extends beyond the domestic market. The same holds true for growth-oriented active management, or some other active approach. The point of linkage is in risk management and extending the boundaries of the selection universe. Regardless of a manager's equity style, their common denominators can be addressed in the OTC market.

OTC OPTIONS AND WARRANTS

OTC options can be classified as first-generation and second-generation options. The latter are called exotic options. We describe each type of OTC option below.

First-Generation OTC Options

The basic type of first-generation OTC options either relaxed or extended the standardized structure of an existing listed option or created an option on stocks, stock baskets, or stock indexes without listed options or futures. Thus, OTC options were first used to modify one or more of the features of listed options: the strike price, maturity, size, exercise type (American or European), and delivery mechanism. The terms were tailored to the specific needs of the investor. For example, the strike price can be any level, the maturity date at any time, the contract of any size, the exercise type American or European, the underlying can be a stock, a stock portfolio, or an equity index or a foreign equity index, and the settlement can be physical, in cash or a combination. An example of how OTC options can differ from listed options is exemplified by an Asian option. Listed options are either European or American in structure relating to the timing of exercise. Flex options are listed options that go beyond standard European or American styles. One example is to provide a capped structure. Asian options are options with a payout that is dependent on the average price of the spot price over the life of the option. Due to the averaging process involved, the volatility of the spot price is reduced. Thus, Asian options are cheaper than similar European or American options.

The first generation of OTC options offered flexible solutions to investment situations that listed options did not. For example, hedging strategies using the OTC market allow the investor to achieve customized total risk protection for a specific time horizon. The first generation of OTC options allow investors to fine-tune their traditional equity investment strategies through customizing strike prices, and maturities, and choosing any underlying equity security or portfolio of securities. Investors could now improve the management of risk through customized hedging strategies or enhance returns through customized buy writes. In addition, investors could invest in foreign stocks without the need to own them, profit from an industry downturn without the need to short stocks, or implement an intermediate asset allocation strategy through the purchase of a warrant.

Exotics: Second-Generation OTC Options

The second generation of OTC equity options includes a set of products that have more complex payoff characteristics than standard American or European call and put options. These second-generation options are sometimes referred to as "exotic" options and are essentially options with specific rules that govern the payoff. Exotic option structures can be created on a stand-alone basis or as part of a broader financing package such as an attachment to a bond issue.

Some OTC option structures are path dependent, which means that the value of the option to some extent depends on the price pattern of the underlying asset over the life of the option. In fact, the survival of some options, such as barrier options, depends on this price pattern. Other examples of path-dependent options include Asian options, lookback options, and reset options. Another group of OTC option structures has properties similar to step functions. They have fixed singular payoffs when a particular condition is met. Examples of this include digital or binary options and contingent options. A third group of options is classified as multivariate because the payoff is related to more than one underlying asset. Examples of this group include a general category of rainbow options such as spread options and basket options.

Competitive market makers are now prepared to offer investors a broad range of derivative products that satisfy the specific requirements of investors. The fastest growing portion of this market pertaining to equities involves products with option-like characteristics on major stock indexes or stock portfolios. It is derived from investor demand for long-dated European options and for options with more complex option structures. The real attractiveness of this market is that there is virtually no limit to the types of payouts.

In this section we review a few selective OTC product structures that can be used as management tools for traditional equity investment strategies. Table 15.2 provides

Table 15.2 Description of Some Basic Exotic Options

Option Structure	Description	Use	Pricing Comment
Knockout call	One of a class of barrier options Option is canceled if the spot price violates barrier target price	Overwriting	Less expensive than standard call option
Knockout put	One of a class of barrier options Option is canceled if the spot price goes above barrier target	Hedging	Less expensive than standard call option
Compound option	Option on an option, call on a put Gives owner the option: to buy the put	Hedging Speculating	Less expensive than standard call option
Spread option	Payout depends on the difference in performance between two assets	Asset allocation	Large risk premium due to correlation
Lookback option	Option that gives the right to holder to buy or sell underlying at best price attained over the life of the option	Equity exposure to volatile sectors	More expensive than standard options Market timing
Quanto option	Quantity-adjusted option Payout depends on underlying price and size in proportion to price	Access to foreign markets with currency hedge.	Pricing depends on correlation of exchange rate and spot price
Chooser option	Holder must choose to set the option as a call or put at some specific time	Similar to straddle	Less expensive than straddle
Asian option	Payout depends on average price of the underlying over a specified time period	Allows participation on average return	Less expensive than standard options Liability management
Basket option	Similar to index options, option written on basket of stocks	Hedging custom equity portfolios	Less expensive than portfolio of options
Binary option	Cash or nothing	Market timing Asset or nothing	Less expensive than standard option.

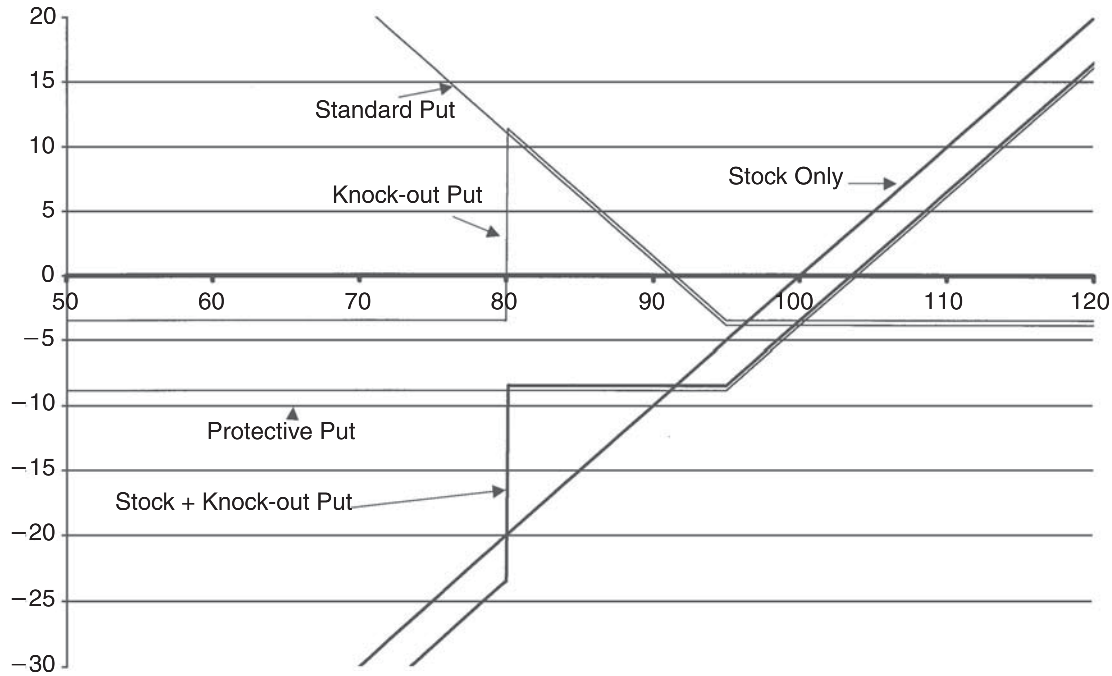


Figure 15.1 Knock-Out Put Option versus Standard Put Option

a partial listing of exotic options together with a brief description, an accompanying equity investment strategy application, and a comment on pricing. For an extensive discussion of exotic option products see, Nelken (1996) and Francis, Toy, and Whittaker (1995). We provide a basic description to accompany Table 15.2. The list of option structures is hardly exhaustive and is intended only to provide an introduction to some of the more common structures.

Barrier Options

Barrier options are path dependent options whose value and survival depends on the path or price pattern of the underlying asset over the life of the option. Moreover, the survival of the option is dependent on whether a "barrier" or predetermined price is crossed by the price of the underlying asset. *Knock-out* and *knock-in options* are examples of barrier options. Knock-out options go out of existence or do not survive when the barrier price is reached or exceeded. Knock-in options, however, come into existence when a barrier is reached or exceeded. The option can be a call or a put and the barrier can be below or above the current underlying asset value. In practice, when a knock-out option is terminated, a rebate is given to the holder of the option. Conversely, if a knock-in option comes into existence, a rebate is paid to the option writer. The options are priced assuming the rebates are paid at expiration or when the barrier is reached. Barrier options can be structured as European, American or Bermudan with regard to when they may be exercised. A Bermudan style option only allows the holder to exercise at specific, discrete times over the life of the option. Thus, a Bermudan option is somewhere between a European and an American. The difference in exercise style can have a dramatic impact on pricing.

Figure 15.1 is payout diagram that compares a standard protective put strategy with one that uses a knock-out put. The knock-out put terminates at a price of 80 and offers no rebate. The consequences are obvious from the exhibit. As long as the barrier is not breached, the strategy behaves similar to a protective put. The savings makes sense as long the investor is satisfied with the downside risk.

Compound Options

A *compound option* is an option written against another option. In other words, the underlying asset of a compound option is an option itself. There are 16 different types of compound options based on the exercise provisions of both the option and the underlying option and whether each is a call or a put. Thus, a call on a put would allow the holder of the call option to purchase a put option. The call could be a European type and the underlying put could be an American type. The pricing of compound options depends on the exercise style of both options and is less expensive than standard options.

Compound options are an alternative way of paying a premium upfront for the right to purchase an option at a later date should the need arise. Often market timing strategies are contingent on new information that enters the market. This could be Federal Reserve policy changes, earnings information or other events that influence financial asset prices. Investors who follow market trends may be engaged in a decision-making process to determine market exposure. The compound option provides an additional layer of choices for the investor in exchange for a premium. Once the decision is made to reduce market exposure, the investor can exercise the compound option or buy a standard option meeting their needs.

Rainbow Options

Overperformance options are rarely used despite their apparent attractiveness. The reason is the price is dependent on the correlation between the two assets, which is quite volatile over time. This makes hedging very difficult for market makers. The most common rainbow option is the option to exchange one asset for another, which is another name for a spread option or an overperformance option.

These options are structured to yield a payoff that depends on the relative performance of one asset versus another. An at-the-money call structure would pay off if there is a positive return differential between the two assets. For example, a call spread option on the relative returns of two stocks, A and B, would pay off if the returns to stock A were sufficiently above the returns to stock B over the investment horizon to pay for the cost of the option. The intrinsic value of the option is the difference between the returns since inception of the contract. The usefulness of this contract is that it can pay off even when equity prices are declining.

Lookback Options

A *lookback option* is one that allows the holder to buy or sell the underlying asset at the most favorable price attained over the life of the contract. This is the price that maximizes the value of the option at expiration. For lookback call options with fixed strike prices, this means using the highest price over that time and for a put option it means using the lowest price. For lookback options with floating strike prices, which are the most common, the opposite holds true.

Lookbacks can be expensive, so it is important to use them appropriately. They can be relatively attractive during periods of high volatility, but not in periods of persistent price appreciation. Lookbacks are preferred to standard options when the price differential between the initial stock price and the lowest (highest) price expected over the life of the option exceeds the difference between the lookback option premium and the premium on a standard option.

Quanto Options

"Quanto" is a term applied to option contracts that are "quantity adjusted" for the size of the exposure. This means that the payoff of a *quanto option* depends on the price of the underlying asset just as an ordinary option does, but it also depends on the size of the exposure as a function of the price. Most applications using quanto derivatives involve the purchase or sale of an asset in a currency different from the investor's domestic currency. The payoff is priced in terms of one currency, but made in terms of another. Consequently, the quanto option is automatically hedged for currency risk. Thus, for investors who want foreign market exposure without currency risk, a quanto put option is one alternative.

Chooser Options

A *chooser option* is also called an as-you-like-it option or a pay-now-choose-later option. It is initiated as neither a call nor a put but contains a provision that allows the holder to designate within some prescribed period whether the option will become a call or a put. There are two important types of chooser options: simple chooser and complex chooser. In the case of a simple chooser structure, the call and put alternatives have the same strike price and time to expiration. This is not the case for complex choosers, which can have a call and put alternatives that vary in both strike price and expiration.

Asian Options

Asian options are path-dependent options with a payout based on an average price. They are also known as *average options*. The payout for this type of option is based on the difference between two prices where one is based on the average price of the underlying asset for a set of dates over a prescribed period of time. Either the final spot price or the strike price is replaced by the average price of the underlying. Consequently, Asian options are priced at a discount to otherwise similar standard options and can be used as a way to reduce the cost of an option strategy.

Table 15.3 shows three different paths that both result in a final price of 36. The profit for the European put is 47.5 – 36 – 0.75 is 10.75, which offsets the 14 loss in the stock position. This is a typical protective put. The Asian option profits are path dependent. Path 2 results in no offsetting gain at all while path 1 provides some protection. The standard put option will always do better in down markets, but in up markets and some volatile market, such as that given by path 3, will not perform as well.

Basket Options

A *basket option* is an option structured against a portfolio or basket of assets, which may include a group of stocks

Table 15.3 Alternative Price Paths: Average Option Example

Observation	Path 1	Path 2	Path 3
1	50	50	50
2	48	48	45
3	46	51	40
4	45	52	38
5	44	53	36
6	42	51	42
7	41	50	44
8	40	52	45
9	39	51	46
10	38	50	47
11	37	49	48
12	36	36	49
Average Price	42.16667	49.41667	44.16667

or may include multiple asset classes. For equity baskets, the stocks are selected on the basis of a criterion such as industry group, risk characteristic or other factor that represents the investor's objective. This is comparable to an index option where the price of the option on an equity index is less than the average price of the options on each individual stock that makes up the index. Basket options are particularly appropriate for investors with equity portfolios that do no resemble the indexes that underlie listed index option contracts. These options are suitable for an investor wishing to use options with an underlying asset that exactly reflects their current portfolio holdings.

Binary Options

Binary options make an inherent gamble that pays off if the price of the underlying asset is above or below a particular price at expiration of the option. Binary options are like gambles that pay something when you win and nothing when you lose. The payment can be cash or the asset or nothing. Binary options are also called digital options or all-or-nothing options or cash-or-nothing options. The solution to a cash-or-nothing call option is the present value of the fixed payout times the probability of the stock price ending above the strike price. Binary options can be structured to pay out only if the spot price is higher than the strike price at expiration or if the spot price exceeds the strike price at any time during the life of the option. The size of the move is irrelevant because the payout is all or nothing.

Using Exotics

Before an investor decides to use exotic options, it is important to understand the impact that a specific exotic structure will have on the risk/reward profile of the current investment and the cost of implementing the strategy. For example, a lookback option that guarantees the optimal exercise value of the option seems very attractive. However, due to the expense of such an option, the investor may not be better off than if she had purchased the underlying security. Thus, cost becomes an important consideration in evaluating the impact of using exotics.

In order to accomplish this, investors need to understand the nature of the exotic derivative in question, including the pricing dynamics, the risks, and the expected benefits. Moreover, a complete understanding of what could go wrong is necessary including the potential costs, the tax implications, and the impact on the performance of the investor's portfolio. Consider, for example, a situation where the investor chooses a put option with a barrier structure that is designed to knock out at some level above the current price. If the barrier is hit suddenly and the put option is "knocked out," the risk is that the market reverses just as suddenly leaving the investor unprotected. Therefore, it is crucial that the investor understand that the cost saving of a barrier option compared to a standard put option has a risk component.

As shown in Table 15.3 there are several traditional equity investment strategies that can be facilitated using second generation options. These include hedging, overwrites, asset allocation, sector rotation, and style management exposure. Also as noted in the description in the table, exotic structures are often, but not always, less expensive than standard option solutions.

Despite the apparent applications, the use of exotic options brings a new element into the portfolio management process. Therefore, the use of exotics ought to be carefully considered and should provide a degree of precision to satisfy the investment objective that can only be achieved with an OTC exotic structure. Investment objectives that can be met with equal efficiency using methods that don't involve options need not require the use of exotic options. Nonetheless, OTC options do provide investors with opportunities to fine-tune their risk/reward profiles by providing flexible product structures that meet very specific investor requirements.

Options have risk management, returns management, cost management, and regulatory management applications. The addition of exotics can only add to these applications. We can sum this up by saying that the value of these products is the means they provide in meeting objectives with greater flexibility and efficiency. However, it must be emphasized that exotic structures are not appropriate in all situations. On the one hand, there are investors who are eager to use the latest derivative product whether they need to or not; on the other hand, there are investors who fear derivatives and will not use them regardless of whether it would facilitate meeting their financial objectives. It is crucial to evaluate the investor's investment objectives in terms of risk and return and how these objectives can be efficiently met. When risk management needs can be met using listed markets, it may be prudent to do so. However, for investors with specific needs that cannot be met by the listed market, a derivatives process ought to be developed and a set of criteria established that can be used as guidelines for determining whether or not an exotic structure makes sense. A partial list of conditions that investors should consider before using derivatives, particularly ones with complex rules governing the payoff is provided below:

- Complete understanding of investor objectives.
- Complete evaluation of current risk/reward profile.
- Analysis of current portfolio and targeted portfolio.
- Assessment of all alternative methods to meet objectives.
- Complete understanding of all financial products under consideration.
- Identify all risks associated with any derivative security.
- Develop worse case scenario analysis and a protocol for responding.
- Incorporate a complete strategy into the derivative.
- Consider tax and accounting consequences.
- Reevaluate procedure and explain it to all parties.
- Conduct a cost/benefit analysis.

The first step is to review the investor's objectives, their attitude toward risk and analyze the risk/return profile of the current portfolio and the targeted or benchmark portfolio. Once this is accomplished, then all possible alternative ways of meeting the objective must be explored in order to ascertain whether or not a solution using options is required. For now, it must be clear that the investor must either possess the knowledge and understanding of all the products and markets under consideration or have access to an expert who does. This includes understanding all aspects of the option and most importantly the risks.

Once the decision is made to use options, then the steps that follow involve identifying and fully considering the proposed option transaction. This includes an intimate understanding of the impact the transactions will have on the current portfolio and what could go wrong. Prior to a transaction involving an exotic structure, it is crucial to understand what can go wrong and what are the risks. Are you buying a risk or selling a risk, and is this the risk you want, one consistent with your tolerance for risk and desire for return? Once this is fully understood, even if the investor finds a means of implementing an investment strategy that uses derivatives in the listed market, it may still make sense to examine the OTC market. There are situations where dealers can more aggressively price an OTC product that produces the same risk profile as a portfolio of listed products.

The question of how issuers and dealers can offer OTC exotic products with complex payout structures at a superior price compared to collection of listed products can be explained by at least three factors. These include the need for a very specific risk exposure, the existence of market inefficiencies, and more effective delivery mechanisms. Market inefficiencies can arise when there are fewer payouts than states of the world. These inefficiencies lead to higher costs in the listed market, which reflect greater risks to the market maker.

In the final analysis, investors who are entertaining the possibility of using exotic option structures ought to engage in a serious effort to educate themselves and surround themselves with experts in the field. All alternatives must then be considered and finally a choice must be made to use exotics on the basis of benefits versus costs and the potential for limiting unfavorable outcomes or disaster.

EQUITY-LINKED DEBT INVESTMENTS

Equity-linked debt (ELD) investments are typically privately placed debt instruments. They differ from conventional debt instruments because the principal, coupon payment, or both are linked to the performance of an established equity index, a portfolio of stocks, or an individual stock. Consistent with other OTC equity derivative securities, equity-linked products have extremely flexible structures. For example, the equity component of the product can assume the characteristics of a call or a put or some combination. The payouts can be more complex mixing exotic-type option payouts with a bond.

In addition to providing flexible structures, equity-linked products also offer the investor the potential for higher returns than conventional debt instruments of similar credit risk. Other characteristics include: more volatile cash flows, the principal guaranteed by issuers with investment grade credit, and the avoidance of certain regulatory restrictions that prevent investors from entering into futures contracts, options, or swap agreements. For example, some pension funds are restricted from using derivatives. ELNs are recorded as a debt instrument and circumvent the restriction. Equity-linked products are typically longer-term investments and therefore have limited liquidity.

ELD investments are also referred to as equity-linked notes. Examples of these are equity participation notes (EPNs, offered by Merrill Lynch), stock upside notes (SUNs, offered by Merrill Lynch), structured upside participating equity receipt (SUPER), and synthetic high-income equity-linked security (SHIELDS). Equity-linked notes are issued by banks, corporations, and government sponsored enterprises, and have maturities ranging from 1 year to 10 years. The coupon can be fixed or floating, linked to an equity index, a portfolio of stocks, or a single stock and denominated in any currency. The equity-linked payment is typically equal to 100% of the equity appreciation, and redemption at maturity is the par value of the bond plus the equity appreciation. Equity participation is actually flexible and changes depending on whether the ELD instrument includes a coupon payment.

The conventional ELD instrument is simply a portfolio consisting of a zero-coupon bond and an index call option. This structure can be extended to include a put or an exotic option. The cash flows associated with an ELD structure are as follows. At issuance, the investor purchases the note, which represents the initial cash flow. Periodic cash flows are derived exclusively from the performance of the linked equity index. For example, if the index appreciated 10% for the year and equity participation is 100%, then assuming that the notional amount is $1 million, the investor would receive $100,000 as a periodic cash flow. The final cash flow includes the return of principal and the final equity payment.

Often, however, cash flows are subject to a cap, which limits the upside participation. SUNs, for example, provide 100% of principal at maturity and pay an annual coupon based on 133% of the year-over-year appreciation in the S&P 500 index subject to a cap of 10%. Thus, the maximum appreciation is 13.3% per annum. If an investor believes the S&P 500 will appreciate by more, this not the appropriate investment vehicle.

The use of an ELD is particularly attractive to domestic insurance companies subject to risk-based capital guidelines, which mandate higher capital requirements for investing in equity than for debt. ELDs are carried as debt, but have their performance linked to equity. Thus, insurance companies can maintain the capital requirements associated with debt instruments and still obtain equity market exposure. Pension funds also can benefit by using ELDs to gain access to foreign equity markets. Direct foreign equity investments subject pension funds to withholding taxes. The use of ELD structures, where the equity component is a foreign index, allows pension funds to avoid withholding taxes. The note has the same structure flexibility as conventional ELD instruments and can also include a currency hedge.

An example of such an equity-linked note is one that combines a zero-coupon bond and an at-the-money call option on the FTSE-100 Index. The redemption value of the note is the higher of par value or the product of par value times the ratio of the value of the FTSE-100 Index at maturity and its value when the note is purchased. This equity-linked note creates a debt instrument with payments based on the returns to the FTSE-100 Index, while eliminating all unwanted risks and costs associated with holding U.K. equities.

As in the case of OTC options, ELD structures are extremely flexible. The investor can decide upon the amount of equity participation, whether to include a coupon or not, whether to target levels of equity appreciation over the life of the product when the target is realized, or whether to create a synthetic convertible bond.

EQUITY SWAPS

Equity swaps are similar in concept to interest rate or currency swaps. They are contractual agreements between two counterparties which provide for the periodic exchange of a schedule of cash flows over a specified time period where at least one of the two payments is linked to the performance of an equity index, a basket of stocks, or a single stock. Like options and futures, equity swaps are substitutes for a direct investment in equities. In a standard or plain vanilla equity swap one counterparty agrees to pay the other the total return to an equity index in exchange for receiving either the total return of another asset or a fixed or floating interest rate. All payments are based on a fixed notional amount and payments are made over a fixed time period.

Equity swap structures are very flexible with maturities ranging from a few months to 10 years. The returns of virtually any asset can be swapped for another without incurring the costs associated with a transaction in the cash market. Payment schedules can be denominated in any currency irrespective of the equity asset and payments can be exchanged monthly, quarterly, annually, or at maturity. The equity asset can be any equity index or portfolio of stocks, and denominated in any currency, hedged or unhedged.

Variations of the plain vanilla equity swap include: international equity swaps where the equity return is linked to an international equity index; currency-hedged swaps where the swap is structured to eliminate currency risk; and call swaps where the equity payment is paid only if the equity index appreciates (depreciation will not result in a payment from the counterparty receiving the equity return to the other counterparty because of call protection). In fact, the first equity swap was designed as a means to invest in foreign securities while avoiding the tax consequences of a direct investment.

A basic swap structure is illustrated in Figure 15.2. In this case, the investor owns a short-term credit instrument that yields the London Interbank Offered Rate (LIBOR) plus a spread. (Note the difference in the quotation convention for equity swaps compared to interest rate swaps. For

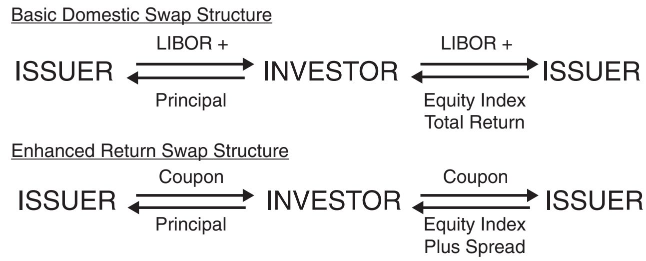


Figure 15.2 Equity Swaps

the latter, the floating rate is quoted flat while the fixed-rate side is quoted as the rate on a comparable maturity Treasury plus the swap spread.) The investor then enters into a swap to exchange LIBOR plus the spread for the total return to an equity index. The counterparty pays the total return to the index in exchange for LIBOR plus a spread. Assuming the equity index is the Nikkei 225, a U.S. investor could swap dollar-denominated LIBOR plus a spread for cash flows from the total return to the Nikkei denominated in yen or U.S. dollars. The index could be any foreign or domestic equity index. A swap could also be structured to generate superior returns if the financing instrument in the swap yields a higher return than LIBOR.

Equity swaps have a wide variety of applications including asset allocation, accessing international markets, enhancing equity returns, hedging equity exposure, and synthetically shorting stocks.

An example of an equity swap is a one-year agreement where the counterparty agrees to pay the investor the total return to the S&P 500 Index in exchange for dollar-denominated LIBOR on a quarterly basis. The investor would pay LIBOR plus a spread $\times$ 91/360 $\times$ notional amount. This type of equity swap is the economic equivalent of financing a long position in the S&P 500 Index at a spread to LIBOR. The advantages of using the swap are no transaction costs, no sales or dividend withholding tax, and no tracking error or basis risk versus the index.

The basic mechanics of equity swaps are the same regardless of the structure. However, the rules governing the exchange of payments may differ. For example, a U.S. investor wanting to diversify internationally can enter into a swap and, depending on the investment objective, exchange payments on a currency-hedged basis. If the investment objective is to reduce U.S. equity exposure and increase Japanese equity exposure, for example, a swap could be structured to exchange the total returns to the S&P 500 Index for the total returns to the Nikkei 225 Index. If, however, the investment objective is to gain access to the Japanese equity market, a swap can be structured to exchange LIBOR plus a spread for the total returns to the Nikkei 225 Index. This is an example of diversifying internationally and the cash flows can be denominated in either yen or dollars. The advantages of entering into an equity swap to obtain international diversification are that the investor exposure is devoid of tracking error, and the investor incurs no sales tax, custodial fees, withholding fees, or market impact associated with entering and exiting a market. This swap is the economic equivalent of being long the Nikkei 225 financed at a spread to LIBOR at a fixed exchange rate.

There are numerous applications of equity swaps, but all assume the basic structure outlined above. Investors can virtually swap any financial asset for the total returns to an equity index, a portfolio of stocks, or a single stock. Market makers are prepared to create structures that allow an investor to exchange the returns of any two assets. The schedule of cash flows exchanged is a function of the assets. For example, an investor wanting to outperform an equity benchmark may be able to accomplish this by purchasing a particular bond and swapping the cash flows for the S&P 500 total return minus a spread.

Equity swaps are a useful means of implementing an asset allocation strategy. One example is an asset swap of the S&P 500 total returns for the total returns to the DAX index. The investor can reduce U.S. equity exposure and increase German equity exposure through an equity swap, thereby avoiding the costs associated with cash market transactions.

SUMMARY

OTC options are privately negotiated contractual agreements between an investor and an issuing dealer. The structure of the option is completely flexible in terms of strike price, expiration, and payout features. OTC warrants are long-term options on equity indexes, basket of stocks, or an individual stock and have the same flexible structure capability as OTC options. Equity-linked debt is a debt instrument with principal or coupon payments linked to the performance of an established equity index, a basket of stocks, or a single stock. Equity swaps are similar in structure to interest rate or currency swaps. They are contractual agreements between two counterparties providing for the periodic exchange of a schedule of cash flows over a specified time period where at least one of the two payments is linked to the performance of an equity index, a basket of stocks, or a single stock.

OTC equity derivatives can provide investors with a means of lowering transaction costs, including commissions and market impact costs. In addition, OTC equity derivatives may be useful for tax management purposes by delaying capital gains. From the standpoint of pension funds, equity derivatives may provide a vehicle for reducing management fees and custodian fees. There are also a variety of legal and regulatory barriers that can be overcome using OTC equity derivatives. These features add to the payoff possibilities available from structured products. Investors can hedge any risk or assume any risk in ways that are only limited by their ability to characterize their desired financial objectives.

REFERENCES

Boyle, P., and Boyle, F. (2001). *Derivatives: The Tools That Changed Finance*. London: Risk Books.

Chance, D. M. (2004). *An Introduction to Derivatives*, 6th edition. Mason, OH: Thomson South-Western.

Collins, B., and Fabozzi, F. J. (1999). *Derivatives and Equity Portfolio Management*. Hoboken, NJ: John Wiley & Sons.

Dubofsky, D. A., and Miller, T. W. (2003). *Derivatives: Valuation and Risk Management*. New York: Oxford University Press.

Francis, J. C., Toy, W., and Whittaker, G. (1995). *The Handbook of Equity Derivatives*. Burr Ridge, IL: Irwin Professional Publishing.

Gastineau, G., and Kritzman, M. (1997). *Dictionary of Financial Risk Management*. Hoboken, NJ: John Wiley & Sons.

Hull, J. (2006). *Options, Futures, and Other Derivative Securities*. Upper Saddle River, NJ: Prentice Hall.

Kat, H. (2001). *Structured Equity Derivatives*. Chichester, England: John Wiley & Sons.

McLaughlin, R. M. (1999). *Over-the-Counter Derivative Products*. New York: McGraw-Hill.

Nelken, I. (ed.) (1996). *The Handbook of Exotic Options*. Burr Ridge, IL: Irwin Professional Publishing.

Rubinstein, M. (1991). Exotic options. Working paper, March.

CHAPTER 16

Volatility Derivatives

ROBERT WHALEY, PhD
Valere Blair Potter Professor of Management, The Owen Graduate School of Management, Vanderbilt University

Abstract: A relatively new stock index product is the *volatility derivative*. The Chicago Board Options Exchange (CBOE) contemplated launching trading volatility options as early as 1993. On January 19, 1998, the Deutsche Terminborse (DTB) became the first exchange in the world to list volatility futures. The CBOE launched trading of VIX futures on its CBOE Futures Exchange on March 26, 2004, with contracts on three-month realized variance being launched on May 18, 2004. The CBOE launched VIX options on February 24, 2006. It was not until the Long-Term Capital Management (LTCM) fiasco in late 1998 that the market finally began to recognize the value of trading stock market volatility as a separate asset class.

Keywords: volatility derivatives, volatility swap, realized future volatility, realized volatility derivative contract, variance swap, implied volatility derivatives contract

Volatility derivative contracts are written not only on stock indexes, but also interest rates, currencies, and commodities like crude oil. Prior to the advent of volatility derivatives, stock market volatility risk was managed using options written on the underlying index. The problem with doing so is that it is expensive. Options have two sources of price risk—risk associated with movements in the underlying index level and risk associated with movements in the market's perception of expected future volatility rate. The only way to isolate the volatility exposure is by trading the options and delta-hedging using the underlying index, index futures, and other index options.

This chapter describes volatility derivative contracts and their uses. We focus primarily on stock market volatility since stock market volatility contracts are the most actively traded. The discussion has two parts. First, we discuss realized volatility contracts and their applications, and then we turn to implied volatility contracts.

REALIZED VOLATILITY DERIVATIVE CONTRACTS

At the outset, we need to correct a misnomer. Industry has come to refer to realized volatility contracts as volatility swaps. A *volatility swap* is not a swap; it is a forward contract. They have traded in OTC markets for more than five years, and are now also exchange traded.

A volatility forward (or swap) is written on the *realized future volatility* of an asset (say, the S&P 500 index). At expiration, its payoff is based on the statistical formula for the annualized standard deviation of index return, that is,

$$\sigma_{\text{realized}} = \sqrt{\frac{\sum_{t=2}^{n_T}\left[\ln\left(\frac{S_t}{S_{t-1}}\right) - \text{mean}\right]^2}{n_T - 2}} \times \sqrt{\text{no, of time intervals in a year}} \tag{16.1}$$

where n_T is the number of price observations used in the computation, and S_t is the index level.

Volatility forwards are usually based on daily closing prices, however, since they are traded primarily in the OTC market, any frequency (e.g., hourly, weekly) is possible. The contract also specifies the source from which for the prices will be obtained. Volatility forwards are sometimes based on squared returns, and sometimes on squared deviations. Formula (16.1) shows squared deviations. The formula for squared returns is the special case where the mean term in the squared brackets of (16.1) is set equal to zero and the adjustment in the numerator is increased to $n_T - 1$. Finally, the volatility is annualized. For daily prices, the last term on the right-hand side is usually $\sqrt{252}$, that is, the square root of the typical number of business days in a year. For weekly prices, the last term is $\sqrt{52}$.

The value represented by formula (16.1) is the price of the asset underlying the forward contract at expiration. The only difference is the underlying asset is not tradable; it is simply a computation of realized volatility. At inception, the buyer and seller agree to a fixed delivery price (quoted as an annualized volatility), σ_X, on the expiration date, T. As expiration approaches, the forward's settlement price becomes more and more certain because some of the prices used in (16.1) have been realized already. On the last day before expiration, only the index level on expiration day remains unknown. Upon settlement, the buyer receives

$$\text{Notional} \times (\sigma_{\text{realized}} - \sigma_X) \tag{16.2}$$

that is, the notional amount of the swap times the difference between the realized and contracted volatility. The seller receives the opposite amount. Sometimes the volatility derivatives are written on the square of volatility, or variance. The buyer of a *variance swap* receives the payoff,

$$\text{Notional} \times (\sigma^2_{\text{realized}} - \sigma^2_X) \tag{16.3}$$

Illustration of the Computation of the Settlement Price of a Realized Volatility Swap

Suppose that on Friday, August 1, 2003, an investor bought a 13-week volatility forward from an OTC derivatives dealer. Its price was 0.12, and its notional amount was \$100 million. The Friday closing index levels over the period were as follows:

Friday Close	S&P 500 Index
20030801	980.15
20030808	977.59
20030815	990.67
20030822	993.06
20030829	1008.01
20030905	1021.39
20030912	1018.63
20030919	1036.30
20030926	996.85
20031003	1029.85
20031010	1038.06
20031017	1039.32
20031024	1028.91
20031031	1050.71

The first step in calculating the settlement price and settlement proceeds is to compute the weekly returns. Next compute the mean weekly return, and the squared returns and deviations. Compute the sum of squares and the annualized volatility. To annualize weekly returns, use the factor, $\sqrt{52}$. The calculations are shown in Table 16.1. The cash settlement proceeds are \$1.27 million

Table 16.1 Computations for Settlement Price and Settlement Proceeds for a Realized Volatility Swap

Friday Close	S&P 500 Index	S&P 500 Return	Squared Returns	Squared Deviations
20030801	980.15			
20030808	977.59	−0.00262	0.00000684	0.00006340
20030815	990.67	0.01329	0.00017665	0.00006310
20030822	993.06	0.00241	0.00000581	0.00000863
20030829	1008.01	0.01494	0.00022327	0.00009206
20030905	1021.39	0.01319	0.00017388	0.00006145
20030912	1018.63	−0.00271	0.00000732	0.00006485
20030919	1036.30	0.01720	0.00029577	0.00014044
20030926	996.85	−0.03881	0.00150634	0.00195002
20031003	1029.85	0.03257	0.00106068	0.00074097
20031010	1038.06	0.00794	0.00006305	0.00000672
20031017	1039.32	0.00121	0.00000147	0.00001709
20031024	1028.91	−0.01007	0.00010134	0.00023759
20031031	1050.71	0.02097	0.00043958	0.00024395
	Mean	0.00535		
	Total		0.00033850	0.00030752
	Annualized volatility		0.13267	0.12646
	Notional amount		\$100,000,000	\$100,000,000
	Forward price		0.120	0.120
	Cash settlement value		\$1,267,275	\$645,649

Table 16.2 Selected Terms of S&P 500 Three-Month Variance Future Contract

Exchange	CBOE Futures Exchange (CFE)
Ticker symbol	VT
Contract unit	\$50 per variance point
Tick size	0.5 of one variance point
Tick value	\$25
Trading hours	8:30 AM to 3:15 PM CST
Contract months	Up to four contract months on the March cycle (Mar., Jun., Sep., Dec.)
Last day of trading	Close of trading on business day before final settlement date.
Final settlement date	Third Friday of contract month.
Final settlement price	Final settlement price is based on the standardized calculation of the realized variance of the S&P 500. This calculation uses continuously compounded daily returns for a three-month period assuming a mean daily return of zero. The calculated variance is then annualized assuming 252 business days per year. The final settlement price is this annualized, calculated variance multiplied by 10,000.

for the squared returns contract and \$.65 million for squared deviations. The difference is unusually large because the S&P 500 index level rose abnormally during this 13-week period, at least relative to historical standards. The rate of return of the S&P 500 index was about 7.2%—nearly 30% on an annualized basis.

CBOE Futures Exchange Realized Volatility Futures Contract

The CBOE Futures Exchange (CFE) launched its three-month realized volatility futures contract on May 18, 2004. The CFE is an all-electronic exchange that was created by the Chicago Board Options Exchange (CBOE) in March 2004. The CFE's realized volatility contract is based on S&P 500 return variance rather than return standard deviation, and its product specifications are provided in Table 16.2 The contract denomination is \$50 per variance point. A price quotation of 633.50, for example, means the contract value is \$31,675. Up to four contracts may trade simultaneously. The contracts are on the March quarterly expiration cycle (March, June, September, December). The final settlement date is the third Friday of the contract month. Trading stops at the close on the preceding business day.

The final settlement price is a variance number and assumes the mean return is zero. Hence, the realized volatility formula (16.1) becomes

$$\sigma^2_{\text{realized}} = \frac{\sum_{t=1}^{n_d} \left[\ln\left(\frac{S_t}{S_{t-1}} \right) \right]^2}{n_e - 1} \times 252 \tag{16.4}$$

where n_a is the actual number of trading days in the three-month interval, and n_e is the expected number of days in the three-month interval. Normally, n_a and n_e are equal. In the event of a market disruption during the contract's life, however, n_a will be less than n_e. Generally speaking, a "market disruption event," as determined by the CFE, occurs when trading on the primary exchanges of a significant number of S&P 500 stocks is suspended or limited in some way or when the primary exchange on which index stocks unexpectedly closes early (or does not open) on a particular day. For each market disruption event, the value of n_a is reduced by one.

Volatility versus Variance Contracts

The industry has come to define volatility as the standard deviation of the natural logarithm of the price ratios (This is consistent with the Black-Scholes model's use of continuously compounded returns). If the forward is defined in terms of variance (that is, volatility squared) rather than volatility, the payoff structure is quite different. Consider Figures 16.1 and 16.2 which plot the payoffs of a volatility forward contract versus a variance forward contract for long and short positions. Since the horizontal axis is defined in terms of volatility, its terminal payoffs are a linear function of volatility. The variance forward, however, is nonlinear. The long variance position (the dotted line in Figure 16.1) has convexity. As volatility falls, the terminal payoff of the long variance position decreases, but at a decreasing rate. At the same time, as volatility increases, the terminal payoff of the long variance forward increases at an increasing rate. Indeed, the variance payoffs loosely resemble a long call position, while the variance payoffs of the short variance futures resemble a short call position.

Expected Return/Risk Management Applications

At first blush, the volatility forward contract seems to be purely a speculative instrument. Traders who believe future volatility will be high relative to the forward price will go long the swap, and those who believe that the market will be very calm will go short. But, the hedging possibilities using realized volatility forwards are many. In the normal course of operation, for example, some market participants become inherently short volatility. Consider the ill-fated index option strategy of Long-Term Capital Management (LTCM). Because index option implied volatilities were as high as they had been anytime since the October 1987 market crash, LTCM sold both index calls and puts with the belief that implied volatility would return to normal levels. Unfortunately, a problem arose when implied volatility continued to rise and their positions were marked-to-market. The cash drain was enormous. Buying realized volatility forwards would have hedged this exposure, at least in part. The same is true for index option market makers who are short market volatility as a result of selling index puts to portfolio

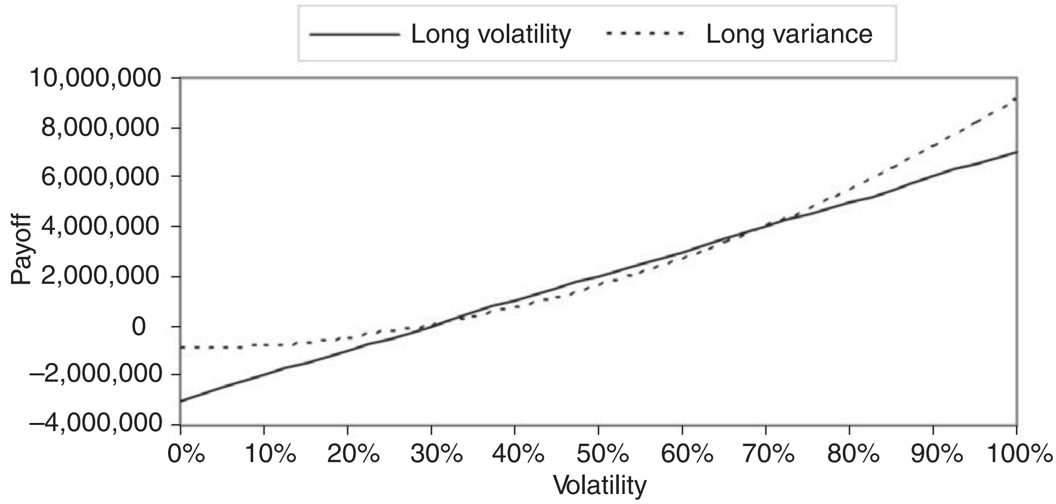


Figure 16.1 Payoff Structure of Volatility and Variance Forward Contracts: Long Positions

insurers (On a typical day. S&P 500 put option volume (and open interest) is nearly double that of S&P 500 calls).

Another hedging possibility is for risk arbitrageurs. Immediately after a merger is announced, risk arbitrageurs step in and buy shares of a target firm and sell the shares of bidder. Because the probability that the merger will be successful is not known, the prices of the target and the bidder will not fully reflect the terms of the offer. If the merger is successful, the spread between the prices will converge. Before the deal is consummated, however, market volatility may increase, making the merger less likely, thereby causing the spread to widen. Buying a realized volatility forward contract can hedge this type of risk exposure.

Yet another application is for individuals or portfolio managers who attempt to track some sort of benchmark index. During periods of high volatility, the portfolio may require more frequent rebalancing and greater transaction cost expenses. Again, buying a realized forward contract on volatility can hedge this exposure.

IMPLIED VOLATILITY DERIVATIVES CONTRACTS

The CFE also lists a futures contract written on the implied return volatility of the S&P 500 index. The CBOE Marker Volatility Index or VIX is constructed in such a way that it represents the implied volatility of an at the money S&P 500 index option with exactly thirty calendar days to expiration. It is sometimes called the "investor fear gauge" because it is set by investors and expresses their consensus view about expected future stock market volatility. The specific details of its construction are contained in appendix to this chapter. What is interesting about its construction is that the index can be created using

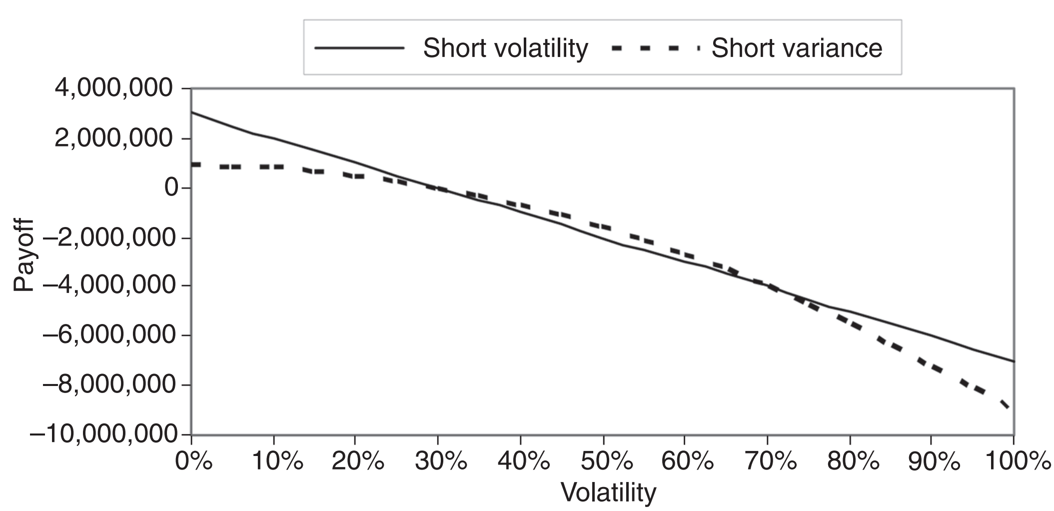


Figure 16.2 Payoff Structure of Volatility and Variance Forward Contracts: Short Positions

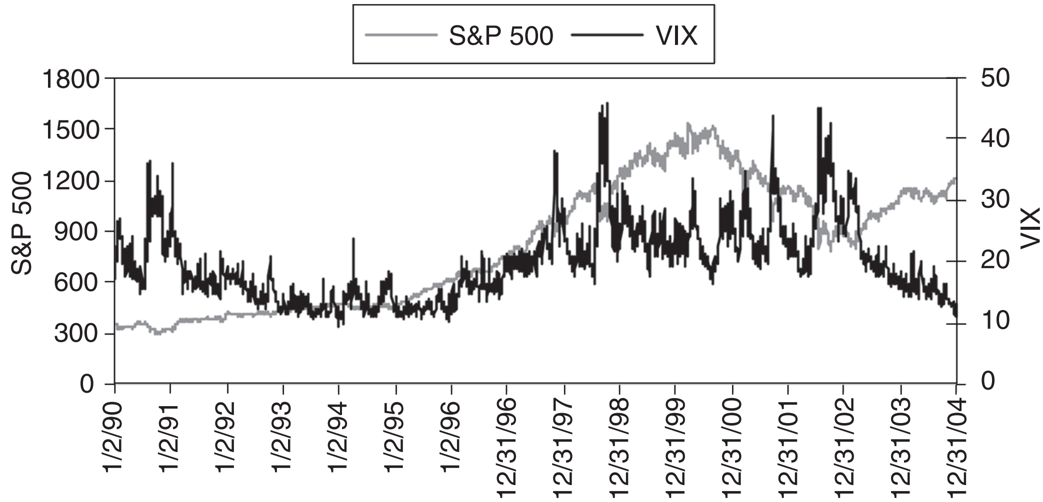


Figure 16.3 Daily Levels of the S&P 500 Index and the VIX during the Period January 1990 through December 2004

a static portfolio of SPX options. This is important since arbitrage between the VIX futures and the underlying VIX index promotes liquidity in both markets.

The relation between the movements of the VIX and the movements of the S&P 500 index are important to understand. Figure 16.3 shows the daily levels of the S&P 500 index and the VIX during the period January 1990 through December 2004. A number of interesting patterns appear.

First, note that the VIX level (that is, the dark line) is more jagged than the S&P 500 index level. What this means is that the volatility of the volatility of the S&P 500 index is greater than the volatility of the index itself. Time-series variation in the expected volatility of stock indexes has been documented in a number of studies. Day and Lewis (1992), for example, demonstrate that the expected variance of the S&P 100 index follows a mean-reverting process. They also show that implied volatilities from S&P 100 index options (OEX) explain a significant amount of the changes in expected variance. In a related paper, Fleming (1998) finds that OEX implied volatilities are good (but not perfect) forecasts of future volatility.

Second, there tends to be an inverse relation between the level of VIX and the level of the S&P 500 index—as the stock market goes up, volatility tends to fall. During 2003 and 2004, for example, the S&P 500 is systematically increasing while the level of VIX falls. Third, the inverse correlation is not perfect. During 1996 and 1997, for example, the level of market volatility is increasing while the stock market is also increasing. All of these phenomenon contribute to making futures contracts on the VIX a potentially new and useful expected return/risk management tool, as we will see in the illustration that follows.

The CFE VIX futures contract has, as its underlying, the VIX. The futures contract specifications are given in Table 16.3. Its denomination is \$100 times the increased-value VIX. The "increased-value VIX" (ticker symbol VBI) is simply the level observed in the marketplace times ten ($VBI = VIX \times 10$). The tick size of the contract is 0.1 of one VBI point or \$10. The available contract months

Table 16.3 Selected Terms of Market Volatility Index (VIX) Futures Contract

Exchange	CBOE Futures Exchange (CFE)
Ticker symbol	VX
Contract unit	\$100 times Increased-Value VIX[a]
Tick size	0.1 of one VBI point
Tick value	\$10
Trading hours	8:30 AM to 3:15 PM CST
Contract months	Two near-term contract months plus two contract months on the February quarterly cycle (Feb., May, Aug., and Nov.)
Expiration day	Third Friday of the contract month.
Last day of trading	Tuesday prior to the third Friday of the expiring month.
Final settlement date	Wednesday prior to the third Friday of the expiring month.
Final settlement price	Cash settled. Final settlement price for VIX futures shall be 10 times a Special Opening Quotation (SOQ) of VIX calculated from the options used to calculate the index on the settlement date. The opening price for any series in which there is no trade shall be the average of that option's bid price and ask price as determined at the opening of trading. The final settlement price will be rounded to the nearest 0.10.

[a] Increased-Value VIX (VBI) is 10 times the VIX index level.

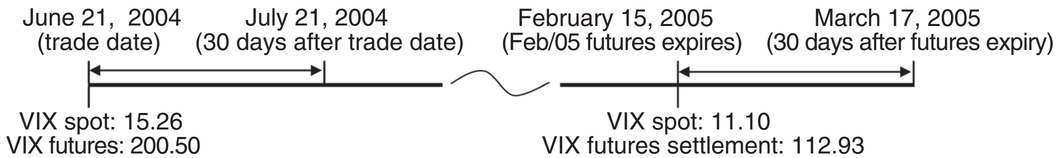


Figure 16.4 VIX index and February 2005 VIX futures assuming futures was traded on June 21, 2004

include the two near-term contract months plus two contract months on the February quarterly cycle (February, May, August, and November). The expiration day is the third Friday of the contract month, although trading stops on the preceding Tuesday. The contract is cash-settled on the Wednesday preceding the third Friday, at a special opening quotation (SOQ).

To understand the distinction between the VIX and the VIX futures, consider Figure 16.4. The figure assumes that we traded the February 2005 VIX futures on June 21, 2004. At the close on June 21, the VIX level was 15.26, and the Feb/05 VIX was at 200.50. Recall that the futures is scaled by 10, so the futures price represents a volatility rate of 20.05%. As the figure illustrates, the level of VIX reflects the market's expected future volatility over the next thirty calendar days (from June 21 to July 21, 2004), while the VIX futures reflects the expected future market volatility during a 30-calendar day period beginning when the Feb/05 futures contract expires and ending 30 calendar days later (February 15 to March 17, 2005). In other words, the VIX futures is a one-month forward volatility rate that begins some time in the future. As it turns out, on February 15, 2005, the Feb/05 VIX was cash settled in the morning at ten times the spot level of VIX, 112.93. By the end of the day, the level of VIX had fallen to 11.10.

The convergence of the Feb/05 VIX futures to the VIX index over the period June 21, 2004 through February 15, 2005 is shown in Figure 16.5. The VIX is multiplied by 10 to put it on the same scale as the futures price. Where the two prices were about 50 points apart in June 2004, they slowly and steadily converged to the same level at expiration. Figure 16.6 shows the open interest of the Feb/05 VIX futures contract. In June 2004, the Feb/05 futures was a distant contract maturity and did not have much open interest. Through time, as the shorter contract maturities expired, the open interest in the Feb/05 contract rose, reaching a peak above 6,000 contracts in January 2005. Like most cash-settled futures, open interest remained high until contract settlement (Futures contracts with physical delivery are generally unwound before contract maturity to avoid the costs of transportation. With cash settlement, no such costs exist.)

To get a sense for how VIX futures contracts are priced, let us assume that we are considering the variance of S&P 500 index returns over the next 60 calendar days (that is, two months). If the returns of the index are independent through time, we can write

$$\bar{\sigma}^2_{1-60}\left(\frac{60}{365}\right) = \bar{\sigma}^2_{1-30}\left(\frac{30}{365}\right) + \bar{\sigma}^2_{31-60}\left(\frac{60-30}{365}\right) \tag{16.5}$$

In (16.5), $\bar{\sigma}^2_{1-60}$ and $\bar{\sigma}^2_{1-60}$ can be considered spot rates of variance, that is, the expected variance rates over the next 30 calendar days and 60 calendar days, respectively. The term,

$$\bar{\sigma}^2_{31-60}\left(\frac{30}{365}\right)$$

however, is a forward variance, that is, the average variance rate that we can expect to observe over a 30-day period beginning 30 days from now. To determine the forward volatility rate, we can rearrange (16.5) to yield

$$\bar{\sigma}_{31-60} = \sqrt{\frac{\bar{\sigma}^2_{1-60}\left(\frac{60}{365}\right) - \bar{\sigma}^2_{1-30}\left(\frac{30}{365}\right)}{\frac{60-30}{365}}} \tag{16.6}$$

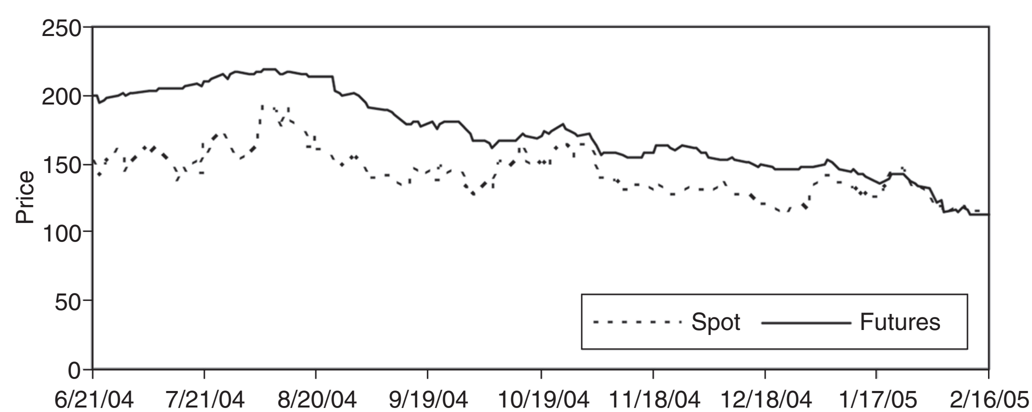


Figure 16.5 Convergence of February 2005 VIX Futures Price to VIX Spot Price (10 times observed VIX) over the Period June 21, 2004, through February 16, 2005

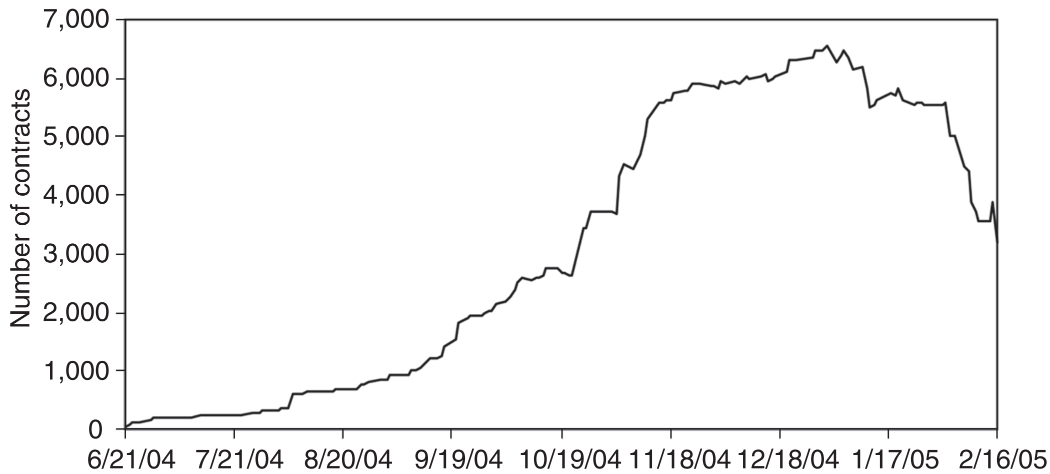


Figure 16.6 Open Interest of February 2005 VIX Futures over Its Life (June 21, 2004, through February 16, 2005).

Equation (16.6) provides us with the insight we need in understanding how to value the VIX futures. The rate on the left-hand side of (16.6) can be thought of as the VIX futures price. In order to estimate its value, we need to know the two variance rates in the numerator on the right-hand side. One way to get these values is to request quotes on 30-day and 60-day variance forwards from an OTC swap dealer. Another is to use S&P 500 index options to imply the variance rates of 30- and 60-day intervals. Note that, in this particular instance, the rate $\bar{\sigma}_{1-30}$ is also the current level of the VIX because the forward period begins in exactly 30 calendar days. Whether the forward price exceeds the current spot price, as it did for the Feb/05 VIX futures, depends upon whether the term structure of realized variance swaps is upward- or downward-sloping. In an upward-sloping environment, the forward price will exceed the spot price, and vice versa. Given that volatility tends to follow a mean-reverting process, the forward rate will be equal to the spot rate on average.

Illustration of How to Estimate VIX Futures Price

Suppose that an investor is given the assignment of determining the fair value of the VIX futures where the contract expires in exactly 15 days. The investor has contacted an OTC derivatives dealer, and the dealer quoted the investor rates of 400 and 420 on 15-day and 45-day realized variance swaps.

The quoted realized variance swap rates straddle the forward period corresponding to the VIX futures. Hence, the fair value of the VIX futures can be determined by

$$\text{VIX futures} = \sqrt{\frac{420\left(\frac{45}{365}\right) - 400\left(\frac{15}{365}\right)}{\frac{45-15}{365}}} = 20.74$$

expressed in VIX points, or 207.40 expressed in VBI points.

Expected Return/Risk Management Applications

Exchange-traded futures on volatility also offer a number of new expected return/risk management strategies. VIX futures can be regarded as a new asset class and can potentially improve the expected return/risk opportunity set. Indeed, because the returns of the S&P 500 portfolio and the returns of the VIX are inversely correlated, the diversification effects can well surpass other strategies such as diversifying across countries (Stock returns in different countries tend to be positively correlated. A major economic shock in one market is usually felt across markets.) VIX futures can also be used to manage individual stock volatility. Individual stock volatility can be thought of as the sum of two components: stock market volatility and firm-specific volatility. Market volatility products allow investors to hedge the stock market volatility component to develop selected exposures in the idiosyncratic risk of individual stocks (Whaley [1993] demonstrates that, for large market capitalizaton firms, nearly 50% of movement in individual stock volatility rate is explained by movements in the market volatility rate.)

One caveat is necessary, however. Many stock market volatility hedging needs are long term. The VIX futures contract, however, is on the stock market volatility rate in a 30 day forward period. Consequently, in order to effectively hedge a short volatility position over a long period of time, it may be necessary to buy a strip of VIX futures so that the volatility rate over the entire hedge interval may be captured.

CBOE VIX Options

The Chicago Board Options Exchange (CBOE) launched VIX option contracts on Friday, February 24, 2006. Like the VIX futures, the CBOE's VIX options contract has, as its underlying, the VIX. The option contract specifications are given in Table 16.4. Its ticker symbol is "VIX," and its denomination is $100 times the level of the CBOE's

Table 16.4 Selected Terms of Market Volatility Index (VIX) Option Contract

Exchange	Chicago Board Options Exchange (CBOE)
Ticker symbol	VIX
Contract unit	100 times CBOE Market Volatility index
Exercise price increments	2-1/2 point increments
Exercise style	European
Tick size	0.05 point up to $3 premiums; 10 point over $3
Tick value	$5; $10
Trading hours	8:30 AM to 15:15 PM CST
Contract months	Two near-term contract months plus two contract months on the February quarterly cycle (Feb., May, Aug., and Nov.)
Expiration day	Wednesday that is 30 days prior to the third Friday of the calendar month immediately following the expiring month.
Last day of trading	Tuesday prior to expiration date each month.
Final settlement price	Cash settled. Exercise settlement value shall be a Special Opening Quotation (SOQ) of VIX calculated from the sequence of opening prices of options used to calculate the index on the settlement date. The opening price for any series in which there is no trade shall be the average of that option's bid price and ask price as determined at the opening of trading. Exercise will result in the delivery of cash on the business day following expiration. The exercise settlement amount is equal to the difference between the exercise-settlement value and the exercise price of the option times $100.

Market Volatility index. The tick size (value) of the contract is 0.05 ($5) for option premiums below $3.00 ($300), and 0.10 ($100) for premiums greater than $3 ($300). The available contract months include the two near-term contract months plus two contract months on the February quarterly cycle (February, May, August, and November). The expiration day is the Wednesday that is 30 days before the third Friday of the calendar month following the expiring month. Trading stops on the Tuesday before the expiration day. The contract is cash-settled on the day after the expiration at a special opening quotation (SOQ). The exercise settlement amount equals the difference between the exercise-settlement value and the exercise price times $100.

SUMMARY

There are two types of volatility derivative contracts—contracts on realized volatility and contracts on volatility implied by index option prices. In this chapter we describe the different volatility contract specifications and explain how the CBOE's Market Volatility Index (VIX) can be constructed from a portfolio of S&P 500 index options. We also note that volatility derivatives can be used as an alternative investment in an asset allocation framework.

APPENDIX: CONSTRUCTION OF THE CBOE'S MARKET VOLATILITY INDEX (VIX)

The purpose of this appendix is to describe the algorithm with which the CBOE's Market Volatility Index (VIX) is computed.

The procedure for calculating VIX is described in CBOE (2003). The theory underlying the procedure is based on the Breeden and Litzenberger (1978) result that the probability density function of asset price can be inferred from the prices of options written on that asset, where the options have a common expiration date and continuum of exercise prices. Demeterfi, Derman, Kamal, and Zou (1999) apply this result in a discretized form to arrive at an equation for the volatility of asset price.

The VIX is the expected future volatility of the S&P 500 index over the next 30 days. It is an implied volatility in that it is based on S&P 500 index option prices. Unlike the implied volatilities from the Black-Scholes-Merton (BSM) option valuation model, however, the VIX does not depend on a particular return distribution (The BSM model assumes a log-normal asset price distribution at the option's expiration.)

To compute the VIX, an eight-step procedure is used.

Step 1: Collect relevant information. The information needed to compute the VIX is (1) the bid/ask price quotes of all nearby and second nearby call and put options traded on the S&P 500 index; and (2) the risk-free interest rate corresponding to each expiration date. For each option series, the bid/ask midpoint is computed. The difference between the call midpoint and put midpoint at each exercise price is computed.

Step 2: Compare the time to expiration in minutes and then years from the current time until option expiration. The time to expiration in minutes is the sum of three components (Time to expiration is computed in minutes to conform to industry practice.) First, we must compute the number of minutes from the current time until midnight on the same day. We next compute the number of minutes from midnight today until midnight on the day before expiration. Finally, we must compute the number of minutes from midnight on the day before expiration until cash settlement at the open on expiration day. The last number is, of course, a constant. The time of cash settlement is at 8:30 AM on expiration day. The number of minutes from midnight on the day before expiration until the time of expiration is therefore

$$8.5 \text{ hours} \times 60 \text{ minutes per hour} = 510 \text{ minutes}$$

The first and second components depend upon the time of day and the number of days to expiration, respectively.

To illustrate, assume that we are computing the level of VIX at 8:38 AM (CST) on October 6, 2003. The number of minutes to midnight on October 6 is

$$22 \text{ minutes} + 15 \text{ hours} \times 60 \text{ minutes per hour}$$
$$= 922 \text{ minutes}$$

On October 6, 2003, the nearby and second expirations of the S&P 500 index options are the October 17, 2003, and November 21, 2003, respectively, and the number of days to expiration are 12 and 47 days inclusive of the current date and the expiration date. The current date and expiration date are already incorporated, however. The number of minutes until midnight on the current date is 922, and the number of minutes from midnight on the day before expiration until time of expiration on the expiration day is 510. Thus, we reduce the number of days to expiration for the nearby and second nearby expirations to 10 and 45 and compute the number of minutes. With 1,440 minutes in each 24-hour day, the number of minutes for the second component of the nearby contract is

$$10 \text{ days} \times 1,440 \text{ minutes per day} = 14,400$$

and the number of minutes for the second component of the second nearby contract is

$$45 \text{ days} \times 1,440 \text{ minutes per day} = 64,800$$

The total numbers of minutes for the two contract expirations are therefore

$$\text{Nearby contract} : 922 + 14,400 + 510 = 15,832$$

and

$$\text{Second nearby contract} : 922 + 64,800 + 510 = 66,232$$

The times to expiration in years are then computed as

$$T_1 = 15,832/525,600 = 0.0301217656$$

and

$$T_2 = 66,232/525,600 = 0.1260121766$$

where 525,600 in the number of minutes in a calendar year (that is, 1,440 minutes per day times 365 days).

Step 3: Compute the interest accumulation factor for each option expiration. The interest accumulation factor is defined as the terminal amount that $1 will accumulate to by the option's expiration if invested at the risk-free rate of interest. On October 6, 2003, the risk-free rate corresponding to the nearby expiration was 0.920% on an annualized basis, and the risk-free rate corresponding to the second nearby expiration was 0.850% (On this particular day, the yield curve of the risk-free rate was inverted at short maturities.) The accumulation factors for the nearby and second nearby contracts were

$$e^{r_1 T_1} = e^{0.00920(0.03012177)} = 1.0002772$$

and

$$e^{r_2 T_2} = e^{0.00850(0.12601218)} = 1.0010717$$

respectively.

Step 4: Identify the at-the-money options for each option expiration. To identify the at-the money options for each expiration, we must first compute the bid/ask midpoints for all calls and puts with the nearby and second nearby contract expirations. This is shown in Tables 16.5 and 16.6. For each exercise price for which a call price and put price are available, compute the absolute difference between the call price and put price. Note that the calls and puts with zero bid prices are excluded for consideration. Such options appear in bold face. The exercise price with the lowest absolute difference is defined as the at-the-money option. On October 6, 2003, the nearby at-the-money exercise price is 1030 (as is shown in Table 16.5 and the second nearby exercise price is 1035 (as is shown in Table 16.6).

Step 5: Compute the forward index level for each contract expiration. With the identity of the at-the-money options known, we compute the implied forward index level using the forward value version of put-call parity, that is,

$$F_i = X_i + e^{r_i T_i}(G_i - P_i)$$

For the nearby at-the-money options, the forward price is

$$F_1 = 1030 + 1.0002771586449(13.500 - 12.400) = 1031.10$$

For the second nearby at-the-money options, the forward price is

$$F_2 = 1025 + 1.0010716773370(29.400 - 26.600) = 1029.99$$

Step 6: Identify the option series used in the computation of the VIX. In computing the VIX, only the prices of out-of-the-money calls and puts are used. To distinguish between in-the-money and out-of-the-money options, the exercise price just below the implied forward price ($X_{i,0}$) is used. The out-of-the-money calls are those with exercise prices greater than or equal $X_{i,0}$, and the out-of-the-money puts are those with exercise prices less than or equal to $X_{i,0}$. If any of these option series have a bid price equal to zero, they are eliminated from consideration (In the event that the bid prices of two calls (puts) at adjacent exercise prices are equal to zero, all call (put) option series with higher (lower) exercise prices are eliminated even though they may nonzero bid pries). For the nearby and second nearby option series in the illustration, the exercise prices just below the forward index levels are $X_{1,0} = 1030$ and $X_{2,0} = 1035$. Since this procedure identifies two options (a call and a put) at exercise price $X_{i,0}$, the arithmetic average of the call and put prices is used.

Step 7: Compute the implied variance for each contract expiration. The formula for computing the implied variance for the nearby contract is

$$\sigma_1^2 = \frac{2}{T_1} \sum_{i=1}^{n_1} \frac{\Delta X_{1,i}}{X_{1,i}^2} e^{r_1 T_1} O(X_{1,i}) - \frac{1}{T_1}\left(\frac{F_1}{X_{1,0}} - 1\right)^2$$

where T_1 is the nearby contract month's time to expiration expressed in years, n_1 is the number of out-of-the-money option series for the nearby contract month, $X_{1,i}$ is the exercise price of the ith option, r_1 is the interest rate corresponding to option's expiration date, F_1 is the forward index level implied by the at-the-money call and put prices, $O(X_{1,i})$ is the bid/ask price midpoint of the

Table 16.5 Nearby S&P 500 Index Option Prices Used in the Computation of the VIX on October 6, 2003, at 8:38 AM (CST)

Nearby Contract Expiration: 10/17/2003							
	Call Price Quotes			Put Price Quotes			
Exercise Price	Bid	Ask	Midpoint	Bid	Ask	Midpoint	Absolute Difference
725	304.10	307.10	305.600	0.00	0.50		
750	279.10	282.10	280.600	0.00	0.50		
775	254.10	257.10	255.600	0.00	0.50		
800	229.10	232.10	230.600	0.00	0.40		
825	204.10	207.10	205.600	0.00	0.25		
850	179.10	182.10	180.600	0.05	0.20	0.125	180.475
875	154.20	157.20	155.700	0.10	0.20	0.150	155.550
890	139.20	142.20	140.700	0.00	0.50		
900	129.30	132.30	130.800	0.20	0.40	0.300	130.500
910	119.40	122.40	120.900	0.00	0.50		
915	114.40	117.40	115.900	0.05	0.50	0.275	115.625
925	104.50	107.50	106.000	0.25	0.60	0.425	105.575
930	100.00	102.60	101.300	0.30	0.70	0.500	100.800
935	95.10	97.10	96.100	0.50	0.60	0.550	95.550
940	90.20	92.20	91.200	0.45	0.90	0.675	90.525
945	85.30	87.30	86.300	0.40	0.90	0.650	85.650
950	80.40	82.40	81.400	0.65	1.00	0.825	80.575
955	75.80	77.80	76.800	0.75	1.10	0.925	75.875
960	70.90	72.90	71.900	0.80	1.30	1.050	70.850
970	61.30	63.30	62.300	1.10	1.60	1.350	60.950
975	56.50	58.50	57.500	1.50	1.90	1.700	55.800
980	51.80	53.80	52.800	1.70	2.20	1.950	50.850
985	47.20	49.20	48.200	2.00	2.50	2.250	45.950
990	42.60	44.60	43.600	2.30	3.10	2.700	40.900
995	38.20	40.20	39.200	3.00	3.70	3.350	35.850
1005	29.50	31.50	30.500	4.40	5.20	4.800	25.700
1010	25.50	27.50	26.500	5.40	6.40	5.900	20.600
1015	21.80	23.80	22.800	6.60	7.60	7.100	15.700
1020	18.50	19.50	19.000	8.00	9.00	8.500	10.500
1025	16.00	16.90	16.450	9.90	10.90	10.400	6.050
1030	13.00	14.00	13.500	11.60	13.20	12.400	1.100
1035	10.10	11.50	10.800	14.00	15.60	14.800	4.000
1040	8.00	9.00	8.500	16.80	18.40	17.600	9.100
1045	6.10	7.00	6.550	19.90	21.50	20.700	14.150
1050	4.70	5.50	5.100	23.20	25.20	24.200	19.100
1055	3.40	4.20	3.800	26.90	28.90	27.900	24.100
1060	2.50	3.30	2.900	30.90	32.90	31.900	29.000
1065	1.90	2.40	2.150	35.20	37.20	36.200	34.050
1070	1.30	1.80	1.550	39.60	41.60	40.600	39.050
1075	0.90	1.40	1.150	44.20	46.20	45.200	44.050
1100	0.10	0.20	0.150	68.60	70.60	69.600	69.450
1115	0.00	0.50		83.40	85.40	84.400	84.400
1125	0.00	0.15		93.40	95.40	94.400	94.400
1135	0.00	0.50		102.90	105.90	104.400	104.400
1150	0.00	0.10		117.80	120.80	119.300	119.300
1175	0.00	0.50		142.80	145.80	144.300	144.300
1200	0.00	0.50		167.80	170.80	169.300	169.300
1225	0.00	0.50		192.80	195.80	194.300	194.300
1250	0.00	0.50		217.80	220.80	219.300	219.300
1275	0.00	0.50		242.80	245.80	244.300	244.300
1300	0.00	0.50		267.80	270.80	269.300	269.300
1325	0.00	0.50		292.80	295.80	294.300	294.300
1350	0.00	0.50		317.70	320.70	319.200	319.200
1375	0.00	0.50		342.70	345.70	344.200	344.200

Table 16.6 Second Nearby S&P 500 Index Option Prices Used in the Computation of the VIX on 6, 2003, at 8:38 AM (CST)

Second Nearby Contract Expiration: 11/21/2003							
Exercise Price	Call Price Quotes			Put Price Quotes			Absolute Difference
	Bid	Ask	Midpoint	Bid	Ask	Midpoint	
600	427.70	430.70	429.200	0.00	0.30		
625	402.70	405.70	404.200	0.00	0.50		
650	377.80	380.80	379.300	0.00	0.50		
675	352.80	355.80	354.300	0.00	0.50		
700	327.90	330.90	329.400	0.00	0.50		
725	303.00	306.00	304.500	0.00	0.50		
750	278.10	281.10	279.600	0.00	0.50		
775	253.30	256.30	254.800	0.10	0.60	0.350	254.450
800	228.50	231.50	230.000	0.30	0.80	0.550	229.450
825	203.90	206.90	205.400	0.60	1.10	0.850	204.550
850	179.40	182.40	180.900	1.10	1.60	1.350	179.550
875	155.00	158.00	156.500	1.70	2.20	1.950	154.550
895	135.80	138.80	137.300	2.30	3.10	2.700	134.600
900	131.20	134.20	132.700	2.60	3.30	2.950	129.750
925	107.70	110.70	109.200	3.90	4.70	4.300	104.900
950	85.40	87.40	86.400	6.00	7.00	6.500	79.900
975	64.00	66.00	65.000	9.50	10.50	10.000	55.000
980	60.00	62.00	61.000	10.20	11.80	11.000	50.000
985	56.00	58.00	57.000	11.20	12.80	12.000	45.000
990	52.10	54.10	53.100	12.30	13.90	13.100	40.000
995	48.30	50.30	49.300	13.50	15.10	14.300	35.000
1005	41.20	43.20	42.200	16.80	17.90	17.350	24.850
1010	37.80	39.80	38.800	17.90	19.50	18.700	20.100
1015	34.50	36.50	35.500	19.70	21.30	20.500	15.000
1020	31.40	33.40	32.400	21.30	23.30	22.300	10.100
1025	28.40	30.40	29.400	23.30	25.30	24.300	5.100
1035	22.90	24.90	23.900	27.90	29.90	28.900	5.000
1050	16.20	17.80	17.000	35.90	37.90	36.900	19.900
1060	12.40	14.00	13.200	42.10	44.10	43.100	29.900
1065	10.70	12.30	11.500	45.40	47.40	46.400	34.900
1070	9.50	10.00	9.750	48.90	50.90	49.900	40.150
1075	8.20	9.20	8.700	52.50	54.50	53.500	44.800
1080	7.00	8.00	7.500	56.30	58.30	57.300	49.800
1100	3.50	4.30	3.900	73.00	75.00	74.000	70.100
1125	1.40	1.90	1.650	95.70	97.70	96.700	95.050
1150	0.60	0.90	0.750	119.20	122.20	120.700	119.950
1175	0.00	0.50		143.80	146.80	145.300	145.300
1200	0.00	0.50		168.60	171.60	170.100	170.100
1225	0.00	0.50		193.50	196.50	195.000	195.000
1250	0.00	0.50		218.40	221.40	219.900	219.900
1275	0.00	0.50		243.40	246.40	244.900	244.900

nearby option with an exercise price of $X_{1,i}$ and $X_{1,0}$ is the exercise price just below the implied nearby forward price. The summation term also includes the at-the-money options. For the at-the-money options, the average of the call and put midpoints is used as $O(X_{1,i})$. Finally, the term $\Delta X_{1,i}$ is the average of the exercise prices that straddle option i's exercise price. At the highest and lowest exercise prices, $\Delta X_{1,i}$, is the absolute difference between option i's exercise price and the adjacent exercise price. The last term on the right-hand side is called the displacement factor.

The same procedure is used to compute the second nearby implied variance,

$$\sigma_2^2 = \frac{1}{T_2}\sum_{i=1}^{n_1} \frac{\Delta X_{2,i}}{X_{2,i}^2} e^{r_2 T_2} O(X_{2,i}) - \frac{1}{T_2}\left(\frac{F_2}{X_{2,0}} - 1\right)^2$$

To illustrate the mechanics of these computations, first compute the values of the last term on the right-hand side (that is, the displacement factors) of the nearby and second nearby contracts. For the nearby contract,

$$\frac{1}{T_1}\left(\frac{F_1}{X_{1,0}} - 1\right)^2 = \frac{1}{0.03012177}\left(\frac{1031.10}{1030} - 1\right)^2$$
$$= 0.3789 \times 10^{-4}$$

and, for the second nearby contract,

$$\frac{1}{T_2}\left(\frac{F_2}{X_{2,0}} - 1\right)^2 = \frac{1}{0.12601218}\left(\frac{1027.80}{1035} - 1\right)^2$$
$$= 0.00018843$$

Next, take the sum in the first term on the right-hand side. Table 16.7 shows the values of each of the n_1 terms

Table 16.7 Nearby S&P 500 Index Option Prices Computation to the composition of the VIX on October 6, 2003, at 8:38 AM (CST)

		Nearby Contract Expiration: 10/17/2003			
C/P	**Exercise Price**	**Price Midpoint**	ΔX_i	**Weight**	**Weight Times Forward Option Price**
P	850	0.125	25	0.0000346021	0.0000043265
P	875	0.150	25	0.0000326531	0.0000048993
P	900	0.300	20	0.0000246914	0.0000074095
P	915	0.275	12.5	0.0000149303	0.0000041070
P	925	0.425	7.5	0.0000087655	0.0000037264
P	930	0.500	5	0.0000057810	0.0000028913
P	935	0.550	5	0.0000057194	0.0000031465
P	940	0.675	5	0.0000056587	0.0000038207
P	945	0.650	5	0.0000055989	0.0000036403
P	950	0.825	5	0.0000055402	0.0000045719
P	955	0.925	5	0.0000054823	0.0000050725
P	960	1.050	7.5	0.0000081380	0.0000085473
P	970	1.350	7.5	0.0000079711	0.0000107640
P	975	1.700	5	0.0000052597	0.0000089440
P	980	1.950	5	0.0000052062	0.0000101548
P	985	2.250	5	0.0000051534	0.0000115985
P	990	2.700	5	0.0000051015	0.0000137779
P	995	3.350	7.5	0.0000075756	0.0000253852
P	1005	4.800	7.5	0.0000074256	0.0000356526
P	1010	5.900	5	0.0000049015	0.0000289267
P	1015	7.100	5	0.0000048533	0.0000344680
P	1020	8.500	5	0.0000048058	0.0000408610
P	1025	10.400	5	0.0000047591	0.0000495081
X_0	1030	12.950	5	0.0000047130	0.0000610500
C	1035	10.800	5	0.0000046676	0.0000504235
C	1040	8.500	5	0.0000046228	0.0000393045
C	1045	6.550	5	0.0000045786	0.0000299985
C	1050	5.100	5	0.0000045351	0.0000231357
C	1055	3.800	5	0.0000044923	0.0000170753
C	1060	2.900	5	0.0000044500	0.0000129085
C	1065	2.150	5	0.0000044083	0.0000094805
C	1070	1.550	5	0.0000043672	0.0000067710
C	1075	1.150	15	0.0000129800	0.0000149311
C	1100	0.150	25	0.0000206612	0.0000031000
				Sum	0.0005943786

for the nearby contract, and Table 16.8 shows the values of each of the n_2 terms of the second nearby contract. The first term in the nearby contract's summation is

$$\frac{\Delta X_{1,1}}{X_{1,1}^2} e^{r_1 T_1} O(X_{1,1}) = \frac{25}{850^2} \times 1.0002772 \times 0.125$$
$$= 0.433 \times 10^{-5}$$

as is shown in Table 16.7. Note that the option price used in the expression is the forward price (that is, the current price carried forward until the end of the contract's life). The sum of the weighted average of the forward option prices is 0.0005943786 for the nearby contract and 0.0025376773 for the second nearby contract. The variance of the nearby contract is therefore

$$\sigma_1^2 = \frac{2}{0.03012177} \times 0.00059438 - 0.3789 \times 10^{-4}$$
$$= 0.03942717$$

and the variance of the second nearby contract is

$$\sigma_2^2 = \frac{2}{0.12601218} \times 0.00253768 - 0.00018843$$
$$= 0.04008827$$

Step 8: Compute the annualized volatility over the next 30 calendar days. The variances of the nearby and second nearby contracts correspond to times to expiration of T_1 years and T_2 years, respectively. VIX, however, maintains a constant time to expiration of 30 days or $30/365 = 0.0821917808$ years. To find the variance over the 30 calendar-day interval, we must interpolate between the variances of the nearby and second nearby contracts, that is,

$$\sigma_{30-\text{day}}^2 = \left(\frac{T_2 - T_{30-\text{day}}}{T_2 - T_1}\right)\sigma_1^2 T_1 + \left(\frac{T_{30-\text{day}} - T_1}{T_2 - T_1}\right)\sigma_2^2 T_2$$
$$= 0.00328583$$

To compute the level of VIX, we annualize the 30-day variance and take the square root, that is,

$$VIX = \sqrt{\sigma_{30-\text{day}}^2\left(\frac{1}{T_{30-\text{day}}}\right)}$$
$$= \sqrt{0.03997755\left(\frac{1}{0.08219178}\right)} = 19.99\%$$

This is precisely the level of VIX reported by the CBOE at 8:38 AM (CST) on October 6.

Table 16.8 Second Nearby S&P 500 Index Option Prices Computation to the Computation of the VIX on October 6, 2003, at 8:38 AM (CST)

Second Nearby Contract Expiration: 11/21/2003					
C/P	**Exercise Price**	**Price Midpoint**	ΔX_i	**Weight**	**Weight Times Forward Option Price**
P	775	0.350	25	0.0000416233	0.0000145838
P	800	0.550	25	0.0000390625	0.0000215074
P	825	0.850	25	0.0000367309	0.0000312548
P	850	1.350	25	0.0000346021	0.0000467629
P	875	1.950	22.5	0.0000293878	0.0000573675
P	895	2.700	12.5	0.0000156050	0.0000421787
P	900	2.950	15	0.0000185185	0.0000546882
P	925	4.300	25	0.0000292184	0.0001257738
P	950	6.500	25	0.0000277008	0.0001802484
P	975	10.000	15	0.0000157791	0.0001579600
P	980	11.000	5	0.0000052062	0.0000573292
P	985	12.000	5	0.0000051534	0.0000619076
P	990	13.100	5	0.0000051015	0.0000669015
P	995	14.300	7.5	0.0000075756	0.0001084467
P	1005	17.350	7.5	0.0000074256	0.0001289715
P	1010	18.700	5	0.0000049015	0.0000917559
P	1015	20.500	5	0.0000048533	0.0000995995
P	1020	22.300	5	0.0000048058	0.0001072852
X_0	1025	26.850	7.5	0.0000071386	0.0001918770
C	1035	23.900	12.5	0.0000116689	0.0002791852
C	1050	17.000	12.5	0.0000113379	0.0001929503
C	1060	13.200	7.5	0.0000066750	0.0000882041
C	1065	11.500	5	0.0000044083	0.0000507497
C	1070	9.750	5	0.0000043672	0.0000426258
C	1075	8.700	5	0.0000043267	0.0000376823
C	1080	7.500	12.5	0.0000107167	0.0000804617
C	1100	3.900	22.5	0.0000185950	0.0000725984
C	1125	1.650	25	0.0000197531	0.0000326275
C	1150	0.750	25	0.0000189036	0.0000141929
				Sum	0.0025376773

REFERENCES

Bollen, N.P.B., and Whaley, R. E. (2004). Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance* 59, 2: 711–754.

Chicago Board Options Exchange (2003). VIX: CBOE volatility Index. Working paper.

Day, T.,E., and Lewis, C. M. (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics* 52: 267–287.

Demeterfi, K., Derman, E., Kamal, M., and Zou, J. (1999). More than you ever wanted to know about volatility swaps. Quantitative Strategies Research Notes, Goldman Sachs, working paper.

Dunbar, N. (2000). *Inventing Money: The Story of Long-Term Capital Management and the Legends Behind It*. Chichester, UK: John Wiley & Sons.

Fleming, J. (1998). The quality of market volatility forecasts implied by S&P 100 index option prices. *Journal of Empirical Finance* 5: 317–345.

Grossman, S. (1988). An analysis of the implications for stock and futures price volatility of program trading and dynamic hedging strategies. *Journal of Business* 61: 275–298.

Heston, S. L., and Nandi, S. (2000). Derivatives on volatility: Some simple solutions based on observables. Federal Reserve Bank of Atlanta working paper.

Lowenstein, R. C. (2000). *When Genius Failed: The Rise and Fall of Long-Term Capital Management*. New York: Random House.

Sulima, C. L. (2001). Volatility and variance swaps. *Capital Market News Federal Reserve Bank of Chicago*, March: 1–4.

Whaley, R. E. (1993). Derivatives on market volatility: Hedging tools long overdue. *Journal of Derivatives* 1: 71–84.

Whaley, R. E. (2000). The investor fear gauge. *Journal of Portfolio Management* 26, 3: 12–17.

Whaley, R. E. (2002). Return and risk of CBOE buy-write monthly index. *Journal of Derivatives* 10, 2: 35–42.

Whaley, R. E. (2006). *Derivatives: Markets, Valuation, and Risk Management*. Hoboken, NJ: John Wiley & Sons.

PART 3

Fixed Income Instruments

Basics

Nonmortgage Related Fixed Income Securities and Money Market Instruments

Structured Products

Fixed Income and Inflation Derivatives

Bond Market

CHAPTER 17

Bonds: Investment Features and Risks

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: Bonds are debt instruments that are issued by a wide-range of entities throughout the world. Unlike the investor in common stock who hopes to share in the good fortunes of a corporation through increased dividends and price appreciation in the stock's price, an investor in a bond has agreed to accept a fixed contractual interest rate. The features that may be included in a bond affect both the performance of a bond when market interest rates change and its risk characteristics. An investor in a bond is exposed to one or more of the following risks: interest rate risk, call and prepayment risk, credit risk, liquidity risk, exchange rate or currency risk, and inflation or purchasing power risk. There are various yield measures that are quoted for bonds: current yield, yield to maturity, yield to call, yield to put, yield to worst, and cash flow yield.

Keywords: bond, global bond market, domestic bond market, foreign bond market, international bond market, offshore bond market, Eurobond market, Eurobonds, maturity, term to maturity, money market instruments, money market, principal, face value, redemption value, maturity value, coupon rate, coupon, step-up notes, zero-coupon bonds, floating-rate securities, fixed-rate bond market, floating-rate bond market, coupon reset date, inverse floaters, reverse floaters, cap, floor, range note, accrued interest, full price, clean price, traded flat, trade date, settlement date, value date, day count conventions, bullet maturity, serial bonds, amortizing securities, sinking fund provision, deferred call, first call date, refunding, first par call date, prepayment, prepayment option, balloon maturity, accelerated sinking fund provision, convertible bond, current yield, yield to maturity, bond-equivalent yield, yield to call, yield to first call, yield to next call, a yield to first par call, yield to refunding, yield to put, yield to worst, cash flow yield, interest rate risk, call risk, prepayment risk, credit risk, default risk, rating agencies, credit spread risk, downgrade risk, liquidity risk. exchange rate risk, currency risk, inflation risk, purchasing power risk

In its simplest form, a bond is a financial obligation of an entity that promises to pay a specified sum of money at specified future dates. The entity that promises to make the payment is called the issuer of the security or the borrower. Some examples of issuers are the U.S. government or a foreign government, a state or local government entity, a domestic or foreign corporation, and a supranational government such as the World Bank. The investor who purchases a bond is said to be the lender or creditor. The promised payments that the issuer agrees to make at the specified dates consist of two components: interest payments and repayment of the amount borrowed.

The purpose of this chapter is to explain the investment features of bonds, the various measures of yield quoted for bonds, and the risks that investors face when investing in bonds.

SECTORS OF THE BOND MARKET

There are many ways to classify the bond market. One way is in terms of the taxability of the interest at the federal income tax level. In the United States, most securities issued by state and local governments and by entities that they establish, referred to as municipal bonds or municipal securities, are exempt from federal income taxation. While there are reasons why some issuers of municipal bonds will issue taxable bonds, the municipal bond market is generally viewed as the market for tax-exempt securities. As such, the primary attraction to investors is this tax feature.

The largest part of the bond market is the taxable market. There are various ways to describe this sector. Investment banking firms that have developed bond market indexes use various classifications. The most popular indexes are those published by Lehman Brothers and within the group of indexes it publishes, the one followed most closely by investors in the United States is the U.S. Aggregate Index. That index contains the six sectors shown in Table 17.1 along with the percentage of each sector in terms of market value as of July 20, 2007. We'll review each of the sectors later in this chapter.

Another way of classifying bond markets is in terms of the *global bond market.* One starts by partitioning a given country's bond market into a national bond market and an international bond market. In turn, a country's national bond market can be divided into a domestic bond market and a foreign bond market with the distinction being the domicile of the issuer. The *domestic bond market* is the market where bond issues of entities domiciled within that country are issued and then traded; the *foreign bond market* is the market where bond issues of nondomiciled entities of that country are issued and then subsequently traded within the country. Each country has a nickname for foreign bonds. For example, in the United States, "Yankee bonds" are bonds issued by non-U.S. entities and then traded in the U.S. market. In the United Kingdom, foreign bonds are called "bulldog bonds." The *international bond market,* also referred to as the *offshore bond market,* is the market where bonds are issued and then traded outside of the country and not regulated by the country.

An important sector of the international bond market is the market for bonds that are underwritten by an international syndicate, issued simultaneously to investors in a number of countries, and issued outside of the jurisdiction of any single country. This market is popularly referred to as the *Eurobond market* and the bonds are called *Eurobonds.* Unfortunately, the name is misleading. The currency in which Eurobonds are denominated can be any currency, not just euros. In fact, Eurobonds are classified according to the denomination of the currency (e.g., Eurodollar bonds and Euroyen bonds). Nor are Eurobonds traded in just Europe. *Global bonds* from the perspective of a country are bonds that are not only traded in that country's foreign bond market but also in the Eurobond market.

While the U.S. bond market is the largest bond market in the world, there are other bond markets in which U.S. investors participate. These are described in Chapters 24, 25, 26 and 31 of Volume I.

Table 17.1 Sectors of the Lehman Brothers U.S. Aggregate Index

Sector	Percent of Market Value (as of July 20, 2007)
Treasury	23.49%
Agency	10.60
Mortgage Pass-through	37.90
Commercial MBS	4.89
Asset-Backed Securities	1.12
Credit	22.02

Source: Data obtained from Lehman Brothers, *Global Relative Value,* Fixed Income Research, July 23, 2007.

FEATURES OF BONDS

The promises of the issuer and the rights of the bondholders are set forth in great detail in the indenture. Bondholders would have great difficulty in determining from time to time whether the issuer was keeping all the promises made in the indenture. This problem is resolved for the most part by bringing in a trustee as a third party to the contract. The indenture is made out to the trustee as a representative of the interests of the bondholders; that is, a trustee acts in a fiduciary capacity for bondholders. A trustee is a bond or trust company with a trust department whose officers are experts in performing the functions of a trustee.

Maturity

Unlike common stock which has a perpetual life, bonds have a date on which they mature. The number of years over which the issuer has promised to meet the conditions of the obligation is referred to as the *term to maturity.* The *maturity* of a bond refers to the date that the debt will cease to exist, at which time the issuer will redeem the bond by paying the amount borrowed. The maturity date of a bond is always identified when describing a bond. For example, a description of a bond might state "due 12/15/2025."

The practice in the bond market is to refer to the "term to maturity" of a bond as simply its "maturity" or "term." Despite sounding like a fixed date in which the bond matures, there are provisions that may be included in the indenture that grants either the issuer or bondholder the right to alter a bond's term to maturity. These provisions, which will be described later in this chapter, include call provisions, put provisions, conversion provisions and accelerated sinking fund provisions.

The maturity of a debt instrument is used for classifying two sectors of the market. Debt instruments with a maturity of 1 year or less are referred to as *money market instruments* and trade in the *money market.* What we typically refer to as the "bond market" is debt instruments with a maturity greater than one year. The bond market is then categorized further based on the debt instrument's term to maturity: short-term, intermediate-term, and long-term. The classification is somewhat arbitrary and varies amongst market participants. A common classification is that short-term bonds have a maturity of from 1 to 5 years, intermediate-term bonds have a maturity from 5 to 12 years, and long-term bonds have a maturity that exceeds 12 years.

Typically, the maturity of a bond does not exceed 30 years. There are, of course, exceptions. For example, Walt Disney Company issued 100-year bonds in July 1993 and the Tennessee Valley Authority issued 50-year bonds in December 1993.

The term to maturity of a bond is important for two reasons in addition to indicating the time period over which the bondholder can expect to receive interest payments and the number of years before the principal will be paid in full. The first reason is that the yield on a bond depends on it. At any given point in time, the relationship between the yield and maturity of a bond (called the yield curve) indicates how bondholders are compensated for investing in bonds with different maturities. The second reason is that the price of a bond will fluctuate over its life as interest rates in the market change. The degree of price volatility of a bond is dependent on its maturity. More specifically, all other factors constant, the longer the maturity of a bond, the greater the price volatility resulting from a change in interest rates.

Par Value

The *par value* of a bond is the amount that the issuer agrees to repay the bondholder by the maturity date. This amount is also referred to as the *principal, face value, redemption value,* or *maturity value.*

Because bonds can have a different par value, the practice is to quote the price of a bond as a percentage of its par value. A value of 100 means 100% of par value. So, for example, if a bond has a par value of $1,000 and is selling for $850, this bond would be said to be selling at 85. If a bond with a par value of $100,000 is selling for $106,000, the bond is said to be selling for 106.

Coupon Rate

The annual interest rate that the issuer agrees to pay each year is called the *coupon rate.* The annual amount of the interest payment made to bondholders during the term of the bond is called the *coupon* and is determined by multiplying the coupon rate by the par value of the bond. For example, a bond with a 6% coupon rate and a par value of $1,000 will pay annual interest of $60.

When describing a bond issue, the coupon rate is indicated along with the maturity date. For example, the expression "5.5s of 2/15/2024" means a bond with a 5.5% coupon rate maturing on 2/15/2024.

For bonds issued in the United States, the usual practice is for the issuer to pay the coupon in two semiannual installments. Mortgage-backed securities and asset-backed securities typically pay interest monthly. For bonds issued in some markets outside the United States, coupon payments are made only once per year.

In addition to indicating the coupon payments that the investor should expect to receive over the term of the bond, the coupon rate also affects the bond's price sensitivity to changes in market interest rates. All other factors constant, the higher the coupon rate, the less the price will change in response to a change in market interest rates.

There are securities that have a coupon rate that increases over time according to a specified schedule. These securities are called *step-up notes* because the coupon rate "steps up" over time. For example, a 5-year step-up note might have a coupon rate that is 5% for the first two years and 6% for the last three years. Or, the step-up note could call for a 5% coupon rate for the first two years, 5.5% for the third and fourth years, and 6% for the fifth year. When there is only one change (or step up), as in our first example, the issue is referred to as a single step-up note. When there is more than one increase, as in our second example, the issue is referred to as a multiple step-up note.

Not all bonds make periodic coupon payments. *Zero-coupon bonds,* as the name indicates, do not make periodic coupon payments. Instead, the holder of a zero-coupon bond realizes interest at the maturity date. The aggregate interest earned is the difference between the maturity value and the purchase price. For example, if an investor purchases a zero-coupon bond for 63, the aggregate interest at the maturity date is 37, the difference between the par value (100) and the price paid (63). The reason why certain investors like zero-coupon bonds is that they eliminated one of the risks that we will discuss later, reinvestment risk. The disadvantage of a zero-coupon bond is that the accrued interest earned each year is taxed despite the fact that no actual cash payment is made.

There are issues whose coupon payment is deferred for a specified number of years. That is, there is no coupon payment for the deferred period and then a lump sum payment at some specified date and coupon payments until maturity. These securities are referred to as *deferred interest securities.*

A coupon-bearing security need not have a fixed interest rate over the term of the bond. These are bonds that have an interest rate that is as variable. These bonds are referred to as *floating-rate securities.* In fact, another way to classify bond markets is the *fixed-rate bond market* and the *floating-rate bond market.* Floating-rate securities appeal to institutional investors such as depository institutions (banks, savings and loan associations, and credit unions) because it provides a better match against their funding costs which are typically floating-rate debt. Typically, the interest rate is adjusted on specific dates, referred to as the *coupon reset date.* There is typically a formula for the new coupon rate that has the following generic formula:

Reference rate + Quoted margin

The quoted margin is the additional amount that the issuer agrees to pay above the reference rate. The most common reference rate is the London Interbank Offered Rate (LIBOR). LIBOR is the interest rate at which major international banks offer each other on Eurodollar certificates of deposit with given maturities. The maturities range from overnight to five years. Suppose that the reference rate is one-month LIBOR and the index spread is 80 basis points. (A basis point is equal to 0.0001 or 0.01%. Thus, 100 basis points are equal to 1%.) Then the coupon reset formula is:

One-month LIBOR + 80 basis points

So, if one-month LIBOR on the coupon reset date is 4.6%, the coupon rate is reset for that period at 5.4% (4.6% plus 80 basis points).

The quoted margin need not be a positive value. It could be subtracted from the reference rate. For example, the reference rate could be the yield on a five-year Treasury security and the coupon rate could reset every six months based on the following coupon reset formula:

Five-year Treasury yield − 50 basis points

While the reference rate for most floating-rate securities is an interest rate or an interest rate index, there are some issues where this is not the case. Instead, the reference rate can be some financial index such as the return on the Standard & Poor's 500 index or a nonfinancial index such as the price of a commodity or the consumer price index.

Typically, the coupon reset formula on floating-rate securities is such that the coupon rate increases when the reference rate increases, and decreases when the reference rate decreases. There are issues whose coupon rate moves in the opposite direction from the change in the reference rate. Such issues are called *inverse floaters* or *reverse floaters.* A general coupon reset formula for an inverse floater is:

$K - L \times$ (Reference rate)

For example, suppose that for a particular inverse floater K is 10% and L is 1. Then the coupon reset formula would be:

10% − Reference rate

Suppose that the reference rate is one-month LIBOR, then the coupon reset formula would be:

10% − one-month LIBOR

If in some month one-month LIBOR at the coupon reset date is 5%, the coupon rate for the period is 5%. If in the next month one-month LIBOR declines to 4.5%, the coupon rate increases to 5.5%.

A floating-rate security may have a restriction on the maximum coupon rate that will be paid at a reset date. The maximum coupon rate is called a *cap.* Because a cap restricts the coupon rate from increasing, a cap is an unattractive feature for the investor. In the case of an inverse floater, one can see from the general formula that the maximum interest rate would be K. This occurs when the reference rate is zero. In contrast, there could be a *floor* which is the minimum coupon rate specified and this is an attractive feature for the investor.

Not all floating-rate notes have the generic formula given above. Some have a coupon rate that depends on the range for a reference rate. This type of floating-rate security, called a *range note,* has a coupon rate equal to the reference rate as long as the reference rate is within a certain range at the reset date. If the reference rate is outside of the range, the coupon rate is zero for that period. For example, a three-year range note might specify that the reference rate is one-year LIBOR and that the coupon rate resets every year. The coupon rate for the year will be one-year LIBOR as long as one-year LIBOR at the coupon reset date falls within the range as specified below:

	Year 1	**Year 2**	**Year 3**
Lower limit of range	4.5%	5.25%	6.00%
Upper limit of range	5.5%	6.75%	7.50%

If one-year LIBOR is outside of the range, the coupon rate is zero. For example, if in year 1 one-year LIBOR is 5% at the coupon reset date, the coupon rate for the year is 5%. However, if one-year LIBOR is 6%, the coupon rate for the year is zero since one-year LIBOR is greater than the upper limit for year 1 of 5.5%.

Accrued Interest

In the United States, coupon interest is typically paid semiannual for government bonds, corporate, agency, and municipal bonds. In some countries, interest is paid annually. For mortgage-backed and asset-backed securities, interest is usually paid monthly. The coupon interest payment is made to the bondholder of record. Thus, if an investor sells a bond between coupon payments and the buyer holds it until the next coupon payment, then the entire coupon interest earned for the period will be paid to the buyer of the bond since the buyer will be the holder of record. The seller of the bond gives up the interest from the time of the last coupon payment to the time until the bond is sold. The amount of interest over this period that will be received by the buyer even though it was earned by the seller is called *accrued interest.*

In the United States and in many countries, the bond buyer must compensate the bond seller for the accrued interest. The amount that the buyer pays the seller is the agreed-upon price for the bond plus accrued interest. This amount is called the *full price.* The agreed-upon bond price without accrued interest is called the *clean price.*

A bond in which the buyer must pay the seller accrued interest is said to be trading *cum-coupon*. If the buyer forgoes the next coupon payment, the bond is said to be trading *ex-coupon*. In the United States, bonds are always traded cum-coupon. There are bond markets outside the United States where bonds are traded ex-coupon for a certain period before the coupon payment date.

There are exceptions to the rule that the bond buyer must pay the bond seller accrued interest. The most important exception is when the issuer has not fulfilled its promise to make the periodic payments. In this case, the issuer is said to be in default. In such instances, the bond's price is sold without accrued interest and is said to be *traded flat.*

When calculating accrued interest, three pieces of information are needed: (1) the number of days in the accrued interest period, (2) the number of days in the coupon period, and (3) the dollar amount of the coupon payment. The number of days in the accrued interest period represents the number of days over which the investor has earned interest. Given these values, the accrued interest (AI), assuming semiannual payments, is calculated as follows:

$$\text{AI} = \frac{\text{Annual coupon}}{2} \times \frac{\text{Days in AI period}}{\text{Days in coupon period}}$$

For example, suppose that (1) there are 50 days in the accrued interest period, (2) there are 183 days in a coupon period, and (3) the annual coupon per \$100 of par value is \$8. Then the accrued interest is:

$$\text{AI} = \frac{\$8}{2} \times \frac{50}{183} = \$1{,}029$$

It is not simple to determine the number of days in the accrued interest period and the number of days in the coupon period. The calculation begins with the determination of three key dates:

- Trade date
- Settlement date
- Value date

The *trade date* is the date on which the transaction is executed. The *settlement date* is the date a transaction is completed. The settlement date varies by the type of bond. Unlike the settlement date, the *value date* is not constrained to fall on a business day.

Interest accrues on a bond from and including the date of the previous coupon up to but excluding the value date. (This is the definition used by the International Securities Market Association [ISMA].) However, this may differ slightly in some non-U.S. markets. For example, in some countries interest accrues up to and including the value date. For a newly issued security, there is no previous coupon payment. Instead, the interest accrues from a date called the dated date.

Day Count Conventions

The number of days in the accrued interest period and the number of days in the coupon period may not be simply the actual number of calendar days between two dates. The reason is that there is a market convention for each type of security that specifies how to determine the number of days between two dates. These conventions are called *day count conventions.*

In calculating the number of days between two dates, the actual number of days is not always the same as the number of days that should be used in the accrued interest formula. The number of days used depends on the day count convention for the particular security. Specifically, there are different day count conventions for Treasury securities than for government agency securities, municipal bonds, and corporate bonds.

For coupon-bearing Treasury securities, the day count convention used is to determine the actual number of days between two dates. This is referred to as the "actual/actual day count convention." For example, consider a coupon-bearing Treasury security whose previous coupon payment was March 1. The next coupon payment would be on September 1. Suppose this Treasury security is purchased with a value date of July 17. The actual number of days between July 17 (the value date) and September 1 (the date of the next coupon payment is 46 days) is shown below:

July 17 to July 31	14 days
August	31 days
September 1	1 day
	46 days

The number of days in the coupon period is the actual number of days between March 1 and September 1, which is 184 days. The number of days between the last coupon payment (March 1) to July 17 is therefore 138 days (184 days − 46 days).

For coupon-bearing agency, municipal, and corporate bonds, a different day count convention is used. It is assumed that every month has 30 days, that any 6-month period has 180 days, and that there are 360 days in a year. This day count convention is referred to as the "30/360 day count convention." For example, consider a security purchased with a value date of July 17, the previous coupon payment on March 1, and the next coupon payment on September 1. If the security is an agency, municipal, or corporate bond rather than a Treasury security, the number of days until the next coupon payment is 44 days as shown below:

July 17 to July 31	13 days
August	30 days
September 1	1 day
	44 days

The number of days from March 1 to July 17 is 136, which is the number of days in the accrued interest period.

Provisions for Paying Off Bonds

The issuer of a bond agrees to repay the principal by the stated maturity date. The issuer can agree to repay the entire amount borrowed in one lump sum payment at the maturity date. That is, the issuer is not required to make any principal repayments prior to the maturity date. Such bonds are said to have a *bullet maturity.*

There are bond issues which consist of a series of blocks of securities maturing in sequence. The blocks of securities are said to be *serial bonds.* The coupon rate for each block can be different. One type of corporate bond in which there are serial bonds is an equipment trust certificate. Municipal bonds are often issued as serial bonds.

Bonds backed by pools of loans (mortgage-backed securities and asset-backed securities) often have a schedule of principal repayments. Such bonds are said to be *amortizing securities.* For many loans, the payments are structured so that when the last loan payment is made, the entire amount owed is fully paid off. Another example of an amortizing feature is a bond that has a *sinking fund provision.* This provision for repayment of a bond may be designed to liquidate all of an issue by the maturity date, or it may be arranged to repay only a part of the total by the maturity date.

A bond issue may have a call provision granting the issuer an option to retire all or part of the issue prior to the stated maturity date. Some issues specify that the issuer must retire a predetermined amount of the issue periodically. Various types of call provisions are discussed below.

Call and Refunding Provisions

An issuer generally wants the right to retire a bond issue prior to the stated maturity date because it recognizes that at some time in the future the general level of interest rates may fall sufficiently below the issue's coupon rate so that redeeming the issue and replacing it with another issue with a lower coupon rate would be economically beneficial. This right is a disadvantage to the bondholder since proceeds received must be reinvested at a lower interest rate. As a result, an issuer who wants to include this right as part of a bond offering must compensate the bondholder when the issue is sold by offering a higher coupon rate, or equivalently, accepting a lower price than if the right is not included.

The right of the issuer to retire the issue prior to the stated maturity date is referred to as a call option. If an issuer exercises this right, the issuer is said to "call the bond." The price which the issuer must pay to retire the issue is referred to as the call price. There may not be a call price but a call schedule which sets forth a call price based on when the issuer can exercise the call option.

When a bond is issued, typically the issuer may not call the bond for a number of years. That is, the issue is said to have a *deferred call.* The date at which the bond may first be called is referred to as the *first call date.* However, not all issues have a deferred call. If a bond issue does not have any protection against early call, then it is said to be a currently callable issue. But most new bond issues, even if currently callable, usually have some restrictions against certain types of early redemption. The most common restriction is that prohibiting the refunding of the bonds for a certain number of years. *Refunding* a bond issue means redeeming bonds with funds obtained through the sale of a new bond issue.

Call protection is much more absolute than refunding protection. While there may be certain exceptions to absolute or complete call protection in some cases, it still provides greater assurance against premature and unwanted redemption than does refunding protection. Refunding prohibition merely prevents redemption only from certain sources of funds, namely the proceeds of other debt issues sold at a lower cost of money. The bondholder is only protected if interest rates decline, and the borrower can obtain lower-cost money to pay off the debt.

Bonds can be called in whole (the entire issue) or in part (only a portion). When less than the entire issue is called, the specific bonds to be called are selected randomly or on a pro rata basis.

Generally, the call schedule is such that the call price at the first call date is a premium over the par value and scaled down to the par value over time. The date at which the issue is first callable at par value is referred to as the *first par call date.* However, not all issues have a call schedule in which the call price starts out as a premium over par. There are issues where the call price at the first call date and subsequent call dates is par value. In such cases, the first call date is the same as the first par call date.

For zero-coupon bonds, there are three types of call schedules that can be used. The first is a call schedule for which the call price is below par value at the first call date and scales up to par value over time. The second type is one in which the call price at the first call date is above par and scales down to par. The third type is a schedule in which the call price is par value at the first call date and any subsequent call date.

The call prices in a call schedule are referred to as the regular or general redemption prices. There are also special redemption prices for debt redeemed through the sinking fund and through other provisions, and the proceeds from the confiscation of property through the right of eminent domain. The special redemption price is usually par value, but in the case of some utility issues it initially may be the public offering price, which is amortized down to par value (if a premium) over the life of the bonds.

Prepayments

For amortizing securities backed by loans and have a schedule of principal repayments, individual borrowers typically have the option to pay off all or part of their loan prior to the scheduled date. Any principal repayment prior to the scheduled date is called a *prepayment.* The right of borrowers to prepay is called the *prepayment option.*

Basically, the prepayment option is the same as a call option. However, unlike a call option, there is not a call price that depends on when the borrower pays off the issue. Typically, the price at which a loan is prepaid is at par value.

Sinking Fund Provision

A *sinking fund provision* included in a bond indenture requires the issuer to retire a specified portion of an issue each year. Usually, the periodic payments required for sinking fund purposes will be the same for each period. A few indentures might permit variable periodic payments, where payments change according to certain prescribed conditions set forth in the indenture. The alleged purpose of the sinking fund provision is to reduce credit risk. This kind of provision for repayment of debt may be designed to liquidate all of a bond issue by the maturity date, or it may be arranged to pay only a part of the total by the end of the term. If only a part is paid, the remainder is called a *balloon maturity.* Many indentures include a provision that grants the issuer the option to retire more than the amount stipulated for sinking fund retirement. This is referred to as an *accelerated sinking fund provision.*

To satisfy the sinking fund requirement, an issuer is typically granted one of following choices: (1) make a cash payment of the face amount of the bonds to be retired to the trustee, who then calls the bonds for redemption using a lottery, or (2) deliver to the trustee bonds purchased in the open market that have a total par value equal to the amount that must be retired. If the bonds are retired using the first method, interest payments stop at the redemption date.

Usually the sinking fund call price is the par value if the bonds were originally sold at par. When issued at a price in excess of par, the call price generally starts at the issuance price and scales down to par as the issue approaches maturity.

There is a difference between the amortizing feature for a bond with a sinking fund provision, and the regularly scheduled principal repayment for a mortgage-backed and an asset-backed security. The owner of a mortgage-backed security and an asset-backed security knows that assuming no default that there will be principal repayments. In contrast, the owner of a bond with a sinking fund provision is not assured that his or her particular holding will be called to satisfy the sinking fund requirement.

Options Granted to Bondholders

A provision in the indenture could grant either the bondholder and/or the issuer an option to take some action against the other party. The most common type of option embedded in a bond is a call option which we discussed above. This option is granted to the issuer. There are two options that can be granted to the bondholder: the right to put the issue and the right to convert the issue.

An issue with a put provision grants the bondholder the right to sell the issue (that is, force the issuer to redeem the issue) at a specified price on designated dates. The specified price is called the put price. Typically, a bond is puttable at par value if it is issued at or close to par value. For a zero-coupon bond, the put price is below par. The advantage of the put provision to the bondholder is that if after the issue date market rates rise above the issue's coupon rate, the bondholder can force the issuer to redeem the bond at the put price and then reinvest the proceeds at the prevailing higher rate.

A *convertible bond* is an issue giving the bondholder the right to exchange the bond for a specified number of shares of common stock. Such a feature allows the bondholder to take advantage of favorable movements in the price of the issuer's common stock. An *exchangeable bond* allows the bondholder to exchange the issue for a specified number of shares of common stock of a corporation different from the issuer of the bond. Convertible bonds are described in Chapter 29 of Volume I.

Currency Denomination

The payments that the issuer makes to the bondholder can be in any currency. For bonds issued in the United States, the issuer typically makes both coupon payments and principal repayments in U.S. dollars. However, there is nothing that forces the issuer to make payments in U.S dollars. The indenture can specify that the issuer may make payments in some other specified currency. For example, payments may be made in euros or yen.

An issue in which payments to bondholders are in U.S. dollars is called a dollar-denominated issue. A nondollar-denominated issue is one in which payments are not denominated in U.S. dollars. There are some issues whose coupon payments are in one currency and whose principal payment is in another currency. An issue with this characteristic is called a dual-currency issue.

Some issues allow either the issuer or the bondholder the right to select the currency in which a payment will be paid. This option effectively gives the party with the right to choose the currency the opportunity to benefit from a favorable exchange rate movement.

YIELD MEASURES

When an investor purchases a bond, he or she can expect to receive a dollar return from one or more of the following sources:

1. The coupon interest payments made by the issuer.
2. Any capital gain (or capital loss—a negative dollar return) when the security matures, is called, or is sold.
3. Income from reinvestment of the interim cash flows.

Any yield measure that purports to measure the potential return from a bond should consider all three sources of return described above.

The most obvious source of return is the periodic coupon interest payments. For zero-coupon instruments, the return from this source is zero, although the investor is effectively receiving interest by purchasing a security below its par value and realizing interest at the maturity date when the investor receives the par value.

When the proceeds received when a bond matures, is called, or is sold are greater than the purchase price, a capital gain results. For a bond held to maturity, there will be a capital gain if the bond is purchased below its par value. A bond purchased below its par value is said to be purchased at a discount. For example, a bond purchased for \$94.17 with a par value of \$100 will generate a capital gain of \$5.83 (\$100 − \$94.17) if held to maturity. For a callable bond, a capital gain results if the price at which the bond is called (that is, the call price) is greater than the purchase price. For example, if the bond in our previous example is callable and subsequently called at \$100.5, a capital gain of \$6.33 (\$100.5 − \$94.17) will be realized. If the same bond is sold prior to its maturity or before it is called, a capital gain will result if the proceeds exceed the purchase price. So, if our hypothetical bond is sold prior to the maturity date for \$103, the capital gain would be \$8.83 (\$103 – \$94.17).

A capital loss is generated when the proceeds received when a bond matures, is called, or is sold are less than the purchase price. For a bond held to maturity, there will be a capital loss if the bond is purchased for more than its par value. A bond purchased for more than its par value is said to be purchased at a premium. For example, a bond purchased for \$102.5 with a par value of \$100 will generate a capital loss of \$2.5 (\$102.5 − \$100) if held to maturity. For a callable bond, a capital loss results if the price at which the bond is called is less than the purchase price. For example, if the bond in our previous example is callable and subsequently called at \$100.5, a capital loss of \$2 (\$102.5 − \$100.5) will be realized. If the same bond is sold prior to its maturity or before it is called, a capital loss will result if the sale price is less than the purchase price. So, if our hypothetical bond is sold prior to the maturity date for \$98.5, the capital loss would be \$4 (\$102.5 − \$98.5).

With the exception of zero-coupon instruments, bonds make periodic payments of interest that can be reinvested until the security is removed from the portfolio. There are also instruments in which there are periodic principal repayments that can be reinvested until the security is removed from the portfolio. Repayment of principal prior to the maturity date occurs for amortizing instruments such as mortgage-backed securities and asset-backed securities. The interest earned from reinvesting the interim cash flows (interest and/or principal payments) until the security is removed from the portfolio is called reinvestment income.

There are several yield measures cited in the bond market. These include current yield, yield to maturity, yield to call, yield to put, yield to worst, and cash flow yield. Below we explain how each measure is calculated and its limitations.

Current Yield

The *current yield* relates the annual dollar coupon interest to the market price. The formula for the current yield is:

$$\text{Current yield} = \frac{\text{Annual dollar coupon interest}}{\text{Price}}$$

For example, the current yield for a 7% 8-year bond whose price is \$94.17 is 7.43% as shown below:

$$\text{Current yield} = \frac{\$7}{\$94.17} = 0.0743 = 7.43\%$$

The current yield will be greater than the coupon rate when the bond sells at a discount; the reverse is true for a bond selling at a premium. For a bond selling at par, the current yield will be equal to the coupon rate.

The drawback of the current yield is that it considers only the coupon interest and no other source that will impact an investor's return. No consideration is given to the capital gain that the investor will realize when a bond is purchased at a discount and held to maturity; nor is there any recognition of the capital loss that the investor will realize if a bond purchased at a premium is held to maturity.

Yield to Maturity

The most popular measure of yield in the bond market is the *yield to maturity.* The yield to maturity is the interest rate that will make the present value of the cash flows from a bond equal to its market price plus accrued interest. To find the yield to maturity, we first determine the cash flows. Then an iterative procedure is used to find the interest rate that will make the present value of the cash flows equal to the market price plus accrued interest. In the illustrations presented below, we assume that the next coupon payment will be 6 months from now so that there is no accrued interest.

To illustrate, consider a 7% 8-year bond selling for \$94.17. The cash flows for this bond are (1) 16 payments every 6 months of \$3.50 and (2) a payment 16 6-month periods from now of \$100. The present value using various discount (interest) rates is:

Interest rate	3.5%	3.6%	3.7%	3.8%	3.9%	4.0%
Present value	100.00	98.80	97.62	96.45	95.30	94.17

When a 4.0% interest rate is used, the present value of the cash flows is equal to \$94.17, which is the price of the bond. Hence, 4.0% is the semiannual yield to maturity.

The market convention adopted is to double the semiannual yield and call that the yield to maturity. Thus, the yield to maturity for the above bond is 8% (2 times 4.0%). The yield to maturity computed using this convention—doubling the semiannual yield—is called a *bond-equivalent yield.*

The following relationships between the price of a bond, coupon rate, current yield, and yield to maturity hold:

Bond selling at	**Relationship**				
Par	Coupon rate	=	Current yield	=	YTM
Discount	Coupon rate	<	Current yield	<	YTM
Premium	Coupon rate	>	Current yield	>	YTM

The yield to maturity considers not only the coupon income but also any capital gain or loss that the investor will realize by holding the bond to maturity. The yield to maturity also considers the timing of the cash flows. It does consider reinvestment income; however, it assumes that the coupon payments can be reinvested at an interest rate equal to the yield to maturity. So, if the yield to maturity for a bond is 8%, for example, to earn that yield the coupon payments must be reinvested at an interest rate equal to 8%. The following illustration clearly demonstrates this point.

Suppose an investor has $94.17 and places the funds in a certificate of deposit that pays 4% every 6 months for 8 years or 8% per year (on a bond-equivalent basis). At the end of 8 years, the $94.17 investment will grow to $176.38. Instead, suppose an investor buys the following bond: a 7% 8-year bond selling for $94.17. The yield to maturity for this bond is 8%. The investor would expect that at the end of 8 years, the total dollars from the investment will be $176.38.

Let's look at what the investor will receive. There will be 16 semiannual interest payments of $3.50, which will total $56. When the bond matures, the investor will receive $100. Thus, the total dollars that the investor will receive is $156 by holding the bond to maturity. But this is less than the $176.38 necessary to produce a yield of 8% on a bond-equivalent basis by $20.38 ($176.38 minus $156). How is this deficiency supposed to be made up? If the investor reinvests the coupon payments at a semiannual interest rate of 4% (or 8% annual rate on a bond-equivalent basis), then the interest earned on the coupon payments will be $20.38. Consequently, of the $82.21 total dollar return ($176.38 minus $94.17) necessary to produce a yield of 8%, about 25% ($20.38 divided by $82.21) must be generated by reinvesting the coupon payments.

Clearly, the investor will only realize the yield to maturity that is stated at the time of purchase if (1) the coupon payments can be reinvested at the yield to maturity and (2) the bond is held to maturity. With respect to the first assumption, the risk that an investor faces is that future interest rates will be less than the yield to maturity at the time the bond is purchased. This risk is referred to as reinvestment risk—a risk we explain later in this chapter. If the bond is not held to maturity, it may have to be sold for less than its purchase price, resulting in a return that is less than the yield to maturity. The risk that a bond will have to be sold at a loss is referred to as interest rate risk as explained later in this chapter.

There are two characteristics of a bond that determine the degree of reinvestment risk. First, for a given yield to maturity and a given coupon rate, the longer the maturity the more the bond's total dollar return is dependent on reinvestment income to realize the yield to maturity at the time of purchase (that is, the greater the reinvestment risk). The implication is that the yield to maturity measure for long-term coupon bonds tells little about the potential yield that an investor may realize if the bond is held to maturity. For long-term bonds, in high interest rate environments the reinvestment income component may be as high as 70% of the bond's potential total dollar return.

The second characteristic that determines the degree of reinvestment risk is the coupon rate. For a given maturity and a given yield to maturity, the higher the coupon rate, the more dependent the bond's total dollar return will be on the reinvestment of the coupon payments in order to produce the yield to maturity at the time of purchase. This means that holding maturity and yield to maturity constant, premium bonds will be more dependent on reinvestment income than bonds selling at par. In contrast, discount bonds will be less dependent on reinvestment income than bonds selling at par. For zero-coupon bonds, none of the bond's total dollar return is dependent on reinvestment income. So, a zero-coupon bond has no reinvestment risk if held to maturity.

Yield to Call

When a bond is callable, the practice has been to calculate a *yield to call* as well as a yield to maturity. As explained earlier, a callable bond may have a call schedule. The yield to call assumes that the issuer will call the bond at some assumed call date and the call price is then the call price specified in the call schedule.

Typically, investors calculate a yield to first call or yield to next call, a yield to first par call, and yield to refunding. The *yield to first call* is computed for an issue that is not currently callable, while the *yield to next call* is computed for an issue that is currently callable. *Yield to refunding* is used when bonds are currently callable but have some restrictions on the source of funds used to buy back the debt when a call is exercised. The refunding date is the first date the bond can be called using lower-cost debt.

The procedure for calculating any yield to call measure is the same as for any yield calculation: determine the interest rate that will make the present value of the expected cash flows equal to the price plus accrued interest. In the case of yield to first call, the expected cash flows are the coupon payments to the first call date and the call price. For the yield to first par call, the expected cash flows are the coupon payments to the first date at which the issuer can call the bond at par and the par value. For the yield to refunding, the expected cash flows are the coupon payments to the first refunding date and the call price at the first refunding date.

To illustrate the computation, consider a 7% 8-year bond with a maturity value of $100 selling for $106.36. Suppose that the first call date is 3 years from now and the call price is $103. The cash flows for this bond if it is called in 3 years are (1) six coupon payments of $3.50 and (2) $103 in six 6-month periods from now. The process for finding the yield to first call is the same as for finding the yield to maturity. It can be shown that a semiannual interest rate of 2.8% makes the present value of the cash flows equal to the price is 2.8%. Therefore, the yield to first call on a bond-equivalent basis is 5.6%.

Let's take a closer look at the yield to call as a measure of the potential return of a security. The yield to call does consider all three sources of potential return from owning a bond. However, as in the case of the yield to maturity, it assumes that all cash flows can be reinvested at the yield to call until the assumed call date. As we just demonstrated, this assumption may be inappropriate. Moreover, the yield to call assumes that (1) the investor will hold the bond to the assumed call date and (2) the issuer will call the bond on that date.

These assumptions underlying the yield to call are often unrealistic. They do not take into account how an investor will reinvest the proceeds if the issue is called. For example, consider two bonds, M and N. Suppose that the yield to maturity for bond M, a 5-year noncallable bond, is 7.5%, while for bond N the yield to call assuming the bond will be called in 3 years is 7.8%. Which bond is better for an investor with a 5-year investment horizon? It's not possible to tell for the yields cited. If the investor intends to hold the bond for 5 years and the issuer calls bond N after 3 years, the total dollars that will be available at the end of 5 years will depend on the interest rate that can be earned from investing funds from the call date to the end of the investment horizon.

Yield to Put

When a bond is puttable, the yield to the first put date is calculated. The yield to put is the interest rate that will make the present value of the cash flows to the first put date equal to the price plus accrued interest. As with all yield measures (except the current yield), yield to put assumes that any interim coupon payments can be reinvested at the yield calculated. Moreover, the yield to put assumes that the bond will be put on the first put date.

Yield to Worst

A yield can be calculated for every possible call date and put date. In addition, a yield to maturity can be calculated. The lowest of all these possible yields is called the yield to worst. For example, suppose that there are only four possible call dates for a callable bond and that a yield to call assuming each possible call date is 6%, 6.2%, 5.8%, and 5.7%, and that the yield to maturity is 7.5%. Then the yield to worst is the minimum of these values, 5.7% in our example.

The yield to worst measure holds little meaning as a measure of potential return.

Cash Flow Yield

Mortgage-backed securities and asset-backed securities are backed by a pool of loans. The cash flows for these securities include principal repayment as well as interest. The complication that arises is that the individual borrowers whose loans make up the pool typically can prepay their loan in whole or in part prior to the scheduled principal repayment date. Because of prepayments, in order to project the cash flows it is necessary to make an assumption about the rate at which prepayments will occur. This rate is called the prepayment rate or prepayment speed.

Given the cash flows based on the assumed prepayment rate, a yield can be calculated. The yield is the interest rate that will make the present value of the projected cash flows equal to the price plus accrued interest. A yield calculated in this way is called a *cash flow yield.*

Typically, the cash flows for mortgage-backed and asset-backed securities are monthly. Therefore, the interest rate that will make the present value of the projected principal repayment and interest payments equal to the market price plus accrued interest is a monthly rate. The bond-equivalent yield is found by calculating the effective 6-month interest rate and then doubling it. That is:

$$\begin{aligned}&\text{Cash flow yield on a bond-equivalent basis}\\&\quad(\text{if monthly pay}) = 2[(1+\text{Monthly yield})^6 - 1]\end{aligned}$$

For example, if the monthly yield is 0.5%, then

$$\begin{aligned}&\text{Cash flow yield on a bond-equivalent basis}\\&\quad = 2[(1.005)^6 - 1] = 6.08\%\end{aligned}$$

As we have noted, the yield to maturity has two shortcomings as a measure of a bond's potential return: (1) it is assumed that the coupon payments can be reinvested at a rate equal to the yield to maturity, and (2) it is assumed that the bond is held to maturity. These shortcomings are equally present in application of the cash flow yield measure: (1) the projected cash flows are assumed to be reinvested at the cash flow yield, and (2) the mortgage-backed or asset-backed security is assumed to be held until the final payoff of all the loans based on some prepayment assumption. The importance of reinvestment risk—the risk that the cash flows will be reinvested at a rate less than the cash flow yield—is particularly important for mortgage-backed and asset-backed securities since payments are typically monthly and include principal repayments (scheduled and prepayments), as well as interest. Moreover, the cash flow yield is dependent on realization of the projected cash flows according to some prepayment rate. If actual prepayments differ significantly from the prepayment rate assumed, the cash flow yield will not be realized.

RISKS ASSOCIATED WITH INVESTING IN BONDS

Bonds expose an investor to one or more of the following risks: (1) interest rate risk; (2) call and prepayment risk; (3) credit risk; (4) liquidity risk; (5) exchange rate or currency risk; and (6) inflation or purchasing power risk.

Interest Rate Risk

The price of a typical bond will change in the opposite direction from a change in interest rates. That is, when interest rates rise, a bond's price will fall; when interest rates fall, a bond's price will rise. For example, consider a 6% 20-year bond. If the yield investors require to buy this bond is 6%, the price of this bond would be $100. However, if the required yield increased to 6.5%, the price of this bond would decline to $94.4479. Thus, for a 50-basis-point increase in yield, the bond's price declines by 5.55%. If, instead, the yield declines from 6% to 5.5%, the bond's price will rise by 6.02% to $106.0195.

The reason for this inverse relationship between price and changes in interest rates or changes in market yields is as follows. Suppose investor X purchases our hypothetical 6% coupon 20-year bond at par value ($100). The yield for this bond is 6%. Suppose that immediately after the purchase of this bond two things happen. First, market interest rates rise to 6.50% so that if an investor wants to buy a similar 20-year bond a 6.50% coupon rate would have to be paid by the bond issuer in order to offer the bond at par value. Second, suppose investor X wants to sell the bond. In attempting to sell the bond, investor X would not find an investor who would be willing to pay par value for a bond with a coupon rate of 6%. The reason is that any investor who wanted to purchase this bond could obtain a similar 20-year bond with a coupon rate 50 basis points higher, 6.5%. What can the investor do? The investor cannot force the issuer to change the coupon rate to 6.5%. Nor can the investor force the issuer to shorten the maturity of the bond to a point where a new investor would be willing to accept a 6% coupon rate. The only thing that the investor can do is adjust the price of the bond so that at the new price the buyer would realize a yield of 6.5%. This means that the price would have to be adjusted down to a price below par value. The new price must be $94.4469. While we assumed in our illustration an initial price of par value, the principle holds for any purchase price. Regardless of the price that an investor pays for a bond, an increase in market interest rates will result in a decline in a bond's price.

Suppose instead of a rise in market interest rates to 6.5%, they decline to 5.5%. Investors would be more than happy to purchase the 6% coupon 20-year bond for par value. However, investor X realizes that the market is only offering investors the opportunity to buy a similar bond at par value with a coupon rate of 5.5%. Consequently, investor X will increase the price of the bond until it offers a yield of 5.5%. That price is $106.0195.

Since the price of a bond fluctuates with market interest rates, the risk that an investor faces is that the price of a bond held in a portfolio will decline if market interest rates rise. This risk is referred to as *interest rate risk* and is a major risk faced by investors in the bond market.

Bond Features that Affect Interest Rate Risk

The degree of sensitivity of a bond's price to changes in market interest rates depends on various characteristics of the issue, such as maturity and coupon rate. Consider first maturity. All other factors constant, the longer the maturity, the greater the bond's price sensitivity to changes in interest rates. For example, we know that for a 6% 20-year bond selling to yield 6%, a rise in the yield required by investors to 6.5% will cause the bond's price to decline from $100 to $94.4479, a 5.55% price decline. For a 6% 5-year bond selling to yield 6%, the price is $100. A rise in the yield required by investors from 6% to 6.5% would decrease the price to $97.8944. The decline in the bond's price is only 2.11%.

Now let's turn to the coupon rate. A property of a bond is that all other factors constant, the lower the coupon rate, the greater the bond's price sensitivity to changes in interest rates. For example, consider a 9% 20-year bond selling to yield 6%. The price of this bond would be $112.7953. If the yield required by investors increases by 50 basis points to 6.5%, the price of this bond would fall by 2.01% to $110.5280. This decline is less than the 5.55% decline for the 6% 20-year bond selling to yield 6%. An implication is that zero-coupon bonds have greater price sensitivity to interest rate changes than same-maturity bonds bearing a coupon rate and trading at the same yield.

Because of default or credit risk (discussed later), different bonds trade at different yields, even if they have the same coupon rate and maturity. How, then, holding other factors constant, does the level of interest rates affect a bond's price sensitivity to changes in interest rates? As it turns out, the higher the level of interest rates that a bond trades at, the lower the price sensitivity.

To see this, we can compare a 6% 20-year bond initially selling at a yield of 6%, and a 6% 20-year bond initially selling at a yield of 10%. The former is initially at a price of $100, and the latter carries a price of $65.68. Now, if the yield on both bonds increases by 100 basis points, the first bond trades down by 10.68 points (10.68%). After the assumed increase in yield, the second bond will trade at a price of $59.88, for a price decline of only 5.80 points (or 8.83%). Thus, we see that the bond that trades at a lower yield is more volatile in both percentage price change and absolute price change, as long as the other bond characteristics are the same. An implication of this is that, for a given change in interest rates, price sensitivity is lower when the level of interest rates in the market is high, and price sensitivity is higher when the level of interest rates is low.

We can summarize these three characteristics that affect the bond's price sensitivity to changes in market interest rates as follows:

Characteristic 1: For a given maturity and initial yield, the lower the coupon rate the greater the bond's price sensitivity to changes in market interest rates.

Characteristic 2: For a given coupon rate and initial yield, the longer the maturity of a bond the greater the bond's price sensitivity to changes in market interest rates.

Characteristic 3: For a given coupon rate and maturity, the lower the level of interest rates the greater the bond's price sensitivity to changes in market interest rates.

A bond's price sensitivity bond will also depend on any options embedded in the issue. This is explained below when we discuss call risk.

Interest Rate Risk for Floating-Rate Securities

The change in the price of a fixed-rate coupon bond when market interest rates change is due to the fact that the bond's coupon rate differs from the prevailing market interest rate. For a floating-rate security, the coupon rate is reset periodically based on the prevailing value for the reference rate plus the contractually specified index spread. The index spread is set for the life of the security. The price of a floating-rate security will fluctuate depending on the following three factors.

First, the longer the time to the next coupon reset date, the greater the potential price fluctuation. For example, consider a floating-rate security whose coupon resets every six months and the coupon formula is 6-month

LIBOR plus 20 basis points. Suppose that on the coupon reset date 6-month LIBOR is 5.8%. If the next day after the coupon is reset, 6-month LIBOR rises to 6.1%, this means that this security is offering a 6-month coupon rate that is less than the prevailing 6-month rate for the remaining six months. The price of the security must decline to reflect this. Suppose instead that the coupon resets every month at 1-month LIBOR and that this rate rises right after a coupon rate is reset. Then, while the investor would be realizing a submarket 1-month coupon rate, it is for only a month. The price decline will be less than for the security that resets every six months.

The second reason why a floating-rate security's price will fluctuate is that the index spread that investors want in the market changes. For example, consider once again the security whose coupon reset formula is 6-month LIBOR plus 20 basis point. If market conditions change such that investors want an index spread of 30 basis points rather than 20 basis points, this security would be offering a coupon rate that is 10 basis points below the market rate. As a result, the security's price will decline.

Finally, as noted earlier, a floating-rate security may have a cap. Once the coupon rate as specified by the coupon reset formula rises above the cap rate, the security offers a below market coupon rate and its price will decline. In fact, once the cap is reached, the security's price will react much the same way to changes in market interest rates as that of a fixed-rate coupon security.

Measuring Interest Rate Risk

Investors are interested in estimating the price sensitivity of a bond to changes in market interest rates. The measure commonly used to approximate the percentage price change is duration. Duration gives the approximate percentage price change for a 100 basis point change in interest rates. Chapters 13 and 14 of Volume III explains the concept of duration and its measurement.

The duration for the 6% coupon 5-year bond trading at par to yield 6% is 4.27. Thus, the price of this bond will change by approximately 4.27% if interest rates change by 100 basis points. For a 50 basis point change, this bond's price will change by approximately 2.14% (4.27% divided by 2). As explained above, this bond's price would actually change by 2.11%. Thus, duration does a good job of approximating the percentage price change. It turns out that the approximation is good the smaller the change in interest rates. The approximation is not as good for a large change in interest rates.

Call and Prepayment Risk

As explained earlier, a bond may include a provision that allows the issuer to retire or call all or part of the issue before the maturity date. From the investor's perspective, there are three disadvantages to call provisions. First, the cash flow pattern of a callable bond is not known with certainty. Second, because the issuer will call the bonds when interest rates have dropped, the investor is exposed to reinvestment risk; that is, the investor will have to reinvest the proceeds when the bond is called at relatively lower interest rates. Finally, the capital appreciation potential of a bond will be reduced, because the price of a callable bond may not rise much above the price at which the issuer will call the bond. Because of these disadvantages faced by the investor, a callable bond is said to expose the investor *to call risk*. The same disadvantages apply to bonds that can prepay. In this case the risk is referred to as *prepayment risk.*

Credit Risk

In general, one thinks of credit risk as the risk that the debtor will fail to satisfy its obligation to the lender (that is, timely payment of principal and/or interest). That is in fact one form of risk referred to as *default risk.* Default risk is gauged by credit ratings assigned by three nationally recognized statistical rating companies: Moody's Investors Service, Standard & Poor's Corporation, and Fitch Ratings. These organizations are popularly referred to as *rating agencies.* We discuss these ratings assigned in Chapter 24 of Volume III.

Bonds with default risk trade in the market at a price that is lower than comparable U.S. government securities, which are considered free of default risk. In other words, a non-U.S. government taxable bond will trade in the market at a higher yield than a U.S. government taxable bond that is otherwise comparable in terms of maturity and coupon rate.

Except in the case of the lowest-rated securities, known as "high-yield" or "junk bonds," an investor is normally more concerned with the changes in the perceived default risk than with the actual event of default. Even though the actual default of an issuer may be highly unlikely, an investor is concerned about the impact that a change in perceived default risk can have on a bond's price. If the perceived default risk increases, the market will require a higher yield for the security. As a result, a bond's price will decline. This risk is referred to as *credit spread risk.* A decline in the price of a bond will also occur if an issue's credit rating is lowered. By a lower credit rating, it is meant the issue is "downgraded." This risk is referred to as *downgrade risk.* Credit spread risk and downgrade risk are discussed in Chapter 24 of Volume III.

Liquidity Risk

When an investor wants to sell a bond prior to the maturity date, he or she is concerned whether the price that can be obtained from dealers is close to the true value of the issue. For example, if recent trades in the market for a particular issue have been between 97.25 and 97.75 and market conditions have not changed, an investor would expect to sell the bond somewhere in the 97.25 to 97.75 area.

Liquidity risk is the risk that the investor will have to sell a bond below its true value where the true value is indicated by recent transactions. The primary measure of liquidity is the size of the spread between the bid price (the price at which a dealer is willing to buy a security) and the ask price (the price at which a dealer is willing to sell a security). The wider the bid-ask spread, the greater the liquidity risk.

A liquid market can generally be defined by "small bid-ask spreads which do not materially increase for large transactions" (Gerber, 1997, p. 278). Bid-ask spreads, and therefore liquidity risk, change over time.

For investors who plan to hold a bond until maturity and need not mark a position to market, liquidity risk is not a major concern. An institutional investor that plans to hold an issue to maturity but is periodically marked to market is concerned with liquidity risk. By marking a position to market, it is meant that the security is revalued in the portfolio based on its current market price. For example, mutual funds are required to mark to market at the end of each day the holdings in their portfolio in order to compute the net asset value (NAV). While other institutional investors may not mark to market as frequently as mutual funds, they are marked to market when reports are periodically sent to clients or the board of directors or trustees.

Exchange Rate or Currency Risk

For a U.S. investor, non-dollar-denominated bond (that is, a bond whose payments are not in U.S. dollars) has unknown U.S. dollar cash flows. The dollar cash flows are dependent on the exchange rate at the time the payments are received. For example, suppose a U.S. investor purchases a bond whose payments are in euros. If the euro depreciates relative to the U.S. dollar, then fewer dollars will be received. The risk of this occurring is referred to as *exchange rate risk* or *currency risk.* Of course, should the euro appreciate relative to the U.S. dollar, the investor will benefit by receiving more dollars.

Inflation or Purchasing Power Risk

Inflation risk or *purchasing power risk* arises because of the variation in the value of cash flows from a security due to inflation, as measured in terms of purchasing power. For example, if an investor purchases a bond with a coupon rate of 7%, but the rate of inflation is 8%, the purchasing power of the cash flow has declined. For all but floating-rate securities, an investor is exposed to inflation risk because the interest rate the issuer promises to make is fixed for the life of the issue. To the extent that interest rates reflect the expected inflation rate, floating-rate securities have a lower level of inflation risk.

SUMMARY

Basically, a bond is a financial obligation of an entity (the issuer) who promises to pay a specified sum of money at specified future dates. In this chapter we have described the basic features of bonds and their investment characteristics.

Bond prices are quoted as a percentage of par value, with par value equal to 100. The coupon rate is the interest rate that the issuer agrees to pay each year; the coupon is the annual amount of the interest payment and is found by multiplying the par value by the coupon rate. Zero-coupon bonds do not make periodic coupon payments; the bondholder realizes interest at the maturity date equal to the difference between the maturity value and the price paid for the bond. A step-up note is a security whose coupon rate increases over time.

A floating-rate security is an issue whose coupon rate resets periodically based on some formula; the typical coupon reset formula is some reference rate plus an index spread. A floating-rate security may have a cap which sets the maximum coupon rate that will be paid at a reset date; a cap is a disadvantage to the bondholder while a floor is an advantage to the bondholder. An inverse floater is an issue whose coupon rate moves in the opposite direction from the change in the reference rate.

Accrued interest is the amount of interest accrued since the last coupon payment and in the United States (as well as in many countries), the bond buyer must pay the bond seller the accrued interest. The full price of a security is the agreed-upon price plus accrued interest; the clean price is the agreed-upon price without accrued interest. Interest accrues on a bond from and including the date of the previous coupon up to but excluding the value date; the value date is usually, but not always, the same as the settlement date.

A bond issue may have a call provision granting the issuer an option to retire all or part of the issue prior to the stated maturity date. A call provision is an advantage to the issuer and a disadvantage to the bondholder. When a callable bond is issued, typically the issuer may not call the bond for a number of years; that is, there is a deferred call. Most new bond issues, even if currently callable, usually have some restrictions against refunding. For an amortizing security backed by a pool of loans, the borrowers typically have the right to prepay in whole or in part prior to the scheduled principal repayment date; this provision is called a prepayment option.

A puttable bond is one in which the bondholder has the right to sell the issue back to the issuer at a specified price on designated dates. A convertible bond is an issue giving the bondholder the right to exchange the bond for a specified number of shares of common stock.

The sources of return from holding a bond to maturity are the coupon interest payments, any capital gain or loss, and reinvestment income. Reinvestment income is the interest income generated by reinvesting coupon interest payments and any principal repayments from the time of receipt to the bond's maturity. The current yield relates the annual dollar coupon interest to the market price and fails to recognize any capital gain or loss and reinvestment income. The yield to maturity is the interest rate that will make the present value of the cash flows from a bond equal to the price plus accrued interest. This yield measure will only be realized if the interim cash flows can be reinvested at the yield to maturity and the bond is held to maturity. The yield to call is the interest rate that will make the present value of the expected cash flows to the assumed call date equal to the price plus accrued interest. Yield measures for callable bonds include yield to first call, yield to next call, yield to first par call, and yield to refunding. The yield to worst is the lowest yield from among all possible yield to calls, yield to puts, and the yield to maturity. For mortgage-backed and asset-backed securities, the cash flow yield based on some prepayment rate is the interest rate that equates the present value of the projected principal and interest payments to the price plus accrued interest. The cash flow yield assumes that all cash flows (principal payments and interest payments) can be reinvested at the calculated yield and that the prepayment rate will be realized over the security's life.

Bonds expose an investor to various risks. The price of a bond changes inversely with a change in market interest rates. Interest rate risk refers to the adverse price movement of a bond as a result of a change in market interest rates; for the owner of a bond it is the risk that interest rates will rise. The coupon rate and maturity of a bond affect its price sensitivity to changes in market interest rate. The duration of a bond measures the approximate percentage price change for a 100-basis-point change in interest rates.

Call risk and prepayment risk refer to the risk that a security will be paid off before the scheduled principal repayment date. From an investor's perspective, the disadvantages to call and prepayment provisions are (1) the cash flow pattern is uncertain, (2) reinvestment risk because proceeds received will have to be reinvested at a relatively lower interest rate, and (3) the capital appreciation potential of a bond will be reduced.

Credit risk consists of three types of risk: default risk, credit spread risk, and downgrade risk. Default risk is gauged by the ratings assigned by the nationally recognized statistical rating organizations (rating agencies).

Liquidity risk depends on the ease with which an issue can be sold at or near its true value and is primarily gauged by the bid-ask spread quoted by a dealer. From the perspective of a U.S. investor, exchange rate risk is the risk that a currency in which a security is denominated will depreciate relative to the U.S. dollar. Inflation risk or purchasing power risk arises because of the variation in the value of cash flows from a security due to inflation.

REFERENCES

Choudhry, M., and Fabozzi, F. J. (eds.) (2004). *Handbook of European Fixed Income Securities*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (1999). *Duration, Convexity, and Other Bond Risk Measures*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (ed.) (2000). *Investing in Asset-Backed Securities*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (2002). *Fixed Income Securities*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (ed.) (2005). *Handbook of Fixed Income Securities*, 7th edition. New York: McGraw-Hill.

Fabozzi, F. J. (2006). *Bond Markets, Analysis, and Strategies*, 6th edition. Upper Saddle River, NJ: Prentice Hall.

Fabozzi, F. J., Bhattacharya, A. K., and Berliner W. S. (2007). *Mortgage-Backed Securities: Products, Structuring, and Analytical Techniques*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., and Mann, S.V. (2001). *Floating Rate Securities*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., Mann, S. V., and Choudhry, M. (2002). *Global Money Markets*. Hoboken, NJ: John Wiley & Sons.

Gerber, R. I. (1997). A user's guide to buy-side bond trading. In F. J. Fabozzi (ed.), *Managing Fixed Income Portfolios* (pp. 277–290). Hoboken, NJ: John Wiley & Sons.

Wilson, R. W., and Fabozzi, F. J. (1996). *Corporate Bonds: Structures and Analysis*. Hoboken, NJ: John Wiley & Sons.

CHAPTER 18

Residential Mortgages

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

ANAND K. BHATTACHARYA, PhD
Managing Director, Countrywide Securities Corporation

WILLIAM S. BERLINER
Executive Vice President, Countrywide Securities Corporation

Abstract: The mortgage market in the United States has emerged as one of the world's largest asset classes. The growth of the mortgage market is attributable to a variety of factors. Most notably, strong sales and price growth in the domestic real estate markets and the increased acceptance of new loan products on the part of the consumer has dovetailed with the acceptance of a variety of loan products as collateral for securitizations. Due to a variety of reasons such as product innovation, technological advancement, and demographic and cultural changes, the composition of the primary mortgage market is evolving at a rapid rate—older concepts are being updated, while a host of new products is also being developed and marketed. Consequently, the mortgage-lending paradigm continues to be refined in ways that have allowed lenders to offer a large variety of products designed to appeal to consumer needs and tastes. This evolution has been facilitated by sophistication in pricing that has allowed for the quantification of the inherent risks in such loans. At the same time, structures and techniques that allow the burgeoning variety of products to be securitized have been created and marketed, helping to meet the investment needs of a variety of market segments and investor clienteles.

Keywords: mortgage, servicers, lien status, prime loans, subprime loans, alternative-A loans, credit scores, loan-to-value ratio (LTV), documentation, fixed-rate mortgages, adjustable-rate mortgages (ARMs), hybrid ARM, interest-only (IO) mortgage, government guarantees, conventional loans, guaranty fee, prepayments, refinancing, curtailment, negative convexity, delinquencies, defaults, loss severity

The purpose of this chapter is to explain mortgage products and their investment characteristics. The chapter introduces the basic tenets of the primary mortgage market and mortgage lending, and summarizes the various product offerings in the sector. This chapter provides a framework for understanding mortgage-backed securities (MBSs).

OVERVIEW OF MORTGAGES

In general, a *mortgage* is a loan that is secured by underlying assets that can be repossessed in the event of default. For the purposes of this chapter, a mortgage is defined as a loan made to the owner of a one- to four-family residential dwelling and secured by the underlying property (both the land and the structure or "improvement"). After issuance, loans must be managed (or "serviced") by units that, for a fee, collect payments from borrowers and pass them on to investors. In addition to managing and tracking payments, *servicers* are also responsible for interfacing with borrowers if they become delinquent on their payments, and also manage the disposition of the loan and the underlying property if the loan goes into foreclosure.

Key Attributes that Define Mortgages

There are a number of key attributes that define the instruments in question that can be characterized by the following dimensions:

- Lien status, original loan term
- Credit classification
- Interest rate type
- Amortization type
- Credit guarantees
- Loan balances
- Prepayments and prepayment penalties

We discuss each below.

Lien Status

The *lien status* dictates the loan's seniority in the event of the forced liquidation of the property due to default by the obligor. A first lien implies that a creditor would have first call on the proceeds of the liquidation of the property if it were to be repossessed. Borrowers often utilize second lien or junior loans as a means of liquefying the value of a home for the purpose of expenditures such as medical bills or college tuition or investments such as home improvements.

Original Loan Term

The great majority of mortgages are originated with a 30-year original term. Loans with shorter stated terms are also utilized by those borrowers seeking to amortize their loans faster and build equity in their homes more quickly. The 15-year mortgage is the most common short-amortization instrument, although issuance of loans with 20- and 10-year terms has grown in recent years.

Credit Classification

The majority of loans originated are of high-credit quality, where the borrowers have strong employment and credit histories, income sufficient to pay the loans without compromising their creditworthiness, and substantial equity in the underlying property. These loans are broadly classified as *prime loans*, and have historically experienced low incidences of delinquency and default.

Loans of lower initial credit quality, which are more likely to experience significantly higher levels of default, are classified as *subprime loans*. Subprime loan underwriting often utilizes nontraditional measures to assess credit risk, as these borrowers often have lower income levels, fewer assets, and blemished credit histories. After issuance, these loans are typically serviced by special units designed to closely monitor the payments of subprime borrowers. In the event that subprime borrowers become delinquent, the servicers move immediately to either assist the borrowers in becoming current or mitigate the potential for losses resulting from loan defaults.

Between the prime and subprime sector is a somewhat nebulous category referenced as *alternative-A loans* or, more commonly, alt-A loans. These loans are considered to be prime loans (the "A" refers to the A grade assigned by underwriting systems), albeit with some attributes that either increase their perceived credit riskiness or cause them to be difficult to categorize and evaluate.

Mortgage credit analysis employs a number of different metrics, including the following.

Credit Scores Several firms collect data on the payment histories of individuals from lending institutions and use sophisticated models to evaluate and quantify individual creditworthiness. The process results in a credit score, which is essentially a numerical grade of the credit history of the borrower. There are three different credit-reporting firms that calculate credit scores: Experian (which uses the Fair Isaacs or FICO model), Transunion (which supports the Emperica model), and Equifax (whose model is known as Beacon). While each firm's credit scores are based on different data sets and scoring algorithms, the scores are generically referred to as FICO scores.

Loan-to-Value Ratios The *loan-to-value ratio (LTV)* is an indicator of borrower leverage at the point when the loan application is filed. The LTV calculation compares the face value of the desired loan to the market value of the property. By definition, the LTV of the loan in the purchase transaction is a function of both the down payment and the purchase price of the property. In a refinancing, the LTV is dependent on the requested amount of the new loan and the market value of the property as determined by an appraisal. If the new loan is larger than the original loan, the transaction is referred to as a cash-out refinancing, while a refinancing where the loan balance remains unchanged is described as a rate-and-term refinancing or no-cash refinancing.

The LTV is important for a number of reasons. First, it is an indicator of the amount that can be recovered from a loan in the event of a default, especially if the value of the property declines. The level of the LTV also has an impact on the expected payment performance of the obligor, as high LTVs indicate a greater likelihood of default on the loan. Another useful measure is the combined LTV (or CLTV), which accounts for the existence of second liens. A $100,000 property with an $80,000 first lien and a $10,000 second lien will have an LTV of 80% but a CLTV of 90%.

Income Ratios In order to ensure that borrower obligations are consistent with their income, lenders calculate income ratios that compare the potential monthly payment on the loan to the applicant's monthly income. The most common measures are called front and back ratios. The front ratio is calculated by dividing the total monthly payments on the home (including principal, interest, property taxes, and homeowners insurance) by pretax monthly income. The back ratio is similar, but adds other debt payments (including auto loan and credit card payments) to the total payments. In order for a loan to be classified as prime, the front and back ratios should be no more than 28% and 36%, respectively. (Because consumer debt figures can be somewhat inconsistent and nebulous, the front ratio is generally considered the more reliable measure, and accorded greater weight by underwriters.)

Documentation Lenders traditionally have required potential borrowers to provide data on their financial status, and support the data with documentation. Loan officers typically required applicants to report and document income, employment status, and financial resources (including the source of the down payment for the transaction). Part of the application process routinely involved compiling documents such as tax returns and bank statements for use in the underwriting process. However, a growing number of loan programs have more flexible documentation requirements, and lenders typically offer programs with a variety of documentation standards. Such programs include programs where pay stubs and tax returns are not required (especially in cases where existing customers refinance their loans), as well as "stated" programs (where income levels and asset values are provided, but not independently verified).

Characterizing Prime versus Subprime Loans

The primary attribute used to characterize loans as either prime or subprime is the credit score. Prime (or A-grade) loans generally have FICO scores of 660 or higher, income ratios with the previously noted maximum of 28% and 36%, and LTVs less than 95%. Alt-A loans may vary in a number of important ways. Alt-A loans typically have lower degrees of documentation, are backed by a second home or investor property, or have a combination of attributes (such as large loan size and high LTV) that make the loan riskier. While subprime loans typically have FICO scores below 660, the loan programs and grades are highly lender-specific. One lender might consider a loan with a 620 FICO to be a B-rated loan, while another lender would grade the same loan higher or lower, especially if the other attributes of the loan (such as the LTV) are higher or lower than average levels.

Interest Rate Type

Fixed-rate mortgages have an interest rate (or *note rate*) that is set at the closing of the loan (or, more accurately, when the rate is "locked"), and is constant for the loan's term. Based on the loan's balance, interest rate, and term, a payment schedule effective over the life of the loan is calculated to amortize the principal balance.

Adjustable-rate mortgages (ARMs), as the name implies, have note rates that change over the life of the loan. The note rate is based on both the movement of an underlying rate (the index) and a spread over the index (the margin) required for the particular loan program. A number of different indexes can be used as a reference rate in determining the loan's note rate when the loan "resets," including the London Interbank Offered Rate (LIBOR), one-year Constant Maturity Treasury (CMT), or the 12-month Moving Treasury Average (MTA), a rate calculated from monthly averages of the one-year CMT. An ARM's note rate resets at the end of the initial period and subsequently resets periodically, subject to caps and floors that limit how much the loan's note rate can change. ARMs most frequently are structured to reset annually, although some products reset on a monthly or semiannual basis. Since the loan's rate and payment can (and often does) reset higher, the borrower can experience "payment shock" if the monthly payment increases significantly.

Traditionally, ARMs had a one-year initial period where the start rate was effective, often referred to as the "teaser" rate (since the rate was set at a relatively low rate in order to entice borrowers.) The loans reset at the end of the teaser period, and continued to reset annually for the life of the loan. One-year ARMs, however, are no longer popular products, and have been replaced by loans that have features more appealing to borrowers.

There are two broad types of ARM loans. One is the fixed-period ARM or *hybrid ARM*, which have fixed initial rates that are effective for longer periods of time (3-, 5- 7-, and 10-years) after funding. At the end of the initial fixed-rate period, the loans reset in a fashion very similar to that of more traditional ARM loans. Hybrid ARMs typically have three rate caps: initial cap, periodic cap, and life cap. The initial cap and periodic cap limit how much the note rate of the loans can change at the end of the fixed period and at each subsequent reset, respectively, while the life cap dictates the maximum level of the note rate.

At the opposite end of the spectrum is the payment-option ARM or negative amortization ARM. Such products are structured as monthly-reset ARMs that begin with a very low teaser rate. While the rate adjusts monthly, the minimum payment is adjusted only on an annual basis and is subject to a payment cap that limits how much the loan's payment can change at the reset. In instances where the payment made is not sufficient to cover the interest due on the loan, the loan's balance increases in a phenomenon called "negative amortization." (The mechanics of negative amortization loans are addressed in more depth later in this chapter.)

Amortization Type

Traditionally, both fixed and adjustable rate mortgages were fully amortizing loans, indicating that the obligor's principal and interest payments are calculated in equal increments to pay off the loan over the stated term. Fully amortizing, fixed rate loans have a payment that is constant over the life of the loan. Since the payments on ARMs adjust periodically, their payments are recalculated at each reset for the loan's remaining balance at the new effective rate in a process called recasting the loan.

A recent trend in the market, however, has been the growing popularity of nontraditional amortization schemes. The most straightforward of these innovations is the *interest-only* or *IO product*. These loans require only interest to be paid for a predetermined period of time. After the expiration of the interest-only or lockout period, the loan is recast to amortize over the remaining term of the loan. The inclusion of principal to the payments at that point amortized over the remaining (and shorter) term of the loan causes the loan's payment to rise significantly after the recast, creating payment shock analogous to that experienced when an ARM resets.

The interest-only product was introduced in the hybrid ARM market, where the terms of the interest-only and fixed-rate periods were contiguous. A by-product of the interest-only ARM can be large changes in the borrower's monthly payment, the result of the combination of post-reset rate increases and the introduction of principal amortization. However, fixed-rate, interest-only products have recently grown in popularity. These are loans with a 30-year maturity that have a fixed rate throughout the life of the loan, but have a fairly long interest-only period (normally 10 years, although 15-year interest-only products are also being produced.) The loans subsequently amortize over their remaining terms. These products were designed to appeal to borrowers seeking the lower payments of interest-only products without the rate risk associated with adjustable rate products.

Another recent innovation is the noncontiguous interest-only hybrid ARM, where the interest-only period is different from the duration of the fixed rate period. As an example, a 5/1 hybrid ARM might have an interest-only period of 10 years. When the fixed period of a hybrid ARM is concluded, the loan's rate resets in the same fashion as other ARMs. However, only interest is paid on the loan until the recast date. These products were developed to spread out the payment shock that occurs when ARM loans reset and recast simultaneously.

Credit Guarantees

The ability of mortgage banks to continually originate mortgages is heavily dependent upon the ability to create fungible assets from a disparate group of loans made to a multitude of individual obligors. These assets are then sold (in the form of loans or, more commonly, MBS) into the capital markets, with the proceeds being recycled into new lending. Therefore, mortgage loans can be further classified based upon whether a credit guaranty associated with the loan is provided by the federal government or quasi-governmental entities, or obtained through other private entities or structural means.

Loans that are backed by agencies of the federal government are referred to under the generic term of government loans. As part of housing policy considerations, the Department of Housing and Urban Development (HUD) oversees two agencies, the Federal Housing Administration (FHA) and the Department of Veterans Affairs (often referred to simply as the Veterans Administration or VA), that support housing credit for qualifying borrowers. The FHA provides loan guarantees for those borrowers who can afford only a low down payment and generally also have relatively low levels of income. The VA guarantees loans made to veterans, allowing them to receive favorable loan terms. These guarantees are backed by the U.S. Department of the Treasury, thus providing these loans with the "full faith and credit" backing of the U.S. government. Government loans are securitized largely through the aegis of the Government National Mortgage Association (GNMA or Ginnie Mae), an agency also overseen by HUD.

So-called *conventional loans* have no explicit guaranty from the federal government. Conventional loans can be securitized either as "private-label" structures or as pools guaranteed by the two government-sponsored enterprises (GSEs), namely Freddie Mac (FHLMC) and Fannie Mae (FNMA). The GSEs are shareholder-owned corporations that were created by Congress in order to support housing activity. While neither enterprise has an overt government guaranty, market convention has always reflected the presumption that the government would provide assistance to the GSEs in the event of financial setbacks that threaten their viability. As we will see later in this chapter, the GSEs insure the payment of principal and interest to investors in exchange for a *guaranty fee*, paid either out of the loan's interest proceeds or as a lump sum at issuance.

Conventional loans that are not guaranteed by the GSEs can be securitized as private-label transactions. Traditionally, loans were securitized in private-label form because they were not eligible for GSE guarantees, either because of their balance or their credit attributes. A recent development is the growth of private-label deals backed either entirely or in part by loans where the balance conforms to the GSEs' limits. In such deals, the originator finds it more economical to enhance the loans' credit using the mechanisms of the private market (most commonly through subordination) than through the auspices of a GSE.

Loan Balances

The agencies have limits on the loan balance that can be included in agency-guaranteed pools. The maximum loan sizes for one- to four-family homes effective for a calendar year are adjusted late in the prior year. The year-over-year percentage change in the limits is based on the October-to-October change in the average home price (for both new and existing homes) published by the Federal Housing Finance Board. Since their inception, Freddie Mac and Fannie Mae pools have had identical loan limits, because the limits are dictated by the same statute. For 2006, the single-family limit is $417,000; the loan limits are 50% higher for loans made in Alaska, Hawaii, Guam, and the U.S. Virgin Islands.

Loans larger than the conforming limit (and thus ineligible for inclusion in agency pools) are classified as "jumbo" loans and can only be securitized in private label transactions (along with loans that do not meet the GSEs' required credit or documentation standards, irrespective of balance). While the size of the private-label sector is significant (as of the second quarter of 2006, approximately \$1.7 trillion in balance was outstanding), it is much smaller than the market for agency pools. Moreover, as the conforming balance limits have risen due to robust real estate appreciation, the market share of agency pools relative to private label deals has grown.

Prepayments and Prepayment Penalties

Mortgage loans can prepay for a variety of reasons. All mortgage loans have a "due on sale" clause, which means that the remaining balance of the loan must be paid when the house is sold. Existing mortgages can also be refinanced by the obligor if the prevailing level of mortgage rates declines, or if a more attractive financing vehicle is proposed to them. In addition, the homeowner can make partial prepayments on their loan, which serve to reduce the remaining balance and shorten the loan's remaining term. As we will discuss later in this chapter, prepayments strongly impact the returns and performance of MBS, and investors devote significant resources to studying and modeling them.

To mitigate the effects of prepayments, some loan programs are structured with prepayment penalties. The penalties are designed to discourage refinancing activity, and require a fee to be paid to the servicer if the loan is prepaid within a certain amount of time after funding. Penalties are typically structured to allow borrowers to partially prepay up to 20% of their loan each year the penalty is in effect, and charge the borrower six months of interest for prepayments on the remaining 80% of their balance. Some penalties are waived if the home is sold, and are described as "soft" penalties; hard penalties require the penalty to be paid even if the prepayment occurs as a result of the sale of the underlying property.

MORTGAGE LOAN MECHANICS

As described above, mortgage loans traditionally are structured as fully amortizing debt instruments, with the principal balance being paid off over the term of the loan. For a fixed rate product, the loan's payment is constant over the term of the loan, although the payment's breakdown into principal and interest changes each month. An amortizing fixed rate loan's monthly payment can be calculated by first computing the mortgage payment factor using the following formula:

$$\text{Mortgage payment factor} = \frac{\text{Interest rate}(1+\text{Interest rate})^{\text{Loan term}}}{(1+\text{Interest rate})^{\text{Loan term}}-1}$$

Note that the interest rate in question is the monthly rate, that is, the annual percentage rate divided by 12. The monthly payment is then computed by multiplying the mortgage payment factor by the loan's balance (either original or, if the loan is being recast, the current balance).

As an example, consider the following loan:

Loan balance:	\$100,000
Annual rate:	6.0%
Monthly rate:	0.50% = 0.005
Loan term:	30 Years (360 Months)

The monthly payment factor is calculated as

$$\frac{0.05(1.005)^{360}}{(1.005)^{360}-1} = 0.0059955$$

Therefore, the monthly payment on the subject loan is \$100,000 × 0.0059955, or \$599.55.

An examination of the allocation of principal and interest over time provides insights with respect to the buildup of owner equity. As an example, Figure 18.1 shows the total payment and the amount of principal and interest for

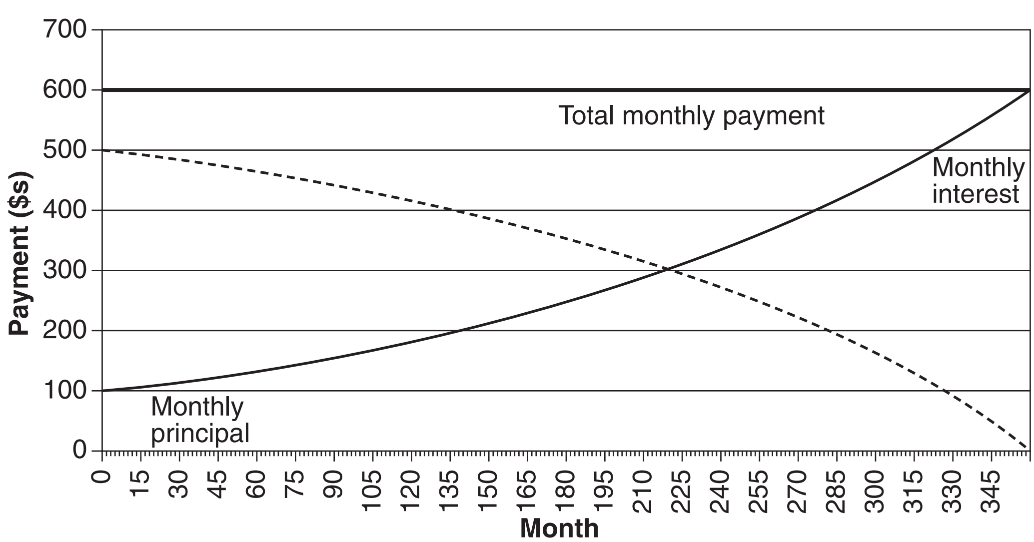


Figure 18.1 Monthly Payment Breakdown for a \$100,000 Fixed-Rate Loan at 6.0% Rate with a 30-Year Term (fixed payment of \$599.55 per month)

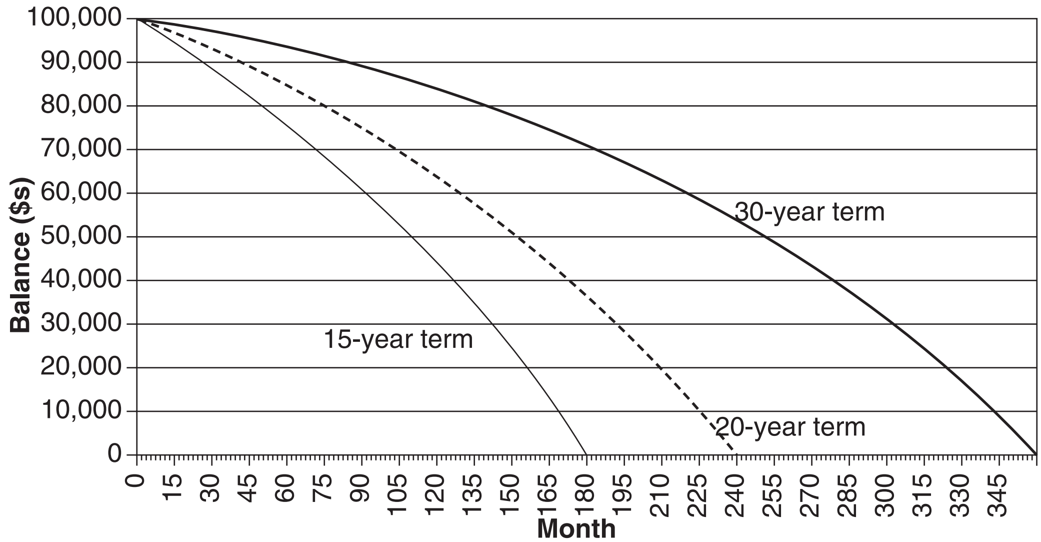


Figure 18.2 Balances for $100,000 6.0% Fixed-Rate Loan over Different Original Terms

the $100,000 loan with a 6.0% interest rate (or note rate, as it is often called) for the life of the loan.

The exhibit shows that the payment is comprised mostly of interest in the early period of the loan. Since interest is calculated from a progressively declining balance, the amount of interest paid declines over time. In this calculation, since the aggregate payment is fixed, the principal component consequently increases over time. In fact, the exhibit shows that the unpaid principal balance in month 60 is $93,054, which means that only $6,946 of the $35,973 in payments made by the borrower up to that point in time consisted of principal. However, as the loan seasons, the payment is increasingly allocated to principal. The crossover point in the example (that is, where the principal and interest components of the payment are equal) for this loan occurs in month 222.

Loans with shorter amortization schedules (e.g., 15-year loans) allow for buildup of equity at a much faster rate. Figure 18.2 shows the outstanding balance of a $100,000 loan with a 6.0% note rate using 30-, 20-, and 15-year amortization terms. In contrast to the $93,054 remaining balance on the 30-year loan, the remaining balances on 20- and 15-year loan in month 60 are $84,899 and $76,008, respectively. In LTV terms, if the purchase price of the home is $125,000 (creating an initial LTV of 80%), the LTV in month 60 on the 15-year loan is 61% (versus 74% for the 30-year loan). Finally, while 50% of the 30-year loan balance is paid off in month 252, the halfway mark is reached in month 154 with a 20-year term, and month 110 for a 15-year loan.

Patterns of borrower equity accumulation due to amortization are important in understanding the attributes of interest-only loans. Figure 18.3 compares the remaining balances over time for the previously described fully amortizing $100,000 loan with a 6% rate, versus an interest-only loan with the same rate and term. A fully amortizing loan would have a monthly payment of $599.95, and would have reduced its principal balance by $6,946 at the end of five years. The interest-only loan, by definition, would amortize none of the principal over the same period. It would have an initial monthly payment at the 6% rate of $500, which would increase to $644 when the loan recasts in month 60. The 29% increase in the payment results from the loan's balance being amortized over the remaining term of 300 months. As Figure 18.3 indicates, the remaining balance of the interest-only loan amortizes faster than the fully amortizing loan because of the higher payment, although the interest-only loan's remaining balance is greater than that of the amortizing loan. The LTV of the amortizing loan (assuming a purchase price of $125,000 and an original LTV of 80%) declines to roughly 74% by month 60 and 72% in month 80. The interest-only loan has an 80% LTV through the first 60 months after issuance, but by month 80 the LTV declines to 77.5%.

For amortizing ARM loans, the initial payment is calculated at the initial note rate for the full 360-month term. At the first reset, and at every subsequent adjustment, the loan is recast, and the monthly payment schedule is recalculated using the new note rate and the remaining term of the loan. For example, payments on a five-year hybrid ARM with a 5.5% note rate would initially be calculated as a 5.5% loan with a 360-month term. If the loan resets to a 6.5% rate after five years (based on both the underlying index and the loan's margin), the payment is calculated using a 6.5% note rate, the remaining balance in month 60, and a 300-month term. In the following year, the payment would be recalculated again using the remaining balance and prevailing rate (depending on the performance of the index referenced by the loan) and a 288-month term. In this case, the loan's initial monthly payment would be $568; in month 60, the loan's payment would change to $624, or the payment at a 6.5% rate for 300 months on a $92,460 remaining balance. (Note that all rate changes are subject to caps that limit the amount that the rate can change over a designated period of time.)

The payments on an interest-only hybrid ARM are similar to those of a fixed rate, interest-only loan. Using the rate structure described above, an interest-only 5/1

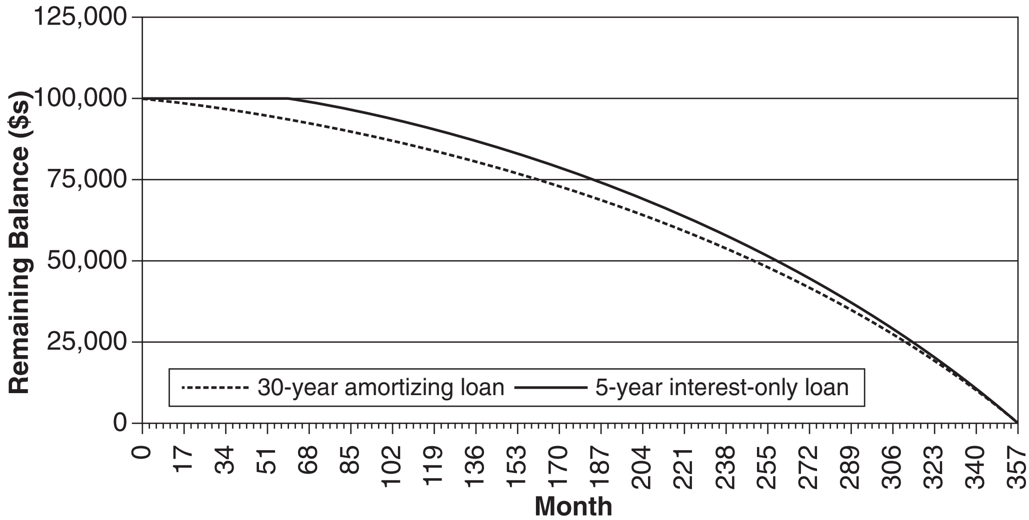


Figure 18.3 Remaining Principal Balance Outstanding for $100,000 6% Loan, Fully Amortizing versus Five-Year Interest-Only Loans

hybrid ARM would have an initial payment of $458. After the 60-month fixed rate, interest-only period, the monthly payments would reset at $675, an increase of roughly 47%. This increase represents the payment shock discussed previously. Depending on the loan's margin and the level of the reference index, borrowers seeking to avoid a sharp increase in monthly payments often refinance their loans into cheaper available products. The desire to mitigate payment shock is also largely responsible for the growth in hybrid ARMs with noncontiguous resets. Since these loans essentially separate the rate reset and payment recast, the payment increases are spread over two periods, reducing the impact of a large one-time increase in payment.

The payment structure for negative-amortization ARM loans is different and complex. The most commonly issued form of products that allow negative amortization are so-called payment-option loans, which incorporate the features of a number of different ARM products. The loans have an introductory rate that is effective for a short period of time (either one or three months). After the initial period, the loan's rate changes monthly, based on changes in the reference index. The borrower's minimum or "required" payment, however, does not change until month 13. The initial or teaser payment is initially calculated to fully amortize the loan over 30 years at the introductory rate. After a year, and in one-year intervals thereafter, the loan is recast. The minimum payment is recalculated based on the loan's margin, the index level effective at that time, and the remaining balance and term on the loan. However, the increase in the loan's minimum monthly payment is subject to a 7.5% cap. (Note that this cap functions differently than those in the hybrid market, which are based on changes in the loan's rate rather than payment.)

The minimum payment may not be sufficient to fully pay the loan's interest, based on its effective rate. This may occur if the loan's index and margin are such that the minimum payment is lower than the interest payment, or if the minimum payment is constrained by the 7.5% payment cap. In that event, the loan undergoes negative amortization, where the unpaid or "deferred" amount of interest is added to the principal balance. Negative amortization is typically limited to 115% of the original loan balance (or 110% in a few states). If this threshold is reached, the loan is immediately recast to amortize the current principal amount over the remaining term of the loan. Under all circumstances, the loan is automatically recast periodically, with payments calculated based on the current loan balance and the remaining term of the loan. At this point, the payment change is not subject to the 7.5% payment cap—a condition that also holds true if the loan recasts because the negative amortization cap is reached. (The first mandatory recast is generally at the beginning of either year 5 or 10; in either case, the loan will subsequently recast every five years thereafter.)

RISKS ASSOCIATED WITH MORTGAGES AND MORTGAGE PRODUCTS

Holders of fixed income investments ordinarily deal with interest rate risk, or the risk that changes in the level of market interest rates will cause fluctuations in the market value of such investments. However, mortgages and associated mortgage products have additional risks associated with them that are unique to the products and require additional analysis. We conclude this chapter with a discussion of these risks.

Prepayment Risk

In a previous section, we noted that obligors have the ability to prepay their loans before they mature. For the holder of the mortgage asset, the borrower's prepayment option creates a unique form of risk. In cases where the obligor refinances the loan in order to capitalize on a drop in market rates, the investor has a high-yielding asset pay off, and it can be replaced only with an asset carrying a lower yield. Prepayment risk is analogous to "call risk" for corporate and municipal bonds in terms of its impact on returns, and also creates uncertainty with respect to the timing of investors' cash flows. In addition, changing prepayment "speeds" due to interest rate moves cause variations in the cash flows of mortgages and securities collateralized by mortgage products, strongly influencing their relative performance and making them difficult and expensive to hedge.

Prepayments are phenomena resulting from decisions made by the borrower and/or the lender, and occur for the following reasons:

- The sale of the property (due to normal mobility, as well as death and divorce).
- The destruction of the property by fire or other disaster.
- A default on the part of the borrower (net of losses).
- Curtailments (that is, partial prepayments).
- Refinancing.

Prepayments attributable to reasons other than refinancings are referred to under the broad rubric of "turnover." Turnover rates tend to be fairly stable over time, but are strongly influenced by the health of the housing market, specifically the levels of real estate appreciation and the volume of existing home sales. Refinancing activity, however, generally depends on being able to obtain a new loan with either a lower rate or a smaller payment, making this activity highly dependent on the level of interest rates, the shape of the yield curve (since short rates strongly influence ARM pricing), and the availability of alternative loan products. In addition, the amount of refinancing activity can change greatly as the result of seemingly small changes in rates.

The paradigm in mortgages is thus fairly straightforward. Mortgages with low note rates (that are "out-of-the-money," to borrow a term from the option market) normally prepay fairly slowly and steadily, while loans carrying higher rates (and are "in-the-money") are prone to experience spikes in prepayments when rates decline. Clearly, this paradigm is dependent on the level of mortgage rates.

It is important to understand how changes in prepayment rates impact the performance of mortgages and MBS. Since prepayments increase as bond prices rise and market yields are declining, mortgages shorten in average life and duration when the bond markets rally, constraining their price appreciation. Conversely, rising yields cause prepayments to slow and bond durations to extend, resulting in a greater drop in price than experienced by more traditional (that is, option-free) fixed income products. As a result, the price performance of mortgages and MBS tends to lag that of comparable fixed maturity instruments (such as Treasury notes) when the prevailing level of yields increases.

This phenomenon is generically described as *negative convexity*. The effect of changing prepayment speeds on mortgage durations, based on movements in interest rates, is precisely the opposite of what a bondholder would desire. (Fixed income portfolio managers, for example, extend durations as rates decline, and shorten them when rates rise.) The price performance of mortgages and MBS is, therefore, decidedly nonlinear in nature, and the product will underperform assets that do not exhibit negatively convex behavior as rates decline.

Figure 18.4 shows a graphic representation of this behavior. Investors are generally compensated for the lagging price performance of MBS through higher base-case yields. However, the necessity of managing negative convexity and prepayment risk on the part of investors involves fairly active management of MBS portfolios, and creates both higher hedging costs and the possibility of losses due to estimation and modeling error. In turn, this creates the desire on the part of some investors to limit their exposure to prepayments by investing in bonds where prepayment risk is transferred within the structure. This type of risk mitigation is central to the structured MBS market.

Credit and Default Risk

Analysis of the credit exposure in the mortgage sector is different from the assessment of credit risk in most other fixed income instruments because it requires:

- Quantifying and stratifying the characteristics of the thousands of loans that underlie the mortgage investment.
- Estimating how these attributes will translate into performance based on standard metrics, and the evaluation of reasonable best-, worst-, and likely-case performance.
- Calculating returns based on these scenarios.

In a prior section, some of the factors (credit scores, LTVs, etc.) that are used to gauge the creditworthiness of borrowers and the likelihood of a loan to result in a loss of principal were discussed. Many of the same measures are also used in evaluating the creditworthiness of a mortgage pool. For example, weighted average credit scores and LTVs are routinely calculated, and stratifications of these characteristics (along with documentation styles and other attributes) are used in the credit evaluation of the pool. In addition to these characteristics of the loans, the following metrics are also utilized in the a posteriori evaluation of a mortgage pool or security.

Delinquencies

These measures are designed to gauge whether borrowers are current on their loan payments or, if they are late, stratifying them according to the seriousness of the delinquency. The most common convention for classifying delinquencies is one promulgated by the Office of Thrift Supervision; this "OTS" method classifies loans as follows:

- Payment due date to 30 days late: *Current*
- 30–60 days late: *30 days delinquent*
- 60–90 days late: *60 days delinquent*
- More than 90 days late: *90+ days delinquent*

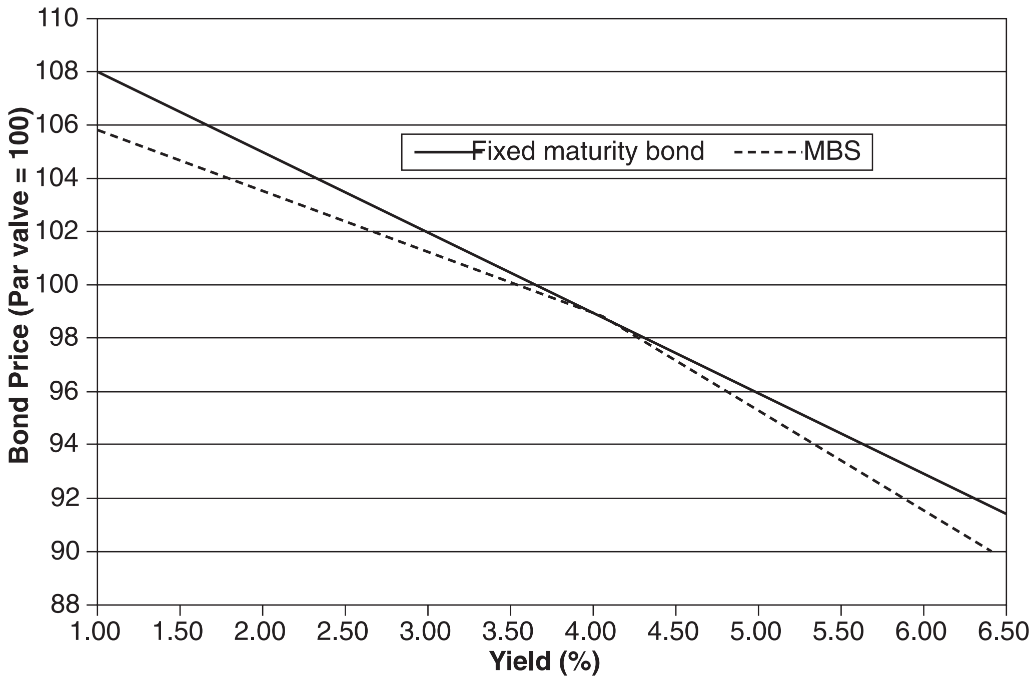


Figure 18.4 Performance Profile of Hypothetical Fixed Maturity Bond versus MBS

Defaults

At some point in their existence, many loans that are associated with delinquencies become current or "cure," as the condition leading to the delinquency (e.g., job loss, illness, etc.) resolves itself. However, some portion of the delinquent loan universe ends up in default. By definition, default is the point where the borrower loses title to the property in question. Default generally occurs for loans that are 90+ days delinquent, although loans where the borrower goes into bankruptcy may be classified as defaulted at an earlier point in time.

Loss Severity

Since the lender has a lien on the borrower's property, much of the value of the loan can be recovered through the foreclosure process. Loss severity measures the face value of the loss on a loan after foreclosure is completed. Depending on the type of loan, loss severities can average in the area of 20% to 40%, and can be heavily influenced by the loan's LTV (since a high LTV loan leaves less room for a decline in the value of the property in the event of a loss). However, in the event of a default, loans with relatively low LTVs can also result in losses, generally for two reasons:

- The appraised value of the property may be high relative to the property's actual market value.
- There are costs and foregone income associated with the foreclosure process.

In light of these factors, the process of evaluating the credit-adjusted performance of either a group of loans involves first gauging the expected delinquencies, defaults, and loss severities of the pool or security based on its credit characteristics. Subsequently, loss-adjusted yields and returns can be generated. It should be noted that investors in some segments of the MBS market do not engage in detailed credit analysis; buyers of agency pools, for example, generally rely on the guaranty of the agency in question. In addition to buyers of mortgages in whole-loan form, credit analysis is primarily undertaken by investors in the subordinate tranches of private label deals. As one might expect, the performance of subordinates is highly sensitive to the credit performance of the collateral pool. This is both because of their role in protecting the senior classes from losses, as well as the sequential nature of loss allocations within the subordinate classes.

SUMMARY

Mortgage products can be defined by a number of critical attributes. These include lien status, loan term, credit classification, interest rate type, and amortization scheme. Many loan products are based on a mix of attributes; an example might be an adjustable-rate loan with an interest-only feature. Loans can be securitized either by using the guarantees of a government agency or quasi-governmental entity (that is, the GSEs), or by utilizing a so-called private-label structure that incorporates credit enhancement through mechanisms such as subordination. An important characteristic defining loans refers to their interest rate classification. Fixed-rate loans have an interest rate fixed for the life of the loan, while adjustable-rate mortgages (or ARMs) reset periodically to a rate based on a reference index. The borrower's payment on an ARM will typically change at the reset date. Depending on the shape of the yield curve and the level of the index, the rate may increase or decrease. An increased payment may create a "payment shock" in which the borrower's payment increases significantly over its initial level.

Mortgage loans have traditionally been issued as fully-amortizing obligations with (for fixed-rate products) a constant payment over their term. However, loans with alternative payment schemes (such as interest-only loans) have become increasingly popular. When the interest-only lockout expires, loans structured as interest-only products will recast in order to amortize the loan over the remaining term. This event creates another form of payment shock, and for interest-only ARMs will exacerbate the payment shock experienced at the reset date. Mortgage loans and mortgage-backed securities have fairly unique risk profiles. Their performance can suffer from changes in prepayment speeds (creating "negative convexity") as well as, for loans in raw form and subordinated MBS, exposure to credit risk.

REFERENCES

Bhattacharya, A. K., Berliner, W. S., and Lieber, J. (2006). Alt-A mortgages and MBS. In F. J. Fabozzi (ed.), *Handbook of Mortgage-Backed Securities*, 6th edition. New York: McGraw-Hill, pp. 187–206.

Bhattacharya, A. K., Banerjee, S., Horowicz, R., and Wang, W. (2006). Hybrid adjustable-rate mortgages (ARMs). In F. J. Fabozzi (ed.), *Handbook of Mortgage-Backed Securities*, 6th edition (pp. 259–286). New York: McGraw-Hill.

Fabozzi, F. J. (ed.) (2006). *Handbook of Mortgage-Backed Securities*, 6th edition. New York: McGraw-Hill.

Fabozzi, F. J. (2006). *Fixed Income Mathematics: Analytical and Statistical Techniques*, 4th edition. New York: McGraw-Hill Publishing.

Fabozzi, F. J., Bhattacharya, A. K., and Berliner, W. S. (2007). *Mortgage-Backed Securities: Products, Structuring, and Analytical Techniques*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., and Modigliani, F. (1992). *Mortgage and Mortgage-Backed Securities Markets*. Boston: Harvard Business School Press.

Liu, D. (2006). Interest-only ARMs. In F. J. Fabozzi. (ed.), *Handbook of Mortgage-Backed Securities*, 6th edition (pp. 333–362). New York: McGraw-Hill.

Mansukhani, S. (2006). Exploring the MBS/ABS continuum: The growth and tiering of the Alt-A hybrid sector. In F. J. Fabozzi (ed.), *Handbook of Mortgage-Backed Securities*, 6th edition (pp. 171–186). New York: McGraw-Hill.

Mansukhani, S., Budhram, A., and Qubbaj, M. (2006). Fixed-rate Alt-A MBS. In F. J. Fabozzi. (ed.), *Handbook of Mortgage-Backed Securities*, 6th edition (pp. 207–258). New York: McGraw-Hill.

CHAPTER 19

Reverse Mortgages

LAURIE S. GOODMAN, PhD
Co-head of Global Fixed Income Research and Manager of U.S. Securitized Products Research, UBS

Abstract: Reverse mortgages are instruments designed for older homeowners; they allow the equity in a home to be monetized. At the initiation of the reverse mortgage, a homeowner elects to take out either a lump-sum payment, a fixed monthly payment for life, or can have access to a line of credit. The reverse mortgage loan must be paid in full when the last surviving borrower dies, moves, or sells the home.

Keywords: reverse mortgage, home equity conversion mortgage (HECM)

Reverse mortgages are a type of residential mortgage instrument which allows homeowners to monetize the equity in their home. This chapter describes the characteristics of this type of mortgage. Reverse mortgages, while still a very small part of the U.S. mortgage market, has shown robust growth. Moreover, with changing U.S. demographics, their popularity is expected to increase.

HOW A REVERSE MORTGAGE WORKS

In a *reverse mortgage,* the homeowner receives cash, either as an up-front payment, as a monthly payment, or as a line of credit. That money is not taxable (technically, it is considered a loan advance, not income), and can be used to live on. It does not impact Medicare or Social Security benefits. It could potentially impact Medicaid benefits. The loan amount will depend on the age of the borrower (younger borrowers receive less money), the appraised home value, current interest rates, and the lending limit in a particular area, if applicable. The mortgages can be prepaid at any time.

Reverse mortgage loans are generally payable in full when the last surviving borrower dies or sells the home. The mortgage may also come due if:

- The borrower permanently moves to a new principal residence.
- The last surviving borrower fails to live in the home for 12 months in a row due to physical or mental illness.
- The property deteriorates, except for reasonable wear and tear, and the borrower fails to correct the problem.
- The borrower fails to pay property taxes or hazard insurance or violates any other borrower obligation.

A simplified example, shown in Table 19.1 will make this more clear. A 75-year-old borrower has taken out a reverse mortgage. Assume the house is now worth $250,000; the borrower has $200,000 of equity and a mortgage for $50,000. This simplified example neglects up-front fees, mortgage insurance (when applicable), assumes a fixed interest rate (when interest rates on this product are generally variable), and gives the borrower a choice of only two payment options: taking the cash all up front or taking the cash in the form of a monthly payment up front. If the borrower chooses to take out a payment up front in lump-sum form (option 1), then based on age, value of home, and interest rate, the maximum amount that can be received up front is $144,590. The borrower must then apply $50,000 to pay down the first mortgage. That frees the borrower from making further interest payments on the mortgage, and leaves the borrower with $94,590 of cash. By contrast, if the borrower chooses the monthly payment option (option 2), the mortgage will be paid off, and the borrower will receive a monthly payment of $668. (Note that if there was no first lien, the borrower would have received $1,021 per month.)

Table 19.1 Reverse Mortgages: A Simple Example

Borrower Age: 75 years old
Recent Appraised Home Value: $250,000
Mortgage: $50,000
Equity: $200,000
Estimated Annual Home Price Appreciation (HPA): 5%
Estimated Annual Interest Rate on Loan: 7% (.5654% per month)

Option 1: Lump Sum Payment
Loan Amount: $144,590
Cash Available: $94,590

Option 2: Monthly Payments
Monthly Loan Payment (without lien): $1021
Monthly Payment (after lien paydown): $668

	Borrower Chooses Option 1: Dies at 80	**Borrower Chooses Option 1: Dies at 90**
Proceeds from Sale of House: 5% HPA	319,070 = 250,000 × [(1.05)^{5}]	519,732 = 250,000 × [(1.05)^{15}]
Less: Accreted Value of Loan at time of death	202,793 = 144,590 × [(1.005654)^{60}]	398,917 = 144,590 × [(1.005654)^{180}]
Money Back to Estate	116,277	120,815
	Borrower Chooses Option 2: Dies at 80	**Borrower Chooses Option 2: Dies at 90**
Proceeds from Sale of House: 5% HPA	319,070 = 250,000 × [(1.05)^{5}]	519,732 = 250,000 × [(1.05)^{15}]
Less: Accreted Value of $50,000 loan (to pay off 1st lien)	70,126 = 50,000 × [(1.005654)^{60}]	137,947 = 50,000 × [(1.005654)^{180}]
Less: Future Value of Monthly Payments @ $668/month	47,828	208,990
Money Back to Estate	201,116	172,795

Table 19.1 first looks at the case in which the borrower elects to take option 1, the up-front payment, and dies at 80, exactly 5 years (60 months) after taking out the loan. In this illustration it is assumed the house has appreciated by 5% per annum, and is now worth $319,070. The loan ($144,590) must be repaid with interest. Assuming an annual interest rate of 7% (equivalent to a monthly rate of 0.5654%), the estate must repay $202,793 (144,590 × ((1.005654)^{60})). Thus, the $116,277 difference ($310,070 house value – $202,793 loan repayment) flows back to the estate. The far right section of Table 19.1 shows the scenario in which the borrower elects to take the up-front payment, and dies at age 90, exactly 15 years (180 months) after taking out the reverse mortgage. We assume the value of the house is $519,732 (reflecting 15 years of 5% home price appreciation) and the accreted value of the loan is $398.917. Thus, $120,815 reverts to the estate.

Now let us consider the case in which the borrower elects the monthly payment option (option 2). The borrower must take $50,000 up front in order to pay off the first lien. Given the borrower's age, appraised home value, interest rate, and so on, our borrower can receive a payment of $668 per month. Thus, when the borrower dies, both the accreted value of the loans plus the future value of the monthly payments must be subtracted from the terminal value of the property. If the borrower dies at 80, exactly 60 months after the mortgage was taken out, the borrower would owe $47,828, the future value of 5 years (60 months) of monthly payments, plus an accreted loan amount of $70,126 (on the original $50,000 loan). Thus, $201,116 reverts back to the estate ($319,070 from the sale of the house—$70,126 to pay back the cash lien—$47,828, the future value of the monthly payment stream). If the borrower dies at 90, the value of the property would be higher, but the future value of the original $50,000 loan would be higher, and a lot more monthly payments have been made to the borrower. Under these circumstances, Table 19.1 shows the estate is then left with $172,795.

PROGRAMS

There are three basic types of reverse mortgage programs:

1. *Home equity conversion mortgage (HECM)*
2. Fannie Mae Home Keeper (FMHK) mortgages
3. Proprietary reverse mortgage products

HECM Program

The HECM program is offered by the Federal Housing Administration (FHA), a division of the Department of Housing and Urban Development, and all HECM mortgages are FHA insured. This is by far the largest of the reverse mortgage programs, and has experienced very robust growth. In fiscal year 2003, there were only 18,097 reverse mortgage loans endorsed. This number doubled to 37,829 loans in fiscal year 2004, then increased rapidly to 43,131 in fiscal year 2005, and 76,351 in fiscal year 2006. In the first nine months of fiscal year 2007, 80,425 reverse mortgage loans were endorsed—more than for the entire year 2006.

To be eligible for an HECM loan, a borrower must be aged 62 or over and live in the home as a principal residence. Empirically, we find that most of the borrowers that take out HECM loans are considerably older than the minimum age. In fact, HUD reports that the median age of an HECM borrower was 75; the median age of all elderly homeowners is 72. The home must be a single-family residence in a one- to four-unit dwelling, a condominium, or part of a planned unit development (PUD). Some manufactured housing is eligible. The overwhelming majority (approximately 85%) of these are single-family properties.

Payment Options

An HECM mortgage can be taken out in any of the following forms:

- **Tenure:** Equal monthly payments as long as at least one borrower lives and continues to occupy the property as a principal residence.
- **Term:** Equal monthly payments for a fixed number of months selected.
- **Line of credit:** Unscheduled payments or installments, drawn down at times and in amounts of the borrower's choosing until the line is exhausted. This line will grow over time.
- **Modified tenure:** Combination of line of credit and monthly payments for as long as the borrower remains in the home.
- **Modified term:** Combination of a line of credit and monthly payments for a fixed period.

In practice, the line of credit is the most popular option. If a borrower wants to take out money immediately (to pay off a first lien, or take a vacation) it is considered to be a drawdown on the line of credit. In our simplified example in Table 19.1, we wanted to show deterministically how the reverse mortgage would impact the estate. Thus, we included only two options: a line of credit in which the amount was drawn down immediately, and a modified tenure, in which the up-front payment was used to pay off the first lien. In reality, borrowers have far more flexibility than we indicated in our example.

Under the conditions of a reverse mortgage, when the house is sold or no longer used a primary residence, the borrower or their heirs will repay the drawn portion of the credit line (or monthly payments) plus interest to the lender. The remaining value of the house belongs to the estate. It is possible that home price appreciation might be low enough, and the borrower might live long enough, that the price of the house is less than the accreted value of the outstanding loans. For example, assume in Table 19.1 that our 75-year-old borrower selected option 2, receiving a $50,000 up-front payment to cover the first lien, and additional payments of $668 per month. If that borrower died at 90, $346,937 ($137,947 from the up-front loan and $208,990 from the monthly payments) would be owed. If the house value had appreciated 2% per annum (rather than the 5% we assumed in Table 19.1), the appreciated house value would be $336,467, or approximately $10,500 less than the amount due on the reverse mortgage. From on investor's point of view, this is not an issue: Government insurance would cover this, as the loans are FHA insured. From a borrower's point of view, it is also not a concern, as the loans are nonrecourse.

The interest rate on the HECM loan is generally reset either monthly or annually, based on the following reset formula: one-year CMT + a margin, where CMT is the constant maturity Treasury.

In addition to the interest expense, the borrower must pay a mortgage insurance premium (MIP) for the FHA insurance. This premium is equal to 2% of the up-front amount plus an annual premium equal to 0.5% of the loan amount. The MIP is meant to guarantee that if the loan servicer goes bankrupt, the government will step in and make future payments. The MIP also guarantees that if there is any shortfall between sales price and repayment amount, the government will make up the difference. In addition to the MIP, reverse mortgages also carry application fees, origination fees, and often a monthly servicing fee. These charges are generally paid by the reverse mortgage, and the costs are added to the principal and paid at the end, when the loan is due.

Amount that Can Be Borrowed

The amount that can be borrowed depends on a borrower's age, the current interest rate, and the appraised value of the home. Moreover, the maximum size of an HECM mortgage will depend on the maximum HUD loan limit. This varies by county and is adjusted annually. Currently, the maximum is $362,790 for single-family homes in high-cost areas and $200,160 for rural areas. That is, the limit in high-cost areas is 87% of the conventional limit of $417,000. It is 48% of the conventional limit in low-cost areas. The implications of these limits are clear—if two borrowers of the same age applied for a loan at the same time, one with a home value of $362,790 and another with a home value of $1 million, they would both receive exactly the same HECM loan. When there is more than one borrower, the loan amount in an HECM mortgage is determined solely by the age of the younger borrower.

Table 19.2 illustrates the amount that can be drawn out under the HECM program. These calculations assume the home is located in a high-cost area. Thus, if a home were appraised for $250,000, the top section of Table 19.2 indicates a 65-year-old borrower would have a credit line available for $116,568; the credit line would be $144,590 if the borrower were 75, and $174,894 if the borrower were 85. Similarly, if the tenure option were selected, a 65-year-old borrower would receive $744/month, a 75-year-old borrower would receive $1,021/month, and an 85-year-old borrower would receive $1,568/month. If the home appraised for $1 million, the FHA loan limits would be binding, limiting the amount the borrower could receive. Thus, as indicated in the bottom section of Table 19.2 a 65-year-old borrower would have a credit line of $172,286, which is only 46% higher than that available on a $250,000 home.

Table 19.2 HECM Reverse Mortgage Options

	City/State:	**Montclair, NJ**
	County:	**Essex**
	Home Value:	**$250,000**
	Liens:	**0**

	Birth Year **Age**	**1944** **62**	**1941** **65**	**1936** **70**	**1931** **75**	**1926** **80**	**1921** **85**	**1916** **90**
1	Cash Available	108,971	116,568	130,034	144,590	159,775	174,894	189,273
	Loan-to-Value	43.6%	46.6%	52.0%	57.8%	63.9%	70.0%	75.7%
2	Monthly Income Available	684	744	864	1,021	1,236	1,568	2,190
3	Line of Credit:							
	Creditline Available	108,971	116,568	130,034	144,590	159,775	174,894	189,273
	Annualized Growth Rate	7.22%	7.22%	7.22%	7.22%	7.22%	7.22%	7.22%
	Creditline Value in 5 Years	154,403	165,168	184,247	204,873	226,389	247,811	268,185
	Creditline Value in 10 Years	218,777	234,030	261,064	290,288	320,775	351,129	379,997

	City/State:	**Montclair, NJ**
	County:	**Essex**
	Home value:	**$1,000,000**
	Liens:	**0**

	Birth Year **Age**	**1944** **62**	**1941** **65**	**1936** **70**	**1931** **75**	**1926** **80**	**1921** **85**	**1916** **90**
1	Cash Available	161,305	172,286	191,730	212,715	234,555	256,216	276,686
	Loan-to-Value	16.1%	17.2%	19.2%	21.3%	23.5%	25.6%	27.7%
2	Monthly Income Available	1,012	1,100	1,275	1,501	1,815	2,297	3,201
3	Line of Credit:							
	Creditline Available	161,305	172,286	191,730	212,715	234,555	256,216	276,686
	Annualized Growth Rate	7.22%	7.22%	7.22%	7.22%	7.22%	7.22%	7.22%
	Creditline Value in 5 Years	228,557	244,116	271,666	301,401	332,346	363,037	392,042
	Creditline Value in 10 Years	323,847	345,893	384,929	427,061	470,908	514,395	555,492

Fannie Mae Home Keeper

The Fannie Mae Home Keeper mortgage program is Fannie Mae's conventional market alternative to the HECM product. It works much like an HECM; the borrower can receive fixed monthly payment for life (that is, for as long as the borrower occupies the home as his/her principal residence), a line of credit, or any combination of monthly payments or a line of credit. However, the Fannie Mae Home Keeper can be used for a broader array of alternatives, including condominiums that are not FHA-approved and new home purchases. The latter is particularly important, as HECM Mortgages require that borrowers have been in their home for at least a year. Let us assume a 75-year-old man wants to sell his home in Philadelphia, with a value of $150,000, and buy a $200,000 home in Florida. To avoid a mortgage payment on the new home (as the borrower's income is very limited), the borrower would have to use the entire $150,000 proceeds from the sale of the Philadelphia home, plus another $50,000 in savings. If the borrower does not have the $50,000, he could not buy the new home (unless he qualifies for and is able to obtain a regular mortgage). But the borrower could seek an FMHK reverse mortgage, which can be used to bridge the $50,000 difference.

Note that even though the loan limits are higher for Fannie Mae programs than the FHA programs, the amount that can be drawn out under the Fannie Mae Home Keeper program is usually less than what can be drawn out under the HECM program. This is illustrated in Table 19.3. A 65-year-old borrower with a home worth $250,000 could draw out $116,568 under the HECM program, while the Fannie Mae Home Keeper would allow only $42,817.

The interest rate on the Home Keeper mortgage is determined as a spread above an index rate—the current weekly average of the one-month secondary market CD rate, which is published by the Federal Reserve. The rate on the Fannie Mae Home Keeper mortgages adjusts monthly.

Proprietary Products

There are a number of lenders that offer proprietary mortgage products. As on the HECM and FMHK products, the interest rates are variable. These proprietary products generally build in additional protections to make sure the accreted value of the loans will not be higher than the home value. First, these proprietary products do not have a tenure option, as the lenders are unwilling to absorb

Table 19.3 Program Comparison

	Home Value:	**$ 250,000**			**$ 250,000**		
	Birth Year:	**1941**			**1921**		
	Age	**65**			**85**		
		FHA/HUD Monthly	**Fannie Mae HomeKeeper**	**Proprietary Reverse Mortgage Program**	**FHA/HUD Monthly**	**Fannie Mae HomeKeeper**	**Proprietary Reverse Mortgage Program**
1	Cash Available	116,568	42,817	42,709	174,894	136,990	115,734
	Loan-to-Value	46.6%	17.1%	17.1%	70.0%	54.8%	46.3%
2	Monthly Income Available	744	336	N/A	1,568	1,334	N/A
3	Line of Credit:						
	Creditline Available	116,568	42,817	42,709	174,894	136,990	115,734
	Annualized Growth Rate	7.22%	N/A	5.00%	7.22%	N/A	5.00%
	Creditline Value in 5 Years	165,168	42,817	54,508	247,811	136,990	147,709
	Creditline Value in 10 Years	234,030	42,817	69,568	351,129	136,990	188,518
	Home Value:	**$ 1,000,000**			**$ 1,000,000**		
	Birth Year:	**1941**			**1921**		
	Age	**65**			**85**		
		FHA/HUD Monthly	**Fannie Mae HomeKeeper**	**Proprietary Reverse Mortgage Program**	**FHA/HUD Monthly**	**Fannie Mae HomeKeeper**	**Proprietary Reverse Mortgage Program**
1	Cash Available	172,286	74,901	178,134	256,216	231,524	470,234
	Loan-to-Value	17.2%	7.5%	17.8%	25.6%	23.2%	47.0%
2	Monthly Income Available	1,100	587	N/A	2,297	2,255	N/A
3	Line of Credit:						
	Creditline Available	172,286	74,901	178,134	256,216	231,524	470,234
	Annualized Growth Rate	7.22%	N/A	5.00%	7.22%	N/A	5.00%
	Creditline Value in 5 Years	244,116	74,901	227,349	363,037	231,524	600,151
	Creditline Value in 10 Years	345,893	74,901	290,161	514,395	231,524	765,961

the risk that the borrower will live long enough that total payments may be higher than the value of the house. Second, the growth rate on the line of credit may be freely altered by the lender. These protections are important to the investor, as there is no government guarantee on these loans.

From the borrower's perspective, the big advantage of the proprietary products is that they do not have a loan limit. Thus, for a home with a high appraised value, the borrower is often better off with a proprietary product. This can be seen in Table 19.7, which compares the HECM product with a proprietary reverse mortgage offering from one major originator of reverse mortgages. Note that for a home valued at $250,000, the HECM product gives the borrower (regardless of age) a much larger line of credit than the proprietary product offered. For a home valued at $1 million, a 65-year-old borrower can have a marginally larger line of credit using the proprietary product ($178,134 versus $172,286). An 85-year-old borrower would have a huge advantage using a proprietary product versus an HECM ($470,234 versus $256,216).

SUMMARY

This chapter provides a brief introduction to the reverse mortgage market. Reverse mortgage products have become much more popular and will continue to grow in importance. This growth will be further aided by changing demographics. Moreover, securitization activity is building, allowing for broader mix of investors to hold these instruments. As reverse mortgage products grow in popularity, we expect the number of originators offering them, as well as their securitization volumes, to increase.

REFERENCES

Davidoff, T. (2004). Maintenance and the home equity of the elderly. Fisher Center for Real Estate and Urban Economics, paper no. 03–288. Berkeley, CA: University of California, Berkeley—Haas School of Business.

Davidoff, T., and Welke, G. (2007). Selection and moral hazard in the reverse mortgage market. Haas School of Business working paper, 1–39.

DiVenti, T. R., and Herzog, T. N. (1992). Modeling home equity conversion mortgages. *Transactions of the Society of Actuaries* 43: 101–115.

Fitch Ratings. (2005). *Repay My Mortgage? Over My Dead Body! – Fitch's Reverse Mortgage Criteria*. New York: Fitch Ratings.

Mayer, C., and Simons, K. (1994). Reverse mortgages and the liquidity of housing wealth. *Journal of the American Real Estate and Urban Economics Association* 22, 2: 235–255.

Miceli, T. J., and Sirmans, C. F. (1994). Reverse mortgages and borrower maintenance risk. *Real Estate Economics* 22, 2: 257–299.

Rodda, D. T., Lam, K., and Youn, A. (2004). Stochastic modeling of federal housing administration home equity conversion mortgages with low-cost refinancing. *Real Estate Economics* 32, 4: 589–617.

Shiller, R., and Weiss, A. (2000). Moral hazard in home equity conversion. *Real Estate Economics* 28, 1: 1–31.

Stucki, B. R. (2006). Using reverse mortgages to manage the financial risk of long-term care. *North American Actuarial Journal* 10, 4: 13–15.

Szymanoski, E. J. (1994). Risk and the home equity conversion mortgage. *Journal of the American Real Estate and Urban Economics Association* 22, 2: 347–366.

Szymanoski, E. J., Enriquez, J. C., and DiVenti, T. R. (2007). Home equity conversion mortgage terminations: Information to enhance the developing secondary market. *Cityscape: A Journal of Policy Development and Research* 9, 1: 5–45.

Venti, S., and Wise, D. (2000). Aging and housing equity. National Bureau of Economic Research, working paper 7882.

Zhai, D. H. (2000). Reverse mortgage securitizations: Understanding and gauging the risk. Moody's Investors Service, June 23.

CHAPTER 20

U.S. Treasury Securities

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: The securities issued by the U.S. Department of the Treasury include bills, notes, and bonds. The U.S. Treasury market is a closely watched market by market participants throughout the world because it plays a prominent role in the global financial market for two reasons. First, because the securities are backed by the full faith and credit of the U.S. government, they are viewed as default-free securities and therefore the yields on these securities are viewed as benchmark risk-free interest rates. Second, because of the large size of the market and the large size of each individual issue, the Treasury market is the most active and liquid sector of the global financial market.

Keywords: Treasury securities, Treasuries, or government bonds, fixed-principal securities, Treasury bills, Treasury coupon securities, Treasury notes, Treasury bonds, cash management bills, Treasury inflation-protected securities (TIPS), real rate, inflation-adjusted principal, reopening of an issue, noncompetitive bid, competitive bid, stop-out yield, high yield, bid-to-cover ratio, single-price auction, Dutch auction, on-the-run issue, current issue, off-the-run issue, when-issued market, interdealer brokers, bank discount basis, bond-equivalent yield, CD equivalent yield, money market equivalent yield, coupon stripping, corpus, principal strips, reconstitution

The securities issued by the U.S. Department of the Treasury (U.S. Treasury hereafter) are called *Treasury securities, Treasuries,* or *U.S. government bonds.* Because they are backed by the full faith and credit of the U.S. government, market participants throughout the world view them as having no credit risk. Hence, the interest rates on Treasury securities are the benchmark default-free interest rates.

Treasury securities are classified as nonmarketable and marketable securities. The former securities include savings bonds that are sold to individuals and state and local government series (SLGS) securities that are sold to state and local government issuers of tax-exempt securities. Market securities can be bought, sold, or transferred after they are issued. The Public Debt Act of 1942 grants the U.S. Treasury considerable discretion in deciding on the terms for a marketable security. An issue may be sold on an interest-bearing or discount basis and may be sold on a competitive or other basis, at whatever prices the Secretary of the Treasury may establish.

In this chapter, the different types of marketable Treasury securities are explained as well their primary and secondary markets. The relationship between the interest rate on Treasury securities and maturity is referred to as the Treasury yield curve. This relationship and how it is used in valuing fixed income securities is described in Chapters 36, 37, and 38 of Volume III. There are derivative instruments in which Treasury securities are the underlying. These contracts are described in Chapter 39 of Volume I.

TYPES OF MARKETABLE TREASURY SECURITIES

There are two types of marketable Treasury securities issued: fixed-principal securities and inflation-indexed securities.

Fixed-Principal Treasury Securities

The U.S. Treasury issues two types of *fixed-principal securities:* discount securities and coupon securities. Discount securities are called Treasury bills; coupon securities are called Treasury notes and Treasury bonds.

Treasury bills are issued at a discount to par value, have no coupon rate, and mature at face value. Generally, Treasury bills can be issued with a maturity of up to two years. The U.S. Treasury typically issues only certain maturities. As of year-end 2007, the practice of the U.S. Treasury is to issue Treasury bills with maturities of 4 weeks, 13 weeks, 26 weeks, and 6 months. At one time, the U.S. Treasury issued a 1-year Treasury bill. In addition, the Treasury Department issues *cash management bills* that can have a maturity from 1 to 7 days, depending on its borrowing needs.

As discount securities, Treasury bills do not pay coupon interest. Instead, Treasury bills are issued at a discount from their face value; the return to the investor is the difference between the face value and the purchase price.

The U.S. Treasury issues securities with initial maturities of two years or more as coupon securities. Coupon securities are issued at approximately par and, in the case of fixed-principal securities, mature at par value. They are not callable. *Treasury notes* are coupon securities issued with original maturities of more than two years but no more than 10 years. As of year-end 2007, the U.S. Treasury issues a 2-year note, a 5-year note, and a 10-year. At one time the U.S. Treasury issued a 3-year notes and 7-year notes. Treasuries with original maturities greater than 10 years are called *Treasury bonds.* As of year end 2007, the U.S. Treasury issues a 30-year bond. The U.S. Treasury had stopped issuing 30-year bonds in October 2001 but resumed issuing them in February 2006. At one time the U.S. Treasury issued 20-year bonds but ceased doing so in January 1986.

Treasury Inflation-Protected Securities

The U.S. Treasury issues coupon securities that provide inflation protection. They do so by having the principal increase or decrease based on the rate of inflation such that when the security matures, the investor receives the greater of the principal adjusted for inflation or the original principal. These Treasury securities, first introduced in January 1997, are popularly referred to as *Treasury inflation-protected securities* (TIPS). As of year-end 2007, the U.S. Treasury issues a 5-year TIPS, a 10-year TIPS, and a 20-year TIPS.

TIPS work as follows. The coupon rate on an issue is set at a fixed rate, the rate being determined via the auction process described later in this chapter. The coupon rate is referred to as the *real rate* because it is the rate that the investor ultimately earns above the inflation rate. The inflation index used for measuring the inflation rate is the nonseasonally-adjusted U.S. City Average All Items Consumer Price Index for All Urban Consumers (CPI-U).

The adjustment for inflation is as follows: The principal that the U.S. Treasury will base both the dollar amount of the coupon payment and the maturity value on is adjusted semiannually. This is called the *inflation-adjusted principal.* For example, suppose that the coupon rate for a TIPS is 3.5% and the annual inflation rate is 3%. Suppose further that an investor purchases on January 1 $100,000 par value (principal) of this issue. The semiannual inflation rate is 1.5% (3% divided by 2). The inflation-adjusted principal at the end of the first six-month period is found by multiplying the original par value by one plus the semiannual inflation rate. In our example, the inflation-adjusted principal at the end of the first six-month period is $101,500. It is this inflation-adjusted principal that is the basis for computing the coupon interest for the first six-month period. The coupon payment is then 1.75% (one-half the real rate of 3.5%) multiplied by the inflation-adjusted principal at the coupon payment date ($101,500).The coupon payment is therefore $1,776.25.

Let's look at the next six months. The inflation-adjusted principal at the beginning of the period is $101,500. Suppose that the semiannual inflation rate for the second six-month period is 1%.Then the inflation-adjusted principal at the end of the second six-month period is the inflation-adjusted principal at the beginning of the six-month period ($101,500) increased by the semiannual inflation rate (1%). The adjustment to the principal is $1,015 (1% times $101,500). So, the inflation-adjusted principal at the end of the second six-month period (December 31 in our example) is $102,515 ($101,500 + $1,015). The coupon interest that will be paid to the investor at the second coupon payment date is found by multiplying the inflation-adjusted principal on the coupon payment date ($102,515) by one-half the real rate (that is, one-half of 3.5%). That is, the coupon payment will be $1,794.01.

As can be seen, part of the adjustment for inflation comes from the coupon payment since it is based on the inflation-adjusted principal. However, the U.S. government has decided to tax the adjustment each year. This feature reduces the attractiveness of TIPS as investments in accounts of tax-paying entities.

Because of the possibility of disinflation (that is, price declines), the inflation-adjusted principal at maturity may turn out to be less than the original par value. However, the Treasury has structured TIPS so that they are redeemed at the greater of the inflation-adjusted principal and the original par value.

An inflation-adjusted principal must be calculated for a settlement date if an issue is sold prior to maturity. The inflation-adjusted principal is defined in terms of an index ratio, which is the ratio of the reference CPI for the settlement date to the reference CPI for the issue date. The reference CPI is calculated with a three-month lag. For example, the reference CPI for May 1 is the CPI-U reported in February. The U.S. Department of the Treasury publishes and makes available on its web site a daily index ratio for an issue.

TREASURY AUCTION PROCESS

Treasury securities are sold in the primary market through an auction process. Each auction is announced several days in advance by means of a Treasury Department press release or press conference. The announcement provides details of the offering, including the offering amount and the term and type of security being offered, and describes some of the auction rules and procedures. Treasury auctions are open to all entities.

The U.S. Treasury makes the determination of the procedure for auctioning new Treasury securities, when to auction them, and what maturities to issue. There are periodic changes in the auction cycles and the maturity of the issues auctioned.

While the Treasury regularly offers new securities at auction, it often offers additional amounts of outstanding securities. This is referred to as a *reopening of an issue.* The Treasury has established a regular schedule of reopenings for certain maturities. To maintain the sizes of its new issues and help manage the maturity of its debt, the Treasury launched a debt buyback program. Under the program, the Treasury redeems outstanding unmatured Treasury securities by purchasing them in the secondary market through reverse auctions.

The auction for Treasury securities is conducted on a competitive bid basis. There are two types of bids that may be submitted by a bidder: noncompetitive bids and competitive bids. A *noncompetitive bid* is submitted by an entity that is willing to purchase the auctioned security at the yield that is determined by the auction process. When a noncompetitive bid is submitted, the bidder specifies only the quantity sought. The quantity in a noncompetitive bid may not exceed a specified amount. A *competitive bid* specifies both the quantity sought and the yield at which the bidder is willing to purchase the auctioned security.

The auction results are determined by first deducting the total noncompetitive tenders and nonpublic purchases (such as purchases by the Federal Reserve) from the total securities being auctioned. The remainder is the amount to be awarded to the competitive bidders. The competitive bids are then arranged from the lowest yield bid to the highest yield bid submitted. (This is equivalent to arranging the bids from the highest price to the lowest price that bidders are willing to pay.) Starting from the lowest yield bid, all competitive bids are accepted until the amount to be distributed to the competitive bidders is completely allocated. The highest yield accepted by the Treasury is referred to as the *stop-out yield* (or *high yield*). Bidders whose bid is higher than the stop-out yield are not distributed any of the new issue (that is, they are unsuccessful bidders). Bidders whose bid was the stop-out yield (that is, the highest yield accepted by the Treasury) are awarded a proportionate amount for which they bid. For example, suppose that $4 billion was tendered for at the stop-out yield, but only $1 billion remains to be allocated after allocating to all bidders who bid lower than the stop-out yield. Then each bidder who bid the stop-out yield will receive 25% of the amount for which they tendered. So, if an entity tendered for $12 million, then that entity would be awarded only $3 million.

The results announced by the U.S Treasury include the stop-out yield, the associated price, and the proportion of securities awarded to those investors who bid exactly the stop-out yield. Also announced is the quantity of noncompetitive tenders, the median-yield bid, and the ratio of the total amount bid for by the public to the amount awarded to the public (called the *bid-to-cover ratio*). For notes and bonds, the announcement includes the coupon rate of the new security. The coupon rate is set to be that rate (in increments of one-eighth of 1%) that produces the price closest to, but not above, par when evaluated at the yield awarded to successful bidders.

Now we know how the winning bidders are determined and the amount that successful bidders will be allotted, the next question is the yield at which they are awarded the auctioned security. All U.S. Treasury auctions are *single-price auctions.* In a single-price auction, all bidders are awarded securities at the highest yield of accepted competitive tenders (that is, the high yield). This type of auction is called a *Dutch auction.*

SECONDARY MARKET

The secondary market for Treasury securities is an over-the-counter (OTC) market where a group of U.S. government securities dealers offers continuous bid and ask prices on outstanding Treasuries. There is virtual 24-hour trading of Treasury securities. The three primary trading locations are New York, London, and Tokyo. The normal settlement period for Treasury securities is the business day after the transaction day ("next day" settlement).

The most recently auctioned issue is referred to as the *on-the-run issue* or the *current issue.* A security that is replaced by the on-the-run issue is called an *off-the-run issue.* At a given point in time there may be more than one off-the-run issue with approximately the same remaining maturity as the on-the-run issue. Treasury securities are traded prior to the time they are issued by the Treasury. This component of the Treasury secondary market is called the *when-issued market,* or WI market. When-issued trading for both bills and coupon securities extends from the day the auction is announced until the issue day.

Government dealers trade with the investing public and with other dealer firms. When they trade with each other, it is through intermediaries known as *interdealer brokers.* Dealers leave firm bids and offers with interdealer brokers who display the highest bid and lowest offer in a computer network tied to each trading desk and displayed on a monitor. Dealers use interdealer brokers because of the speed and efficiency with which trades can be accomplished.

Price Quotes for Treasury Bills

Bids and offers on Treasury bills are quoted in a different way than Treasury coupon securities. Unlike Treasury notes and bonds that pay interest semiannually, Treasury bills prices are quoted on a *bank discount basis* using the following formula:

$$\text{Annualized yield on a bank discount basis} = (\text{Dollar discount/Face value}) \times (360/\text{Days to maturity})$$

where dollar discount is the difference between the face value and the price.

As an example, consider a Treasury bill with 43 days to maturity, a face value of \$1 million, and selling for \$993,908.33. The dollar discount is \$6,091.67. The annualized yield on a bank discount basis is then

$$\text{Yield on a bank discount basis} = (\$6{,}091.67/\$1{,}000{,}000) \times (360/43) = 5.1\%$$

The price of a Treasury bill can be determined from the yield on a bank discount basis by using the following formula:

$$\text{Price} = \text{Face value} - [\text{Yield on a bank discount basis} \times \text{Face value} \times (\text{Days to maturity}/360)]$$

For example, consider again the 43-day Treasury bill. If the yield on a bank discount basis is 5.1%, then the price is

$$\text{Price} = \$1{,}000{,}000 - [0.051 \times \$1{,}000{,}000 \times (43/360)] = \$993{,}908.33$$

As a yield measure, the yield on a bank discount basis is flawed for two reasons. First, the measure is based on a face-value investment rather than on the actual dollar amount invested. Second, the yield is annualized according to a 360-day rather than a 365-day year, making it difficult to compare Treasury bill yields with Treasury notes and bonds, which pay interest on a 365-day basis. The use of 360 days for a year is a money market convention. Despite its shortcomings as a measure of return, this is the method that dealers have adopted to quote Treasury bills.

The yield measure employed by market participants to make the quotes on Treasury bills comparable to Treasury notes and bonds is called the *bond-equivalent yield.* The *CD equivalent yield,* also called the *money market equivalent yield,* makes the quoted yield on a Treasury bill more comparable to yield quotations on other money market instruments that pay interest on a 360-day basis. This is achieved by taking into consideration the price of the Treasury bill rather than its face value. The formula for the CD equivalent yield is

$$\text{CD equivalent yield} = \frac{360 \times \text{Yield on a bank discount basis}}{360 - (\text{Days to maturity} \times \text{Yield on a bank discount basis})}$$

As an illustration, consider a 123-day Treasury bill with a face value of \$1 million, selling for \$982,916.67, and offering a yield on a bank discount basis of 5%. Then

$$\text{CD equivalent yield} = \frac{360 \times 0.05}{360 - (123 \times 0.05)} = 5.09\%$$

Quotes on Treasury Coupon Securities

Treasury coupon securities are quoted on a price basis in points. One point is equal to 1% of par. The points are split into units of 32nds, so that a price of 97–14, for example, refers to a price of 97 and 14 32nds, or 97.4375 per 100 of par value. The 32nds are themselves often split by the addition of a plus sign or a number. A plus sign indicates that half a 32nd (or a 64th) is added to the price, and a number indicates how many eighths of 32nds (or 256ths) are added to the price. A price of 97–14+, therefore, refers to a price of 97 plus 14 32nds plus one 64th, or 97.453125, and a price of 97–142 refers to a price of 97 plus 14 32nds plus 2 256ths, or 97.4453125.

The buyer of a Treasury coupon security must compensate the seller of the bond for accrued interest (that is, the coupon interest earned from the time of the last coupon payment to the settlement date of the bond). In general, when calculating accrued interest for any bond, three pieces of information are needed: (1) the number of days in the accrued interest period, (2) the number of days in the coupon period, and (3) the dollar amount of the coupon payment. The number of days in the accrued interest period is the number of days over which the seller has earned interest before selling the security. For Treasury coupon securities, the convention used is to determine the actual number of days between two dates, referred to as the actual/actual day count convention.

The calculation of the actual number of days in the accrued interest period and the number of days in the coupon period begins with the determination of three key dates: trade date, settlement date, and date of the previous coupon payment. The trade date is the date on which the transaction is executed. The settlement date is the date a transaction is completed. For Treasury securities, settlement is the next business day after the trade date. Interest accrues on a Treasury coupon security from and including the date of the previous coupon payment up to but excluding the settlement date.

Given these values, the accrued interest for Treasury coupon securities is

$$\text{Accrued interest} = (\text{Annual dollar coupon}/2) \times \frac{\text{No. of days in accrued interest period}}{\text{No. of days in coupon period}}$$

STRIPPED TREASURY SECURITIES

The U.S. Treasury does not issue zero-coupon notes or bonds. However, because of the demand for zero-coupon instruments with no credit risk, the private sector has created such securities using a process called *coupon stripping.*

To illustrate the process, suppose that \$2 billion of a 10-year fixed-principal Treasury note with a coupon rate of

5% is purchased by a dealer firm to create zero-coupon Treasury securities. The cash flow from this Treasury note is 20 semiannual payments of $50 million each ($2 billion times 0.05 divided by 2) and the repayment of principal (also called the *corpus*) of $2 billion 10 years from now. As there are 21 different payments to be made by the U.S. Treasury for this note, a security representing a single payment claim on each payment is issued, which is effectively a zero-coupon Treasury security. The amount of the maturity value or a security backed by a particular payment, whether coupon or corpus, depends on the amount of the payment to be made by the U.S. Treasury on the underlying Treasury note. In our example, 20 zero-coupon Treasury securities each have a maturity value of $50 million, and one zero-coupon Treasury security, backed by the corpus, has a maturity value of $2 billion. The maturity dates for the zero-coupon Treasury securities coincide with the corresponding payment dates by the U.S. Treasury.

Zero-coupon Treasury securities are created as part of the U.S. Treasury's Separate Trading of Registered Interest and Principal of Securities (STRIPS) program to facilitate the stripping of designated Treasury securities. Today, all Treasury notes and bonds (fixed-principal and inflation-indexed) are eligible for stripping. The zero-coupon Treasury securities created under the STRIPS program are direct obligations of the U.S. government. Moreover, the securities clear through the Federal Reserve's book-entry system.

On dealer quote sheets and vendor screens, STRIPS, or simply, strips, are identified by whether the cash flow is created from the coupon (denoted *ci*), principal from a Treasury bond (denoted *bp*), or principal from a Treasury note (denoted *np*). Strips created from the coupon are called *coupon strips* and strips created from the principal are called *principal strips.*

A disadvantage of a taxable entity's investing in stripped Treasury securities is that accrued interest is taxed each year even though interest is not paid. Thus, these instruments are negative cash flow instruments until the maturity date. They have negative cash flow because tax payments on interest earned but not received in cash must be made. One reason for distinguishing between coupon strips and principal strips is that some foreign buyers have a preference for principal strips. This preference is due to the tax treatment of the interest in their home country. The tax laws of some countries treat the interest from a principal strip as a capital gain, which receives a preferential tax treatment (that is, lower tax rate) compared with ordinary interest income if the stripped security was created from a coupon strip.

A market participant can purchase in the market a package of zero-coupon Treasury securities such that the cash flow of the package of securities replicates the cash flow of a mispriced Treasury coupon security. By doing so, the market participant will realize a yield higher than the yield on the Treasury coupon security. This process is called *reconstitution.*

SUMMARY

The U.S. Treasury issues Treasury bills (a discount security) and Treasury notes and bonds (coupon securities) via a competitive bidding auction process according to a regular auction cycle. Treasury coupon securities include fixed-principal and inflation-protected principal securities. Securities issued by the U.S. Treasury are viewed as free of credit risk and therefore serve as a benchmark for the risk-free interest rates in the market. While the U.S. Treasury does not issue zero-coupon notes and bonds, these instruments are created through the U.S. Treasury's STRIPS program via a coupon stripping process by dealers.

REFERENCES

Fabozzi, F. J. (1998). *Treasury Securities and Derivatives* (John Wiley & Sons, 1998).

Fabozzi, F. J., and Fleming, M. J. (2005). U.S. Treasury and agency securities. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Securities* (pp. 229–250). New York: McGraw-Hill.

Fleming, M. J. (1997). The round-the-clock market for U.S. Treasury securities. *Federal Reserve Bank of New York Economic Policy Review*, July: 9–32.

Fleming, M. J. (1999). Price formation and liquidity in the U.S. Treasury market: The response to public information. *Journal of Finance* 54, 5: 1901–1915.

Fleming, M. J. (2002). Are larger Treasury issues more liquid? Evidence from bill reopenings. *Journal of Money, Credit and Banking* 34, 3: 707–735.

Fleming, M. J. (2003). Meauring Treasury market liquidity. *Federal Reserve Bank of New York Economic Policy Review*, September: 83–108.

Fleming, M. J. (2007). Who buys Treasury securities at auction? *Current Issues in Economics and Finance* 13, 1: 1–17.

CHAPTER 21

Federal Agency Securities

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

GEORGE P. KEGLER
President, Cassian Consultants

Abstract: The federal agency market includes the debt of various entities chartered by Congress to provide funding support for the housing and agricultural sectors of the U.S. economy and specific funding projects of the U.S. government. The largest issuers are also known as government-sponsored enterprises (GSEs). GSEs are either public or government owned shareholder corporations (Fannie Mae, Freddie Mac, and Tennessee Valley Authority [TVA]) or the funding entities of federally chartered bank lending systems (Federal Home Loan Banks and the Federal Farm Credit Banks). The debt of the GSEs is not guaranteed by the U.S. government. Certain smaller federal agencies have been created by Congress to address the funding of specific projects and can have partial or total government "full faith and credit" guarantees (Financing Corporation [FICO], Resolution Funding Corporation [REFCORP], Export-Import [Ex-Im] Bank, U.S.A.I.D.S., Private Export Funding Corporation [PEFCO], and the Small Business Administration [SBA]). Many of the smaller federal agencies have limited or chosen to not issue their own debt, but have used the Federal Financing Bank for their nonappropriated funding needs. Over 97% of the outstanding federal agency market debt is issued by the GSEs. In 2007, the outstanding non-mortgage-backed debt of the GSEs and all federal agencies represented 9.4% of the U.S. debt market. The debt obligations of the federal agency issuers include a broad scope of maturities, structures, liquidity, and size. The variety of issuance practices, from the smallest to the largest multibillion, multicurrency calendar programs, high credit ratings, and market liquidity gives the federal agency market a broad domestic and global base of investors and dealers. Federal agency debt occupies a unique and important place on the "efficient frontier" of the debt market curve.

Keywords: federal agency securities, federally related institutions, government-sponsored enterprises (GSEs), debentures, mortgage-backed

securities, asset-backed securities, Tennessee Valley Authority (TVA), Fannie Mae, Freddie Mac, Federal Agricultural Mortgage Corporation, Federal Farm Credit Bank System (FFCBS), Federal Home Loan Bank System, Financing Corporation, Resolution Trust Corporation (RTC), the Farm Credit Assistance Corporation

Federal agency securities can be classified by the type of issuer—federally related institutions and government-sponsored enterprises (GSEs). Federal agencies that provide credit for certain sectors of the credit marker issue two types of securities: debentures and mortgage-backed/asset-backed securities. Our focus here is on the former securities.

FEDERALLY RELATED INSTITUTIONS

Federally related institutions are arms of the federal government and generally do not issue securities directly in the marketplace. Federally related institutions include the Export-Import Bank of the United States, the Tennessee Valley Authority, the Commodity Credit Corporation, the Farmers Housing Administration, the General Services Administration, the Government National Mortgage Association, the Maritime Administration, the Private Export Funding Corporation, the Rural Electrification Administration, the Rural Telephone Bank, the Small Business Administration, and the Washington Metropolitan Area Transit Authority.

All federally related institutions are exempt from SEC registration. With the exception of securities of the Tennessee Valley Authority and the Private Export Funding Corporation, the securities are backed by the full faith and credit of the U.S. government. Interest income on securities issued by federally related institutions is exempt from state and local income taxes.

Since the federally related institution that has issued securities in recent years is the *Tennessee Valley Authority (TVA)*, we discuss these securities.

Tennessee Valley Authority

Established by Congress in 1933 primarily to provide flood control, navigation, and agricultural and industrial development, and to promote the use of electric power in the Tennessee Valley region, the TVA is the largest public power system in the United States. The TVA primarily finances its capital requirements through internally generated funds and by issuing debt. The TVA issues a variety of debt securities in U.S. dollars and other currencies (British pounds and euros). The debt obligations issued by the TVA may be issued only to provide capital for its power program or to refund outstanding debt obligations.

TVA debt obligations are not guaranteed by the U.S. government. However, the securities are rated triple A by Moody's and Standard & Poor's. The rating is based on the TVA's status as a wholly owned corporate agency of the U.S. government and the view of the rating agencies of the TVA's financial strengths. These strengths include (1) the requirements that bondholders of power bonds are given a first pledge of payment from net power proceeds, and (2) electricity rates charged by the TVA are sufficient to ensure both the full payment of annual debt service and operating and capital costs.

According to the TVA's annual report, as of September 30, 2006 TVA had 87 long-term public debt issues outstanding, totaling $20.51 billion. There are issues targeted to individual investors (retail debt offerings) and institutional investors (nonretail offerings).

For retail offerings, there are standard callable bonds (2000 Series A through Series E and 1998 Series A Estate Features), with one interesting investment feature. There is an "estate feature" that allows the bonds to be redeemed at par value plus accrued interest upon the death of the bondholder. The Putable Automatic Rate Reset Securities (PARRS) bonds (1999 Series A and 1998 Series D) are noncallable but have two interesting features. First, they have a fixed coupon rate for the first 5 years. Then there is an annual reset provision that provides for a reduction in the issue's coupon rate under certain conditions. The reduction is tied to the 30-year Treasury Constant Maturity (CMT). Second, the bondholder has the right to put the bond at par value plus accrued interest if and when the coupon rate is reduced. More recently, the TVA has issued "electronotes." The retail bonds (as well as electronotes) just described are referred to as "power bonds." There are retail bonds that are "subordinated debt." That is, they are subordinated to the power bonds. The only outstanding issue is the 1996 Series A Quarterly Income Debt Securities (QIDS).

For institutional investors, the TVA has 23 issues of global bonds outstanding; 16 are U.S. dollar–denominated, noncallable notes, and two are U.S. dollar, callable notes. Additionally, TVA has issued three noncallable global notes denominated in pounds sterling (1998 Series H, 2001 Series B, and 2003 Series A) and issues of putable notes that may not be called (2000 Series F Put, 1997 Series C Exchange, and 1996 Series A Double Put).

GOVERNMENT-SPONSORED ENTERPRISES

Government-sponsored enterprises (GSEs) are privately owned, publicly chartered entities. They were created by Congress to reduce the cost of capital for certain borrowing sectors of the economy deemed to be important enough to warrant assistance. The entities in these sectors include farmers, homeowners, and students. The enabling legislation dealing with a GSE is reviewed periodically.

GSEs issue securities directly in the marketplace. The market for these securities, while smaller than that of Treasury securities, has in recent years become an active and important sector of the bond market. Since 1998, a number of the GSEs have initiated programmatic debt issuance platforms, which will be discussed in more detail, in addition to more traditional funding methodologies. GSEs are also issuers of foreign currency denominated and U.S. dollar global bonds.

There are five GSEs that currently issue debentures: Freddie Mac, Fannie Mae, Federal Home Loan Bank System, Federal Farm Credit System and the Federal Agricultural Mortgage Corporation. Fannie Mae, Freddie Mac, and Federal Home Loan Bank are responsible for providing credit to the housing sectors. The Federal Agricultural Mortgage Corporation provides the same function for agricultural mortgage loans. The Federal Farm Credit Bank System is responsible for the credit market in the agricultural sector of the economy.

The interest earned on obligations of the Federal Home Loan Bank System, the Federal Farm Credit System, and the Student Loan Marketing Association are exempt from state and local income taxes. In addition to the debt obligations issued by these five GSEs, there are issues outstanding by one-time GSE issuers that have been dismantled. These GSEs include the Financing Corporation, Resolution Trust Corporation, and the Farm Credit Assistance Corporation.

The price quotation conventions for GSE securities will vary between types of debt. Short term GSE discount notes are quoted on a yield basis, the same as for Treasury bills. The most liquid programmatic GSE issues are generally quoted on two primary bases. One is a price basis, like Treasury securities. That is, the bid and ask price quotations are expressed as a percentage of par plus fractional 32nds of a point. Two is a spread basis, as an indicated yield spread in basis points, off a choice of proxy curves or issue. The Treasury market is the most popular bellwether proxy from which most GSE debt is quoted. The less liquid GSE securities types, such as callable debt, that contain some form of optionality, may be quoted on a yield spread basis off either Treasuries, U.S. dollar interest rate swaps curve or a yield curve referencing GSE debt or a particular GSE issue.

A third quotation convention was introduced to the GSE debt market in 2001, when Freddie Mac began its Reference Note auctions. In preauction trading the issues have been quoted on a "When Issued" (WI) basis (see Bond Market Association, 2001) a straight yield basis, such as used in trading Treasury WI issues. This quotation convention is used until the issue is priced at auction, at which point the price quotes usually return to a yield spread basis. Some GSE issues trade with almost the same liquidity as Treasury securities. Other issues that are supported only by a few dealers trade much like off-the-run corporate bonds.

Types and Features of GSE Securities

In general, GSEs issue two types of debt: debentures and discount notes. Debentures can be either notes or bonds. GSE-issued notes, with minor exceptions, have 1- to 20-year maturities and bonds have maturities longer than 20 years. There are issues with bullet maturities and those with call provisions. GSEs also issue structured notes. The variety of notes issued by the GSEs will be discussed in greater detail later in this chapter.

Discount notes are short-term obligations with maturities ranging from overnight to 360 days. As with Treasury bills, no coupon interest is paid. Instead, the investor earns interest by buying the note at a discount.

Programmatic GSE Issuance Platforms

In 1998, Fannie Mae and Freddie Mac began issuing respectively, Benchmark and Reference Notes and Bonds. These programmatic platforms incorporated preannounced funding calendars and large minimum sized issues to introduce greater transparency in their funding programs and to promote greater liquidity for the issued debt. In 1999, both GSEs included Benchmark Bills and Reference Bills, respectively, in weekly auction formats to augment their short-term discount note funding programs. Subsequently, the Federal Home Loan Banks and the Federal Farm Credit Banks, through their respective funding entities, the Federal Home Loan Banks Office of Finance and the Federal Farm Credit Funding Corporation, initiated programmatic debt platforms. Federal Home Loan Banks issue Federal Home Loan TAPs and Federal Farm Credit Banks issue Farm Credit Designated Notes. Whereas the funding needs of Freddie Mac and Fannie Mac are derived from single corporate entities, which allows for more exact issuance calendar announcements, the demands of funding separate bank balance sheets within the Federal Home Loan Bank and Federal Farm Credit Bank systems has limited the amount of programmatic funding for these GSEs. Freddie Mac, Fannie Mae, and the Federal Home Loan Banks utilize auctions when issuing most of their short-term debt. Though varying in size and scope between the GSEs, the auctioned maturities include regular 1-, 2-, 3-, 6-, and 12-month maturities. Freddie Mac has also incorporated auctions in the issuance of 2-, 3-, and 5-year Reference Notes. This has allowed the "When Issued" (WI) trading of GSE coupon debt for the first time, a significant milestone in the transparency and liquidity of the GSE securities market.

Both Freddie Mac and Fannie Mae will periodically announce repurchase and/or exchange transactions involving their programmatic issued securities.

Description of GSEs and Securities Issued

The five GSEs that currently issue securities and the three GSEs that have outstanding issues can be briefly described as follows.

Fannie Mae

The residential mortgage debt market in the United States represents the largest mortgage debt market in the world. The problem the U.S. government faces is to attract investors to invest in residential mortgages. At one time, savings and loan associations were the primary investors, especially with special inducements the government provided. But since there was not an active market where these debt instruments traded, mortgages were illiquid and financial institutions that invested in them were exposed to liquidity risk.

In the 1930s, Congress figured out a way to handle this problem. It created a federally related institution, the Federal National Mortgage Association, now officially known as "*Fannie Mae* which was charged with the responsibility to create a liquid secondary market for mortgages. Fannie Mae was to accomplish this objective by buying and selling mortgages. Fannie Mae needed a funding source in case it faced a liquidity squeeze. Congress provided this by giving Fannie Mae a credit line with the Treasury.

Despite the presence of Fannie Mae, the secondary mortgage market did not develop to any significant extent. During periods of tight money, Fannie Mae could do little to mitigate a housing crisis. In 1968, Congress divided Fannie Mae into two entities: (1) the current Fannie Mae and (2) the Government National Mortgage Association (popularly known as "Ginnie Mae"). Ginnie Mae's function is to use the "full faith and credit of the U.S. government" to support the market for government-insured mortgages. While starting out as a federally related institution, today Fannie Mae is a GSE.

Fannie Mae issues Benchmark Bills, Benchmark Notes and Benchmark Bonds, Callable Benchmark Notes, Subordinated Benchmark Notes, Investment Notes, callable securities, and structured notes. Benchmark Notes and Benchmark Bonds are noncallable instruments. The minimum issue size is $4 billion for Benchmark Notes and $2 billion for Benchmark Bonds. Issued quarterly are 2-year or 3-year, 5-year, 10-year, and 30-year maturities.

Freddie Mac

In 1970, two years after Congress divided Fannie Mae into the now current Fannie Mae and Ginnie Mae, Congress created Freddie Mac (at one time called the Federal Home Loan Mortgage Corporation). The reason for the creation of *Freddie Mac* was to provide support for conventional mortgages. These mortgages are not guaranteed by the U.S. government.

Freddie Mac issues Reference Bills, discount notes, medium-term notes, Reference Notes and Bonds, Callable Reference Notes, Euro Reference Notes (debt denominated in euros) and global bonds. Reference Bills and discount notes are issued with maturities of 1 year or less. Reference Notes and Bonds have maturities of 2 to 30 years and Callable Reference Notes have maturities of 2 to 10 years. Freddie Mae will issue and/or reopen Reference Bills, Reference Notes, 30-year Reference Bonds, and Euro Reference Notes according to a published issuance calendar and within minimum issue size guidelines. Freddie Mac Reference Notes and Reference Bonds are eligible for stripping.

Both Freddie Mac and Fannie Mae issue bullet and callable medium-term notes (MTNs) and structured notes, which are customized based on demand (reverse inquiry) from institutional investors. The structured notes issued have been various floating-rate, zero-coupon, and step up securities. There are securities denominated in U.S. dollars as well as issues denominated in a wide range of foreign currencies.

Freddie Mac and Fannie Mae issue subordinated securities, in the form of Freddie SUBS and Fannie Mae Subordinated Benchmark Notes, respectively. These are unsecured subordinated obligations of the separate corporations that rank junior in right of payment to all of Freddie Mac's and Fannie Mae's existing and future obligations. The payment structure is as follows. Separately the effected corporation must defer payment of interest on all outstanding subordinated debt if certain conditions are realized. Deferral of interest is not permitted for more than five consecutive years nor beyond the maturity date. Accrual of interest is compounded at the issue's coupon rate. During any deferral period, the effected GSE may not declare or pay dividends on, or redeem, purchase, or acquire its common stock or its preferred stock. The first separate offerings of Freddie Mac and Fannie Mae subordinated debt issues were in 2001, both receiving an Aa2 from Moody's Investors Service and AA– from Standard & Poor's.

Federal Home Loan Bank System (FHL Banks)

The *Federal Home Loan Bank System* (FHL Banks) consists of the 12 district Federal Home Loan Banks and their member banks. The Federal Home Loan Bank Board was originally responsible for regulating all federally chartered savings and loan associations and savings banks, as well as state-chartered institutions insured by the Federal Savings and Loan Insurance Corporation. These responsibilities have been curtailed since 1989.

The major source of debt funding for the Federal Home Loan Banks is the issuance of consolidated debt obligations, which are joint and several obligations of the 12 Federal Home Loan Banks. Consolidated FHL Bank discount notes with maturities from 1 to 360 days are issued daily. Discount notes are also auctioned twice weekly in 4-, 9-, 13-, and 26-week maturities. Because FHL Bank bond issuance is directly related to member bank needs, there is no debt calendar in the traditional sense. Bullets, callables, and floaters are issued on a daily basis. The FHL-Banks have several Programs to facilitate the issuance of certain bond types. The TAP Issue program was launched in 1999. This program aggregates FHL Bank demand for six common (1.5-, 2-, 3-, 5-, 7-, and 10-year) bullet maturities, and then offers them daily through competitive auctions. These issues feature standardized terms and are reopened via auction for 3-month periods, enabling them to reach multibillion dollar size. TAP Issues can also be reopened as they roll down the curve. Callable bonds are issued daily, primarily as customized issues from reverse inquiry of institutional investors. The FHL Bank Global Bond Program will periodically offer larger sized ($1 billion minimum for callable and $3 billion minimum for bullet maturities) with standardized term and are targeted to foreign investors in either U.S. dollars or other currencies.

The Federal Agricultural Mortgage Corporation (Farmer Mac)

The *Federal Agricultural Mortgage Corporation* (Farmer Mac) provides a secondary market for first mortgage agricultural real estate loans. It was created by Congress in 1998 to improve the availability of mortgage credit to farmers, ranchers, and rural homeowners, businesses, and communities. It does so by purchasing qualified loans from lenders in the same way as Freddie Mac and Fannie Mae.

Farmer Mac raises funds by selling debentures and mortgage-backed securities backed by the loans purchased. The latter securities are called agricultural mortgage-backed securities (AMBSs). The debentures that are issued include discount notes and medium-term notes.

Federal Farm Credit Bank System (Farm Credit)

The purpose of the *Federal Farm Credit Bank System (FFCBS)* is to facilitate adequate, dependable credit and related services to the agricultural sector of the economy. The Farm Credit System consists of three entities: the Federal Land Banks, Federal Intermediate Credit Banks, and Banks for Cooperatives. Before 1979, each entity issued securities in its own name. Starting in 1979, they began to issue debt on a consolidated basis as "joint and several obligations" of the FFCBS. All financing for the FFCBS is arranged through the Federal Farm Credit Banks Funding Corporation (FFCBFC), which issues consolidated obligations.

The FFCBFC issues debt through five formats. Discount notes are offered daily through posted rates. Calendar Bonds of 3- and 6-month maturities are offered monthly. Designated Bonds of typically 2-year maturities can be offered twice monthly as either a new issue ($1 billion minimum) or reopening ($100 million minimum). Unscheduled bonds are issued throughout the month in varying sizes and structures either by competitive bidding or negotiated reverse inquiry by institutional investors. FFCB Master Notes are issued as individually tailored daily investment agreements usually designed for a single investor.

Sallie Mae

Sallie Mae provides liquidity for private lenders participating in the Federal Guaranteed Student Loan Program, the Health Education Assistance Loan Program, and the PLUS loan program (a program that provides loans to the parents of undergraduate students). In 2004 Sallie Mae completed unwinding its status as a GSE. The outstanding debt issued by Sallie Mae as a GSE has been "grandfathered" as GSE until maturity. Currently, Sallie Mae issues unsecured short-term debt (discount notes, six month floating-rate notes), and structured asset-backed issues securitizing portions of the 95% to 97% government guaranteed loans that Sallie Mae has processed.

Financing Corporation (FICO)

The deposits of savings and loans were once insured by the Federal Savings and Loan Insurance Corporation (FSLIC), overseen by the Federal Home Loan Bank Board. When difficulties encountered in the savings and loan industry raised concerns about FSLIC's ability to meet its responsibility to insure deposits, Congress passed the Competitive Equality and Banking Act in 1987. This legislation included provisions to recapitalize FSLIC and establish a new government-sponsored agency, the Financing Corporation (FICO), to issue debt in order to provide funding for FICO. FICO issued its first bonds in September 1987—a 30-year noncallable $500 million issue. The principal of these bonds is backed by zero-coupon Treasury securities. The legislation permitted FICO to issue up to $10.825 billion but not more than $3.75 billion in any 1 year. FICO was legislated to be dismantled in 2026, or after all securities have matured, whichever came sooner.

Resolution Trust Corporation (REFCORP)

The 1987 legislation that created FICO did not go far enough to resolve the problems facing the beleaguered savings and loan industry. In 1989, Congress passed more comprehensive legislation, the Financial Institutions Reform, Recovery and Enforcement Act (FIRREA). This legislation had three key elements. First, it transferred supervision of savings and loans to a newly created Office of Thrift Supervision. Second, it shifted the FSLIC insurance function to a Savings Association Insurance Fund, placed under the supervision of the Federal Deposit Insurance Corporation. Third, it established the *Resolution Trust Corporation (RTC)* as a GSE charged with the responsibility of liquidating or bailing out insolvent savings and loan institutions. The RTC obtained its funding from the Resolution Funding Corporation (REFCORP), which was authorized to issue up to $40 billion of long-term bonds. The principal of this debt is backed by zero-coupon Treasury bonds. REFCORP has issued both 30-year and 40-year bonds.

Farm Credit Financial Assistance Corporation (FACO)

In the 1980s, the FFCBS faced financial difficulties because of defaults on loans made to farmers. The defaults were caused largely by high interest rates in the late 1970s and early 1980s and by depressed prices on agricultural products. To recapitalize the Federal Farm Credit Bank System, Congress created the Farm Credit Financial Assistance Corporation (FACO) in 1987. This federally sponsored agency was authorized to issue debt to assist the FFCBS. FACO bonds, unlike the debt of other GSEs, are backed by the Treasury.

Repo Transactions Market in GSE Debt Collateral

Due to the high credit characteristics (all senior debt issued by the GSEs is rated Aaa by Moody's) discussed in more detail below and the steady increase in secondary trading activity, an active "repo" market has developed in GSE debt. Based on data published by the Securities Industry and Financial Markets Association (SIFMA), the

Fixed Income Clearing Corporation's Government Securities Division (GDS) reported that repo transactions in nonmortgage GSE collateral for the year 2006 totaled $35.2 trillion and represented 8.4% of the total U.S. government securities market repo volume. For the year 2005, repo transactions, in the same collateral, totaled $37.2 trillion and was 9.3% of the total U.S. government securities repo market. GSE debt is also accepted collateral for monetary policy–related temporary reserve operations conducted by the Federal Reserve Bank of New York's Open Market desk. GSE collateral will typically have a 2% extra margin requirement (known as a security "haircut," which varies depending on credit, maturity, and type of security used) in repurchase or reverse repurchase transactions as compared with 5% to 7% haircuts on investment-rated corporate debt.

Credit Risk

With the exception of the securities issued by the Farm Credit Financial Assistance Corporation, GSE securities are not backed by the full faith and credit of the U.S. government, as is the case with Treasury securities. Consequently, investors purchasing GSEs are exposed to credit risk. The yield spread between these securities and Treasury securities of comparable maturity reflects differences in perceived credit risk and liquidity. The spread attributable to credit risk reflects any financial difficulty faced by the issuing GSEs and the likelihood that the federal government will allow the GSE to default on its outstanding obligations.

Two examples will illustrate this point. In late 1981 and early 1982, the net income of the Fannie Mae weakened, causing analysts to report that the securities of this GSE carried greater credit risk than previously perceived. As a result, the yield spread over Treasuries on its debt rose from 91 basis points (on average) in 1981 to as high as 150 basis points. In subsequent years, the Fannie Mae's net income improved, and its yield spread to Treasuries narrowed. As another example, in 1985 the yield spread on securities of the Farm Credit Bank System rose substantially above those on comparable-maturity Treasuries because of this GSE's financial difficulties. The spread between 1985 and 1986 varied with the prospects of Congressional approval of a bailout measure for the system. More recently, certain issues on the accounting treatment of portfolio hedges affected both Fannie Mae and Freddie Mac; during this period, there was some minor volatility in their debt spreads to the Treasury curve.

Yield Spreads

Because of credit risk and liquidity, GSEs will trade at a yield premium to comparable-maturity Treasury securities. The yield spread will differ for each issuing entity, the maturity and structure of the security and the program through which it has been issued. Most GSE issues are priced on a spread basis to the London Interbank Offered Rate (LIBOR). Larger programmatic issues will usually trade at a slight premium to their smaller issue–sized counterparts. There will be larger spread variances between noncallable and callable issues with similar final maturities, as there will be between GSE issues with shorter call protection versus longer call protection with like final maturity dates. The longer the time before the issue can be called, the less valuable the embedded option in the call. As a result, the longer the noncall period, the tighter the yield spread.

SUMMARY

Created by Congress, the GSEs are chartered to support the funding and secondary liquidity for the housing and agricultural sectors of the U.S. economy. The variety and size of the GSE funding programs touches on all points of the yield curve and has helped to create a liquid market in both noncall and callable bond structures. The Fed wire payment and delivery advantage and high credit rating gives federal agency debt additional collateral uses and a broad investor and dealer base. The GSEs' use of traditional selling group, syndicate, and auctions for pricing and the initial distribution of new issues has helped create a transparent and active funding calendar in their debt. Within the federal agency market, the GSEs represent about 97% of the outstanding debt. Freddie Mac, Fannie Mae, Federal Home Loan Banks, Federal Farm Credit Banks, and TVA are the largest, most active issuers of federal agency debt. Smaller federal agencies primarily use the Federal Financing Bank as their source of nonappropriated funding.

REFERENCES

Bond Market Association (2001). *The Bond Market Association's Practice Guidelines for When Issued Trading in GSE Auctioned Securities*. New York: Bond Market Association.

Congressional Budget Office (2006). *Measuring the capital positions of Fannie Mae and Freddie Mac*. CBO paper, June.

Moe, R. C., and Stanton, T. H. (1989). Government-sponsored enterprises as federal instrumentalities: Reconciling private management with public accountability. *Public Administration Review* 49, 4: 321–329.

Silber, W. L. (1974). The market for federal agency securities: Is there an optimum size of issue? *The Review of Economics and Statistics* 56, 1: 14–22.

Stanton, T. H. (2002). *Government-Sponsored Enterprises: Mercantilist Companies in the Modern World*. Washington, DC: AEI Press.

CHAPTER 22

Municipal Securities

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: Debt obligations are issued by state and local governments and by entities that they establish. These securities are referred to as municipal securities or municipal bonds. The two general types of municipal bond structures are tax-backed securities and revenue bonds. There are also municipal bonds with special bond structures. The primary attractiveness of municipal bonds is that the interest earned is exempt from federal income taxation. While not all municipal securities are exempt from federal income taxation, tax-exempt municipal bonds are the largest component of the market.

Keywords: municipal bonds, tax-backed debt obligations, revenue bonds, general obligation debt, Municipal Securities Rulemaking Board (MSRB), taxable municipal bonds, original-issue discount (OID) bond, alternative minimum taxable income (AMTI), alternative minimum tax (AMT), official statement, municipal notes, variable rate demand obligations (VRDOs), utility revenue bonds, transportation revenue bonds, housing revenue bonds, higher education revenue bonds, health care revenue bonds, seaport revenue bonds, insured bonds, letter of credit (LOC), bank-backed municipal bonds, refunded bonds, equivalent taxable yield, tender option bonds (TOBs), inverse floater, residual, effective marginal tax rate, structure risk, tax risk

In this chapter, we discuss the types of debt obligations issued by states, municipal governments, and public agencies and their instrumentalities and the investment characteristics of these financial instruments.

ISSUERS AND ISSUANCE PROCEDURES

Issuers of *municipal bonds* include municipalities, counties, towns and townships, school districts, and special service system districts. Included in the category of municipalities are cities, villages, boroughs, and incorporated towns that received a special state charter. Counties are geographical subdivisions of states whose functions are law enforcement, judicial administration, and construction and maintenance of roads.

As with counties, towns and townships are geographical subdivisions of states and perform similar functions as counties. A special purpose service system district, or simply special district, is a political subdivision created to foster economic development or related services to a geographical area. Special districts provide public utility services (water, sewers, and drainage) and fire protection services. Public agencies or instrumentalities include authorities and commissions.

The number of municipal bond issuers is remarkable: more than 60,000. Even more noteworthy is the number of different issues: more than 1.3 million. There are more than 50,000 bonds that are priced in the Standard & Poor's Investortools Main Municipal Bond Index (see Garrett, 2008).

The *Municipal Securities Rulemaking Board (MSRB)* regulates various aspects of the municipal bond market including municipal securities brokers and dealers. The MSRB was established in 1975 by the Securities and Exchange Commission (SEC) as a self-regulatory organization pursuant to a Congressional directive. It adopts rules to (1) prevent fraudulent and manipulative acts and practices, (2) promote fair and equitable principles for the trading of municipal securities, and (3) protect investors and the public interest. (For a detailed discussion of the MSRB, see Maco and Taffe [2008].)

Municipal bonds are issued in one of three ways: negotiated sale, competitive bidding, or private placement. In a negotiated sale, an investment banker is retained by the issuer to underwrite the issue and then sell the bonds to the public. In a competitive bidding, investment bankers bid on an issue, the winning bidder being the investment bank that bids the lowest interest rate (or equivalent, the highest price). The investment bank or syndicate that wins the issue then distributes the securities to the public. In a private placement, a method typically reserved for small-size bond issues and accounting for less than 1% of all new issuance, the issue is placed directly with one or more institutional investors. An issuer may not have a choice as to which method to use. A state may mandate certain types of bonds be issued using a particular method. For example, some states will mandate that the state's *general obligation bonds* be sold via competitive bidding (see Peng and Brucato, 2001). Peng, Kriz, and Neish (2008) provide a detailed description of the factors municipal issues consider in selecting between a competitive bidding and negotiated sale.

TAX-EXEMPT AND TAXABLE MUNICIPAL SECURITIES

There are both tax-exempt and taxable municipal securities. "Tax-exempt" means that interest on a municipal security is exempt from federal income taxation. The tax exemption of municipal securities applies to interest income, not capital gains. The exemption may or may not extend to taxation at the state and local levels. The state tax treatment depends on (1) whether the issue from which the interest income is received is an "in-state issue" or an "out-of-state issue," and (2) whether the investor is an individual or a corporation. The treatment of interest income at the state level will be one of the following:

1. Taxation of interest from municipal issues regardless of whether the issuer is in state or out of state.
2. Exemption of interest from all municipal issues regardless of whether the issuer is in state or out of state.
3. Exemption of interest from municipal issues that are in state but some form of taxation where the source of interest is an out-of-state issuer.

However, the differential tax treatment of interest from in-state and out-of-state bond issues has been challenged. In Kentucky, a state court ruled that the state apply the same tax treatment to interest from both types of issuers. In August 2006, the Kentucky Supreme Court let that ruling stand. Kentucky has appealed the decision to the United States Supreme Court, which considered the case in October 2007. If the ruling stands, then only 1 and 2 above will be permitted by states.

Most municipal securities that have been issued are tax-exempt. Municipal securities are commonly referred to as tax-exempt securities although taxable municipal securities have been issued and are traded in the market. Municipalities issue *taxable municipal bonds* to finance projects that do not qualify for financing with tax-exempt bonds. An example is a sports stadium. The most common types of taxable municipal bonds are *industrial revenue bonds* and *economic development bonds*. Since there are federally mandated restrictions on the amount of tax-exempt bonds that can be issued, a municipality will issue taxable bonds when the maximum is reached. There are some issuers who have issued taxable bonds in order to take advantage of demand outside of the United States.

There are other types of tax-exempt bonds. These include bonds issued by nonprofit organizations. Such organizations are structured so that none of the income from the operations of the organization benefit an individual or private shareholder. The designation of a nonprofit organization must be obtained from the Internal Revenue Service. Since the tax-exempt designation is provided pursuant to Section 501(c)(3) of the Internal Revenue Code, the tax-exempt bonds issued by such organizations are referred to as 501(c)(3) obligations. Museums and foundations fall into this category. Tax-exempt obligations also include bonds issued by the District of Columbia and any possession of the United States—Puerto Rico, the U.S. Virgin Islands, Guam, American Samoa, and the Northern Mariana Islands. The interest income from securities issued by U.S. territories and possessions is exempt from federal, state, and local income taxes in all 50 states.

TAX PROVISIONS AFFECTING MUNICIPAL SECURITIES

Federal tax rates and the treatment of municipal interest at the state and local levels affect municipal security values and strategies employed by investors. There are provisions in the Internal Revenue Code that investors in municipal securities should recognize. These provisions deal with original issue discounts, the alternative minimum tax, and the deductibility of interest expense incurred to acquire municipal securities.

Treatment of Original-Issue Discount

If at the time of issuance the original-issue price is less than its maturity value, the bond is said to be an *original-issue discount (OID) bond.* The difference between the par value and the original-issue price represents tax-exempt interest that the investor realizes by holding the issue to maturity.

For municipal bonds there is a complex treatment that investors must recognize when purchasing OID municipal bonds. The Revenue Reconciliation Act of 1993 specifies that any capital appreciation from the sale of a municipal bond that was purchased in the secondary market after April 30, 1993, could be either (1) free from any federal income taxes, (2) taxed at the capital gains rate, (3) taxed at the ordinary income rate, or (4) taxed at a combination of the two rates.

The key to the tax treatment is the rule of de minimis for any type of bond. The rule states that a bond is to be discounted up to 0.25% from the par value for each remaining year of a bond's life before it is affected by ordinary income taxes. The discounted price based on this rule is called the market discount cutoff price. The relationship between the market price at which an investor purchases a bond, the market discount cutoff price, and the tax treatment of the capital appreciation realized from a sale is as follows. If the bond is purchased at a market discount, but the price is higher than the market discount cutoff price, then any capital appreciation realized from a sale will be taxed at the capital gains rate. If the purchase price is lower than the market discount cutoff price, then any capital appreciation realized from a sale may be taxed as ordinary income or a combination of the ordinary income rate and the capital gains rate. (Several factors determine what the exact tax rate will be in this case.)

The market discount cutoff price changes over time because of the rule of de minimis. The price is revised. An investor must be aware of the revised price when purchasing a municipal bond because this price is used to determine the tax treatment.

Alternative Minimum Tax

Alternative minimum taxable income (AMTI) is a taxpayer's taxable income with certain adjustments for specified tax preferences designed to cause AMTI to approximate economic income. For both individuals and corporations, a taxpayer's liability is the greater of (1) the tax computed at regular tax rates on taxable income and (2) the tax computed at a lower rate on AMTI. This parallel tax system, the *alternative minimum tax (AMT),* is designed to prevent taxpayers from avoiding significant tax liability as a result of taking advantage of exclusions from gross income, deductions, and tax credits otherwise allowed under the Internal Revenue Code.

One of the tax preference items that must be included is certain tax-exempt municipal interest. As a result of AMT, the value of the tax-exempt feature is reduced. However, the interest of not all municipal issues is subject to the AMT. Under the current tax code, tax-exempt interest earned on all private activity bonds issued after August 7, 1986 must be included in AMTI. There are two exceptions. First, interest from bonds that are issued by 501(c)(3) organizations (that is, not-for-profit organizations) is not subject to AMTI. The second exception is interest from bonds issued for the purpose of refunding if the original bonds were issued before August 7, 1986. The AMT does not apply to interest on governmental or nonprivate activity municipal bonds. An implication is that those issues that are subject to the AMT will trade at a higher yield than those exempt from AMT.

For investors in mutual funds that invest in municipal bonds, the prospectus will disclose whether the fund's manager is permitted to invest in AMT bonds and if it permitted, the maximum amount. Usually, when a mutual fund allows investments in AMT bonds, the maximum is 20%. The year-end 1099 form provided to investors in mutual funds will show the percentage of the income of the fund must be included in AMTI.

Deductibility of Interest Expense Incurred to Acquire Municipals

Ordinarily, the interest expense on borrowed funds to purchase or carry investment securities is tax deductible. There is one exception that is relevant to investors in municipal bonds. The Internal Revenue Code specifies that interest paid or accrued on "indebtedness incurred or continued to purchase or carry obligations, the interest on which is wholly exempt from taxes," is not tax deductible. It does not make any difference if any tax-exempt interest is actually received by the taxpayer in the taxable year. In other words, interest is not deductible on funds borrowed to purchase or carry tax-exempt securities.

Special rules apply to commercial banks. At one time, banks were permitted to deduct all the interest expense incurred to purchase or carry municipal securities. Tax legislation subsequently limited the deduction first to 85% of the interest expense and then to 80%. The 1986 tax law eliminated the deductibility of the interest expense for bonds acquired after August 6, 1986. The exception to this nondeductibility of interest expense rule is for bank-qualified issues. These are tax-exempt obligations sold by small issuers after August 6, 1986 and purchased by the bank for its investment portfolio.

An issue is bank qualified if (1) it is a tax-exempt issue other than private activity bonds, but including any bonds issued by 501(c)3 organizations, and (2) it is designated by the issuer as bank qualified and the issuer or its subordinate entities reasonably do not intend to issue more than $10 million of such bonds. A nationally recognized and experienced bond attorney should include in the opinion letter for the specific bond issue that the bonds are bank qualified.

TYPES OF MUNICIPAL SECURITIES

Municipal securities are issued for various purposes. Short-term notes typically are sold in anticipation of the receipt of funds from taxes or receipt of proceeds from the sale of a bond issue, for example. Proceeds from the sale of short-term notes permit the issuing municipality to cover seasonal and temporary imbalances between outlays for expenditures and inflows from taxes. Municipalities issue long-term bonds as the principal means for financing both (1) long-term capital projects such as schools, bridges, roads, and airports; and (2) long-term budget deficits that arise from current operations.

An *official statement* describing the issue and the issuer is prepared for new offerings. Municipal securities have legal opinions that are summarized in the official statement. The importance of the legal opinion is twofold. First, bond counsel determines if the issue is indeed legally able to issue the securities. Second, bond counsel verifies that the issuer has properly prepared for the bond sale by having enacted various required ordinances, resolutions, and trust indentures and without violating any other laws and regulations.

There are basically two types of municipal security structures: tax-backed debt and revenue bonds. We describe each type, as well as variants.

Tax-Backed Debt

Tax-backed debt obligations are secured by some form of tax revenue. The broadest type of tax-backed debt obligation is the general obligation debt. Other types that fall into the category of tax-backed debt are appropriation-backed obligations, debt obligations supported by public credit enhancement programs, and short-term debt instruments.

General Obligation Debt

General obligation pledges include unlimited and limited tax general obligation debt. The stronger form is the unlimited tax general obligation debt (also called an ad valorem property tax debt) because it is secured by the issuer's unlimited taxing power (corporate and individual income taxes, sales taxes, and property taxes) and is said to be secured by the full faith and credit of the issuer. A limited tax general obligation debt (also called a limited ad valorem tax debt) is a limited tax pledge because for such debt there is a statutory ceiling on the tax rates that may be levied to service the issuer's debt.

There are general obligation bonds that are secured not only by the issuer's general taxing powers to create revenues accumulated in a general fund, but also secured by designated fees, grants, and special charges from outside the general fund. Due to the dual nature of the revenue sources, bonds with this security feature are referred to as double-barreled in security. As an example, special purpose service systems issue bonds that are secured by a pledge of property taxes, a pledge of special fees/operating revenue from the service provided, or a pledge of both property taxes and special fees/operating revenues.

Appropriation-Backed Obligations

Bond issues of some agencies or authorities carry a potential state liability for making up shortfalls in the issuing entity's obligation. While the appropriation of funds must be approved by the issuer's state legislature, and hence they are referred to as appropriation-backed obligations, the state's pledge is not binding. Because of this nonbinding pledge of tax revenue, such issues are referred to as moral obligation bonds. An example of the legal language describing the procedure for a moral obligation bond that is enacted into legislation is as follows:

> In order to further assure the maintenance of each such debt reserve fund, there shall be annually apportioned and paid to the agency for deposit in each debt reserve fund such sum, if any, as shall be certified by the chairman of the agency to the governor and director of the budget as necessary to restore such reserve fund to an amount equal to the debt reserve fund requirement. The chairman of the agency shall annually, on or before December 1, make and deliver to the governor and director of the budget his certificate stating the sum or sums, if any, required to restore each such debt reserve fund to the amount aforesaid, and the sum so certified, if any, shall be apportioned and paid to the agency during the then current state fiscal year.

The reason for the moral obligation pledge is to enhance the creditworthiness of the issuing entity. The first moral obligation bond was issued by the Housing Finance Agency of the state of New York. Historically, most moral obligation debt has been self-supporting; that is, it has not been necessary for the state of the issuing entity to make an appropriation. In those cases in which state legislatures have been called on to make an appropriation, they have. For example, the states of New York and Pennsylvania did this for bonds issued by their Housing Finance Agency; the state of New Jersey did this for bonds issued by the Southern Jersey Port Authority.

Another type of appropriation-backed obligation is lease-backed debt. There are two types of leases. One type is basically a secured long-term loan disguised as lease. The "leased" asset is the security for the loan. In the case of a bankruptcy, the court would probably rule such an obligation as the property of the user of the leased asset and the debt obligation of the user. In contrast, the second type of lease is a true lease in which the user of the leased asset (called the lessee) makes periodic payments to the leased asset's owner (called the lessor) for the right to use the leased asset. For true leases, there must be an annual appropriation by the municipality to continue making the lease payments.

Dedicated Tax-Backed Obligations

States and local governments have issued increasing amounts of bonds where the debt service is to be paid from so-called dedicated revenues such as sales taxes, tobacco settlement payments, fees, and penalty payments. Many are structured to mimic asset-backed securities.

Let's look at one type of such security. Tobacco settlement revenue (TSR) bonds are backed by the tobacco settlement payments owed to the state or local entity resulting from the master settlement agreement between most of the states and the four major U.S. tobacco companies (Philip Morris Inc., R. J. Reynolds Tobacco Co., Lorillard

Tobacco Co., and Brown & Williamson Tobacco Corp.) in November 1998. The states that are parties to the settlement have subsequent to the settlement issued $36.5 billion of tax-exempt revenue bonds. There are unique risks associated with TSR bonds having to do with structural risk, the credit risk of the four tobacco companies, cash flow risk, and litigation risk. (These risks are discussed in Lian [2008] and Ellis [2008].) The initial credit ratings when these bonds were first issued was typically within the A or AA range; however, by mid-2007, their credit ratings were generally in the BBB range, reflecting these risks.

Debt Obligations Supported by Public Credit Enhancement Programs

Unlike a moral obligation bond, there are bonds that carry some form of public credit enhancement that is legally enforceable. This occurs when there is a guarantee by the state or a federal agency or when there is an obligation to automatically withhold and deploy state aid to pay any defaulted debt service by the issuing entity. It is the latter form of public credit enhancement that is employed for debt obligations of a state's school systems.

Short-Term Debt Instruments

Short-term debt instruments issued by municipalities include notes, commercial paper, variable-rate demand obligations, and a hybrid of the last two products.

Municipal Notes Usually, *municipal notes* are issued for a period of 12 months, although it is not uncommon for such notes to be issued for periods as short as 3 months and for as long as 3 years. Municipal notes include bond anticipation notes (BANs) and cash flow notes. BANs are issued in anticipation of the sale of long term bonds. The issuing entity must obtain funds in the capital market to pay off the obligation.

Cash flow notes include tax anticipation notes (TANs) and revenue anticipation notes (RANs). TANs and RANs (also known as TRANs) are issued in anticipation of the collection of taxes or other expected revenues. These are borrowings to even out irregular flows into the treasury of the issuing entity. The pledge for cash flow notes can be either a broad general obligation pledge of the issuer or a pledge from a specific revenue source. The lien position of cash flow noteholders relative to other general obligation debt that has been pledged the same revenue can be either (1) a first lien on all pledged revenue, thereby having priority over general obligation debt that has been pledged the same revenue, (2) a lien that is in parity with general obligation debt that has been pledged the same revenue, or (3) a lien that is subordinate to the lien of general obligation debt that has been pledged the same revenue.

Commercial Paper Commercial paper is also used by municipalities to raise funds on a short-term basis ranging from 1 day to 270 days. There are two types of commercial paper issued, unenhanced and enhanced. Unenhanced commercial paper is a debt obligation issued based solely on the issuer's credit quality and liquidity capability.

Enhanced commercial paper is a debt obligation that is credit enhanced with bank liquidity facilities (e.g., a *letter of credit*), insurance, or a bond purchase agreement. The role of the enhancement is to reduce the risk of nonrepayment of the maturing commercial paper by providing a source of liquidity for payment of that debt in the event no other funds of the issuer are currently available.

Provisions in the 1986 tax act restricted the issuance of tax-exempt commercial paper. Specifically, the act limited the new issuance of municipal obligations that are tax exempt, and as a result, every maturity of a tax-exempt municipal issuance is considered a new debt issuance. Consequently, very limited issuance of tax-exempt commercial paper exists. Instead, issuers use one of the next two products to raise short-term funds.

Variable-Rate Demand Obligations *Variable-Rate Demand Obligations (VRDOs)* are floating-rate obligations that have a nominal long-term maturity but have a coupon rate that is reset either daily or every 7 days. The investor has an option to put the issue back to the trustee at any time with 7 days notice. The put price is par plus accrued interest. There are unenhanced and enhanced VRDOs.

Commercial Paper/VRDO Hybrid The commercial paper/VRDO hybrid is a product that is customized to meet the investor's cash flow needs. There is flexibility in structuring the maturity as with commercial paper because there is a remarketing agent who establishes interest rates for a range of maturities. While there may be a long stated maturity for such issues, they contain a put provision as with a VRDO. The range of the put period can be from 1 day to more than 360 days. On the put date, the investor has two choices. The first is to put the bonds to the issuer; by doing so, the investor receives principal and interest. The second choice available to the investor is to extend the maturity at the new interest rate and put date posted by the remarketing agent at that time.

Revenue Bonds

Revenue bonds are the second basic type of security structure found in the municipal bond market. These bonds are issued for enterprise financings that are secured by the revenues generated by the completed projects themselves, or for general public-purpose financings in which the issuers pledge to the bondholders the tax and revenue resources that were previously part of the general fund. This latter type of revenue bond is usually created to allow issuers to raise debt outside general obligation debt limits and without voter approval.

The trust indenture for a municipal revenue bond details how revenue received by the enterprise will be distributed. This is referred to as the flow-of-funds structure. In a typical revenue bond, the revenue is first distributed into a revenue fund. It is from that fund that disbursements for expenses are made. The typical flow-of-fund structure provides for payments in the following order into other funds: operation and maintenance fund, sinking fund, debt service reserve fund, renewal and replacement fund, reserve maintenance fund, and surplus fund.

Revenue bonds can be classified by the type of financing. These include *utility revenue bonds, transportation revenue bonds, housing revenue bonds, higher education revenue bonds, health care revenue bonds, seaport revenue bonds*, sports complex and convention center revenue bonds, and industrial development revenue bonds. We discuss these revenue bonds as follows. Revenue bonds are also issued by Section 501(c)3 entities (museums and foundations).

Utility Revenue Bonds

Utility revenue bonds include water, sewer, and electric revenue bonds. Water revenue bonds are issued to finance the construction of water treatment plants, pumping stations, collection facilities, and distribution systems. Revenues usually come from connection fees and charges paid by the users of the water systems. Electric utility revenue bonds are secured by revenues produced from electrical operating plants. Some bonds are for a single issuer who constructs and operates power plants and then sells the electricity. Other electric utility revenue bonds are issued by groups of public and private investor-owned utilities for the joint financing of the construction of one or more power plants.

Also included as part of utility revenue bonds are resource recovery revenue bonds. A resource recovery facility converts refuse (solid waste) into commercially saleable energy, recoverable products, and residue to be landfilled. The major revenues securing these bonds usually are (1) fees paid by those who deliver the waste to the facility for disposal, (2) revenues from steam, electricity, or refuse-derived fuel sold to either an electric power company or another energy user, and (3) revenues from the sale of recoverable materials such as aluminum and steel scrap.

Transportation Revenue Bonds

Included in the category of transportation revenue bonds are toll road revenue bonds, highway user tax revenue bonds, airport revenue bonds, and mass transit bonds secured by fare-box revenues. For toll road revenue bonds, bond proceeds are used to build specific revenue-producing facilities such as toll roads, bridges, and tunnels. The pledged revenues are the monies collected through tolls. For highway-user tax revenue bonds, the bondholders are paid by earmarked revenues outside of toll collections, such as gasoline taxes, automobile registration payments, and driver's license fees. The revenues securing airport revenue bonds usually come from either traffic-generated sources—such as landing fees, concession fees, and airline fueling fees—or lease revenues from one or more airlines for the use of a specific facility such as a terminal or hangar. Muller (2008) provides a discussion of how to analyze toll road bonds. The analysis of airport revenue bonds is provided by Oliver and Clements (2008); case studies of airport revenue bonds are provided by Spiotto (2008) and Oliver (2008).

Housing Revenue Bonds

There are two types of housing revenue bonds: single-family mortgage revenue bonds and multifamily housing revenue bonds.

Single-family revenue bonds are issued by state and local housing finance agencies in order to obtain funds to assist low- to middle-income individuals purchase their first home. This assistance is accomplished by using the proceeds from the bond sale to acquire the newly originated mortgages and pooling them. More specifically, the loans are 1-to-4-single-family home, 30-year fixed-rate mortgages. While the primary source of repayment for these bonds are the mortgage payments on the pool of loans, there are several other layers of credit protection. These include (1) overcollateralization of the loan pool (that is, from 102% to as much as 110% of the bonds outstanding), (2) for loans in the pool with a loan-to-value ratio of 80% or greater, primary mortgage insurance is required (either Federal Housing Administration or Veteran's Administration or private mortgage insurance with a rating of at least double A), and (3) the housing finance agency of many states will provide their general obligation pledge. (See Van Kuller [2008a] for more details on these credit enhancements as well as how to analyze single-family bonds.)

As with mortgage-backed securities issued in the taxable sector, investors in single-family mortgage revenue bonds are exposed to prepayment risk. (See Chapter 32 of Volume I.) This is the risk that borrowers in the mortgage pool will prepay their loans when interest rates decline below their loan rate. The disadvantage to the investor is twofold. First, the proceeds received from the prepayments must be reinvested at a lower rate. Second, a property of bonds with prepayment or call options is that their price performance is adversely affected when interest rates decline compared to noncallable bonds.

Multifamily revenue bonds are usually issued for a variety of housing projects involving tenants who qualify as low-income families and senior citizens. There are various forms of credit enhancement for these bonds. Some of these are what is found in commercial mortgage-backed securities where the underling is multifamily housing: overcollateralization, senior-subordinated structure, private and agency mortgage insurance (state insurance for some issues), bank letters of credit, and cross-collateralization and cross default provisions in pools. In addition, there may be credit enhancement in the form of moral obligations or an appropriation obligation of the state or city issuing the bonds. Van Kuller (2008b) explains the structures of multifamily housing revenue bonds and to analyze their credit risk.

Higher Education Revenue Bonds

There are two types of higher education revenue bonds: college and university revenue bonds and student loan revenue bonds. The revenues securing public and private college and university revenue bonds usually include dormitory room rental fees, tuition payments, and sometimes the general assets of the college or university. For student loan revenue bonds, the structures are very similar to what is found in the student loan sector of the taxable asset-backed securities market. For a discussion of how to analyze the credit risk of higher education revenue bonds, see Mincke (2008).

Health Care Revenue Bonds

Health care revenue bonds are issued by private, not-for-profit hospitals (including rehabilitation centers, children's hospitals, and psychiatric institutions) and other health care providers such as health maintenance organizations (HMOs), continuing care retirement communities and nursing homes, cancer centers, university faculty practice plans, and medical specialty practices. The revenue for health care revenue bonds usually depends on federal and state reimbursement programs (such as Medicaid and Medicare), third-party commercial payers (such as Blue Cross, HMOs, and private insurance), and individual patient payments. Cavallaro (2008) explains how to analyze hospital revenue bonds.

Seaport Revenue Bonds

The security for seaport revenue bonds can include specific lease agreements with the benefiting companies or pledged marine terminal and cargo tonnage fees.

Special Bond Structures

Some municipal securities have special security structures. These include *insured bonds*, *bank-backed municipal bonds*, and *refunded bonds*. We describe these three special security structures as follows.

Insured Bonds

Municipal bonds can be credit enhanced by an unconditional guarantee of a commercial insurance company. The insurance cannot be canceled and typically is in place for the term of the bond. The insurance provides for the insurance company writing the policy to make payments to the bondholders of any principal and/or coupon interest that is due on a stated maturity date but that has not been paid by the bond issuer. The insurer's payment is not an advance of the payments due by the issuer but is rather made according to the original repayment schedule obligation of the issuer.

As of 2007, it has been estimated that there were almost $600 billion of insured municipal bonds outstanding and that more than 50% of newly issued municipal bonds were insured (Cirillo, 2008). The track record on municipal bonds is unblemished. Since the first introduction of municipal bond insurance in 1971, no insurer has failed to make payments on any insured municipal bond as of year end 2007. That said, as of early 2008, the major bond insurers faced potential downgrading because of their commitments in the subprime mortgage market.

The insurers of municipal bonds are typically monoline insurance companies that are primarily in the business of providing guarantees. They include the following triple-A-rated monoline insurers as of year end 2007: Ambac Assurance Corp. (AMBAC or Ambac), Assured Guaranty Corp., CIFG Financial Guaranty, Financial Guaranty Insurance Corp. (FGIC), Financial Security Assurance Inc. (FSA), MBIA Insurance Corp. (MBIA), and XL Capital Assurance, Inc. (XL). These are the same insurance companies that provide an insurance wrap for asset-backed securities. There are lower-rated insurers, and they are used by some municipalities when a rating below triple A is sought. These monoline insurers include Radian Asset Insurance, Inc. and ACA Financial Guaranty, double A and single A rated insurers, respectively, as of year end 2007.

Not all bonds in a series issued by a municipality may be covered by insurance. The cover of the official statement must clearly identify which bonds in the series are insured. If there are both insured and non-insured bonds in a series, that must clearly be disclosed in the official statement. In addition, the name of the bond insurer(s) must be clearly shown on the cover of the official statement.

Bonds trading in the secondary market that do not carry insurance can be insured through a personalized insurance policy for a negotiated premium. The insured bond lot only will continue to carry the insurance. An investor can check if a bond lot is insured by contacting the secondary market desk of an insurer. Closed-end funds and unit investment trusts can obtain insurance for a group of bonds. However, once these entities sell the insured bonds, the insurance does not carry over to the new owner.

By obtaining municipal bond insurance, the issuer obviously reduces the credit risk for the investor. Typically, it is bonds issued by smaller governmental units that are not widely known in the financial community, bonds that have a sound though complex and confusing security structure, and bonds issued by infrequent local-government borrowers that do not have a general market following among investors that find it advantageous to obtain municipal bond insurance. Cirillo (2008) provides a more thorough discussion of issuers of insuring bonds.

It should be noted that the credit quality considerations of bond insurers in evaluating whether to insure an issue are more stringent than that used by rating agencies when assigning a rating to an issue. The reason is simple: The bond insurer is making a commitment for the life of the issue. Rating agencies only assign a rating that would be expected to be downgraded in the future if there is credit deterioration of the issuer. Put simply, rating agencies can change a rating but bond insurers cannot change their obligation. Consequently, bond insurers typically insure bonds that would have received an investment-grade rating (at least triple B) in the absence of any insurance.

Bank-Backed Bonds

Municipal issuers have increasingly used various types of facilities provided by commercial banks to credit enhance and thereby improve the marketability of issues. There are three basic types of bank support: letter of credit, irrevocable line of credit, and revolving line of credit.

A *letter of credit (LOC)* is the strongest type of support available from a commercial bank. The parties to a LOC agreement are (1) the bank that issues the LOC (that is, the

LOC issuer), (2) the municipal issuer who is requesting the LOC in connection with a security (the LOC-backed bonds), and (3) the LOC beneficiary who is typically the trustee. The municipal issuer is obligated to reimburse the LOC issuer for any funds it draws down under the agreement.

There are two types of LOCs: direct-pay LOC and standby LOC. With a direct-pay LOC, typically the issuer is entitled to draw upon the LOC in order to make interest and principal payment if a certain event occurs. The LOC beneficiary receives payments from the LOC issuer with the trustee having to request a payment. In contrast, with a standby LOC, the LOC beneficiary typically can only draw down on the agreement if the municipal issuer fails to make interest and/principal payments at the contractual due date. The LOC beneficiary must first request payment from the municipal issuer before drawing upon the LOC. When a LOC is issued by a smaller local bank, there may be a second LOC in place issued by a large national bank. This type of LOC is called a confirming LOC and is drawn upon only if the primary LOC issuer (the smaller local bank) fails to pay a draw request. For a further discussion of LOCs, see Zerega (2008).

An irrevocable line of credit is not a guarantee of the bond issue, though it does provide a level of security. A revolving line of credit is a liquidity-type credit facility that provides a source of liquidity for payment of maturing debt in the event no other funds of the issuer are currently available. Because a bank can cancel a revolving line of credit without notice if the issuer fails to meet certain covenants, bond security depends entirely on the creditworthiness of the municipal issuer.

Refunded Bonds

Municipal bonds are sometimes refunded. An issuer may refund a bond issue for the same reasons that a corporate treasurer may seek to do so: (1) reducing funding costs after taking into account the costs of refunding, (2) eliminating burdensome restrictive covenants, and (3) altering the debt maturity structure for budgetary reasons.

Often, a refunding takes place when the original bond issue is escrowed or collateralized by direct obligations guaranteed by the U.S. government. By this it is meant that a portfolio of securities guaranteed by the U.S. government are placed in a trust. The portfolio of securities is assembled such that the cash flows from the securities match the obligations that the issuer must pay. For example, suppose that a municipality has a 5% \$200 million issue with 15 years remaining to maturity. The bond obligation therefore calls for the issuer to make payments of \$5 million every 6 months for the next 15 years and \$200 million 15 years from now. If the issuer wants to refund this issue, a portfolio of U.S. government obligations can be purchased that has a cash flow that matches that liability structure: \$5 million every 6 months for the next 15 years and \$200 million 15 years from now.

Once this portfolio of securities whose cash flows match those of the municipality's obligation is in place, the refunded bonds are no longer general obligation or revenue bonds. Instead, the issue is supported by the cash flows from the portfolio of securities held in an escrow fund. Such bonds, if escrowed with securities guaranteed by the U.S. government, have little, if any, credit risk and are therefore the safest municipal bonds available.

The escrow fund for a refunded municipal bond can be structured so that the refunded bonds are to be called at the first possible call date or a subsequent call date established in the original bond indenture. Such bonds are known as prerefunded municipal bonds. While refunded bonds are usually retired at their first or subsequent call date, some are structured to match the debt obligation to the retirement date. Such bonds are known as escrowed-to-maturity bonds. For a further discussion of refunded municipal bonds, see Feldstein (2008).

MUNICIPAL BOND YIELDS

Interest rates on municipal bonds reflect not only the risks associated with corporate bonds but also reflect the tax advantage of tax-exempt municipal bonds, including the impact of the AMT and state and local tax treatment. A commonly used yield measure when comparing the yield on a tax-exempt municipal bond with a comparable taxable bond is the equivalent taxable yield and is computed as follows:

$$\text{Equivalent taxable yield} = \text{Tax-exempt yield}/(1 - \text{Effective marginal tax rate})$$

The equivalent taxable yield shows the approximate yield that an investor would have to earn on a taxable bond in order to realize the same yield after taxes.

The *effective marginal tax rate* must take into account both the exemption of interest income from federal income taxes and the effective tax rate applied at the state level if one applies. In computing the effective state marginal tax rate, consideration is given to the deductibility of state taxes for determining federal income taxes. To do so, the following formula can be used to calculate the effective state marginal tax rate:

$$\text{Effective state marginal tax rate} = (1 - \text{Federal marginal tax rate}) \times \text{State marginal tax rate}$$

For example, in 2007 the Pennsylvania tax rate was flat at 3.07% for a taxpayer who does not reside in the city of Philadelphia. Thus, the state marginal tax rate is 3.07%. Assuming an investor is faces a 35% federal marginal tax rate, then the effective state marginal tax rate is

$$(1 - 0.35) \times (0.0307) = 0.019955 \text{ or roughly } 2\%$$

In a state that does not tax municipal interest from either in-state or out-of-state issuers, the state marginal tax rate is obviously zero. In comparing the yield offered on in-state and out-of-state issuers, this adjustment is important.

The federal marginal tax rate in the above formula is the benefit received from being able to deduct state taxes in determining federal income taxes. For investors who do not itemize deductions or whose income is such that state tax deductions have minimal value, the federal marginal tax rate has a benefit is zero and the effective state marginal tax rate is therefore the state marginal tax rate.

The effective marginal tax rate that is used in the formula for the equivalent taxable yield is then the sum of the federal marginal tax rate plus the effective state marginal tax rate. In our example, an investor facing a 35% federal marginal tax rate and an effective state marginal tax rate of 2% would have an effective marginal tax rate of 37%. Suppose, for example, a yield on a municipal bond being considered for acquisition is 6%. Then the equivalent taxable yield is $5\%/(1 - 0.37) = 0.794$ or 7.94%.

A convention in the bond market is to quote yields on municipal bonds relative to some benchmark taxable bond yield such as a comparable maturity Treasury security or as a percentage of the London Interbank Offered Rate (LIBOR) from the swap yield curve. This ratio is referred to as the yield ratio, and it is normally less than 100% because municipal bonds offer a yield that is less than the yield on a comparable taxable bond.

FLOATING-RATE MUNICIPAL SECURITIES

As in the taxable bond market, municipal bonds may have a fixed or floating interest rate. There are two types of floating-rate municipal bonds. The first has the traditional floating-rate formula that calls for the resetting of the issue's coupon rate based on a reference rate plus a quoted margin. The quoted margin is fixed over the life of the bond issue. In the corporate bond market, the reference rate is typically a Treasury rate or some short-term money market rate or swap. In the municipal bond market, the reference rate is typically some percentage of a taxable reference rate (e.g., 75% of six-month LIBOR) or a standard industry reference rate such as the Securities Industry and Financial Markets Association (SIFMA) Municipal Swap Index (formerly The Bond Market Association/PSA Municipal Swap Index). The index, created by the Municipal Market Data, serves as the reference rate in a municipal swap and is calculated weekly.

The other type of floating-rate municipal bond is an inverse floating-rate bond or *inverse floater.* As the name suggests, for an inverse floater the coupon rate changes in the opposite direction of the change in interest rates. That is, if interest rates increase (decrease) since the previous reset of the coupon rate, the new coupon rate decreases (increase). Inverse floaters in the municipal market are created by a sponsor who deposits a fixed-rate municipal security into a trust. The trust then creates two classes of floating-rate securities. The first is a short-term floating-rate security. This floating-rate security can be tendered for redemption at par value on specified dates (typically every week) and are referred to as *tender option bonds (TOBs).* The interest on the TOBs is determined through an auction process that is conducted by a remarketing agent. The second bond class created is the inverse floater. The interest paid to this bond class is the residual interest from the fixed-rate municipal bonds placed in the trust after paying the floating-rate security bondholders and the expenses of the trust. For this reason, the inverse floater is sometimes called the *residual.* When reference rates rise (fall) and the floating-rate security receives a greater (lesser) share of the interest from the fixed-rate municipal security in the trust, the inverse floater investor receives less (more) interest. The holders of the inverse floater have the option to collapse the trust. They can do so by requiring the trustee to pay off the floating-rate securities outstanding and instructing the trustee to give them the fixed-rate securities placed in the trust.

RISKS ASSOCIATED WITH INVESTING IN MUNICIPAL BONDS

Investors in municipal bonds face the typical risk associated with investing in bonds: credit risk, interest rate risk, call risk, and liquidity risk.

Credit risk includes credit default risk, credit spread risk, and downgrade risk. Credit default risk is gauged by the ratings assigned by Moody's, Standard & Poor's, and Fitch. Interest rate risk is typically measured by the duration of a bond: the approximate percentage price change of a bond for a 100-basis-point change in interest rates. Call risk arises for callable bonds and the adverse consequences associated when interest rates decline were mentioned earlier in this chapter. An investor in single-family housing revenue bonds is exposed to a form of call risk, prepayment risk.

There are two risks that are to some extent unique to investors in the municipal bond market. The first is *structure risk.* This is the risk that the security structure may be legally challenged. This may arise in new structures, with the best example being the Washington Public Supply System (WPPS) bonds in the 1980s.

The second risk is *tax risk.* This risk comes in two forms. The first is the risk that the federal income tax rate will be reduced. To understand this risk, note that in the formula for the equivalent taxable yield, the yield is lower the smaller the effective marginal tax rate. A reduction in the effective marginal tax rate therefore reduces the equivalent taxable yield and so that the yield on municipal bonds can stay competitive with taxable bonds, the price of municipal bonds will decline. The second type of tax risk is related to legal risk. The Internal Revenue Service may declare a bond issued as a tax-exempt as taxable. This may be the result of the issuer not complying with IRS regulations. A loss of the tax-exemption feature will cause the municipal bond to decline in value in order to provide a yield comparable to similar taxable bonds.

SUMMARY

Municipal securities are issued by state and local governments as well as authorities created by them. While there are both tax-exempt and taxable municipal securities, the market is dominated by the former. Tax exemption refers to the exemption of interest income from taxation at the federal level. Because of the importance of the tax advantage, investors must be aware of federal income tax provisions affecting municipal securities: treatment of original-issue discount, alternative minimum tax rules, and deductibility of interest in acquiring municipal securities with borrowed funds. The treatment of interest income at the state and local levels varies.

There are basically two types of municipal security structures: tax-backed debt and revenue bonds. There are also municipal bonds with special structures (insured bonds, bank-backed bonds, and refunded bonds).

Because of their tax advantage, yields offered on municipal bonds are normally less than that on comparable taxable bonds. To compare a tax-exempt municipal bond's yield with that of a taxable bond, the equivalent taxable yield can be computed. This yield depends on the investor's effective marginal tax rate. The effective marginal tax rate depends on the investor's federal marginal tax rate and effective state marginal tax rate.

Municipal bonds expose investors to the usual risks associated with bond investing—credit risk, interest rate risk, call risk, and liquidity risk—as well as some unique risks—structure risk and tax risk.

REFERENCES

Cavallaro, L. (2008). Hospital bond analysis. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 845–858), Hoboken, NJ: John Wiley & Sons.

Cirillo, D. K. (2008). How to analyze the municipal bond insurers and the bonds they insure. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 1085–1099), Hoboken, NJ: John Wiley & Sons.

Feldstein, S. G. (2008). How to analyze refunded municipal bonds. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 1035–1038), Hoboken, NJ: John Wiley & Sons.

Ellis, G. (2008). Analysis of tobacco revenue settlement bonds: Assessing cigarette consumption decline estimates. *Journal of Structured Finance* (Winter): 60–79.

Garrett, D. J. (2008). Discovering relative value using custom indices. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 571–575), Hoboken, NJ: John Wiley & Sons.

Lian, G. (2008). How to analyze tobacco bonds. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 957–978), Hoboken, NJ: John Wiley & Sons.

Maco, P. S., and Taffe, J. W. (2008). The Municipal Securities Rulemaking Board. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 383–396), Hoboken, NJ: John Wiley & Sons.

Mincke, B. (2008). How to analyze higher education bonds. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 1055–1075), Hoboken, NJ: John Wiley & Sons.

Muller, R. H. (2008). Toll road analysis. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 981–993), Hoboken, NJ: John Wiley & Sons.

Oliver, W. E. (2008). Aruba Airport Authority airport revenue bonds. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 1127–1133), Hoboken, NJ: John Wiley & Sons.

Oliver, W. E., and Clements, D. (2008). How to analyze airport revenue bonds. In S. G. Feldstein and F. J. Fabozzi (Eds.), *The Handbook of Municipal Bonds* (pp. 813–818), Hoboken, NJ: John Wiley & Sons.

Peng, J., and Brucato, P. Jr. (2001). Do competitive-only laws have an impact on the borrowing cost of municipal bonds? *Municipal Finance Journal* 22: 61–76.

Peng, J., Kriz, K. A., and Neish, T. (2008). In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 51–66), Hoboken, NJ: John Wiley & Sons.

Spiotto, J. E. (2008). Tax-exempt airport finance: Tales from the friendly skies. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 1165–1184), Hoboken, NJ: John Wiley & Sons.

Van Kuller (2008a). Single-family housing bonds. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 861–891), Hoboken, NJ: John Wiley & Sons.

Van Kuller (2008b). Multifamily housing bonds. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 893–921), Hoboken, NJ: John Wiley & Sons.

Zerega, T. (2008). The use of letters-of-credit in connection with municipal securities. In S. G. Feldstein and F. J. Fabozzi (eds.), *The Handbook of Municipal Bonds* (pp. 1015–1023), Hoboken, NJ: John Wiley & Sons.

CHAPTER 23

Corporate Fixed Income Securities

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: To obtain financing, a corporation can rely on several sources of funds. In general, the sources of financing can be classified as equity and debt. Equity financing includes the sale of common stock and preferred stock. Because of the investment attributes of preferred stock, it is classified in the financial market as a fixed income security. Debt can be obtained by either borrowing funds from a bank or issuing debt obligations in the nonbank market. The latter includes corporate bonds, medium-term notes, asset-backed securities, and commercial paper. These debt instruments can be issued in the public market or privately placed. Corporate debt obligations are referred to as fixed income securities.

Keywords: term bonds, bonds, serial bonds, indenture, mortgage debt, open-end mortgage, after-acquired property clause, release and substitution of property clause, first mortgage bonds, first refunding bonds, general and refunding mortgage bonds, collateral trust bonds, equipment trust certificates (ETCs), negative pledge clause, make-whole premium provision, yield-maintenance premium provision, make-whole redemption price, currently callable issue, noncallable, nonrefundable, bullet bonds, sinking-fund provision, balloon maturity, mandatory sinking fund, nonmandatory sinking-fund provision, specific sinking fund, nonspecific sinking fund, accelerated sinking-fund provision, speculative-grade bond, high-yield bond, junk bonds, deferred-interest bonds, step-up bonds, payment-in-kind, extendible reset bond, Trade Reporting and Compliance Engine (TRACE), interdealer platforms, dealer-to-customer platforms, multiple dealer-to-customer platforms, single dealer-to-customer platforms, private-placement market, Rule 144A private placement, traditional private-placement market, non-144A securities, 144A securities, medium-term note (MTN), rate offering schedule, structured notes, reverse inquiry, preferred stock, fixed-rate preferred stock, nonparticipating preferred stock, participating preferred stock, prior preferred stock, preference preferred stock, cumulative preferred stock, noncumulative preferred stock, convertible preferred stock, perpetual preferred stock, adjustable-rate preferred stock, auction preferred stock, remarketed preferred stock, intercorporate tax dividend exclusion

Corporations are classified into five general categories by bond information services:

- Utilities
- Transportations
- Industrials
- Banks
- Finance (nonbanks)

Within these five general categories finer breakdowns are often made to create more homogeneous groupings. For example, utilities are subdivided into electric power companies, gas distribution companies, water companies, and communication companies. Transportations are divided further into airlines, railroads, and trucking companies. Industrials are the catchall class and the most heterogeneous of the groupings with respect to investment characteristics because this category includes all kinds of manufacturing, merchandising, and service companies.

Corporations issue in public markets several types of fixed income securities. These include debt instruments and preferred stock. Debt instruments that are publicly issued include corporate bonds, medium-term notes, asset-backed securities, and commercial paper. In this chapter we will describe the general characteristics of the first two types of debt instruments as well as preferred stock. Asset-backed securities and commercial paper are covered in Chapters 53 and 35 of Volume I respectively. A key investment attribute of corporate securities is their credit risk. In Chapter 24 of Volume III we describe the various aspects of credit risk, the credit ratings assigned to corporate debt obligations and preferred stock, and the factors considered by rating agencies in assigning rating.

CORPORATE BONDS

Most corporate bonds are *term bonds*; that is, they run for a term of years and then become due and payable. Any amount of the liability that has not been paid off prior to maturity must be paid off at that time. The term may be long or short. Generally, obligations due in under 10 years from the date of issue are called *notes.* However, it should be pointed out that the word "notes" has been used to describe particular types of securities that can have maturities considerably longer than 10 years. Most corporate borrowings take the form of *bonds* due in 20 to 30 years. Term bonds may be retired by payment at final maturity or retired prior to maturity if provided for in the indenture. Some corporate bond issues are arranged so that specified principal amounts become due on specified dates. Such issues are called *serial bonds.* Equipment trust certificates (discussed later) are structured as serial bonds.

While the prospectus may provide most of the needed information, the indenture is the more important document. The *indenture* sets forth in great detail the promises of the issuer. Here we will look at what indentures of corporate debt issues contain. For corporate debt securities to be publicly sold, they must (with some permitted exceptions) be issued in conformity with the Trust Indenture Act of 1939. This act requires that debt issues subject to regulation by the Securities and Exchange Commission (SEC) have a trustee. Also, the trustee's duties and powers must be spelled out in the indenture.

Secured Debt and Unsecured Debt

A corporate bond can be secured or unsecured. We describe each type next.

Secured Debt

By secured debt it is meant that some form of collateral is pledged to ensure repayment of the debt.

Utility Mortgage Bonds Debt secured by real property such as plant and equipment is called *mortgage debt.* The largest issuers of mortgage debt are electric utility companies. Most electric utility bond indentures do not limit the total amount of bonds that may be issued. This is called an *open-ended mortgage.* The mortgage generally is a first lien on the company's real estate, fixed property, and franchises, subject to certain exceptions or permitted encumbrances owned at the time of the execution of the indenture or its supplement. The *after-acquired property clause* also subjects to the mortgage property acquired by the company after the filing of the original or supplemental indenture.

To provide for proper maintenance of the property and replacement of worn-out plant, maintenance fund, maintenance and replacement fund, or renewal and replacement fund provisions are placed in indentures. These clauses stipulate that the issuer spend a certain amount of money for these purposes. Depending on the company, the required sums may be around 15% of operating revenues. As defined in other cases, the figure is based on a percentage of the depreciable property or amount of bonds outstanding. These requirements usually can be satisfied by certifying that the specified amount of expenditures has been made for maintenance and repairs to the property or by gross property additions. They can also be satisfied by depositing cash or outstanding mortgage bonds with the trustee; the deposited cash can be used for property additions, repairs, and maintenance or in some cases the redemption of bonds.

Another provision for bondholder security is the *release and substitution of property clause.* If the company releases property from the mortgage lien (such as through a sale of a plant or other property that may have become obsolete or no longer necessary for use in the business), it must substitute other property or cash and securities to be held by the trustee, usually in an amount equal to the released property's fair value. It may use the proceeds or cash held by the trustee to retire outstanding bonded debt. Certainly, a bondholder would not let go of the mortgaged property without substitution of satisfactory new collateral or adjustment in the amount of the debt because the bondholder should want to maintain the value of the security behind the bond. In some cases the company may waive the right to issue additional bonds.

Although the typical electric utility mortgage does not limit the total amount of bonds that may be issued, certain issuance tests or bases usually have to be satisfied before the company can sell more bonds. Bonds may also be issued in exchange or substitution for outstanding bonds, previously retired bonds, and bonds otherwise acquired. A further earnings test found often in utility indentures requires interest charges to be covered by pretax income available for interest charges of at least a specified number of times.

Mortgage bonds go by many different names such as *first mortgage bonds* or *first refunding mortgage bonds.* There are instances when a company might have two or more layers of mortgage debt outstanding with different priorities. This situation usually occurs because the companies cannot issue additional first mortgage debt (or the equivalent) under the existing indentures. Often, this secondary debt level is called *general and refunding mortgage bond*s. In reality, this is mostly second mortgage debt.

Other Mortgage Debt Nonutility companies do not offer much mortgage debt nowadays; the preferred form of debt financing is unsecured. In the past, railroad operating companies were frequent issuers of mortgage debt. In the broad classification of industrial companies, only a few have first mortgage bonds outstanding. While electric utility mortgage bonds generally have a lien on practically all of the company's property, mortgage debt of industrials has more limited liens. Mortgages may also contain maintenance and repair provisions, earnings tests for the issuance of additional debt, release and substitution of property clauses, and limited after-acquired property provisions. In some cases, shares of subsidiaries might also be pledged as part of the lien.

Some mortgage bonds are secured by a lien on a specific property rather than on most of a company's property, as in the case of an electric utility.

Other Secured Debt Debt can be secured by many different assets. For example, a debt issue can be secured by a first-priority lien on substantially all of the issuer's real property, machinery, and equipment, and by a second-priority lien on its inventory, accounts receivables, and intangibles.

Collateral trust debentures, bonds, and notes are secured by financial assets such as cash, receivables, other notes, debentures, or bonds, and not by real property. Collateral trust notes and debentures have been issued by companies engaged in vehicle leasing. Protective covenants for these collateralized issues may include limitations on the equipment debt of subsidiaries, on the consolidated debt of the issuer and its subsidiaries, on dividend payments by the issuer and its subsidiaries, and on the creation of liens and purchase money mortgages, among other things. The eligible collateral is held by a trustee and periodically marked to market to ensure that the market value has a liquidation value in excess of the amount needed to repay the entire outstanding bonds and accrued interest. If the collateral is insufficient, the issuer must, within several days, bring the value of the collateral up to the required amount. If the issuer is unable to do so, the trustee would then sell collateral and redeem bonds. Another collateralized structure allows for the defeasance or "mandatory collateral substitution," which provides the investor assurance that it will continue to receive the same interest payments until maturity. Instead of redeeming the bonds with the proceeds of the collateral sale, the proceeds are used to purchase a portfolio of U.S. government securities in such an amount that the cash flow is sufficient to meet the principal and interest payments on the mortgage-backed bond. Because of the structure of these issues, the rating agencies have assigned their highest rating (triple A) to them. The rating is based on the strength of the collateral and the issues' structure, not on the issuers' credit standing.

Equipment Trust Financing Railroads and airlines have financed much of their rolling stock and aircraft with secured debt. The securities go by various names such as *equipment trust certificates (ETCs),* in the case of railroads, and secured equipment certificates, guaranteed loan certificates, and loan certificates in the case of airlines. We look at railroad equipment trust financing first for two reasons: (1) the financing of railway equipment under the format in general public use today goes back to the late nineteenth century, and (2) it has had a superb record of safety of principal and timely payment of interest, more traditionally known as dividends. Railroads probably comprise the largest and oldest group of issuers of secured equipment financing.

Probably the earliest instance in U.S. financial history in which a company bought equipment under a conditional sales agreement (CSA) was in 1845 when the Schuylkill Navigation Company purchased some barges. Over the years, secured equipment financing proved to be an attractive way for railroads—both good and bad credits—to raise the capital necessary to finance rolling stock. Various types of instruments were devised—equipment bonds (known as the New York Plan), conditional sales agreements (also known as the New York CSA), lease arrangements, and the Philadelphia Plan equipment trust certificate. The New York Plan equipment bond has not been used since the 1930s. The Philadelphia Plan ETC is the form used for most, if not all, public financings in today's market.

The ratings for ETCs are higher than on the same company's mortgage debt or other public debt securities. This is due primarily to the collateral value of the equipment, its superior standing in bankruptcy compared with other claims, and the instrument's generally self-liquidating nature. The railroad's actual credit worthiness may mean less for some equipment trust investors than for investors in other rail securities or, for that matter, other corporate paper. However, that is not to say that financial analysis of the issuer should be ignored.

ETCs are issued under agreement that provide a trust for the benefit of the investors. Each certificate represents an interest in the trust equal to its principal amount and bears the railroad's unconditional guarantee of prompt payment, when due, of the principal and dividends (the term dividends is used because the payments represent income from a trust and not interest on a loan). The trustee holds the title to the equipment, which when the certificates are retired, passes to, or vests in, the railroad, but the railroad has all other ownership rights. It can take the depreciation and can utilize any tax benefits on the subject equipment. The railroad agrees to pay the trustee sufficient rental for the principal payments and the dividends due on the certificates, together with expenses of the trust and certain other charges. The railroad uses the equipment in its normal operations and is required to maintain it in good operating order and repair (at its own expense). If the equipment is destroyed, lost, or becomes worn out or unsuitable for use (that is, suffers a "casualty occurrence"), the company must substitute the fair market value of that equipment in the form of either cash or additional equipment. Cash may be used to acquire additional equipment unless the agreement states otherwise. The trust equipment is usually clearly marked that it is not the railroad's property.

Immediately after the issuance of an ETC, the railroad has an equity interest in the equipment that provides a margin of safety for the investor. Normally, the ETC investor finances no more than 80% of the cost of the equipment and the railroad the remaining 20%. Although modern equipment is longer lived than that of many years ago, the ETC's length of maturity is still generally the standard 15 years (there are some exceptions noted as follows).

The structure of the financing usually provides for periodic retirement of the outstanding certificates. The most common form of ETC is the serial variety. It is usually issued in 15 equal maturities, each one coming due annually in years 1 through 15. There are single-maturity (or "bullet-maturity") ETCs. There are also sinking-fund equipment trust certificates where the ETCs are retired through the operation of a normal sinking fund, one-fifteenth of the original amount issued per year.

The standing of railroad or common carrier ETCs in bankruptcy is of vital importance to the investor. Because the equipment is needed for operations, the bankrupt railroad's management will more than likely reaffirm the lease of the equipment because, without rolling stock, it is out of business. Cases of disaffirmation of equipment obligations are very rare indeed, but if equipment debt were to be disaffirmed, the trustee could repossess and then try to release or sell it to others. Any deficiency due the equipment debtholders would still be an unsecured claim against the bankrupt railway company. Standard-gauge, nonspecialized equipment should not be difficult to release to another railroad.

The Bankruptcy Reform Act of 1978 provides specifically that railroads be reorganized, not liquidated, and subchapter IV of Chapter 11 grants them special treatment and protection. It protects the rights of the equipment lenders while giving the trustee the chance to cure any defaults. Railroad bankruptcies usually do not occur overnight but creep up gradually as the result of steady deterioration over the years. New equipment financing capability becomes restrained. The outstanding equipment debt at the time of bankruptcy often is not substantial and usually has a good equity cushion built in.

Airline equipment debt has some of the special status that is held by railroad equipment trust certificates. Of course, it is much more recent, having developed since the end of World War II. Many airlines have had to resort to secured equipment financing, especially since the early 1970s. Like railroad equipment obligations, certain equipment debt of certified airlines, under Section 1110 of the Bankruptcy Reform Act of 1978, is not subject to Sections 362 and 363 of the Act, namely the automatic stay and the power of the court to prohibit the repossession of the equipment. The creditor must be a lessor, a conditional vendor, or hold a purchase money security interest with respect to the aircraft and related equipment. The secured equipment must be new, not used. Of course, it gives the airline 60 days in which to decide to cancel the lease or debt and to return the equipment to the trustee. If the reorganization trustee decides to reaffirm the lease in order to continue using the equipment, it must perform or assume the debtor's obligations, which become due or payable after that date, and cure all existing defaults other than those resulting solely from the financial condition, bankruptcy, insolvency, or reorganization of the airline. Payments resume including those that were due during the delayed period. Thus, the creditor will get either the payments due according to the terms of the contract or the equipment.

The equipment is an important factor. If the airplanes are of recent vintage, well-maintained, fuel efficient, and relatively economical to operate, it is more likely that a company in distress and seeking to reorganize would assume the equipment lease. However, if the outlook for reorganization appears dim from the outset and the airplanes are older and less economical, the airline could very well disaffirm the lease. In this case, releasing the aircraft or selling it at rents and prices sufficient to continue the original payments and terms to the security holders might be difficult. Of course, the resale market for aircraft is on a plane-by-plane basis and highly subject to supply and demand factors. Multimillion-dollar airplanes have a somewhat more limited market than do boxcars and hopper cars.

The lease agreement required the airline to pay a rental sufficient to cover the interest, amortization of principal, and a return to the equity participant. The airline was responsible for maintaining and operating the aircraft, as well as providing for adequate insurance. It must also keep the equipment registered and record the ETC and lease under the Federal Aviation Act of 1958.

In the event of a loss or destruction of the equipment, the company may substitute similar equipment of equal value and in as good operating condition and repair and as airworthy as that which was lost or destroyed. It also has the option to redeem the outstanding certificates with the insurance proceeds.

An important point to consider is the equity owner. If the airline runs into financial difficulty and fails to make the required payments, the owner may step in and make the rental payment in order to protect its investment. The carrier's failure to make a basic rental payment within the stipulated grace period is an act of default but is cured if the owner makes payment. Thus, a strong owner lends support to the financing, and a weak one little.

An investor should not be misled by the title of the issue just because the words secured or equipment trust appear. Investors should look at the collateral and its estimated value based on the studies of recognized appraisers compared with the amount of equipment debt outstanding. Is the equipment new or used? Do the creditors benefit from Section 1110 of the Bankruptcy Reform Act? Because the equipment is a depreciable item and subject to wear, tear, and obsolescence, a sinking fund starting within several years of the initial offering date should be provided if the debt is not issued in serial form. Of course, the ownership of the aircraft is important as just noted. Obviously, one must review the obligor's financials because the investor's first line of defense depends on the airline's ability to service the lease rental payments.

Unsecured Debt

We have discussed many of the features common to secured debt. Take away the collateral and we have unsecured debt.

Unsecured debt, like secured debt, comes in several different layers or levels of claim against the corporation's assets. But in the case of unsecured debt, the nomenclature attached to the debt issues sounds less substantial. For example, "general and refunding mortgage bonds" may sound more important than "subordinated debentures," even though both are basically second claims on the issuing corporation. In addition to the normal debentures and notes, there are junior issues representing the secondary and tertiary levels of the capital structure. The difference in a high-grade issuer may be considered insignificant as long as the issuer maintains its quality. But in cases of financial distress, the junior issues usually fare worse than the senior issues. Only in cases of very well-protected junior issues will investors come out whole—in which case, so would the holders of senior indebtedness. Thus, many investors are more than willing to take junior debt of high-grade companies; the minor additional risk, compared to that of the senior debt of lower-rated issuers, may well be worth the incremental income.

Credit Enhancements

Some debt issuers have other companies guarantee their debt. This is normally done when a subsidiary issues debt and the investors want the added protection of a third-party guarantee. The use of guarantees makes it easier and more convenient to finance special projects and affiliates, although guarantees are extended to operating company debt.

There are also other types of third-party credit enhancements. Some captive finance subsidiaries of industrial companies enter into agreements requiring them to maintain fixed charge coverage at such a level so that the securities meet the eligibility standards for investment by insurance companies under New York State law. The required coverage levels are maintained by adjusting the prices at which the finance company buys its receivables from the parent company or through special payments from the parent company. These supplemental income maintenance agreements, while usually not part of indentures, are very important considerations for bond buyers.

Another type of support can call for an agreement between the company and its parent that stipulates that the parent (1) agrees to cause the subsidiary to maintain a positive tangible net worth in accordance with generally accepted accounting principles; (2) will provide the necessary funds to pay debt service if the subsidiary is unable to meet the obligations when due; and (3) shall own, directly or indirectly, all of the outstanding voting capital stock of the subsidiary throughout the life of the support agreement. In addition, in case of a default by the parent in meeting its obligations under the default agreement, or in the case of default by the subsidiary in the payment of principal and/or interest, the holders of the securities or the trustee may proceed directly against the parent.

Another credit-enhancing feature is the letter of credit (LOC) issued by a bank. An LOC requires the bank to make payments to the trustee when requested so that monies will be available for the bond issuer to meet its interest and principal payments when due. Thus, the credit of the bank under the LOC is substituted for that of the debt issuer.

Insurance companies also lend their credit standing to corporate debt, both new issues and outstanding secondary market issues. While a guarantee or other type of credit enhancement may add some measure of protection to a debtholder, an analysis of both the issuer and the guarantor should be performed. In many cases, only the latter is needed if the issuer is merely a financing conduit without any operations of its own. However, if both concerns are operating companies, it may very well be necessary to analyze both because the timely payment of principal and interest ultimately will depend on the stronger party. A downgrade of the enhancer's claims-paying ability reduces the value of the bonds.

Negative Pledge Clause

One of the important protective provisions for unsecured debtholders is the *negative pledge clause.* This provision, found in most senior unsecured debt issues and a few subordinated issues, prohibits a company from creating or assuming any lien to secure a debt issue without equally securing the subject debt issue(s) (with certain exceptions). Its inclusion in the indenture is designed to prevent other creditors from obtaining a senior position at the expense of existing creditors; however, it is not intended to prevent other creditors from sharing in the position of debenture holders. It is not necessary to have such a clause unless the issuer runs into trouble. But like insurance, it is not needed until the time that no one wants arrives.

Provisions for Paying Off Bonds

There are provisions that may result in all or a portion of a bond issue being paid off prior to the stated maturity date. These include (1) call and refund provisions and (2) sinking-fund provisions. We describe both below.

Call and Refund Provisions

An important question in negotiating the terms of a new bond issue is whether the issuer shall have the right to redeem the entire amount of bonds outstanding on a date before maturity. Issuers generally want this right because they recognize that at some time in the future the general level of interest rates may fall sufficiently below the issue's coupon rate that redeeming the issue and replacing it with another issue with a lower coupon rate would be attractive. This right is a disadvantage to the bondholder.

A company wanting to retire a debt issue prior to maturity usually must pay a premium over the par value for the privilege. The initial call premium on long-term debt traditionally has been the interest coupon plus par or the initial reoffering price (in some cases it is the higher of the two). Thus, a 30-year bond initially priced at 100 with a 7% coupon may have a call price of 107% for the first year, scaled down in relatively equal amounts to par starting in year 21 to maturity.

Instead of a specified fixed premium that must be paid by the issuer if the bond is called, a bond may have a *make-whole premium provision*, also called a *yield-maintenance premium provision.* The provision specifies a formula for determining the premium that the issuer must pay to call an issue and is such that the amount of the premium, when added to the principal amount and reinvested at the redemption date in U.S. Treasury securities having the same remaining life, would provide a yield equal to the original yield. The premium plus the principal at which the issue is called is referred to as the *make-whole redemption price.* The purpose of the make-whole premium is to protect the yield of those investors who purchased the issue at issuance.

If a bond issue does not have any protection against early call, it is said to be a *currently callable issue.* But most new bond issues, even if currently callable, usually have some restrictions against certain types of early redemption. The most common restriction is that prohibiting the refunding of the bonds for a certain number of years. Bonds that are noncallable for the issue's life are more common than bonds that are nonrefundable for life but otherwise callable.

Bonds are sometimes referred to as *noncallable* and *nonrefundable* with the terms used interchangeably. However, technically they have different meanings. Call protection is much more absolute than refunding protection. Although there may be certain exceptions to absolute or complete call protection in some cases (such as sinking funds and the redemption of debt under certain mandatory provisions), it still provides greater assurance against premature and unwanted redemption than does refunding protection. Refunding prohibition merely prevents redemption only from certain sources, namely the proceeds of other debt issues sold at a lower cost of money. The holder is protected only if interest rates decline, and the borrower can obtain lower-cost money to pay off the debt.

A number of industrial companies issued long-term debt with extended call protection, not refunding protection. A number are noncallable for the issue's life. For such issues the prospectus expressly prohibits redemption prior to maturity. These noncallable-for-life issues are referred to as *bullet bonds.*

Bonds can be called in whole (the entire issue) or in part (only a portion). When less than the entire issue is called, the specific bonds to be called are selected randomly or on a pro rata basis.

Sinking-Fund Provision

The indenture may include a *sinking-fund provision.* This provision allows for a debt's periodic retirement or amortization over its life span. This provision for repayment of corporate debt may be designed to liquidate all of a bond issue by the maturity date, or it may be arranged to pay only a part of the total by the end of the term. If only a part is paid, the remainder is called a *balloon maturity.* The purpose of the sinking-fund provision is to reduce credit risk.

A variety of sinking-fund types are found in publicly issued corporate debt. The most common is the *mandatory sinking fund,* requiring the periodic redemption of a certain amount of a specific debt issue. A mandatory sinking fund specifies that the issuer may satisfy the provision in whole or in part, by (1) delivering bonds acquired through open-market purchases or other means or (2) paying cash to the trustee who will call bonds for redemption at 100. This type is found in most longer-term industrial issues and some electric utility bonds.

Another type of sinking-fund provision that is most prevalent in electric utility company issues is the *nonmandatory sinking-fund provision.* This provision allows the issuer to satisfy the sinking-fund provision by the utilization of unfunded property additions or improvements at a certain percentage of their cost. This third alternative is referred to as a property credit. Property credits so utilized cannot be further employed under the mortgage.

A corporate sinking-fund provision may be a specific sinking-fund provision or a nonspecific sinking-fund provision. A *specific sinking fund* applies to just the named issue. A *nonspecific sinking fund,* also known as a funnel, tunnel, blanket, or aggregate sinking fund, is based on the outstanding amount of a company's total bonded indebtedness. In most cases, the redemption price for bonds called under the funnel sinking fund is par.

Usually, the periodic payments required for sinking-fund purposes will be the same for each period. A few indentures might permit variable periodic payments, where payments change according to certain prescribed conditions set forth in the indenture. Many corporate bond indentures include a provision that grants the issuer the option to retire more than the amount stipulated for sinking-fund retirement. This is referred to as an *accelerated sinking-fund provision.*

Usually, the sinking-fund call price is the par value if the bonds were originally sold at par. When issued at a price in excess of par, the call price generally starts at the issuance price and scales down to par as the issue approaches maturity.

Speculative-Grade Bonds

Speculative-grade bonds are those rated below investment grade by the rating agencies (that is, BBB− and lower by Standard & Poor's and Fitch Ratings and Baa3 and lower by Moody's). They may also be unrated, but not all unrated debt is speculative. They are also known as *high-yield bonds* and *junk bonds*.

Types of Issuers

Several types of issuers fall into the less-than-investment-grade high-yield category. These include:

- Original issuers
- Fallen angels
- Restructuring and leveraged buyouts

Original issuers may be young, growing corporations lacking the stronger balance sheet and income statement profile of many established corporations, but often with lots of promise. Also called venture capital situations or growth or emerging market companies, the debt is often sold with a story projecting future financial strength. From this we get the term "story bond." There are also the established operating firms with financials neither measuring up to the strengths of investment-grade corporations nor possessing the weaknesses of companies on the verge of bankruptcy. Subordinated debt of investment-grade issuers may be included here. A bond rated at the bottom rung of the investment-grade category (Baa and BBB) or at the top end of the speculative-grade category (Ba and BB) is known as a "businessman's risk."

Fallen angels are formerly companies with investment-grade-rated debt that have come upon hard times with deteriorating balance sheet and income statement financial parameters. (Companies that have been upgraded to investment-grade status are referred to as rising stars.) They may be in default or near bankruptcy. In these cases, investors are interested in the workout value of the debt in a reorganization or liquidation, whether within or without the bankruptcy courts. Some refer to these issues as "special situations." Over the years they have fallen on hard times; some have recovered and others have not.

General Motors Corporation and Ford Motor Company are examples of fallen angels. From 1954 to 1981, General Motors Corp. was rated AAA by S&P; Ford Motor Co. was rated AA by S&P from 1971 to 1980. In August 2005, Moody's lowered the rating on both automakers to junk bond status.

Restructurings and leveraged buyouts are companies that have deliberately increased their debt burden with a view toward maximizing shareholder value. The shareholders may be the existing public group to which the company pays a special extraordinary dividend, with the funds coming from borrowings and the sale of assets. Cash is paid out, net worth decreased and leverage increased, and ratings drop on existing debt. Newly issued debt gets junk bond status because of the company's weakened financial condition.

In a leveraged buyout (LBO), a new and private shareholder group owns and manages the company. The debt issue's purpose may be to retire other debt from commercial and investment banks and institutional investors incurred to finance the LBO. The debt to be retired is called bridge financing because it provides a bridge between the initial LBO activity and the more permanent financing.

Unique Features of Some Issues

Often actions that are taken by management that result in the assignment of a non-investment-grade bond rating result in a heavy corporate interest payment burden. This places severe cash flow constraints on the firm. To reduce this burden, firms involved with heavy debt burdens have issued bonds with deferred coupon structures that permit the issuer to avoid using cash to make interest payments for a period of 3 to 7 years. There are three types of deferred coupon structures:

- Deferred-interest bonds
- Step-up bonds
- Payment-in-kind bonds

Deferred-interest bonds are the most common type of deferred coupon structure. These bonds sell at a deep discount and do not pay interest for an initial period, typically from 3 to 7 years. (Because no interest is paid for the initial period, these bonds are sometimes referred to as zero-coupon bonds.) *Step-up bonds* do pay coupon interest, but the coupon rate is low for an initial period and then increases ("steps up") to a higher coupon rate. Finally, *payment-in-kind* (PIK) bonds give the issuer an option to pay cash at a coupon payment date or give the bondholder a similar bond (that is, a bond with the same coupon rate and a par value equal to the amount of the coupon payment that would have been paid). The period during which the issuer can make this choice varies from 5 to 10 years.

An *extendible reset bond* structure allows the issuer to reset the coupon rate so that the bond will trade at a predetermined price. The coupon rate may reset annually or even more frequently, or reset only one time over the life of the bond. Generally, the coupon rate at the reset date will be the average of rates suggested by two investment banking firms. The new rate will then reflect (1) the level of interest rates at the reset date and (2) the credit spread the market wants on the issue at the reset date. Notice the difference between an extendible reset bond and a floating-rate issue. In a floating-rate issue, the coupon rate resets according to a fixed spread over the reference rate, with the index spread specified in the indenture. The amount of the index spread reflects market conditions at the time the issue is offered. The coupon rate on an extendible reset bond, in contrast, is reset based on market conditions (as suggested by several investment banking firms) at the time of the reset date. Moreover, the new coupon rate reflects the new level of interest rates and the new spread that investors seek. The advantage to investors of extendible reset bonds is that the coupon rate will reset to the market rate—both the level of interest rates and the credit spread—in principle keeping the issue at par value.

Secondary Market

Historically, the trading of corporate bond trading is done in the over-the-counter (OTC) market conducted via telephone and based on broker-dealer trading desks. In this market, broker-dealer trading desks take principal positions in corporate bonds in order to fulfill buy and sell orders of their customers. There has been a transition away from this traditional form of bond trading and toward electronic trading.

In 2002 the National Association of Securities Dealers (NASD) instituted a mandatory reporting of OTC secondary market transactions for corporate bonds that met specific criteria. The reporting system, the *Trade Reporting and Compliance Engine (TRACE),* requires that all NASD broker/dealers report transactions in corporate bonds to TRACE. At the end of each trading day, market aggregate statistics are published on corporate bond market activity. End of day recap information provided includes (1) the number of securities and total par amount traded, (2) advances, declines, and 52-week highs and lows, and (3) the 10 most active investment-grade, high-yield, and convertible bonds for the day.

Electronic Trading of Corporate Bonds

There are four major advantages of electronic trading over traditional corporate bond trading in the OTC market (see Jones and Fabozzi, 2005):

- Providing liquidity to the markets
- Price discovery (particularly for less liquid markets)
- Use of new technologies
- Trading and portfolio management efficiencies

As an example of the last advantage, a portfolio manager can load buy/sell orders on a web site, trade from these orders, and then clear these orders.

In its 2006 survey of electronic trading transaction systems, the Securities Industry and Financial Markets Association (SIFMA) found that there was a rapid increase in the adoption of electronic execution not only in the United States but globally. The SIFMA categorizes electronic trading systems based on (1) who the participants are and the way in which they conduct trades with each other and (2) the methodology or technology employed by participants for price discovery and trade execution.

The first classification includes two types of platforms:

- Interdealer platforms
- Dealer-to-customer platforms

Interdealer platforms allow dealers to execute transactions electronically with other dealers via the anonymous services of "brokers' brokers." The customers of dealers are not involved in interdealer systems.

Dealer-to-customer platforms support trading between customers and broker-dealers. There are two types of dealer-to-customer platforms. *Multiple dealer-to-customer platforms* typically display to customers the best bid or offer price of those posted by all dealers. The participating dealer usually acts as the principal in the transaction. *Single dealer-to-customer platforms* permit investors to execute transactions directly with the specific dealer desired.

In addition to electronic trading platforms to support trading in the secondary market just described, there are *new issue platforms* that support the sales of newly issued corporate bonds to either institutional investors or broker-dealers or both.

The systems used for price discovery and trade execution include:

- Request-for-quotes systems
- Order-driven systems
- Market-making or crossing matching systems
- Auction systems

Request-for-quotes systems permit buy-side customers to request executable quotes from broker-dealers with whom they have a customer relationship. These systems are used in multiple dealer-to-customer platforms. In an order-driven system, a participant can enter quotations into central order book.

Private-Placement Market for Corporate Bonds

Securities privately placed are exempt from registration with the SEC because they are issued in transactions that do not involve a public offering. The *private-placement market* has undergone a major change since the adoption of SEC Rule 144A in 1990, which allows the trading of privately placed securities among qualified institutional buyers.

Not all private placements are *Rule 144A private placement.* Consequently, the private-placement market can be divided into two sectors. First is the *traditional private-placement market,* which includes *non-144A securities.* Second is the market for *144A securities.*

Rule 144A private placements are now underwritten by investment bankers on a firm commitment basis, just as with publicly issued bonds. The features in these issues are similar to those of publicly issued bonds. For example, the restrictions imposed on the borrower are less onerous than for traditional private-placement issues. For underwritten issues, the size of the offering is comparable to that of publicly offered bonds.

Unlike publicly issued bonds, the issuers of privately placed issues tend to be less well known. In this way, the private-placement market shares a common characteristic with the bank loan market that we will discuss later in this chapter. Borrowers in the publicly issued bond market are typically large corporations. Issuers of privately placed bonds tend to be medium-sized corporations. Those corporations that borrow from banks tend to be small corporations.

Although the liquidity of issues has increased since Rule 144A became effective, it is still not comparable to that of publicly offered issues. Yields on privately placed debt issues are still higher than those on publicly offered bonds. However, one market observer reports that the premium that must be paid by borrowers in the private placement market has decreased as investment banking firms have committed capital and trading personnel to making markets for securities issued under Rule 144A.

MEDIUM-TERM NOTES

A *medium-term note* (MTN) is a corporate debt instrument, with the unique characteristic that notes are offered continuously to investors by an agent of the issuer. Investors can select from several maturity ranges: 9 months to 1 year, more than 1 year to 18 months, more than 18 months to 2 years, and so on up to 30 years. Medium-term notes are registered with the SEC under Rule 415 (the shelf registration rule), which gives a corporation the maximum flexibility for issuing securities on a continuous basis.

The term "medium-term note" to describe this corporate debt instrument is misleading. Traditionally, the term "note" or "medium-term note" was used to refer to debt issues with a maturity greater than 1 year but less than 15 years. Certainly, this is not a characteristic of MTNs because they have been sold with maturities from 9 months to 30 years and even longer. For example, in July 1993, Walt Disney Corporation issued a security with a 100-year maturity off its MTN shelf registration. General Motors Acceptance Corporation first used MTNs in 1972 to fund automobile loans with maturities of five years and less. The purpose of the MTN was to fill the funding gap between commercial paper and long-term bonds. It is for this reason that they are referred to as "medium term." MTNs were issued directly to investors without the use of an agent.

The modern-day MTN was pioneered by Merrill Lynch in 1981. The first MTN issuer was Ford Motor Credit Company. By 1983, GMAC and Chrysler Financial used Merrill Lynch as an agent to issue MTNs. Merrill Lynch and other investment banking firms committed funds to make a secondary market for MTNs, thereby improving liquidity. In 1982, Rule 415 was adopted, making it easier for issuers to sell registered securities on a continuous basis.

Borrowers have flexibility in designing MTNs to satisfy their own needs. They can issue fixed- or floating-rate debt. The coupon payments can be denominated in U.S. dollars or in a foreign currency.

Primary Market

MTN differ from corporate bonds in the manner in which they are distributed to investors when they are initially sold. Although some investment-grade corporate bond issues are sold on a best-efforts basis, typically they are underwritten by investment bankers. Traditionally, MTNs have been distributed on a best-efforts basis by either an investment banking firm or other broker/dealers acting as agents. Another difference between corporate bonds and MTNs when they are offered is that MTNs are usually sold in relatively small amounts on a continuous or an intermittent basis, whereas corporate bonds are sold in large, discrete offerings.

A corporation that wants an MTN program will file a shell registration with the SEC for the offering of securities. Although the SEC registration for MTN offerings is between $100 and $1 billion, after the total is sold, the issuer can file another shelf registration. The registration will include a list of the investment banking firms, usually two to four, that the corporation has arranged to act as agents to distribute the MTNs.

The issuer then posts rates over a range of maturities: for example, 9 months to 1 year, 1 year to 18 months, 18 months to 2 years, and annually thereafter. This is called the *rate offering schedule.* Usually, an issuer will post rates as a spread over a Treasury security of comparable maturity. Rates are not posted for maturity ranges that the issuer does not desire to sell.

The agents will then make the offering rate schedule available to their investor base interested in MTNs. An investor who is interested in the offering will contact the agent. In turn, the agent contacts the issuer to confirm the terms of the transaction. Because the maturity range in the offering rate schedule does not specify a specific maturity date, the investor can choose the final maturity subject to approval by the issuer. The minimum size that an investor can purchase of an MTN offering typically ranges from $1 million to $25 million.

The rate offering schedule can be changed at any time by the issuer either in response to changing market conditions or because the issuer has raised the desired amount of funds at a given maturity. In the latter case, the issuer can either not post a rate for that maturity range or lower the rate.

Structured MTNs

Some issues of MTNs are coupled with transactions in the derivative markets (options, futures/forwards, swaps, caps, and floors) in order to create debt obligations with more risk-return features unavailable in the corporate bond market. Specifically, an issue can be floating-rate over all or part of the life of the security, and the coupon reset formula can be based on a benchmark interest rate, equity index or individual stock price, a foreign exchange rate, or a commodity index. Inverse floaters (that is, floaters whose coupon moves in the opposite direction of the change of a reference interest rate) are created in the structured MTN market. MTNs can have various embedded options included.

MTNs created when the issuer simultaneously transacts in the derivative markets are called *structured notes.* The most common derivative instrument used in creating structured notes is a swap. By using the derivative markets in combination with an offering, borrowers are able to create investment vehicles that are more customized for institutional investors to satisfy their investment objectives. Moreover, it allows institutional investors who are restricted to investing in investment-grade debt issues the opportunity to participate in other asset classes to make a market play. For example, an investor who buys an MTN whose coupon rate is tied to the performance of the S&P 500 is participating in the equity market without owning common stock. If the coupon rate is tied to a foreign stock index, the investor is participating in the equity market of a foreign country without owning foreign common stock. In exchange for creating a structured note, borrowers can reduce their funding costs.

In a typical offering of a corporate bond, the sales force of the underwriting firm will solicit interest in the offering from its customer base. That is, the sales force will make an inquiry. In the structured note market, the process is often quite different. Because of the small size of an offering and the flexibility to customize the offering in the swap market, investors can approach an issuer through its agent about designing a security for their needs. This process of customers inquiring of issuers or their agents to design a security is called a *reverse inquiry.* Transactions that originate from reverse inquiries account for a significant share of MTN transactions.

PREFERRED STOCK

Unlike corporate bond and MTNS, *preferred stock* is a class of stock. It is classified on the balance sheet as equity. An investor in preferred stockholder is entitled to dividends just like the investor in common stock. However, unlike common stock, there is a specified dividend rate. The dividend amount is the product of the dividend rate and the par value of the preferred stock. The dividend rate can be fixed or it can be a floating rate. A preferred stock issue in which the dividend rate is fixed is referred to as *fixed-rate preferred stock.* There are different types of preferred stock where the dividend rate floats that vary as to how the dividend rate is determined. We'll discuss these types below.

While there are occasionally exceptions, preferred stock limits the investor to the dividend amount as specified by the dividend rate. That is, the investor can earn no more than this amount in the form of dividends Thus, most preferred stock is *nonparticipating preferred stock.* Historically, there have been issues entitling the investor in preferred stock to participate in earnings distribution beyond the specified amount (based on some formula). Preferred stock with this feature is referred to as *participating preferred stock.*

It is because most preferred stock is of the nonparticipating variety that we classify preferred stock as a fixed income security. Thus, we can see that not all fixed income securities are debt obligations.

Dividend payments to preferred stockholders have priority over the payment to common stockholders but are paid after debt holders. A company usually has outstanding several preferred stock issues. In such cases, one of the issues is typically designated as having priority in the case of dividends payments over the others and is called *prior preferred stock.* The other preferred stock issues are called *preference preferred stock.* Hence, prior preferred stock has less risk than preference preferred stock and therefore sells for a lower yield in the market.

If the issuer fails to make a preferred stock dividend payment, the preferred stockholders cannot force the issuer into bankruptcy. This is an attribute that preferred stock shares with common stock. When a preferred stock dividend payment is missed, the treatment of the unpaid dividend depends on whether the preferred stock is cumulative preferred stock or noncumulative preferred stock. With *cumulative preferred stock,* the dividend payment accrues until it is fully paid. Preferred stock of this variety shares this feature with a debt obligation. In the case of *noncumulative preferred stock,* the dividend payment is lost and is no longer the obligation of the issuer, as is the case with common stock. Regardless if the issue is cumulative or noncumulative, the failure to make dividend payments may result in preferred stockholders being given temporary voting rights and in the imposition of certain restrictions on certain activities of management.

In the liquidation of a corporation, the distribution of corporate assets to preferred stockholders comes after all debt holders are paid off. Preferred stockholders, as well as debt holders, can only recover up to their par value. Preferred stockholders are preferred to common stockholders in the distribution of corporate assets in a liquidation. As noted earlier, there is usually prior preferred stock and preference preferred stock in a corporation's capital structure. Not only does the former have priority over the latter with respect to dividend payments, but also in the case of a liquidation. Because preferred stock exposes an investor to credit risk, they are rated by the rating agencies.

Almost all preferred stock has a sinking-fund provision. This is the same feature that we described for corporate bonds. Also, as with corporate bonds which may have a conversion feature, a preferred stock may have a conversion feature that allows the investor to convert shares into common stock. Issues with this feature are called *convertible preferred stock.*

Preferred stock may be issued without a maturity date. This type of preferred stock is called *perpetual preferred stock.* There are putable and callable preferred stock issues.

As noted earlier, there are different types of preferred stock that have a floating or adjustable dividend rate. They include adjustable-rate preferred stock, auction preferred stock, and remarketed preferred stock. For *adjustable-rate preferred stock,* the rate is determined by a formula. For *auction preferred stock,* the dividend rate is reset based on the results of an auction. Participants in the auction consist of current holders and potential buyers. The dividend rate that participants are willing to accept reflects current market conditions. In the case of *remarketed preferred stock,* the dividend rate is determined periodically by a remarketing agent, who resets the dividend rate so that any preferred stock can be tendered at par and be resold (remarketed) at the original offering price.

Tax Treatment of Dividends

Payments made to preferred stockholders are treated as a distribution of earnings. Hence, unlike interest payments that are treated as business expenses by a corporation and therefore tax deductible in determining earnings, preferred stock dividend payments are not. While this raises the after-tax cost of funds if a corporation issues preferred stock rather than issuing debt or borrowers via bank loans, there is a provision in the tax code that makes the holding of preferred stock more appealing to corporate treasurers of other corporations and thereby allows a corporation to issue preferred stock at a reduced cost. This provision is the *intercorporate tax dividend exclusion* which exempts 80% of qualified dividends from federal income taxation if the recipient is a qualified corporation. For example, if Corporation A owns the preferred stock of Corporation B, for each \$1 million of dividends received by A, only \$200,000 will be taxed at A's marginal tax rate. The purpose of this provision is to mitigate the effect of double taxation of corporate earnings. This tax provision is the chief reason that the major buyers of preferred stock are corporations who are seeking tax-advantaged investments.

SUMMARY

In this chapter we looked at three types of corporate fixed income securities: corporate bonds, medium-term notes, and preferred stock.

Corporate bonds can represent either secured debt or unsecured debt. Call, refunding, and sinking-fund provisions that may be included in a corporate bond issue allow the issuer to prepay all or a portion of a bond issue prior to the stated maturity date. The corporate bond market can be broken into the investment-grade market and the speculative-grade market, the latter commonly referred to as the high-yield or junk bond market. Unique features of some high-yield bond issues are deferred interest bonds, step-up bonds, and payment-in-kind bonds. The Trade Reporting and Compliance Engine (TRACE) is the NASD trading system for corporate bonds. There are several types of electronic trading systems for corporate bonds. Corporate bonds can be sold through a public offering or placed privately.

Medium-term notes are offered continuously to investors by an agent of the issuer via a rate offering schedule. Investors can select from several maturity ranges. Structured notes are MTNs created when the issuer simultaneously transacts in the derivative markets. The process typically involves a reverse inquiry.

Preferred stock is a form of equity that shares characteristics of both common stock and corporate debt. From an investor perspective, because the dividends and the distribution upon liquidation are limited, preferred stock is classified as a fixed income security. While dividends payments to preferred stockholders are not tax deductible for a corporation, the intercorporate tax dividend exclusion makes investing in preferred stock by corporate treasurers appealing.

REFERENCES

Crabbe, L. E. (2005). Medium-term notes. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Securities*, 7th edition (pp. 339–350). New York: McGraw-Hill.

Crabbe, L. E., and Fabozzi, F. J. (2002). *Managing a Corporate Bond Portfolio*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (2002). *Fixed Income Securities*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (2008). *Bond Markets, Analysis, and Strategies*, 7th edition. Upper Saddle River, NJ: Prentice Hall.

Fabozzi, F. J., and Mann, S. V. (2005). Nonconvertible preferred stock. In F.J. Fabozzi (ed.), *The Handbook of Fixed Income Securities*, 7th edition (pp. 385–394). New York: McGraw-Hill.

Fabozzi, F. J., Mann, S. V., and Wilson, R. S. (2005). Corporate bonds. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Securities*, 7th edition (pp. 305–336). New York: McGraw-Hill.

Jones, F. J., and Fabozzi, F. J. (2005). The primary and secondary bond markets. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Securities*, 7th edition (pp. 31–51). New York: McGraw-Hill.

Wilson, R. S., and Fabozzi, F. J. (1995). *Corporate Bonds: Structures and Analysis*. Hoboken, NJ: John Wiley & Sons.

CHAPTER 24

The Eurobond Market

MOORAD CHOUDHRY, PhD
Head of Treasury, KBC Financial Products, London

Abstract: The integration and globalization of the world's capital markets has been most evident in the Eurobond market. It is an important source of funds for many banks and corporates, as well as sovereign governments. The Eurobond market has benefited from much of the recent advances in financial engineering, and has undergone innovative changes since its inception in the 1960s. It continues to develop new structures, in response to the varying demands and requirements of specific groups of investors. The range of innovations have customized the market to a certain extent and often the market is the only opening for certain types of government and corporate finance. Investors also often look to the Eurobond market due to constraints in their domestic market, and Euro securities have been designed to reproduce the features of instruments that certain investors may be prohibited from investing in their domestic arena. Other instruments are designed for investors in order to provide tax advantages. The key feature of Eurobonds is the way they are issued, internationally across borders

and by an international underwriting syndicate. The method of issuing Eurobonds reflects the cross-border nature of the transaction, and unlike government markets where the auction is the primary issue method, Eurobonds are typically issued under a "fixed price reoffer" method or a "bought deal." There is also a regulatory distinction, as no one central authority is responsible for regulating the market and overseeing its structure.

Keywords: bought deal, Clearstream, covenant, Eurobonds, Euroclear, fixed price reoffer scheme, global note, gray market, lead manager, primary market, secondary market, syndicate, trustee, underwriters

This chapter reviews the Eurobond market in terms of the structure of the market, the nature of the instruments themselves, the market players, the issuing process and technical aspects such as taxation and swap arrangements. We also review the secondary market.

EUROBONDS

A *Eurobond* is a debt capital market instrument issued in a "Eurocurrency" through a *syndicate* of issuing banks and securities houses, and distributed internationally when issued, that is sold in more than one country of issue and subsequently traded by market participants in several international financial centers. The Eurobond market is divided into sectors depending on the currency in which the issue is denominated. For example, U.S. dollar Eurobonds are often referred to as Eurodollar bonds, similar sterling issues are called Eurosterling bonds. The prefix "Euro" was first used to refer to deposits of U.S. dollars in continental Europe in the 1960s. The Euro-deposit now refers to any deposit of a currency outside the country of issue of that currency, and is not limited to Europe. For example, a deposit of Singapore dollars in the Dubai branch of Citigroup is a Euro-deposit. For historical reasons and also due to the importance of the U.S. economy and investor base, the major currency in which Eurobonds are denominated has always been U.S. dollars.

The first ever Eurobond is generally considered to be the issue of \$15 million nominal of 10-year $5\frac{1}{2}\%$ bonds by Autostrada, the Italian state highway authority, in July 1963. (Decovny [1998, p. 68] states that the first Eurobond issue was in 1957, but its identity is not apparent.) The bonds were denominated in U.S. dollars and paid an annual coupon in July each year. This coincides with the imposition in the United States of the interest equalization tax, a withholding tax on domestic corporate bonds, which is often quoted as being a prime reason behind the establishment of overseas deposits of U.S. dollars.

FOREIGN BONDS

At this stage it is important to identify "foreign bonds" and distinguish them from Eurobonds. Foreign bonds are debt capital market instruments that are issued by foreign borrowers in the domestic bond market of another country. As such, they trade in a similar fashion to the bond instruments of the domestic market in which they are issued. They are usually underwritten by a single bank or a syndicate of domestic banks, and are denominated in the currency of the market in which they are issued. For those familiar with the sterling markets the best example of a foreign bond is a Bulldog bond, which is a sterling bond issued in the United Kingdom by a non-U.K. domiciled borrower. Other examples are Yankee bonds in the United States, Samurai bonds in Japan, Rembrandt bonds in the Netherlands, Matador bonds in Spain, and so on. Hence, a U.S. company issuing a bond in the United Kingdom, denominated in sterling and underwritten by a domestic bank, would be issuing a Bulldog bond, which would trade as a gilt (a U.K. government security), except with an element of credit risk attached. In today's integrated global markets, however, the distinction is becoming more and more fine. Many foreign bonds pay gross coupons and are issued by a syndicate of international banks, so the difference between them and Eurobond may be completely eroded in the near future.

The most important domestic market for foreign bond issues has been the U.S. dollar market, followed by euros, Swiss francs, and Japanese yen. There are also important markets in Canadian, New Zealand, and Australian dollars, and minor markets in currencies such as Hong Kong dollars, Kuwaiti dinars, and Saudi Arabian riyals.

EUROBOND INSTRUMENTS

There is a wide range of instruments issued in the Eurobond market, designed to meet the needs of borrowers and investors. We review the main types in this section.

Conventional Bonds

The most common type of instrument issued in the Euro markets is the conventional vanilla bond, with fixed coupon and maturity date. Coupon frequency is on annual basis. The typical face value of such Eurobonds is \$1,000, €1,000, £1000, or so on. The bond is unsecured, and therefore depends on the credit quality of its issuer in order to attract investors. Eurobonds have a typical maturity of 5 to 10 years, although many high-quality corporates have issued bonds with maturities of 30 years or even longer. The largest Eurobond market is in U.S. dollars, followed by issues in euros, Japanese yen, sterling, and a range of other currencies such as Australian, New Zealand, and

Canadian dollars; South African rand; and so on. Issuers will denominate bonds in a currency that is attractive to particular investors at the time, and it is common for bonds to be issued in more "exotic" currencies, such as Middle Eastern, Latin American, and Asian currencies.

Eurobonds are not regulated by the country in whose currency the bonds are issued. They are typically registered on a national stock exchange, usually London or Luxembourg. Listing of the bonds enables certain institutional investors, who are prohibited from holding assets that are not listed on an exchange, to purchase them. The volume of trading on a registered stock exchange is negligible, however; virtually all trading is on an over-the-counter (OTC) basis directly between market participants.

Interest payments on Eurobonds are paid gross and are free of any withholding or other taxes. This is one of the main features of Eurobonds, as is the fact that they are "bearer" bonds; that is, there is no central register. Historically, this meant that the bond certificates were bearer certificates with coupons attached; these days, bonds are still designated "bearer" instruments but are held in a central depository to facilitate electronic settlement.

Floating-Rate Notes

An early innovation in the Eurobond market was the floating-rate note. They are usually short- to medium-dated issues, with interest quoted as a spread to a reference rate. The reference rate is usually the London Interbank Offered Rate (LIBOR), or the Singapore Interbank Offered Rate (SIBOR) for issues in Asia. The euro Interbank Rate (EURIBOR) is also now commonly quoted. The spread over the reference rate is a function of the credit quality of the issuer, and can range from 10 to 20 points over the reference rate for high-quality credits and from 150 to 450 basis points or even higher for low-rated borrowers. Bonds typically pay a quarterly coupon, although semi-annual and monthly coupon bonds are also issued. The first floating-rate note issue was by ENEL, an Italian utility company, in 1970. The majority of issuers are financial institutions such as banks and securities houses.

There are also perpetual, or undated, floating-rate notes, the first issue of which was by National Westminster Bank plc in 1984. They are essentially similar to regular floating-rate notes, except that they have no maturity date and are therefore "perpetual." Most perpetual floating-rate notes are issued by banks, for whom they are attractive because they are a means of raising capital similar to equity but with the tax advantages associated with debt. They also match the payment characteristics of the banks assets. Traditionally, the yield on perpetuals is higher than both conventional bonds and fixed-term floating-rate notes.

Zero-Coupon Bonds

An innovation in the market from the late 1980s was the zero-coupon bond, or pure discount bond, which makes no interest payments. Like zero-coupon bonds initially in government markets, the main attraction of these bonds for investors was that, as no interest was payable, the return could be declared entirely as capital gain, thus allowing the bondholder to avoid income tax. Most jurisdictions including the United States and United Kingdom have adjusted their tax legislation so that the return on zero-coupon bonds now counts as income and not capital gain.

Convertible Bonds

Another instrument that is common in the Eurobond market is the convertible bond. A Eurobond is convertible if it may be exchanged at some point for another instrument, usually the ordinary shares (equity) of the issuing company. The decision to elect to convert is at the discretion of the bondholder. Convertibles are analyzed as a structure comprised of a conventional bond and an embedded option.

The most common conversion feature is an *equity convertible,* which is a conventional bond that is convertible into the equity of the issuer. The conversion feature allows the bondholder to convert the Eurobond, on maturity or at specified times during the bond's life, into a specified number of shares of the issuing company at a set price. In some cases the bond is convertible into the shares of the company that is guaranteeing the bond. The issuing company must release new shares in the event of conversion. The price at which the bond is convertible into shares, known as the exercise price, is usually set at a premium above the market price of the ordinary shares in the market on the day the bond is issued. Investors will exercise their conversion rights only if the market price has risen sufficiently that a gain will be realized by converting. The incorporation of a conversion feature in a bond is designed to make the bond more attractive to investors, as it allows them to gain from a rise in the issuing company's share price. The conversion feature also acts as a floor for the bond price. The advantages of convertibles for borrowers include the following:

- As the bond incorporates an added attraction in the form of the conversion feature, the coupon payable on the bond is lower than it otherwise would be; this enables the borrower to save on interest costs.
- Issuing convertibles is one method by which companies can broaden the geographical base of their equity holders.
- Companies are usually able to raise a higher amount at one issue if the bond is convertible, compared to a conventional bond.
- Against these factors must be weighed certain disadvantages associated with convertibles, which include the following:
  - The investor's insurance against the volatility of share price movements, an attraction of the convertible is gained at the cost of a lower coupon than would be obtained from a conventional bond.
  - Convertibles are often issued by companies that would have greater difficulty placing conventional paper. Convertibles are usually subordinated and are often viewed more as equity rather than debt. The credit and interest rate risk associated with them is consequently higher than for conventional bonds.

There have been variations on the straight convertible bond in the Eurobond market. This includes the convertible preference share. This is a combination of a perpetual debt instrument cash flow with an option to convert into ordinary shares. Sometimes these issues are convertible not into shares of the issuer, but rather into the equity of a company in which the issuer has a significant shareholding.

Another variation is the *equity note,* which is a bond that is redeemed in shares and not in cash. The equity note is not a true convertible, since the conversion feature is not an option for the bondholder but a condition of the bond issue, and is guaranteed to take place. A more accurate description of an equity note would be an "interest-bearing equity future" note.

Eurobonds have also been issued with a feature that allows conversion into other assets such as crude oil or gold, or into other bonds with different payment characteristics. These are known as *asset convertibles.* Examples of such bonds include floating-rate notes that are convertible under specified circumstances into fixed-rate bonds. One version of this was the drop-lock bond, which was first introduced in the early 1980s during a period of high interest rates. Drop-lock bonds are initially issued as floating-rate notes but convert to a fixed-rate bond at the point that the reference rate falls to a preset level. The bond then pays this fixed rate for the remainder of its life. During the 1990s, as interest rate volatility fell to relatively lower levels, drop-locks fell out of favor, and it is now rare to see them issued.

Currency convertibles are bonds that are issued in one currency and are redeemed in another currency or currencies. Often, this at the discretion of the bondholder; other currency convertibles pay their coupon in a different currency to the one they are denominated in. In certain respects, currency convertibles possess similar characteristics to a conventional bond issued in conjunction with a forward contract. The conversion rate is specified at the time of issue, and may be either a fixed-rate option or a floating-rate option. With a fixed-rate option, the exchange rate between the currencies is fixed for the entire maturity of the bond at the time it is issued; with a floating-rate option, the exchange rate is not fixed and is the rate prevailing in the market at the time the conversion is exercised. Initially, most currency convertibles offered a fixed-rate option, so that the foreign exchange risk resided entirely with the issuer. Floating-rate options were introduced in the 1970s when exchange rates began to experience greater volatility.

Eurowarrants

The Eurobond warrant or *Eurowarrant* is essentially a call option attached to a conventional bond. The call option is convertible into either ordinary shares or other bonds of the issuing company or, rarely, another company. A typical Eurobond warrant will be comprised of a conventional bond, issued in denominations of \$1,000 or \$10,000, paying a fixed coupon. The attached warrant will entitle the bondholder to purchase shares (or bonds) at a specified price at set dates, or a set time period, up until maturity of the warrant, whereupon the warrant expires worthless. Warrants are often detached from their host bond and traded separately.

The exercise price of a warrant is fixed at a premium over the market price of the equity at issue. This premium is separate from the premium associated with a warrant in the *secondary market,* which is the total premium cost connected with buying the warrant and immediately exercising it into the equity, and not the cost associated with a purchase of the equity in the open market.

There are several advantages that Eurobond warrants hold for investors. They are composed of two assets that are usually traded separately in the secondary market; indeed, warrants are often attached to bonds as a "sweetener" for investors. Investors have an interest in the performance of the shares of the issuer without having a direct exposure to them. Should the intrinsic value of the warrant fall to zero, there is still time value associated with the warrant up until the maturity of the bond. Warrants typically possess high leverage or "gearing," which is defined as the ratio of the cost of the warrant to the cost of the shares that the warrant holder is entitled to purchase. Borrowers may also gain from attaching warrants to their bond issues. The advantages include being able to pay a lower coupon than might otherwise have been the case. The exercise of a warrant results in the issuer receiving cash for the shares that are purchased (albeit at a below-market rate), compared with a convertible bond, where the issuer receives only bonds that are subsequently canceled. This is a feature of the warrant's gearing, as the value of the warrant is always less than the price at which the company guarantees to issue new equity to the warrant holder. The disadvantage at the time the warrant is exercised is that the company is receiving a below-market price for its shares at a time when they are trading at a historically high level; however, there is a form of compensation for this since the company would have issued the bonds at a lower coupon rate than would have been the case had the warrants not been attached.

THE ISSUING PROCESS: MARKET PARTICIPANTS

When a company raises a bond issue, its main concerns will be the success of the issue and the interest rate that must be paid for the funds borrowed. An issue is handled by an international syndicate of banks. A company wishing to make a bond issue will invite a number of investment banks and securities houses to bid for the role of *lead manager.* The bidding banks will indicate the price at which they believe they can get the issue away to investors, and the size of their fees. The company's choice of lead manager will be based on the bids, but also the reputation and standing of the bank in the market. The lead manager, when appointed, will assemble a syndicate of other banks to help with the issue. This syndicate will often be made up of banks from several different countries. The lead manager has essentially agreed to underwrite the issue, which means that it guarantees to take the paper off the issuer's hands (in return for a fee). If there is an insufficient level of investor demand for the bonds, the lead manager will be left holding ("wearing") the issue, which, in addition to being costly, will not help its name in the market. When we referred to an issuer's assessing the reputation of potential lead managers, this included the company's view on the "placing power" of the bank, its perceived ability to get the entire issue away. The borrowing company would prefer the issue to be oversubscribed, which is when demand outstrips supply.

In many cases the primary issue involves a *fixed price reoffer* scheme. The lead manager will form the syndicate, which will agree on a fixed issue price, a fixed commission, and the distribution among themselves of the quantity of bonds they agreed to take as part of the syndicate. The banks then reoffer the bonds that they have been allotted to the market, at the agreed price. This technique gives the lead manager greater control over a Eurobond issue. It sets the price at which other *underwriters* in the syndicate can initially sell the bonds to investors. The fixed price reoffer mechanism is designed to prevent underwriters from selling the bonds back to the lead manager at a discount to the original issue price, that is, "dumping" the bonds.

Before the bond issue is made, but after its basic details have been announced, it is traded for a time in the *gray market.* This is a term used to describe trading in the bonds before they officially come to the market, mainly market makers selling the bond short to other market players or investors. Activity in the gray market serves as useful market intelligence to the lead manager, which can gauge the level of demand that exists in the market for the issue. A final decision on the offer price is often delayed until dealing in the gray market indicates the best price at which the issue can be gotten away.

Let us now consider the *primary market* participants in greater detail.

The Borrowing Parties

The range of borrowers in the Euromarkets is very diverse. From virtually the inception of the market, borrowers representing corporates, sovereign and local governments, nationalized corporations, supranational institutions, and financial institutions have raised finance in the international markets. The majority of borrowing has been by governments, regional governments, and public agencies of developed countries, although the Eurobond market is increasingly a source of finance for developing country governments and corporates.

Governments and institutions access the Euromarkets for a number of reasons. Under certain circumstances, it is more advantageous for a borrower to raise funds outside its domestic market, due to the effects of tax or regulatory rules. The international markets are very competitive in terms of using intermediaries, and a borrower may well be able to raise cheaper funds in the international markets. Other reasons why borrowers access Eurobond markets include:

- A desire to diversify sources of long-term funding. A bond issue is often placed with a wide range of institutional and private investors, rather than the more restricted investor base that may prevail in a domestic market. This gives the borrower access to a wider range of lenders, and for corporate borrowers this also enhances the international profile of the company.
- For both corporates and emerging-country governments, the prestige associated with an issue of bonds in the international market.
- The flexibility of a Eurobond issue compared to a domestic bond issue or bank loan, illustrated by the different types of Eurobond instruments available.

Against this are balanced the potential downsides of a Eurobond issue, which include the following:

- For all but the largest and most creditworthy of borrowers, the rigid nature of the issue procedure becomes significant during times of interest and exchange rate volatility, reducing the funds available for borrowers.
- Issuing debt in currencies other than those in which a company holds matching assets, or in which there are no prospects of earnings, exposes the issuer to foreign exchange risk.

Generally, though, the Euromarket remains an efficient and attractive market in which a company can raise finance for a wide range of maturities.

The nature of the Eurobond market is such that the ability of governments and corporates to access it varies greatly. Access to the market for a first-time borrower has historically been difficult and has been a function of global debt market conditions. There is a general set of criteria, first presented by van Agtmael (1983), that must be fulfilled initially, which for corporates include the following:

- The company should ideally be domiciled in a country that is familiar to Eurobond issuers, usually as a result of previous offerings by the country's government or a government agency. This suggests that it is difficult for a corporate to access the market ahead of a first issue by the country's government.
- The borrowing company must benefit from a level of name recognition or, failing this, a sufficient quality credit rating.
- The company ideally must have a track record of success and needs to have published financial statements over a sufficient period of time, audited by a recognized and respected firm, and the company's management must make sufficient financial data available at the time of the issue.
- The company's requirement for medium-term or long-term finance, represented by the bond issue, must be seen to fit into a formal strategic plan.

Generally, Eurobond issuers are investment-grade rated, and only a small number, less than 5% according to International Monetary Fund data, are not rated at all.

The Underwriting Lead Manager

Issuers of debt in the Eurobond market select an investment bank to manage the bond issue for them. This bank is known as the underwriter because, in return for a fee, it takes on the risk of placing the bond among investors. If the bond cannot be placed in total, the underwriting bank will take on the paper itself. The issuer will pick an investment bank with whom it already has an existing relationship, or it may invite a number of banks to bid for the mandate. In the event of a competitive bid, the bank will be selected on the basis of the prospective coupon that can be offered, the fees and other expenses that it will charge, the willingness of the bank to support the issue in the secondary market, the track record of the bank in placing similar issues and the reach of the bank's client base. Often, it is a combination of a bank's existing relationship with the issuer and its reputation in the market for placing paper that will determine whether or not it wins the mandate for the issue.

After the mandate has been granted, and the investment bank is satisfied that the issuer meets its own requirements on counterparty and reputational risk, both parties will prepare a detailed financing proposal for the bond issue. This will cover topics such as the specific type of financing, the size and timing of the issue, approximate pricing, fees, and so on. The responsibilities of the lead manager include the following:

- Analyzing the prospects of the bond issue being accepted by the market; this is a function of both the credit quality of the issuer and the market's capacity to absorb the issue.
- Forming the syndicate of banks to share responsibility for placing the issue. These banks are co-lead managers and syndicate banks.
- Assisting the borrower with the prospectus, which details the bond issue and also holds financial and other information on the issuing company.
- Assuming responsibility for the legal issues involved in the transaction, for which the bank's in-house legal team and/or external legal counsel will be employed.
- Preparing the documentation associated with the issue.
- Taking responsibility for the handling of the fiduciary services associated with the issue, which is usually handled by a specialized agent bank.
- If deemed necessary, establishing a pool of funds that can be used to stabilize the price of the issue in the gray market, used to buy (or sell) bonds if required.

These duties are usually undertaken jointly with other members of the syndicate. For first-time borrowers, the prospectus is a very important document, as it is the main communication media used to advertise the borrower to investors. In a corporate issue, the prospectus may include the analysis of the company by the underwriters, financial indicators and balance sheet data, a detailed description of the issue specifications, the members of the underwriting syndicate, and details of placement strategies. In a sovereign issue, the prospectus may cover a general description of the economy of the country's, including key economic indicators such as balance-of-payments figures and export and import levels, the state of the national accounts and budget, a description of the political situation (with an eye on the stability of the country), current economic activity, and a statement of the current external and public debt position of the country.

The Co-lead Manager

The function of the co-lead manager in Eurobond issues developed as a consequence of the distribution of placing ability across geographic markets. For example, as the Eurobond market developed, underwriters who were mainly U.S. or U.K. banks did not have significant client bases in say, the continental European market, and so banking houses that had a customer base there would be invited to take on some of the issue. For a long time the ability to place $5,000,000 nominal of a new Eurobond issue was taken as the benchmark against a potential co-lead manager.

The decision by a lead manager to invite other banks to participate will depend on the type and size of the issue. Global issues such as those by the World Bank, which have nominal sizes of $1 billion or more, have a fairly large syndicate. The lead manager will assess whether it can place all the paper or it, in order to achieve geographic spread (which may have been stipulated by the issuer) it needs to form a syndicate. It is common for small issues to be placed entirely by a single lead manager.

Investors

The structure of the Eurobond market, compared to domestic markets, lends a certain degree of anonymity, if such is desired, to end-investors. This is relevant essentially in the case of private investors. The institutional holders of investors are identical to those in the domestic bond markets, and include institutional investors such as insurance companies, pension funds, investment trusts, commercial banks, and corporations. Other investors include central banks and government agencies; for example, the Kuwait Investment Office and the Saudi Arabian Monetary Agency both have large Eurobond holdings. In the United Kingdom, banks and securities houses are keen holders of floating-rate note Eurobonds, usually issued by other financial institutions.

FEES, EXPENSES, AND PRICING

Traditionally, Eurobond issues were placed in accordance with an accepted broad set of pricing and fee rules.

Fees

The fee structures for placing and underwriting a Eurobond issue are relatively identical for most issues. The general rule is that fees increase with maturity and decreasing credit quality of the issuer, and decrease with nominal size. Fees are not paid directly but are obtained by adjusting the final price paid to the issuer, that is, taken

Table 24.1 Expense Elements, Eurobond Issue

Printing (prospectus, certificates, etc.)	Clearing and bond issuance
Legal counsel (Issuer and investment bank)	Paying agent
Stock exchange listing fee	Trustee
Promotion	Custodian
Underwriters expenses	Common depositary

out of the sale proceeds of the issue. The allocation of fees within a syndicate can be slightly more complex, and in the form of an underwriting allowance. This is usually paid out by the lead manager.

Typical fees will vary according to the type of issue and issuer, and also whether the bond itself is plain vanilla or more exotic. Fees range from 0.25% to 0.75% of the nominal of an issue. Higher fees may be charged for small issues.

Expenses

The expenses associated with the launch of a Eurobond issue vary greatly. Table 24.1 illustrates the costs associated with a typical Eurobond transaction. Not every bond issue will incur every expense, however these elements are common. The expense items in this table do not include the issuer's own expenses with regard to financial accounting and marketing. The reimbursement for underwriters is intended to cover such items as legal expenses, travel, delivery of bonds, and other business expenses.

In general, Eurobonds are listed on either the London or Luxembourg stock exchange. Certain issues in the Asian markets are listed on the Singapore exchange. To enable listing to take place, an issuer will need to employ a listing agent, although this is usually arranged by the lead manager. The function of the listing agent is to (1) provide a professional opinion on the prospectus, (2) prepare the documentation for submission to the stock exchange, and (3) make a formal application and conduct negotiations on behalf of the issuer.

Pricing

One of the primary tasks of the lead manager is the pricing of the new issue. The lead manager faces an inherent conflict of interest between its need to maximize its returns from the syndication process and its obligation to secure the best possible deal for the issuer, its client. An inflated issue price invariably causes the yield spread on the bond to rise as soon as the bond trades in the secondary market. This would result in a negative impression being associated with the issuer, which would affect its next offering. However, too low a price can permanently damage a lead manager's relationship with the client.

For Eurobonds that are conventional vanilla fixed-income instruments, pricing does not present too many problems in theory. The determinants of the price of a new issue are the same as those for a domestic bond offering and include the credit quality of the borrower, the maturity of the issue, the total nominal value, the presence of any option feature, and the prevailing level and volatility of market interest rates. Eurobonds are perhaps more heavily influenced by the target market's ability to absorb the issue, and this is gauged by the lead manager in its preliminary offering discussions with investors. The credit rating of a borrower is often similar to that granted to it for borrowings in its domestic market, although in many cases a corporate will have a different rating for its foreign currency debt compared to its domestic currency debt.

In the gray market, the lead manager will attempt to gauge the yield spread over the reference pricing bond at which investors will be happy to bid for the paper. The reference bond is the benchmark for the maturity that is equivalent to the maturity of the Eurobond. It is commonly observed that Eurobonds have the same maturity date as the benchmark bond that is used to price the issue. As lead managers often hedge their issue using the benchmark bond, an identical maturity date helps to reduce basis risk.

ISSUING THE BOND

The three key dates in a new issue of Eurobonds are the announcement date, the offering day, and the closing day. Prior to the announcement date, the borrower and the lead manager (and co-lead managers if applicable) will have had preliminary discussions to confirm the issue specifications, such as its total nominal size, the target coupon, and the offer price. These details are provisional and may well be different at the time of the closing date. At these preliminary meetings, the lead manager will appoint a fiscal agent or *trustee* and a principal paying agent. The lead manager will appoint other members of the syndicate group, and the legal documentation and prospectus will be prepared.

On the announcement date the new issue is formally announced, usually via a press release. The announcement includes the maturity of the issuer and a coupon rate or range in which the coupon is expected to fall. A telex is also sent by the lead manager to each prospective underwriter, which is a formal invitation to participate in the syndicate. These banks will also receive the preliminary offering circular, a timetable of relevant dates for the issue, and documentation that discloses the legal obligations that they are expected to follow should they decide to participate in the issue. The decision to join is mainly, but not wholly, a function of the bank's clients' interest in the issue, which the bank needs to sound out.

The pricing day signals the end of the subscription period, the point at which the final terms and conditions of the issue are agreed between the borrower and the syndicate group. If there has been a significant change in market conditions, the specifications of the bond issue will change. Otherwise, any required final adjustment of the price is usually undertaken by a change in the price of the bond relative to par. The ability of the lead manager to assess market conditions accurately at this time is vital to the successful pricing of the issue.

Once the final specifications have been determined, members of the syndicate have roughly 24 hours to accept or reject the negotiated terms; the bonds are then formally offered on the offering day, the day after the pricing day, when the issuer and the managing group sign the subscription or underwriting agreement containing the final specifications of the issue. The underwriting syndicate then enters into a legal commitment to purchase the bonds from the issuer at the price announced on the pricing day. A final offering circular is the produced, and the lead manager informs the syndicate of the amount of their allotments. The lead manager may wish to either overallocate or underallocate the number of available bonds, depending on its view on future levels and direction of interest rates. There then begins the stabilization period, when the bonds begin to trade in the secondary period, where Eurobonds trade in an over-the-counter market. About 14 days after the offering day, the closing day occurs. This is when syndicate members pay for bonds they have purchased, usually by depositing funds into a bank account opened and run by the lead manager on behalf of the issuer. The bond itself is usually represented by a *global note,* held in *Euroclear* or *Clearstream,* initially issued in temporary form. The temporary note is later changed to a permanent global note. Tranches of an issue targeted at U.S. investors may be held in the Depository Trust Corporation as a registered note.

The Gray Market

The subscription period of a new Eurobond issue is characterized by uncertainty about potential changes in market conditions. After the announcement of the issue, but before the bonds have been formally issued, the bonds trade in the gray market. The gray market is where bonds are bought and sold for settlement on the first settlement date after the offering day. Gray market trading enables the lead manager to gauge the extent of investor appetite for the issue, and make any adjustment to coupon if required. A gray market that functions efficiently will, at any time, reflect the market's view on where the bond should trade, and what yield the bond should be offered. It enables investors to trade in the primary market possessing information as to the likely price of the issue in the secondary market.

Another principal task of the lead manager is to stabilize the price of the bond issue for a short period after the bond has started trading in the secondary market. This is known as the stabilization period, and the process is undertaken by the lead manager in concert with some or all of the syndicate members. A previously established pool of funds may be used for this purpose. The price at which stabilization occurs is known as the syndicate bid.

Alternative Issue Procedures

In addition to the traditional issue procedure where a lead manager and syndicate offer bonds to investors based on a price set, on pricing day, based on a yield over the benchmark bond, there are a number of other issue procedures that are used. One of these methods is the bought deal, where a lead manager or a managing group approaches the issuer with a firm bid, specifying issue price, amount, coupon, and yield. Only a few hours are allowed for the borrower to accept or reject the terms. If the bid is accepted, the lead manager purchases the entire bond issue from the borrower. The lead manager then has the option of selling part of the issue to other banks for distribution to investors, or doing so itself. In a volatile market, the lead manager will probably parcel some of the issue to other banks for placement. However, it is at this time that the risk of banks dumping bonds on the secondary market is highest; in this respect, lead managers will usually preplace the bonds with institutional investors before the bid is made. The bought deal is focused primarily on institutional rather than private investors. As the syndicate process is not used, the bought deal requires a lead manager with sufficient capital and placement power to enable the entire issue to be placed.

In a prepriced offering, the lead manager's bid is contingent on its ability to form a selling group for the issue. Any alterations in the bid required for the formation of the group must be approved by the borrower. The period allocated for the formation of the group is usually two to four days, and after the group has been formed, the process is identical to that for the bought deal.

Yet another approach is the auction issue, under which the issuer will announce the maturity and coupon of a prospective issue and invite interested investors to submit bids. The bids are submitted by banks, securities houses, and brokers and include both price and amount. The advantages of the auction process is that it avoids the management fees and costs associated with a syndicate issue. However, the issuer does not have the use of a lead manager's marketing and placement expertise, which means it is a method that can be employed only by very high-quality, well-known borrowers.

COVENANTS

Eurobonds are unsecured, and as such, the yield demanded by the market for any particular bond will depend on the credit rating of the issuer. Until the early 1980s, Eurobonds were generally issued without *covenants,* due to the high quality of most issuers. Nowadays, it is common for covenants to be given with Eurobond issues. Three covenants in particular are frequently demanded by investors:

1. A negative pledge
2. An "event risk" clause
3. A gearing ratio covenant

Negative Pledge

A negative pledge is one that restricts the borrowings of the group that ranks in priority ahead of the debt represented by the Eurobond. In the case of an unsecured Eurobond issue, this covenant restricts new secured borrowings by the issuer, as well as new unsecured borrowings by any of the issuer's subsidiaries, since these would rank ahead of the unsecured borrowings by the parent company in the event of the whole group going into receivership.

Disposal of Assets Covenant

This sets a limit on the amount of assets that can be disposed of by the borrower during the tenor (term to maturity) of the debt. The limit on disposals could be, typically, a cumulative total of 30% of the gross assets of the company. This covenant is intended to prevent a breakup of the company without reference to the Eurobond investors.

Gearing Ratio Covenant

This places a restriction on the total borrowings of the company during the tenor of the bond. The restriction is set as a maximum percentage say, 150% to 175% of the company's or group's net worth (share capital and reserves).

TRUST SERVICES

A Eurobond issue requires an agent bank to service it during its life. The range of activities required is detailed below.

Depositary

The depositary for a Eurobond issue is responsible for the safekeeping of securities. In the Euromarkets, well over 90% of investors are institutions, and as a result, issues are made in dematerialized form and are represented by a global note. Trading and settlement is in computerized book-entry form via the two main international clearing systems, Euroclear and Clearstream. Both these institutions have appointed a group of banks to act on their behalf as depositaries for book-entry securities; they are known as common depositaries, because the appointment is common to both Euroclear and Clearstream. Both clearing firms have appointed separately a network of banks to act as specialized depositaries, which handled securities that have been issued in printed note or definitive form.

The common depositary is responsible for:

- Representing Euroclear and Clearstream, and facilitating delivery-versus-payment of the primary market issue by collecting funds from the investors, taking possession of the temporary global note (which allows securities to be released to investors), and making a single payment of funds to the issuer.
- Holding the temporary global note in safe custody, until it is exchanged for definitive notes or a permanent global note.
- Making adjustments to the nominal value of the global note that occur after the exercise of any options or after conversions, in line with instructions from Euroclear or Clearstream and the fiscal agent.
- Surrendering the canceled temporary global note to the fiscal agent after the exchange into definitive certificates or a permanent global note, or on maturity of the permanent global note.

A specialized depositary will hold definitive notes representing aggregate investor positions held in a particular issue; on coupon and maturity dates, it presents the coupons or bond to the paying agent and passes the proceeds on to the clearing system.

Paying Agent

Debt issuance in the Euromarkets requires a fiscal or principal paying agent, or in the case of a program of issuance (e.g., a Euro medium-term note program) an issuing and paying agent. The responsibility of the paying agent is to provide administrative support to the issuer throughout the lifetime of the issue. The duties of a paying agent include:

- Issuing securities upon demand in the case of a debt program.
- Authenticating definitive notes.
- Collecting funds from the issuer and paying these out to investors as coupon and redemption payments.
- In the case of global notes, acting on behalf of the issuer to supervise payments of interest and principal to investors via the clearing systems, and in the case of definitive notes, paying out interest and coupon on presentation by the investor of the relevant coupon or bond to the paying agent.
- Transferring funds to sub-paying agents, where these have been appointed. A security that has been listed in Luxembourg must have a local sub-paying agent appointed for it.
- Maintaining an account of the cash flows paid out on the bond.
- Arranging the cancellation and subsequent payment of coupons, matured bonds, and global notes and sending destroyed certificates to the issuer.

A paying agent will act solely on behalf of the issuer, unlike a trustee, who has an obligation to look after the interests of investors. For larger bond issues, there may be a number of paying agents appointed, of which the principal paying agent is the coordinator. A number of sub-paying agents may be appointed to ensure that bondholders in different country locations may receive their coupon and redemption payments without delay. The term "fiscal agent" is used to describe a paying agent for a bond issue for which no trustee has been appointed.

Registrar

The role of the registrar is essentially administrative, and it is responsible for keeping accurate records of bond ownership for registered securities. As most Eurobonds are issued in bearer form, there is not a great deal of work for registrars in the Euromarket, and the number of holders of registered notes is normally quite low.

The responsibilities of the registrar include:

- Maintaining a register of all bondholders and records of all transfers of ownership.
- Coordinating the registration, transfer, or exchange of bonds.
- Issuing and authenticating new bonds should any transfer or exchange take place.
- Maintaining a record of the outstanding principal value of the bond.
- Undertaking administrative functions relating to any special transfers.

Trustee

An issuer may appoint a trustee to represent the interests of investors. In the event of default, the trustee is required to discharge its duties on behalf of bondholders. In certain markets, a trustee is required by law; for instance, in the United States, a trustee has been a legal requirement since 1939. In other markets, an issuer may appoint a trustee in order to make the bond issue more attractive to investors, as it means that there is an independent body to help look after their interests. This is particularly important for a secured issue, where the trustee sometimes holds collateral for the benefit of investors. Assets that are held by the trustee can be protected from the creditors of the issuer in the event of bankruptcy. A trustee has a variety of powers and discretion, which are stated formally in the issue trust deed, and these include its duties in relation to the monitoring of covenants and duties to bondholders.

Custodian

A custodian provides safekeeping services for securities belonging to a client. The client may be an institutional investor, such as a pension fund, that requires a portfolio of securities in many locations to be kept in secure custody on their behalf. As well as holding securities, the custodian usually manages corporate actions such as dividend payments.

Fiscal Agent

A Eurobond issuer will appoint either a fiscal agent or a trustee; both perform similar roles but under differing legal arrangements. The fiscal agent is appointed by and is the representative of the issuer, so, unlike a trustee, it does not represent the bondholders. The main responsibilities of the fiscal agent are to pay the principal and interest payments, and it performs a number of administrative roles as well, such as the publication of financial information and notices to investors.

Listing Agent

Issuers must appoint a listing agent if they wish to list the bond on the London or Luxembourg stock exchange, as this is a requirement of the rules of the exchange. The listing agent communicates with the exchange on behalf of the issuer, and lodges the required documentation with it. In the United Kingdom, the listing agent must be authorized under financial regulatory legislation and is usually the lead manager for the issue, although it is also common for a fiduciary service provider to be appointed to this role.

FORM OF THE BOND

Eurobonds are issued in temporary global form or permanent global form. If issued in temporary form, the note is subsequently changed into either permanent global form or definitive form, which may be either a bearer note or registered.

Temporary Global Note

On issue, the majority of Eurobonds are in the form of a single document known as a "temporary global bond." This document represents the entire issue, executed by an officer of the issuer and certified by the fiscal agent or principal paying agent. After a period of time, the temporary global bond, as its name suggests, is exchanged for either a permanent global bond or bonds in definitive form, which are separate certificates representing each bond holding.

The main reason bonds are issued in temporary form is time constraints between the launch of issue, when the offer is announced to the market, and closing, when the bonds are actually issued. This period differs according to the type of issue and instrument; for example, for a plain vanilla issue, it can be as little as two weeks, whereas for more exotic issues (such as a securitization), it can be a matter of months. The borrower will be keen to have the periods short as possible, as the financing is usually required quickly. As this results in there being insufficient time to complete the security printing and authentication of the certificates, which represent the final definitive form, a temporary bond is issued to enable the offering to be closed and placed in a clearing system, while the final certificates are produced. Bonds are also issued in temporary form to comply with certain domestic selling regulations and restrictions, for example, a U.S. regulation that definitive bonds cannot be delivered for a 40-day period after issue. This is known as the "lock-up" period.

Permanent Global Note

Like the temporary bond, the permanent global bond is a word-processed document and not a security-printed certificate, issued on the closing date. It represents the entire issue and is compiled by the underwriter's legal representatives. In most cases, it is actually held for safekeeping on behalf of Euroclear and Clearstream by the trust or clearing arm of a bank, known as the "common depositary." Borrowers often prefer to issue notes in permanent global form because this carries lower costs compared to definitive notes, which are security printed.

Definitive Note

Under any circumstances where it is required that investors have legal ownership of the debt obligation represented by a bond issue they have purchased, a borrower is obliged to issue the bond in definitive form. The situations under which this becomes necessary are listed on the permanent global bond document, and include the following:

- Where an investor requires a definitive bond to prove legal entitlement to the bond (s)he has purchased, in the case of any legal proceedings undertaken concerning the bond issue.
- In the event of default, or if investors believe default to have occurred.
- Where for any reason the bonds can no longer be cleared through a clearing system, in which case they must be physically delivered in the form of certificates.

Bonds issued in definitive form may be either bearer or registered securities. A bearer security has similar characteristics to cash money, in that the certificates are documents of value and the holder is considered to be the beneficiary and legal owner of the bond. The bond certificate is security printed and the nature of the debt obligation is detailed on the certificate. Transfer of a bearer security is by physical delivery. Some of the features of traditional bearer securities include:

- Coupons, attached to the side of the certificate, which represent each interest payment for the life of the bond. The holder is required to detach each coupon as it becomes due and send it to the issuer's paying agent.
- A promise to pay, much like a bank note, which confirms that the issuer will pay the bearer the face value of the bond on the specified maturity date.
- In some cases, a talon, which is the right for the bondholder to claim a further set of coupons once the existing set has been used (this applies only to bonds that have more than 27 interest payments during their lifetime, as International Capital Markets Association (ICMA, the professional association of the Eurobond dealing community) rules prohibit the attachment of more than 27 coupons to a bond on issue).

The administrative burdens associated with bearer securities is the main reason why the procedures associated with them are carried out via the clearing systems and paying agents, rather than individually by each investor.

Registered Bonds

Bonds issued in registered from are transferred by an entry on a register held by the issuer or its agent; the promise to pay is made to those names that appear on the register. Most Eurobonds are issued in bearer form for ease in clearing. Issues that are placed wholly or partly in the United States do, however, include an option allowing investors to take the bonds in registered form. This is done as most issues in the United States are sold under private placement, in order to be exempt from SEC selling restrictions, and private placement in that country requires that the bonds are in registered form. In such cases, the issuer will appoint a New York registrar for the issuer, usually the trust arm of a bank.

CLEARING SYSTEMS

The development of the international bond market has taken place alongside the introduction of specialized clearing systems, which are responsible (among other things) for the settlement and safekeeping of Eurobonds. The two main clearing systems are Euroclear and Clearstream.

Euroclear was created by the Morgan Guaranty Trust Company of New York in 1968. Ultimately, ownership passed to a consortium of banks, and it is now run by Euroclear Clearance Systems plc, and operated by a cooperative company in Brussels.

The original Cedel was created in 1970 in Luxembourg and is owned by a consortium of around 100 banks, no one of which may hold more than 5% of the company. The two clearing systems do not restrict their operations to the settlement and custody of Eurobonds.

Both clearing systems exist to avoid the physical handling of bearer instruments, both on issue and in the secondary market. This means that on issue the actual bond certificates, which may be in definitive bearer or global form are passed on to a "trust" bank, known as the depositary for safekeeping. The clearing system will track holdings via a book entry. To participate in the clearing system setup, an investor must have two accounts with it, which may be its own accounts or accounts held by their bank, who will act as a nominee on their behalf; these are a securities clearance account, to which a security is credited, and a cash account, through which cash is received or paid out.

The clearing system will allocate a unique identification code, known as the International Securities Identification Number (ISIN) to each Eurobond issue, and a "common code" is derived from the ISIN. The common code is essentially the identification used for each bond issue whenever an instruction is sent to the clearing agent to deal in it. The ISIN will be in addition to any number issued by a domestic clearing agent, for example, the stock exchange number for London listed securities. Both clearing systems have specific roles in both the primary and secondary markets. In the primary market, they accept a new issue of Eurobonds, and on closing, the required number of bonds are credited to the securities clearance account of the banks that are part of the issue syndicate. Securities are then transferred (electronic book entry) to securities accounts of investors.

The clearance systems keep a record on the coupon payment and redemption dates for each bond, and "present" the bonds for payment on each appropriate date. Investors, therefore, do not need to present any coupons or certificates themselves, which is why the system is now paperless.

SECONDARY MARKET

Most Eurobonds are tradeable, although the liquidity of individual issues is variable. Although in theory transfer is by physical delivery because the bonds are bearer instruments, the great majority of bonds will settle by the Euroclear or Clearstream International ("Clearstream") settlement systems. Liquidity in the market varies over time and for individual issues will be a function of:

- Size of issue.
- Level of investor demand for the paper.
- Commitment of market makers to support the issue.

A large number of Eurobonds are illiquid, and market makers will quote a bid price only. No offer price is made because the market maker (unless he actually owns some of the issue) will be unable to find bonds to deliver to the buyer if it is illiquid. Many Eurobonds issued in second-tier currencies, such as Malaysian ringgit, will have been issued and then immediately asset swapped, and hence there will be no paper available to trade. Many large issuers will issue Eurobonds in a currency other than that which they require, in order to meet a specific customer demand for paper in that currency; after issue, the proceeds are swapped into the desired currency. In the meantime, the bonds will be held to maturity by the investors and usually not traded in the secondary market.

High-quality Eurobond issues will trade almost as government paper. For example, issues by the World Bank or the European Investment Bank (EIB) trade at very low spreads above the same currency government bonds, and at sub-LIBOR in the asset-swap market, and are highly liquid. For example, at times, EIB sterling Eurobonds have traded at only 7 to 9 basis points above the same maturity gilt. At the other end of the spectrum are those Eurobonds issued by infrequent issuers, for which no offer price may be available.

The market in trading Eurobonds is conducted on an OTC basis. In 1998, a number of automated electronic trading system were also introduced. The preeminence of London as the main trading center for the Eurobond market is well established, although Brussels, Frankfurt, Zurich, and Singapore are also important trading centers. The advantages of London as a trading center are generally regarded as being:

- A low level of regulatory interference in the functioning of the market.
- The presence of well-established infrastructure and institutions, as well as experienced human resources.
- The use of the English language as the market's main language of communication.

There are over 40 different market makers registered with the ICMA, and although in theory they are all required to make two-way prices in their chosen markets, the level of commitment is very varied. The bid-offer spread can be as low as 0.10 for very liquid issues such as World Bank and EIB bonds, to no offer price quoted for illiquid issues. In between, there is a range of spread sizes. The normal market size also varies, from £100,000 nominal to £500,000.

The valuation of Eurobonds is usually done on the basis of a yield spread over the relevant government bond yield curve. This yield spread is a function of the credit quality of the bond, its liquidity in the market, and the level of supply and demand. The bonds also move in line with general moves in interest rates, so that if there is a change in the gilt yield curve, a sterling Eurobond will change in yield, irrespective of whether the bond's issuer was perceived as being a weaker or a stronger credit. A market maker wishing to hedge a position in Eurobonds will usually use either the benchmark government against which the bond is priced or, if a noncash option is preferred, will use bond futures contracts to hedge the position.

LEGAL AND TAX ISSUES

Investor and borrowers in the Eurobond market may at any one time fall under the auspicies of a number of countries laws and regulations. These relate to the withholding tax on the bond coupons, income tax, disclosure and prospectus requirements, and restrictions on sales to certain classes of investor. The most important legal considerations for professional participants relate to (1) the possibility that the bonds are eventually distributed to residents in the United States, which is prohibited; and (2) London, as the principal financial center where the sale and trading of bonds takes place. The first consideration means that the market is subject to legislation in the United States that dates from the U.S. Securities Act of 1933 and federal income tax regulations. The second consideration means that the market comes under certain aspects of English law. With regard to taxation, the key features of Eurobonds are that:

- The bonds are "bearer" rather than registered securities.
- Interest and principal payments are not subject to withholding tax at source in the country where the issuer is resident for tax purposes.

The fact that payments of interest and principal on Eurobonds are not subject to any form of withholding tax at source in the country where the borrower is deemed to be resident for tax purposes is the primary feature of Eurobonds for investors, generally cited to be of key importance in making the market attractive for investors across a range of countries. Nonresident investors in Eurobonds are usually subject to the withholding tax requirements of the resident country of the bond issuer when that party repays interest or principal on bonds held by these nonresidents. The tax advantages to an investor from the absence of withholding tax (combined with the fact that the bonds are issued in bearer form) are significant. A large proportion of Eurobonds are held by private investors, and much of this is made anonymously by means of external discretionary accounts, such as those run by Swiss banks. This is a source of some frustration to tax authorities in certain countries. The absence of withholding tax also confers a certain benefit to issuers of Eurobonds. Where a bond issue was subject to withholding tax, an issuer would need to make the terms of the issue more attractive, that is, a higher coupon, in order to make the bond as attractive as the Eurobond issuer. This will carry higher associated costs for the issuer.

EUROBONDS AND SWAP TRANSACTIONS

The issue of new Eurobonds and the use of "asset swaps" in conjunction with issues is a vital part of the market, with investment banks keeping a close observation of the asset swap curve to spot any opportunities that may arise that makes a new issue of paper more attractive. New issues of Eurobonds are often launched to facilitate a swap which has been arranged in advance.

The existence of the currency swap and asset swap market is one of the key reasons for the growth and popularity of the Eurobond market. A borrower can issue bonds in virtually any liquid and convertible currency, according to where there is demand and what the yield curve looks like, and swap the proceeds into the currency that it requires. The cost of borrowing is usually significantly lower than if the borrower had issued bonds in the required currency. Swap driven issues are very common in the Eurobond market, and the key motivator is that borrowing costs will be cheaper. If this cheap borrowing opportunity is not available, it is unlikely that the bond will be issued, because entering into a swap exposes the issuer to additional credit risk. Swap financing will require a borrower to obtain debt initially that has undesirable currency and/or coupon characteristics. If the counterparty to a swap defaults, the borrower will be left with a risk exposure on the original debt. However, swap financing remains attractive because of the opportunity to obtain cheaper borrowing costs, despite the additional exposure to credit risk entailed in the transaction.

The market in swaps is governed by the International Swaps and Derivatives Association (ISDA). In the market, the majority of transactions are plain vanilla in nature and involve one of the following:

- Cross-currency fixed-rate swaps, usually referred to as currency swaps.
- Interest-rate swaps.
- Cross-currency hybrid swaps.
- Basis swaps.

Currency swaps are very common in the market. Under the plain vanilla version, two counterparties issue fixed-rate debt denominated in different currencies. They then exchange the interest (and sometimes) the principal repayments on their respective debt obligations. Under the conventional pattern, the amounts exchanged remain fixed at maturity. We will not cover the mechanics of a currency swap here, as this is reviewed in any number of derivatives texts, as are interest-rate swaps and the concepts of comparative advantage and the fixed- versus floating-rate legs of an interest-rate swap. Swap agreements do not always involve the exchange of debt repayment streams. In certain cases, one of the revenue streams exchanged in a swap can represent the income interest stream on an asset, or conventional security such as a corporate bond. Eurobond issues are frequently brought to the market primarily for the purpose of such "asset swapping." For the investment bank, swapping asset base interest payments is one means by which bond issues can be repackaged.

Other instruments used include basis swaps, which involve the exchange of two floating-rate payments streams, each of which is based on a short-term interest rate. The most common of these instruments have the following reference rates:

- LIBOR versus the U.S. commercial paper rate
- LIBOR versus the prime rate

Basis swaps are not the primary motivators of Eurobond issues, but are often included in more complex swap agreements, which may involve Eurobond borrowing.

For further details on the use of swap arrangements as part of Eurobond transactions, see Choudhry (2004a).

SETTLEMENT

Settlement of Eurobond transactions takes place within 28 days for primary market issues and T + 3 days for secondary market trades. Virtually all trades settle within the two main clearing systems, Euroclear and Clearstream. Euroclear was established in Brussels in 1968 by an international group of banks, the original entity known as Cedel was established in Luxembourg in 1970. Both clearing systems will settle in T + 3 days; however, the facility exists to settle trades in T + 1 if both parties to a trade wish it.

In the Euroclear system bonds are placed in the custody of the clearing system, through a Europe-wide network of depository banks. The transfer of bonds on settlement is undertaken by means of a computer book entry. This was the basic concept behind the introduction of Euroclear, the substitution of book entries for the physical movement of bonds. The actual physical securities to which a trading party has title are not identified in the majority of transactions made through Euroclear. The clearing system is made possible because the terms and conditions of any Eurobond issue are objectively specified, so that all bonds of a particular issue are standardized and therefore fungible for one another. There is no requirement to assign a specific bond serial number to an individual holder, which occurs with registered bonds. Clearstream operates on much the same basis. Participants in either system must be institutions with their own account (they may have an agent settle for them). Settlement takes place through the simultaneous exchange of bonds for cash on the books of the system. An "electronic bridge" connecting the two systems allows transfer of securities from one to the other.

SUMMARY

The Eurobond market has grown into an important sector of the debt capital markets. Originally driven by investor and regulatory restrictions in the United States, it is now the vital conduit through which capital is raised for sovereign and corporate borrowers alike. Bond issues are unsecured, so their relative attraction for investors depends on their formal credit rating and the liquidity of the secondary market. This liquidity differs by issuer name from very liquid to completely illiquid. Bonds are issued by a syndicate of banks and settle in the international clearing systems, Euroclear and Clearstream, by electronic book entry. There is wide variation in the type of bonds that are issued, from plain vanilla to structured and index-linked securities. All bonds issued as part of structured finance transactions are usually issued as Eurobonds.

REFERENCES

Andersen, T. (1982). How the grey market became respectable. *Euromoney*, May: 48–55.

Bowe, A. M. (1989). *Eurobonds*. Homewood, IL: Irwin Professional Publishing.

Choudhry, M. (2004a). *Fixed Income Markets*. Singapore: John Wiley & Sons.

Choudhry, M. (2004b). The European repo market. In F. J. Fabozzi and M. Choudhry (eds), *The Handbook of European Fixed Income Securities* (pp. 307–354). Hoboken, NJ: John Wiley & Sons.

Crawford, A. (1987). Stabilization brings the jitters. *Euromoney*, April: 277.

Decovny, S. (1998). *Swaps*. London: FT Prentice Hall.

Hallak, I. (2003). Courts and sovereign Eurobonds: Credibility of the judicial enforcement of repayment. CFS Working Paper Series, No. 2003/34.

Kerr, I. (1984). *A History of the Eurobond Market*. London: Euromoney Publications.

Munves, D. (2004). The Eurobond market. In F. J. Fabozzi and M. Choudhry (eds), *The Handbook of European Fixed Income Securities* (pp. 167–200). Hoboken, NJ: John Wiley & Sons.

Van Agtmael, A. (1983). Issuance of Eurobonds: Syndication and underwriting techniques and costs. In A. George, and I. Giddy (eds.), *International Financial Handbook* (Section 5.2). Chichester, UK: John Wiley & Sons.

CHAPTER 25

The Euro Government Bond Market

ANTONIO VILLARROYA, MD
Global Head of Rates Derivatives Strategy, Merrill Lynch

Abstract: The European Monetary Union and the introduction of the euro currency went a long way toward making the Eurozone government market the largest bond market in the world. What's more, this status is unlikely to be challenged in the coming years, with the new member countries of the European Union expected to join the single currency over time. Yet despite being integrated in many aspects, it should not be forgotten that the market comprises many issuers, with different credit ratings and issuing techniques, so is not completely homogeneous. These differences in credit status, together with the varying liquidity of their issues, their eligibility for the futures market and other micro factors, are the main drivers of intra-euro rate differentials.

Keywords: European Monetary Union (EMU), Pfandbrief, Maastricht, universal mobile telecommunications system (UMTS) licenses, French OAT, Italian BTP, German OBL, Spanish bonos, U.S. Treasuries, cheapest to deliver (CTD), swap spreads, strips, trading platforms, EuroMTS, tracking error, repo, primary dealers, syndication, Stability and Growth Pact (SGP), Euribor panel

Despite the appearance of many new fixed-income assets, the government bond market continues to be, by far, the largest market in the Eurozone. In this chapter, we analyze the recent trends in this market, its primary and secondary markets, and the key intra-country spread determinants.

THE EUROZONE: THE FASTEST-GROWING GOVERNMENT BOND MARKET

The history of the government bond market in continental Europe is relatively short, as most of the countries in this region did not have a liquid government bond market until the early 1990s. Yet, after several years of steady growth, the key event for the European government market was the culmination of the *European Monetary Union (EMU)* in January 1999. Up to that moment, the excessive fragmentation of the different European bond markets and the embedded exchange-rate risk had prevented the emergence of a large government bond market in Europe. Before this consolidation process, the market could not be considered deep enough to compete with the *U.S. Treasury* market as the asset of choice for investors looking for a liquid "risk-free" asset. The start of the EMU, therefore, made a much deeper government market possible, widening significantly this market's investor base.

Helped by its strong growth in the late 1990s, the euro government market totals more than €3.5 trillion as of January 2008 and has become the world's largest government bond market, helped also by the retreat of the U.S. Treasury market in the late 1990s. In fact, the euro market is around 65% larger than the Treasury market, and its outstanding issuance is 50% bigger than that of Japanese government bonds, accounting for approximately 40% of the world's outstanding government bonds as of mid-2007. This percentage was only around 13% of the combined G4 market in 1990.

The euro market has not only become the largest government bond market in terms of size, but also in terms of number of issues, with nearly 270 liquid issues (over €1 billion outstanding and one-year € maturity), significantly more than the 130 issues in the Treasury market and nearly 10 times more than the 25 liquid issues trading in the U.K. gilt market.

This market's growth rate has been fairly steady since the beginning of the EMU, with 1998 and 1999 registering the largest increases. Yet the pace of growth has decreased since then because of the large windfalls from third-generation telephone (universal mobile telecommunications system [UMTS]) licenses in some euro countries and the limits imposed by the region's *Stability and Growth Pact (SGP)*, registering just single-digit growth rates in the three years from 2000 through 2002. Furthermore, the low interest-rate and steep yield-curve environment have caused a shift in euro government supply toward the Treasury bill market at the expense of the bond market.

The beginning of the EMU also benefited the other euro-denominated fixed income markets, among them quasi-sovereigns, high-grade and high-yield credit bonds and asset-backed securities. Yet, although all these other markets experienced a massive increase in the post-EMU years, they are still quite small compared with the government sector and still very far from reaching the relative size they represent in the U.S. fixed income market. The government market continues to account for nearly 60% of all euro-denominated bonds outstanding, followed by the *Pfandbrief* market (around 7%) and the financial sector bond market, with 10%.

Country Breakdown

Just three countries (Germany, Italy, and France) make up over two-thirds of the total euro government market, increasing this percentage to nearly 90% of total outstanding bonds if the Spanish, Dutch, and Belgian markets are included. These relative weights have remained very stable since the enginning of EMU although it is worth noting how the relative weight of those countries that follow the SGP more strictly diminished compared with those whose deficits have remained closer to the 3% threshold. In general terms, the amounts issued by each country are very close to their respective market weights, with the total amount of fixed-rate bond supply around €470 billion annually in recent years.

It is also interesting to note how the average duration of the different euro government bond markets has been converging in the past few years, with the duration of the lower-rated countries increasing to almost match the stable or declining duration figures of the core euro countries. This process has taken the modified duration of the euro G8 markets to within a 0.4-year range, with a range of just around 0.25 years for the euro G4 countries.

Maturity Breakdown

Due to the large percentage of short-end supply and the time decay of the longer-dated issuance, the bulk of outstanding government debt is concentrated in short-end maturities, with one-third of total debt outstanding maturing in less than three years and two-thirds in less than eight years. The large decline in outstanding terms above 10 years' maturity is also significant, with less than 20% maturing beyond this point. In terms of maturity, the main recent event in the euro market has been the launch of ultra-long government bonds, with France, for instance, issuing a 50-year OAT in 2005. Helped by the flatness of the long end of the euro curve, these bonds have been issued to help European pension funds better hedge their long-dated liabilities. This practice of issuing long-dated nominal or real bonds is more common in the United Kingdom, a market where asset-liability matching issues are more extreme.

EURO GOVERNMENT BOND PRIMARY MARKET

The two main developments in the primary euro government bond market since the inception of the EMU have been the decline in the relative amount of government sector supply within the Eurozone bond market and the increase in competence of the euro debt agencies.

The healthy economic growth and fiscal consolidation seen in the Eurozone in the late 1990s helped to reduce debt-to-GDP ratios in this period, despite an increase in gross terms. This decline was especially obvious in the Mediterranean countries, whose deficit- and debt-to-GDP ratios fell significantly in the second half of the 1990s under the constraints of the *Maastricht Treaty* criteria. Helped also by the sale of third-generation telephone licences (UMTS) in 2000, some of these countries had to undertake buyback programs and/or bond exchanges to be able to provide liquidity to their markets amid their declining funding needs.

Subsequently, the deceleration in growth in the early 2000s took some of these deficit and debt ratios higher, even causing some rating downgrades (Italy and Greece) and showing the pro-cyclical nature of these countries' funding needs in both absolute and relative terms. It also showed how, in general, within a monetary union, growth is good for a specific bond market—especially in a relatively small country—as its effect in terms of reducing funding needs more than offsets the possible rate increase caused by inflation expectations floating higher.

The decline in the amount of government bonds being issued in the late 1990s was partially offset by a sharp increase in corporate bond supply, especially within the high-grade spectrum. Yet, despite its significant growth, this market is still very far from the government market in terms of bonds outstanding.

The broadening of the investor base prompted by the start of the EMU brought about a significant increase in competition between the various Eurozone sovereign issuers, magnified by the single currency and the small difference in the credit risk components of these similarly rated countries. If, before EMU, the currency risk had helped these borrowers to ensure a quasi-monopoly situation in their own markets, with the appearance of the euro currency, all these treasuries had to compete for the same pool of funds. This increase in competition forced the euro debt agencies to improve their transparency, predictability, and relationship with market participants.

Another important factor bought about by the EMU was the standardization of the bond markets, thanks to the beneficial effect on government debt of the exchangeability of that debt, thus increasing foreign investors' preference for these markets. To compete with other non-euro government markets, having a market as homogeneous as possible among all the different euro issuers was second to none. Accordingly, the euro treasuries increased their coordination in terms of the basic characteristics of their instruments, procedures, coupon calculation conventions (actual/actual), and even taxation. This subject has been studied by the Giovannini group for the European Commission, which produced several reports between 1997 and 2000 on the integration of the national treasuries and markets, giving some guidelines for better coordination of debt agencies with a view to achieving a better substitutability of bonds and a more efficient bond market.

The broadening of fixed income managers' mandates since the introduction of the single currency, the disappearance of foreign exchange risk for many investors, the redemption of long-held bonds, and the increase in exchangeability between these markets helped to increase significantly the percentage of sovereign debt held by nonresidents. As an example of a middle-sized market, the percentage of nonresident holdings of *Spanish bonos* increased from 20% before the EMU to well above 50% just four years later.

Another area on which the debt agencies had to increase their focus was their communications policy, as another of the obvious consequences of the above-mentioned loss of the domestic edge was the necessary increase in transparency and predictability, especially in terms of issuance policy. In fact, most euro debt agencies now publish periodical supply calendars, providing as much detail as possible on amounts and maturities to be issued, as well as any other useful information on new bond lines, swap operations, average duration targets, and the like. This information is shared with their respective market makers, and also via periodical bulletins and their web sites or pages on financial news services, such as Bloomberg or Reuters.

Measures to Improve Market Liquidity

Besides this improvement in information provided to the market, the above-mentioned increase in competition has made the euro debt agencies improve as much as possible the liquidity of their bonds. Liquidity and credit ratings are the key drivers of the relative performance of euro countries' bonds and, therefore, the debt agencies will try to improve their bonds' liquidity to decrease their funding costs. In the primary market, this increase in liquidity has been key, as explained below.

Bond Auctions: Sizes, Maturities, and Types of Bond

The broadening of the investor base, together with the desire to enhance liquidity in the secondary market, has been the main driver of the continuous increase in not only the size of bond issuance outstanding, but also the amounts offered at each auction.

This has been more obvious in the largest euro countries, the clearest example being the euro benchmark government bond, the 10-year Bund. In fact, those German 10-year bonds issued in 1998–1999 had an average outstanding value of around €10 billion, but their size increased with the arrival of the euro to reach as much as €27 billion outstanding by 2002, stabilizing thereafter at around €25 billion. In addition, most *Italian BTPs* now reach outstanding amounts of more than €20 billion, while the average size of a *French OAT* is between €15 billion and €20 billion. Accordingly, these outstanding euro government bonds have become much closer to their U.S. counterparts, as some Treasuries reach the $35 billion level.

Although less extreme, a similar pattern has been observed, not only in other maturities of the German curve (current five-year *OBLs* total €20 billion, whereas the pre-EMU ones were between €5 billion and €8 billion), but also in practically all other euro countries. This increase in auctioned and outstanding sizes has been even more dramatic in the smaller countries.

The smaller euro countries, because of their smaller nominal funding needs, used to issue a large number of small bonds before the currency union. Yet the outstanding size of many of the bonds (many of them below €2 billion) did not reach sufficient levels to be considered a liquid asset in which investors could trade large amounts without significantly affecting its price. Therefore, these countries have had to concentrate most of their supply into just a few bonds a year, sometimes having to carry out exchange auctions or buybacks to reach this critical mass. This situation was even more extreme in the high-growth late 1990s period and in the fiscally stricter countries. Nowadays, practically only the euro G4 countries and Greece issue bonds across the entire yield curve, while the rest of the euro countries just launch a couple of bonds every year, tapping them afterwards to reach a minimum amount.

The level that could be considered a minimum for liquidity purposes could be the €5 billion MTS threshold. Below this level, bonds are considered too easy to squeeze and, therefore, their liquidity is much lower, creating a sort of vicious circle. This €5 billion level is actually the target many smaller euro countries have when they launch a new bond, especially when they are issued via syndicate. Otherwise, they tend to try to reach this amount as quickly as possible.

Exchange Auctions and Buybacks

To reach this minimum amount as soon as possible, to reduce the level of their liabilities, smooth their debt's redemption profile, or improve the liquidity of selected issues, many European debt agencies carry out bond exchange auctions and/or buybacks. These operations are even more important for those countries that, due to their small size or strict fiscal policy, have low funding needs.

Bond Exchanges

The bond exchange procedure has been used profusely by many euro debt agencies, such as Spain, France, Italy, Portugal, and Belgium, which have been carrying out frequent bond exchange auctions for many years, either as one-off operations or by opening exchange windows during a specific period of time. As mentioned above, the main target of these exchanges is to provide liquidity to the new bonds as quickly as possible.

These operations have normally been concentrated in the last months of the year, as in a declining rate environment, exchanging old (that is, high-coupon) bonds for new, lower-coupon bonds has a cost because of the difference in price. Accordingly, these debt agencies tend to wait to have as much information as possible on the evolution of their countries' fiscal deficits in order to evaluate the amount of cash they can allocate to these operations. In general terms, these operations are well perceived by the market, as they allow investors to exchange their old, less liquid bonds for the new benchmarks. On top of this, the debt agency can increase the liquidity of its new benchmarks more rapidly than it otherwise could.

Bond Buybacks

The rationale behind bond buybacks is very similar to that behind the exchange auctions (that is, to increase the country's funding needs to allow larger—and faster—issuance of the current benchmark bonds). In fact, a buyback is just the first leg of an exchange auction, the other being the actual bond issuance. The main difference is that buybacks tend to be concentrated in short maturity bonds, thus helping to smooth the redemption profile by limiting upcoming years' redemption payments and, therefore, supply. The procedure for these buybacks could either be via OTC purchases or preannounced buyback windows, normally restricted to *primary dealers.*

Other Key Characteristics of the Primary Markets

Other key features of the primary markets include (1) issuance maturities and techniques, (2) issuing procedure, and (3) primary dealers. Each characteristic is discussed below.

Issuance Maturities and Type of Bonds

Although the introduction of the euro helped to homogenize some characteristics and maturities of the bonds issued, there are still some differences between the euro countries' supplied assets. Euro-denominated fixed-coupon bonds make up the bulk of issuance, but there are also some other types of bond issued by the Eurozone countries.

In general terms, the maturities issued are split between the short-end (two- and three-year), the intermediate sector (five-year), the long-end (10-year), and ultra-long-end bonds. Within this sector, the most frequently tapped maturity used to be the 30-year sector, although some countries also tap their 15-year bonds. In addition, since 2005, some euro countries have started to issue 50-year bonds, because of the low rate environment, the flatness of the long end of the curve—and therefore the low level of the forwards—and the increase in long-dated demand by pension funds and insurance companies, trying to improve the asset-liability match of their portfolios in an increasingly more regulated environment.

Most of these bonds normally pay fixed-rate coupons, the main exception being Italian CCTs, which have a seven-year maturity and pay a floating coupon related to the yield of the Italian six-month Treasury bills. Floating-rate note supply has fallen significantly since 1998–1999, although some countries still issue a small part of their supply in floating-rate notes. Another noticeable exception to fixed-coupon issuance is French TECs. These bonds' coupons, paid on a quarterly basis, are linked to the Tec10 index, an average yield of OATs with a constant maturity of 10 years. Yet their supply has also decreased significantly over the last few years.

Finally, one sector that continues to gain importance, not only in terms of amounts issued, but also investor interest, is the *inflation-linked bond market.* Since 2004, Germany, Greece, and Italy have joined France in issuing this type of asset. The sector continues to gain relevance, and its outstanding issuance is already above €130 billion in France, over 15% of total French debt outstanding.

Issuing Procedure: Syndication versus Auctions

Because of some Eurozone countries' relatively low funding needs and due to the increase in competition for investor preference (and to achieve the above-mentioned critical mass), many countries are increasingly launching their new bonds via *syndication.* This method, used by most national treasuries and debt agencies, allows them to allocate large sums in one go (€5 billion is the usual amount) and reach a broader base of final investors, facilitating the good performance of the bonds after launch. These syndicate issues, also used by quasi-sovereign issuers, such as the EIB, Federal Home Loan Mortgage Corporation (FHLMC), or Kreditanstalt Für Wiederaufbau (KfW), tend to be followed by subsequent taps.

In these syndicate issues, the borrower tends to name several (three to four) lead managers who would allocate most of the expected amount to be issued, with a co-lead group allocating the rest of this target amount. The lead group would, in normal terms, be formed by domestic and foreign banks, usually primary dealers in that market.

Primary Dealers

To ensure the good performance of their bond auctions and regular pricing of their bonds, the government debt agencies establish a group of primary dealers for their bond markets. In general terms, these institutions (normally investment banks) will have to bid in the auctions and quote a certain number of bonds with a maximum predetermined bid-offer spread. However, these banks have access to the second round of the auctions (under better conditions) and should be the main beneficiaries of other deals in these Treasuries, such as swap operations or the above-mentioned syndicate issuance.

SECONDARY MARKET AND INTRA-EURO SPREAD DETERMINANTS

In general terms, within a Monetary Union, the spreads between same maturity bonds from different countries should be determined by the relative liquidity of these bonds and their credit status.

With this in mind, yield differences among Eurozone countries should tend to diminish and almost disappear in the long run. On the one hand, the decline in these countries' financing needs as they strengthen their fiscal positions, forced by the SGP, tends to make their credit ratings converge, albeit slowly. On the other hand, the smaller countries, helped by a broader investor base within the single currency and the above-mentioned enhanced supply mechanisms and *trading platforms,* should see improved liquidity of their bonds, helping to diminish the liquidity component of their spreads to the core euro countries. This reduction in the liquidity premium and the relative creditworthiness of the Eurozone countries should make bond spreads converge in the long run.

Yet it should be taken into account that a large part of this convergence had already taken place before the actual start of the EMU (see Figure 25.1). Once the market had priced in a significant probability of a country qualifying for entry into the euro, investors could put on convergence trades, tightening significantly the peripheral spreads to the euro core countries. These trades had a limited risk, as in most cases the final exchange rate parities were already known (mid-rate of the previous exchange rate mechanism, or ERM, bands).

Sovereign Credit Ratings

Credit-rating agencies (CRAs) try to encapsulate in the qualifications they assign to different sovereign issuers the financial and economic conditions of a specific country, as well as its ability and willingness to pay its obligations. These ratings should, therefore, theoretically, be a good indicator of the financial health of the issuer and should be correlated to the yields and spreads within the Eurozone, as they should measure, to a certain extent, the borrowers' small but positive default probabilities.

It is also worth remembering that although the euro is their domestic currency, euro countries do not have the ability to unilaterally print money anymore and, therefore, the ratings these countries were assigned at the beginning of currency union equal their former foreign currency ratings as opposed to their domestic currency ones, which were better because of their ability to print their own money.

Before January 1999, four of the countries in the euro area already deserved the highest credit rating, according to the major three CRAs (Germany, France, the Netherlands, and Austria). From the start of currency union, three more countries have joined the top-notch club, namely Ireland

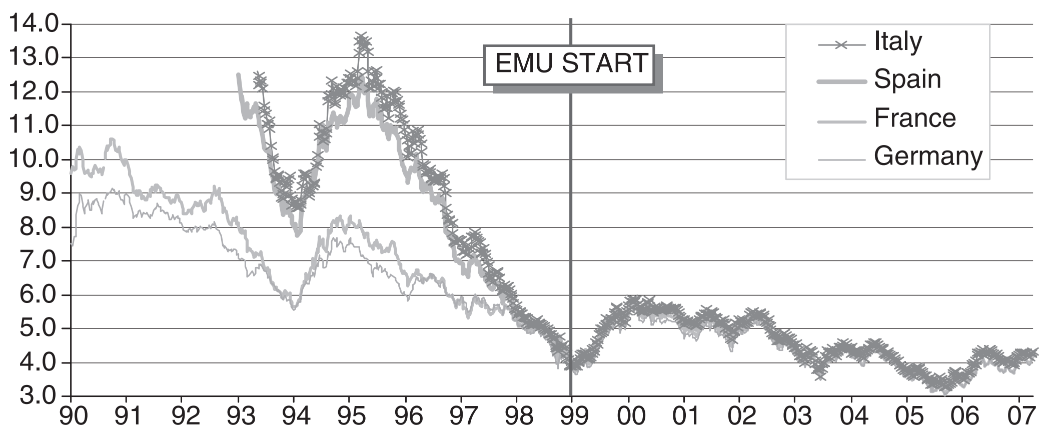


Figure 25.1 German, French, Spanish, and Italian 10-Year Rates Converged at the Beginning of the EMU.
Source: Created from data obtained from Bloomberg.

Table 25.1 Euro Countries' Credit Rating (December 2007)

	Moody's	S&P	Fitch	Last change (Post EMU)	
Germany	Aaa	AAA	AAA		
France	Aaa	AAA	AAA		
Netherlands	Aaa	AAA	AAA		
Austria	Aaa	AAA	AAA		
Ireland	Aaa	AAA	AAA	Oct-01	Upgrade
Finland	Aaa	AAA	AAA	Feb-02	Upgrade
Spain	Aaa	AAA	AAA	Dec-04	Upgrade
Belgium	Aa1	AA+	AA+		
Portugal	Aa2	AA−	AA	Jun-05	Downgrade
Italy	Aa2	A+	AA−	Oct-06	Downgrade
Greece	A1	A	A	Nov-04	Downgrade

Ratings obtained from the respective rating agencies.

(October 2001, S&P), Finland (February 2002, S&P), and Spain (December 2004, S&P). The rest of the countries are still below this category, with Greece being the lowest-rated country in the region (in the EMU since 2001). As seen in Table 25.1, there are no significant divergences between the ratings these three agencies assign to each specific country, although S&P and Fitch appear to be slightly stricter than Moody's in this regard.

As these ratings reflect the ability and willingness of the countries to assume their obligations and, taking it to the extreme, their probability to default, there should be a direct relationship between the countries' ratings and their yields (or spreads to benchmark curve). This relationship is clearly shown in Figure 25.2, which represents each country's rating versus its average 10-year yield spread versus Germany in the five first years of currency union. It seems clear from the exhibit that there is an almost linear relationship between spreads and ratings, with the distance between each country's spread to the regression line being a proxy of each country's liquidity premium. This "liquidity premium" is more evident in the AAA-rated category, where the market clearly differentiated between very liquid and deep markets, such as France, and smaller, less liquid countries, such as Austria. That said, most AAA-rated euro countries now trade very closely to each other, with their spreads normally within 5 basis points of each other for the same maturity.

Changes in the rating of any of these euro countries should, in theory, make its bonds under or outperform the rest of the markets, as was the case when Moody's upgraded the Kingdom of Spain by two notches to Aaa in December 2001. Yet, most of the time, these rating changes have been largely anticipated by the market, either due to the improvement in that country's official rating outlook or just based on previous comments or reports from these agencies. In fact, some well anticipated downgrades, such as the Italian one in July 2004, hardly had any market impact, as investors had been wary of holding large amounts of Italian BTPs prior to the well touted downgrade, and the actual cut to AA– was seen as an all-clear sign for investors who were underweight Italian debt in their portfolios to add some extra yield.

Other Intra-Euro Bond Spread Drivers

Credit ratings and the size and liquidity of each bond market are the main long-term drivers of intra-euro government bond spreads. Yet there are many other smaller and more micro spread drivers that are becoming increasingly more relevant, thanks to the above-mentioned credit and liquidity convergence among these countries.

Supply Dynamics, Fiscal Trends, and Issuance Policy

Although credit rating and fiscal outlook are by far the two most important spread drivers in the Eurozone, the extent of the market impact of these fiscal features depends significantly on the assets chosen to fund those needs. Fiscal

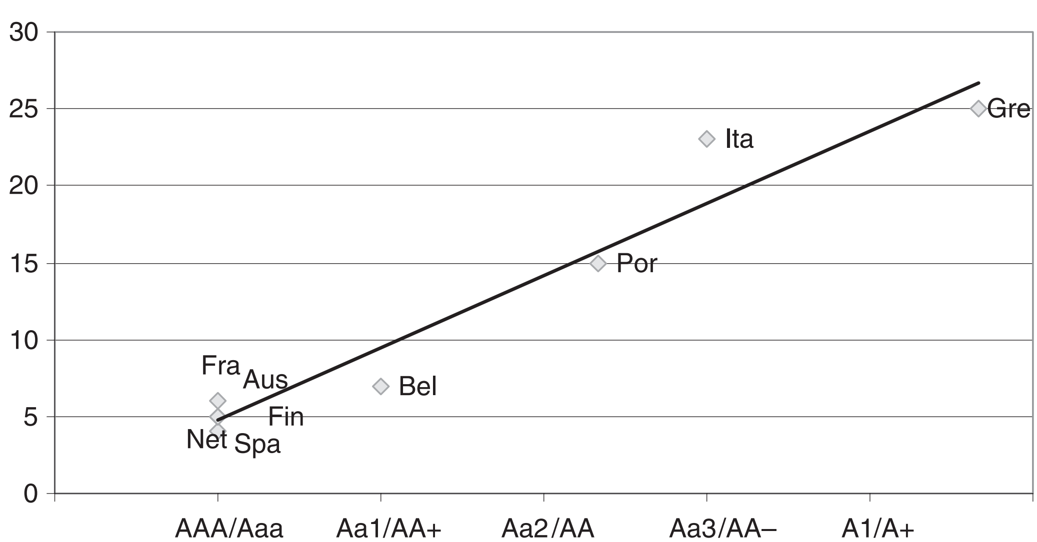


Figure 25.2 10-Year Spread to Germany (Average in the First Five Years of EMU) versus Credit Rating.
Source: Data obtained from Merrill Lynch and BBG.

needs have a noticeable impact on bond markets when these gaps are funded using government bonds, while their market impact is much more limited if this funding is obtained from other sources, such as Treasury bills, loans, privatizations, and so on. Another factor to bear in mind is that these issued amounts are relevant not only in gross, but also in net terms (*ex*-redemptions) as it is this second amount that better reflects each country's financial needs. In addition, it can be assumed that a sizeable part of the bonds being paid down (or bought back) are reinvested in the same market so as to keep unchanged the country composition of the portfolio, helping this market to outperform the rest of its euro counterparts, with a similar effect taking place for coupon payments.

The breakdown of these countries' funding by maturity and type of asset is also affected by market dynamics, as, for instance, steep yield curves favor the increase in Treasury bill issuance (2000–2001), while low-rate and flat yield-curve environments make long-end issuance more interesting, locking in low funding levels for long periods.

The maturity breakdown of government bond issuance can also be key in determining euro government spreads. Accordingly, the announcement of an unexpected supply increase (or decrease) in a specific maturity can significantly affect the spreads and slope of the euro curve. This feature was clearly seen at the end of 2001, when the German debt agency announced its intention to issue just €6 billion in 30-year Bunds in 2002, considerably below market expectations. The initial reaction was not only clear outperformance of the German long end, but also sizeable flattening in the 30-/10-year slope and a widening of long German *swap spreads.* These dynamics underline the importance of accurate forecasting of the amounts and maturity breakdown of each country's upcoming supply. On top of this, when a bond auction takes place, the actual increase in the amount of paper in the market may affect its price simply due to supply-demand conditions, although such an impact can depend on the market conditions of that moment.

Bond Index Tracking and Passive Fund Management

As in many other financial markets, many fixed income fund managers measure their performance against bond indices, made up of the most liquid bonds in each market. So, any noticeable deviation in the characteristics of the managed portfolio from the index tracked means a risk for the asset manager. Therefore, these indexed funds tend to track (although to a different degree, depending on the risk characteristics of the portfolio) the evolution of the indices. In fact, the most passive funds managers try to minimize their *tracking error* by replicating dynamically the characteristics of the index in terms of average duration and country breakdown.

Accordingly, index-tracking fund managers have to anticipate any possible change in these indices to avoid increases in their tracking errors. The indices are usually rebalanced at the end of each month according to the bonds entering or leaving the index, with those months with heavy long-term supply and/or large drops from the index producing significant changes in index duration at month end. Indexed investors, therefore, have to buy or sell bonds around those days to match these duration changes. To minimize tracking error further, these managers have to make their adjustments at the same time as the index is rebalanced, with the obvious consequences for the bond market around that period.

Bond Future Deliverability

Bond futures have become, due to their liquidity and leverage characteristics, the main hedging and investment instruments of many market participants. Their open interest and traded volumes have, therefore, increased sharply in the last few years. As the underlying issues of these futures are specific government bonds, these bonds tend to follow a similar evolution to the future they represent. Accordingly, the bonds included in an exchange-traded future deliverable basket and, especially, the *cheapest to deliver (CTD)* tend to trade rich in their own curve, thanks to the large amount of long and short positions in the future, as well as the possibility of squeezes in the delivery dates.

The degree of its dearness will depend, among other factors, on the outstanding amount of the bond, the open interest of the future, bond-market volatility and the bond's supply dynamics. As discussed next, Eurex's victory in the Eurozone "battle of the futures" has made German deliverable (and CTD) bonds trade richer than other German and euro bonds in their respective maturities.

Bond Swap Spreads and Their Relationship to Peripheral Spreads

The evolution of euro government bond peripheral spreads has always been linked to the performance of swap spreads (and vice versa). Yet this relationship should be taken with a pinch of salt, as, with the German rate on both sides of the equation, any spike in the German Bund market will make this correlation increase spuriously.

That said, there are two reasons why the performance of German swap spreads are related to euro peripheral spreads. The first is that, flows apart, the bond-swap spread reflects the yield differential between a government rate and the composition of a string of Euribor rates (that is, a swap fixed rate). As the average credit quality of the banks in the *Euribor panel* is A to AA, any increase in investor preference for credit quality will make both swap and peripheral spreads widen versus the core euro government rate, thus increasing the correlation between both differentials. Yet, this increase in the correlation is mainly due to the outperformance of the benchmark asset (German bonds in this case) rather than to any similarity between the swap rate and that of the peripheral country.

Market Volatility as a Spread Driver

One recent driver of peripheral versus core spreads has been the sharp decline in financial-market volatility in 2004 through 2007, mainly as a result of abundant global liquidity, as well as the increasing efficiency and transparency of central banks. In this low-volatility environment, the search for any yield pickup becomes crucial, and the extra yield offered by the high-yielding countries becomes even more interesting. Taking it to the extreme, in a world where spread volatility disappears, the yield pickup offered by the peripheral countries becomes a free lunch for investors—even more considering that all euro countries enjoy the same status in terms of eligibility for *repo* operations with the European Central Bank. Accordingly, any model trying to forecast, for instance, German-Italian yield spreads—based, for example, on Bund swap spreads—would need to incorporate the decline in rate volatility to justify the decline in these differentials.

Other Related Markets

Government bond markets are closely related to other fixed income assets and interest rate and bond futures. This market is also increasingly related to the interest rate swap market.

Wholesale Electronic Markets and Trading Platforms

One of the most significant developments since the start of the euro has been the success of *EuroMTS*, an electronic broking system launched in April 1999. Before 1999, most bond markets were telephone based, but this platform has expanded rapidly to cover practically all the government markets and its market share has expanded significantly. The success of these trading platforms has been favored by the broadening of this market with the start of the currency union and they have become very important in increasing investor confidence, market liquidity, and price transparency. This increase in platform trading has not only taken place in the Eurozone, being has also been the case in the United States and other bond markets.

The other advance in bond trading has been dealer-to-customer platforms, where institutions can compare prices from several intermediaries simultaneously, with the obvious benefit for final investors.

Strip Markets

Many euro government bonds can be stripped, breaking them down into the single payments they involve, that is, one flow for each remaining coupon payment and another for the principal. With this procedure, an n-year maturity coupon-bearing bond is transformed into $n + 1$ *strips* (zero-coupon bonds), which can be traded separately in the market. Yet this market is much less liquid in the Eurozone than in the United States.

Repo Markets

Despite the homogenization of euro government bond markets, repo markets have remained largely domestic and unevenly developed throughout the single-currency area, showing hardly any increase in cross-border transactions. Regulatory, legal, and tax-specific issues, as well as different market practices, have been the main reason for the lack of a truly unified repo market in the euro area.

Euro Futures and Options Market

The large increase in the size and number of investors in the euro government bond market has brought about a significant improvement in the depth and liquidity of the bond futures market. In fact, since 1999, Eurex has continued to confirm its status as the most active derivatives exchange globally, ahead of the Chicago Board of Trade (CBOT), while the Bund contract has established itself firmly as the most actively traded futures contract in the world. This 10-year bond-based future is actually the most widely used hedging instrument for all euro-denominated issues.

In this regard, the winner-takes-all characteristic of a futures market (where liquidity is key) sparked a dispute between the Eurex and Matif futures exchanges in the initial years of the monetary union. While their characteristics were very similar (it could even be argued that Matif's future coupon was closer to existing bond yields), the winner of this battle appears to have been the Eurex future, becoming the main reference for all maturities (10-year Bunds, five-year Bobl, two-year Schatz, and even the 30-year Buxl). These contracts include only German bonds in the deliverable baskets, helping to keep this country's deliverable bonds more expensive than the other euro countries, helped also by the existence of an options market linked to these futures.

Interest Rate Swaps as the Benchmark Curve for Eurozone Government Bonds

Given the absence of a single, clearly defined benchmark sovereign yield curve and the continuous expansion of the interest rate swap market since the late 1990s, government bond market participants have increased the use of the swap curve as a reference for the valuation (and hedge) of government and nonsovereign bonds. Another factor that has enhanced the depth and liquidity of the swap market is the enlargement of the Eurozone corporate bond market, as both investors and issuers can use swaps to convert their fixed-rate liabilities into floating-rate ones, or vice versa.

It has, in fact, been argued that interest-rate swaps could eventually replace government bonds in many of their functions, such as extracting information on the future path of short-term rates, or hedging interest-rate risks, their also being a more homogeneous asset. Yet it should not be forgotten that government bonds will remain the key funding vehicle for these sovereign issuers and that their significantly lower credit risk makes these assets a cleaner tool for assessing future rate changes and hedging interest rate risks, while they are the perfect candidate for performing the function of collateral.

SUMMARY

In this chapter, we first analyzed the substantial growth in the Eurozone government bond market and the changes brought about by the European Monetary Union. We then focused on the main drivers of interest rate differentials between these countries, namely, credit and liquidity, as well as financial market volatility. Finally, we focused on integration and the continued differences between the region's issuers, as well as related markets: strips, futures, repos, and swaps.

REFERENCES

European Commission, Maastricht criteria. (1992). *European Community Treaty, Article* 121 (1), February.

Bank of Spain. (2001). *The euro-area government securities markets*. Bank of Spain working paper 0120, October.

Giovannini Group. (2000). *Report on co-ordinated issuance of public debt in the euro area*, November.

Bank for International Settlements. (2006). *Triennial Central Bank survey of foreign exchange and derivatives market activity*, September.

Standard & Poor's. (2006). *Ratings definitions*, December, update.

CHAPTER 26

The German Pfandbrief and European Covered Bonds Market

GRAHAM "HARRY" CROSS
Financial Tutor, 7city Learning

Abstract: Covered bonds are created through a securitization process. The collateral is commercial and residential mortgage loans and/or public sector assets. Unlike an asset-backed or mortgage-backed securities, the investor gets dual protection in the form of a claim against the issuer and a preferential claim over the cover pool in the case of the insolvency of the issuer. There are other differences. The covered bond market is dominated by the German Pfandbriefe. Within the German bond market, these bonds represent more than about one third of all German bonds and the largest uniform asset class within the European market. Globally, it is the largest bond market after the U.S. debt market.

Keywords: covered bonds, Pfandbrief, jumbo Pfandbriefe, traditional Pfandbriefe, structured Pfandbriefe, global Pfandbriefe, Pfandbrief bank, mortgage bank, Obligations Foncières (OFs), Cédulas Hipotecarias, Lettres de Gage, Irish Asset Covered Securities

This chapter describes the German mortgage bond or *Pfandbrief* market, its institutions, and working practice. We also consider other aspects of the European covered bond market. The instruments themselves are essentially plain vanilla bonds, and while they can be analyzed in similar ways to U.S. agency bonds and mortgage-backed bonds, there are also key differences between them, which we highlight in this chapter.

THE PFANDBRIEF MARKET

Pfandbriefe are bonds issued by institutions, which are subject to special governing legislation. These bonds are "covered" or backed by underlying asset pools, equating to at least the same nominal amount of the issue. The assets contained within these pools must be recorded into a cover register, maintained by the *Pfandbrief bank*, to ensure that these are easily identifiable. In this regard, *covered bonds* such as Pfandbriefe are considered highly secure. In the event of the issuer becoming insolvent, the creditors would receive a preferential claim over the assets in the cover pool, which is there solely to protect them.

Pfandbriefe are categorized into two types. Öffentliche Pfandbriefe, which are bonds fully collateralized by loans to public-sector entities (also known as "Public" Pfandbriefe), while Hypotheken Pfandbriefe (Mortgage

Pfandbriefe) are fully collateralized by residential and commercial mortgages, whose loan-to-value ratio must not exceed 60%. The former constitute just over 90% of the overall Pfandbrief market.

The market has become the largest asset class on the European bond market and is ranked the sixth largest in the world as of mid-2007. It is regulated within a stringent legal framework and is under special supervision. Supervision is conducted by the German Financial Supervisory Authority (BAFin). The Pfandbrief banks are, in addition to being bound under the terms of the German Banking Act (KWG), by which all German banks are governed, are also subject to the provisions of the Pfandbrief Act (PfandBG) 2005 as well. All of these factors have in the past assisted issuers in obtaining the highest possible ratings (AAA) for their Pfandbrief paper. This situation has in some cases, however, changed somewhat and this will be discussed in detail later.

Although the German Pfandbrief market has a history dating back well over 200 years, its recognition as an asset class by international investors has only occurred recently in the mid-1990s with the advent of the *jumbo Pfandbrief*. The name "jumbo" is derived from the large issue volume, with the size requirement of €500 million. This sector of the market, founded in the spring of 1995, and geared towards the liquidity criteria of large international investors, has managed to establish itself as Europe's fourth largest bond market. Prior to the arrival of the first issue of this nature, a DM 1 billion Frankfurter Hypothekenbank bond, the Pfandbrief market had been an illiquid and highly fragmented sector comprising of some 17,000 individual issues with a very small average volume of around €80 million. Investors in these "Traditional" Pfandbriefe were almost exclusively domestic.

In this light, the main focus of this chapter shall be on the jumbo sector as this has the most relevance to the investment community at large.

HISTORY OF THE PFANDBRIEF

The origins of the German Pfandbrief system are widely regarded to lie within the "cabinets-ordre" of Frederick II of Prussia, back in August 29, 1767—the basis of which concerned the introduction of the Pfandbrief system in an attempt to remedy the aristocrats' shortage of credit in the areas of Prussia that had been ravaged during the Seven Years War (1756–1763).

On the basis of this royal decree, the Silesian Landschaft, an association of estates belonging to the aristocracy, churches, and monasteries, was set up in 1770. In time, more of these cooperations were set up throughout the individual provinces of Prussia, as compulsory law associations to the aristocratic landowners. These so-called "Landschaften" facilitated the refinancing of loans to their members by issuing debentures. Purchase of this paper ensured the creditor acquired a direct charge over the estate, which the landowner had put up as collateral. In the event of default, the estate named in the Pfandbrief, the Landschaft and all of the landowners belonging to the Landschaft served the Pfandbrief holder as security. Understandably, this paper was also known as "estate Pfandbrief" and largely corresponds with today's Mortgage Pfandbrief.

The Pfandbrief system rapidly gained popularity throughout Europe and the development of the present day format was given a decisive boost from the foundation of organizations outside of Prussia, such as Crédit Foncier de France in 1852. Issuers of this second-generation Pfandbrief were not law associations but private real estate credit institutions, which adopted the system for the refinancing of loans to the public-sector borrowers and loans guaranteed by public-sector institutions and agencies (public-sector loans).

Whereas in the early days Pfandbriefe were used to finance agriculture, this new variation was used to finance the then rapidly expanding towns and cities of Europe. In the latter half of the nineteenth century, one of the major priorities facing European governments was the provision of housing to meet the widespread exodus from rural areas and the corresponding growth in urban population levels. Concentration from the outset was on real estate financing and, above all, the financing of construction of housing and commercial properties. In this respect, today's mortgage banks were among Europe's first large-scale financial intermediaries and can very much be regarded as a by product of the industrial age.

The first German mortgage bank of the type familiar today was established, by the decree of the senate, on December 8, 1962 (Frankfurter Hypothekenbank in Frankfurt). From this moment, numerous other mortgage banks emerged in quick succession in almost all of the German federal states until, by the beginning of the twentieth century, a total of 40 private mortgage banks existed. Throughout the ensuing years of economic "boom and bust," the business sector occupied by these real estate credit institutions, understandably, became one of the biggest sectors in banking.

These developments led to the promulgation of the German Mortgage Banks Act (Hypothekenbankgesetz–HBG) of 1900, which was the first uniform law in the field of banking for the entire German Reich. This Act provided a legally prescribed, uniform organization framework for this group of institutions that has stood the test of time, right up until the present day.

The new generation of Pfandbrief had spread across Europe from France, through Germany to the United Kingdom, Italy, and Spain among others. In this respect, it is interesting to note that in the annex to the preamble to the Act contained the laws from Germany's neighboring countries, evidencing the influence of foreign laws on the lawmakers of Germany. However, during the twentieth century with the onset of two world wars, global economic crisis, inflation and the currency reform in 1948 resulted in a curbing of cross border influence. This in turn caused the mortgage banks throughout Europe to develop in sharply divergent ways. Some countries chose to abandon the whole Pfandbrief concept altogether, whereas others turned the mortgage banks into state monopoly institutions.

In Germany, no other group in the whole of the banking sector was as impacted by these factors as the mortgage banks, which had seen their business volumes fall drastically by the time of the currency reform. Nevertheless, Pfandbriefe proved an invaluable tool in the reconstruction programs that were set up to deal with the aftermath of the war and their popularity grew with each successive decade. The reunification in Germany after the fall of the Berlin Wall highlights this resurgence, as a demand for both commercial and residential property construction as well as for public infrastructure renewal had to be met in the new federal states in East Germany. With the advent of the euro and amendments that were made to the German Mortgage Bank Act, new avenues of cross-border lending were opened up in Germany for the mortgage banks. They now had the ability to market Pfandbriefe internationally.

The market has grown considerably from the lowly position it found itself in, midcentury, a period when mortgage banks reported business volumes down to levels of 5% of those quoted just 30 years earlier, to it current status as one of the largest bond markets in the world.

KEY FEATURES OF INVESTOR INTEREST

Reduction of Credit Risk

The tight legal framework within which the participants of the Pfandbrief market must operate is one of the foremost reasons why Pfandbriefe appeal to both domestic and international investors. In addition to being bound by the general provisions set out in the German Banking Act (KWG), the law by which all German banks are governed, German Pfandbrief banks are also subject to the requirements of the Pfandbrief Act.

The Pfandbrief Act (PfandBG), which came into force July 2005, superseded the Hypothekenbankgesetz (Mortgage Bank Act). Under the old legislation, mortgage banks were bound by the specialist bank principle and were permitted to engage only in public sector and mortgage lending activities. However, with the inception of the new Act, any institution may now issue Pfandbriefe provided it has core capital of at least €25 million and meets the requirements set forth under the Act with regard to the management, monitoring, and control of risks. Furthermore, in order to engage in Pfandbrief business a license is needed from the Federal Financial Supervisory Authority (BaFin) and in doing so, the institution must submit to BaFin a business plan stating that it intends to engage in Pfandbrief business on a regular and sustained basis. Only real estate–secured mortgage loans are eligible as cover assets for mortgage Pfandbriefe and the property serving as cover must be located in an European Economic Area (EEA) state, the United States, Canada, Japan, or Switzerland. A further precondition for inclusion in the mortgage cover pool is proof that the Pfandbrief bank has the necessary expertise in the respective market. For public Pfandbriefe, loans to European Union (EU) member states, the other G7 states, and Switzerland as well as to their regional and local authorities qualify as cover assets. Loans granted to other European Organization for Economic Co-operation and Development (OECD) states may also serve as cover.

In addition to what could be considered low risk fields of activity; a strict regional restriction is added in order to reduce risks in connection with cross-border business. Under this restriction, loans may only be granted to borrowers situated within the member states of the European Union, the EEA, the European OECD countries as well as the non-European G7 countries.

Further security is provided to the investor by the fact that Pfandbrief bonds are required to be covered by assets, which have at least the same value and bear the same interest rate. Moreover, Pfandbrief banks are required under the Net Present Value Directive to keep excess cover of a least 2% of the volume of Pfandbriefe outstanding in the cover pools.

It is necessary for these underlying assets to be segregated into two separate cover pools, one for mortgage loans and the other public sector loans, thus reflecting the two types of business within which the mortgage banks are involved. In the case of mortgage Pfandbriefe, covering assets are "first-charge" mortgages.

In the event of a Pfandbrief bank's becoming insolvent, the Pfandbrief creditor would receive a preferential claim over the assets in the respective cover pool, which is there solely to protect them. They would not be required to participate in the insolvency procedures, but instead have any claim satisfied on schedule in accordance with the terms of the respective issue out of the cover assets. However, if the claim cannot be satisfied on time, in respect of coupon payments and redemptions because the cover pool is insolvent, separate proceedings will then commence in regard to the pool affected.

The Pfandbrief legislation contains further protective measures to safeguard investors in mortgage Pfandbriefe. Namely, a limit imposed on those mortgages being used as cover to a maximum of 60% of the "prudently" calculated mortgage lending value. This provides a safety cushion against the potential cyclical fluctuations in the market value of the cover pool asset.

The comparatively low risk that a portfolio of both residential and commercial mortgages entails is also expressed in the equity weighting of 50% for mortgage loans with a lending limit of up to 60%.

These elements obviously offer exceptional safety to investors in the Pfandbrief market and should therefore limit the impact of any adverse market movements on the back of any detrimental news in regard to the parent companies.

Liquidity

The jumbo Pfandbrief market, on its own, is Europe's fourth largest bond market, surpassed only by the government markets of Italy, Germany, and France. The name is derived from the large issue volume, with the size requirement of €500 million. In comparison, the average size of the *traditional Pfandbrief* is approximately €150 million, which tends to prohibit the trading-oriented investor and favor the "buy-and-hold" types.

The minimum issue size requirement for the jumbos is only of theoretical significance as the majority of issues are launched in considerably larger sizes. Indeed, since August 2006, the minimum issue size is €1 billion. It can therefore be seen that the volume of jumbo Pfandbriefe is equal to that of the bonds brought by medium-sized sovereign issuers within the Eurozone. The overall Pfandbrief market is the biggest bond market within Europe as of mid 2007. Of course, this figure includes *structured Pfandbriefe*, the smaller traditional variety, as well as the jumbo sector.

The market-making obligations further enhance the liquidity of the jumbo Pfandbriefe. Namely, that jumbo Pfandbriefe are syndicated by at least three market makers who pledge to quote two-way (bid/offer) prices simultaneously, for lots up to €15 million during the usual trading hours—9.00 A.M. to 5.00 P.M. (GMT + 1) for the life of the issue. The issuer itself may also perform the function of market maker and should obtain an undertaking from the assigned market makers not to exceed the following bid/offer spreads when quoting:

Up to and including 4 years	—	5 cents
Over 4 years up to and including 6 years	—	6 cents
Over 6 years up to and including 8 years	—	8 cents
Over 8 years up to and including 15 years	—	10 cents
Over 15 years up to and including 20 years	—	15 cents
Over 20 years	—	25 cents

The maximum bid/offer spread is adjusted according to the remaining life of the bond.

There are further nuances to the market that should be noted. Admission to either the official or the regulated market at one German stock exchange is compulsory for jumbo Pfandbriefe; an official listing must be obtained immediately after issue or not later than 30 days after the settlement date. However, only a fraction of Pfandbrief trading is settled through the stock exchange. By far the greater share of trading is executed off the floor, for the most part via the telephone or, to an ever-greater extent, through the numerous electronic trading systems including the EuroCreditMTS. To be eligible to trade on this platform, bonds must fulfill stringent credit criteria. They must have a triple-A rating from either Moody's or S&P and a minimum volume outstanding of €3 billion. Jumbo Pfandbriefe are responsible for more than 80% of the issues traded on EuroCreditMTS.

Over and above this, there are certain recommendations in place regarding the issuance of jumbos:

- The coupon should be expressed in fractions of not less than a quarter percentage point.
- In the event of an issue's being tapped, the tap amount should not be less than €125 million per add-on.
- In the case of new issues or taps, a maximum of five days should separate pricing date and settlement date.

In addition, all jumbo Pfandbriefe with a volume outstanding of €1.25 billion or greater and with a residual life of more than two years are greatly assisted by the market making pledge given by 17 institutions to provide a repo market in these issues.

MARKET INSTRUMENTS

The Pfandbrief market is comprised of several types of issues; in addition to the aforementioned traditional and jumbo Pfandbriefe, there are global issues and a variety of structured issues and the latest enhancements to the product range by the way of medium-term note (MTN) and commercial paper (CP) programs.

As previously discussed, the major difference between traditional and jumbo Pfandbriefe is the issue volume. Further distinctions are also evident in the issuing procedures of the two. *Traditional Pfandbriefe* are brought to the market in tap form and individual series feature within one issue. Jumbos, however, are issued via syndicates using the fixed price reoffer method. To guarantee the liquidity, jumbos must have at least three market makers willing to make prices throughout normal trading hours. Some time ago, a book-building procedure with a premarketing phase was put in place, in line with standard practices within the international markets. A so-called "pot procedure," similar to the auction procedure, has been introduced as well. With this method, syndicate banks can put together an order book from which the respective issuer can decide on allocation. This places the issuer in a position to allot investor demand among the syndicate banks in the run-up to the issue, thus enabling greater control over the book and, of course, more precise pricing.

Traditional Pfandbriefe may be issued in either bearer or registered form, whereas jumbos are only issuable as bearer bonds. For several years now, there has been a considerable shift in favor of the bearer paper, an indication of the growing share of jumbo issues brought by the mortgages banks and their willingness to provide fungible bonds to their investors.

As a rule, Pfandbriefe are issued with maturities of 1 to 10 years, and currently the most predominant incidence of issuance occurs in the medium-term maturities of 5 to 7 years. However, this predominance has been on the wane over the past few years, and more and more bonds are appearing on the market with lives of less than 1 year or more than 10 years.

Global Pfandbriefe

Global Pfandbriefe issues are aimed specifically at the large financial centers around the world. For example, in order to facilitate investor access to the market, particularly in the United States, the first globals were issues almost exclusively in accordance with Securities and Exchange Commission (SEC) Rule 144a. This prevents the need for investors to go through the costly SEC registration procedure and avoids the need for annual accounts in line with U.S. accounting regulations. It does, however, restrict sales to so-called "qualified investors" with a portfolio of at least $100 million. A number of mortgage banks have gained a frequent issuer status in the United States, in accordance with Rule 12g 3-2 (b), which grants exemption from the extensive registration and reporting requirements. Under this rule, the publishing of a separate U.S. prospectus is not required; the standard documents presented in the issuer's home country are sufficient. Despite these helpful measures, the process of marketing Pfandbriefe in the United States is still very much in its infancy and the competition for the attention of investors is huge.

By definition, jumbo Pfandbriefe are always plain vanilla structures: jumbos are fixed-interest bullet bonds, the coupon on which is payable annually in arrears. The calculation of interest accrued is done uniformly using the actual/actual method in line with international practice. While this standardization helps to enhance the transparency of the market, it inhibits the ability for these issues to be targets to an investor's specific needs, and this is where the structured issues come into their own.

Structured Pfandbriefe

Aside from the traditional and jumbo Pfandbriefen, the mortgage banks also offer *structured Pfandbriefe* for those investors who seek a more individually tailored product to suit their portfolios. These products are structured to particularly suit the investors' interest rate expectations and their desired risk/return profiles. Structured Pfandbriefe allow the mortgage banks to combine the asset quality of the Pfandbrief with the advantages offered by derivatives.

MTN and CP Programs

A recent important addition to the range of refinancing tools has arrived in the form of MTN and CP programs. Pfandbriefe issued under these programs offer a greater range of maturities and can be denominated in different currencies. For the mortgage banks they offer a superior degree of flexibility in refinancing, as a variety of bonds can be issued as and when required. They offer a reduction in costs as the workload involved in issuance is much less, and, finally, they open the market to an increased range of investors with specific investment criteria.

Clearing

Transactions in Germany are usually settled through Clearstream Banking AG, Frankfurt, a subsidiary of Deutsche Börse AG, formed as a result of the merger of Deutsche Börse Clearing AG and Cedel International. The remainder are settled via Euroclear or Clearstream International.

KEY DIFFERENCES BETWEEN COVERED BONDS AND ABSs OR MBSs

While covered bonds are often regarded as similar to asset-backed securities (ABSs) and mortgage-backed securities (MBSs), many noteworthy differences exist between them:

- The assets behind the covered bonds assets remain on the originator's balance sheet, even though they may be maintained in distinct pools or lodged in special purpose affiliates. However, in the case of ABSs or MBSs, the assets are segregated from any other assets and are usually off balance sheet and placed in a special purpose vehicle (SPV).
- The covered bond issuer is the source of the principal and interest cash flows, whereas the actual assets provide those payments in the case of the ABS/MBS.
- In certain jurisdictions, covered bondholders have some recourse to "noneligible" assets and, in the case of the special purpose affiliates, may also rely on some form of parental support for the issuer. For ABSs/MBSs, in the event of insufficient proceeds from the pool assets to cover the claim, holders have no recourse above and beyond the collateral contained within the pools and the original ABS/MBS structure.
- Eligible assets for covered bonds are clearly defined by law and are substitutable. Therefore, the asset mix varies over time and is relatively heterogeneous. For ABSs/MBSs, the assets are of the originator's discretion and once the structure is finalized, no asset adjustments can generally be made. The mix of assets can usually be regarded as quite homogeneous.
- Asset quality is a measure of the strengths of the specific structure created for the ABS/MBS. However, it is a function of the issuer and underwriting standards of the covered bond, as well as the features of each issues framework.
- Covered bondholders, in the event of issuer insolvency and provided that the covering assets continue to meet regulator requirements, will still receive interest and principal payments according to the contractual dates (with the exception of Spain). However, certain credit events, such as deterioration in the quality of the underlying assets, would trigger the acceleration of ABS/MBS payments.

MARKET PARTICIPANTS

In the summer of 2005, The Association of German Pfandbrief Banks (Verband Deutscher Pfandbriefbanken, vdp) succeeded The Association of German Mortgage Banks (VDH), in line with the new Pfandbrief legislation. Its membership, understandably, increased in numbers as more banks fell under its remit. The vdp's members, coming from all German banking groups, rank among the most prolific providers of capital for residential and commercial properties as well as for the public sector and its institutions.

THE CREDIT RATING APPROACH TOWARDS PFANDBRIEFE

The two main international ratings agencies, Moody's and Standard & Poor's, adopt different methodologies when approaching Pfandbriefe, and this has caused some confusion among investors.

Moody's approach for Pfandbriefe and covered bonds in general is based on the so-called "joint-default analysis," which takes into account both the credit strength of the issuer and, on "issuer default" (the removal of support from the sponsor bank), the value of the cover pool. The senior unsecured rating of the issuer is the measure by which the credit strength of the issuer is gauged. The credit quality of the cover pool is measured by Moody's "collateral score." The higher the credit quality, the lower the collateral score. The lower the collateral score, the lower the level of losses that will impact the cover pool at time of issuer default in Moody's EL model (expected loss–based analytical model).

The approach applied by S&P is somewhat different. Although they too recognize the link between the creditworthiness of the issuer and its covered bonds, S&P operates on the basis that any potential weakness of the issuer can be overcome by the provision of a higher degree of overcollaterisation. As a result, S&P's ratings are based essentially on an analysis of the collateral pool and therefore tend to be higher than those of Moody's.

THE EUROPEAN COVERED BOND MARKET

As more European countries aim to establish their own covered bond markets with updated legislation, investors are getting a larger choice of Pfandbrief-like products. Most of the laws are based on the established German framework and aim to provide the same high quality of asset, but slight differences still remain. Here, we look at the differences between the main runners in the covered bond arena.

France

The mortgage bond market in France dates back to 1852 when, on February 28, the Decree of 1852 established mortgage banks that were authorized to lend funds to property owners. These loans were repayable by long-term annual instalments. However, it was not until June 1999 when modifications to this law broadened the appeal of *Obligations Foncières* (*OFs*) for international investors. These modifications to the Mortgage Act had two main objectives: to lower refinancing costs for the issuer and to offer investors secure and liquid products.

France had seen Germany's mortgage banks, with the success of the Pfandbrief market, being able to raise refinancing facilities at considerably lower costs than their French counterparts. They wanted quickly to follow suit. The French banks realized that failing to do so could result in their domestic market share being eroded by aggressive competition from across the border.

Another major reason for amendments to the Mortgage Act was to attempt to restore a widespread confidence in the French mortgage-lending sector after the real estate crisis that occurred in the early 1990s. The new requirements set in place were successful in doing just that.

The year 1999 saw the creation of a new type of financial institution in France, the Société de Credit Foncier (SCF) or mortgage loan company, provided for under the new law. Their creation sets the Obligations Foncières aside from other newly created European mortgage-backed sectors such as Spain's Cédulas Hipotecarias by the fact that their issuance is restricted solely to these Sociétés de Credit Foncier.

SCFs have the sole purpose of refinancing eligible assets, mainly through the issuance of OFs. While they have the legal status of banks, they are prohibited from engaging in traditional banking activities and from holding equity stakes in any subsidiaries, which mean that they operate very similarly to an SPV. French issuers also manage only one asset pool comprising both types of loans, and whether the pool consists of public-sector loans, mortgage loans, or a mix of the both depends on the business model of the issuer.

Their bankruptcy remoteness is greatly enhanced through one of the most reassuring features of the French law and that is its specific exclusion of the SCF from any bankruptcy proceedings initiated at the level of its parent(s). The SCF is therefore less vulnerable to the default of its parent credit institution.

However, these legal provisions do not completely isolate the creditworthiness of the SCF from external factors, but only limit the extent to which credit risk contamination may occur. For this reason, Moody's, when granting ratings, begin their analysis by assessing the creditworthiness of the SCF itself. They achieve this by principally analyzing:

- The strategic importance of the SCF to the refinancing of its parent credit institution(s).
- The support extended to the SCF by its shareholder(s) whether in terms of liquidity or capital.
- The nature and quality of the SCF's assets, underpinned by conservative loan-to-value thresholds.
- The capacity of the institution managing the SCF to adequately perform this role.
- Its asset and liability management practices, notably regarding interest rate mismatches.

The fact that the bankruptcy of a parent cannot be extended to an SCF is welcome, however, as a Moody's report published in October 1999 states:

> . . . the fact that OFs are issued by special purpose subsidiaries means that OF holders have no direct recourse to assets outside the SCF although they could reasonably expect some parent support. This is notably different from *Pfandbriefe* where bondholders have an eventual direct recourse to non-eligible assets if cover assets are insufficient to cover their claims and become *pari passu* with other senior unsecured creditors. Along similar lines, in case of insolvency of an originating credit institution, asset replenishment and/or substitution is no longer possible, which leaves the SCF fully exposed to asset quality deterioration and repayment, and ensuing cashflow mismatches.

Although Moody's continues, "We consider that this element of weakness is mitigated by the strong likelihood that the French regulator would exert pressure on an SCF's shareholder(s) to extend support to this subsidiary."

Having thus arrived at a senior unsecured debt rating for the SCF, Moody's then turns its attention to the specific characteristics of the OFs issued by the mortgage loan company. Given that the OFs exhibit a reduced frequency of default, a reflection of the "bankruptcy-remote" element of SCFs in regard to parent(s), and the lower loss potential due to their secured nature, Moody's grant a rating to OFs of "up to three notches above the senior unsecured debt rating of the SCF."

Like Pfandbriefe, Obligations Foncières bondholders retain preferential rights with regard to the event of bankruptcy over any other claims. The similarities do not stop there. Issuers and market makers have agreed that the minimum size of issuance should be €500 million, that the issue is supported by a market making commitment from at least three banks, quoting continuous prices with bid/offer spreads of between 5 and 20 cents. Also, it almost goes without saying, all OFs must be rated by at least two of the internationally recognized ratings agencies.

While currently lacking in size in comparison to its German neighbor, the OFs are rapidly proving to be a worthy competitor.

Spain

The year 1999 also witnessed, again due to a modification of legislation, the debut of the first international issue of the Spanish *Cédulas Hipotecarias* or "mortgage notes."

Like other covered bonds, their initial existence dates back many years previous, in the case of Spain's offering, to 1869. A considerable number of cédulas have been issued in the domestic retail market since that time.

In 1981, the introduction of the "Ley del Mercado Hipotecario" (Mortgage Market Law) and its subsequent amendments allowed Cédulas Hipotecarias to be issued by almost any credit institution.

The first jumbo-style issue was brought to the market in March 1999, and since then 21 more bonds have been launched. However, despite the enthusiastic start, only one bond was issued in 2000 and one of the existing issues was tapped. The year 2001 showed more promise, with a total of five new issues, and the number of issuers increased from two to five.

Spanish cédulas are, so far, exclusively backed by mortgage loans; the legal framework for the issuance of "Cédulas Territoriales," public-sector loans, is still in the preparation stage.

Unlike the OFs, cédulas do not possess the protection of bankruptcy remoteness in regard to their issuing entity; the probability of default between them is inextricably linked. Understandably, the ratings of these issues are therefore determined by the creditworthiness of the issuer and the whole process of rating is conducted on a case-by-case basis, analyzing the issuing institution as well as the specific characteristics of the security itself.

Under Spanish law, the underlying assets for the Cédulas Hipotecarias do not count as special assets. They are not separated from the bankrupt's assets in the event of the issuer becoming insolvent, as is the case with the German and French Pfandbrief-style bonds, and this obviously places the holder of cédulas in a much weaker position by comparison. However this weakness is considered to be largely offset by the fact that cédulas have the highest level of surplus cover (overcollateralization) in Europe of at least 11%, which is imposed by law.

Cédulas have a "bond issuing ceiling" of up to 90% of the volume of "eligible mortgages" (loan-to-value ceiling of a maximum of 70% for commercial properties and 80% for residential properties). Even in the event of a full use of this ceiling, Cédulas Hipotecarias have an overcollateralization of over 11%, as the mortgage loans also serve as collateral, although they cannot be included in the calculation of the maximum volume outstanding because of the higher loan to value levels. If this limit is exceeded at any time, the issuer has to restore the overcollateralization limits by:

- Depositing cash collateral of government bonds with the Bank of Spain within 10 working days.
- Buying back/amortizing early outstanding cédulas.
- Adding new qualifying mortgages to the existing ones (e.g., by purchasing Participaciones Hipotecarias, mortgage participations are used for the securitization of mortgages).

It should be noted that due to the limited use of Cédulas Hipotecarias so far, the actual degree of overcollateralization is at least within triple digits and this mandatory requirement is a major strength of the cédulas system.

The quality and size of the mortgage portfolio and the surplus cover are also subject to regular monitoring by the Bank of Spain.

All in all, the secured nature of this type of product strongly reduces the loss potential in a default scenario and to date, since their inception back in 1869, no Cédulas Hipotecarias has ever defaulted. Whereas the German jumbo Pfandbrief still retains ite number 1 position, the Spanish jumbo market has gained ground in the importance stakes over recent years. Indeed in 2005, the Spanish market actually overtook the jumbo Pfandbrief market in terms of volume of issuance and went on to produce that same feat in the following year.

Luxembourg

In November 1997 the Grand Duchy passed a new law that authorized the creation of a brand new financial entity known as the Banque d'Emission de Lettres de Gage, a mortgage bond–issuing bank.

The Luxembourg law was modelled closely on the German Mortgage Bank Act governing the issuance of Pfandbriefe. Like Germany, the *Lettres de Gage* are subdivided into two categories: one backed by public sector loans (Lettres de Gage Publiques) and the other by mortgages (Lettres de Gage Hypothecaires). The bondholders also enjoy the same preferential rights over the covering assets which rank above all other existing claims, while the matching principal familiar in the German market also applies to the Luxembourg law.

There are, however, some key variances from Germany's mortgage law and perhaps the most important arises from the different geographical restrictions on lending business between the two. In the case of Luxembourg, public-sector loans from the whole OECD area are eligible for refinancing via covered bonds without restrictions.

There are two diametrically opposing views as to the effect this difference has on the security aspect of the

Lettres de Gage; the first is that the Luxembourg could be considered to be more secure than its German counterpart. This is thought to be due to the fact that in their search for diversified assets to use as collateral for their Pfandbrief-like product, Luxembourg banks will diversify their exposure to top-rated OECD sovereigns such as Australia and Japan.

However, competition among the mortgage banks to deliver superior returns on investments will lead them to pursue assets in lower-rated OECD member countries such as Turkey and Mexico.

In the market we observe that German banks are keen to be involved in this wider business opportunity. This is borne out by the German mortgage bank involvement in the three main Luxembourg Pfandbrief banks. Pfandbriefbank International (PBI) is part of the HVB group, Europäische Hypothekenbank S.A is a 100%-owned subsidiary of the Eurohypo group and Erste Europäische Pfandbrief und Kommunalkreditbank (EPB), the third specialist bank to receive a Pfandbrief license is jointly owned by Commerzbank (75%) and a Geneva-based holding company of the financier Dr. Wolfgang Schuppli (25%). The latter also holds a 49% stake in HypoEssen and, through another holding, a 100% stake in Düsseldorfer Hypothekenbank AG.

The Luxembourg market is still relatively small in comparison to its European cousins.

Ireland

The Irish covered market is the most recent in Europe. When Ireland sought to create their covered bond market, they looked at all the relevant laws already in place throughout Europe, and cherry-picked the most attractive factors from an investor's perspective. What made this initiative even more impressive was the fact that Ireland has no history in issuing mortgage bonds.

Toward the end of 2001, the Irish Asset Covered Securities Act was passed allowing banks recognized by the Central Bank of Ireland as "Designated Credit Institutions" (DCIs) to issue *Irish Asset Covered Securities*.

When the legal framework was first put forward in early 2000, some of the proposed features of these issues were considered to be unique attractions from an investor's perspective. Their impact, however, has been somewhat nullified by progresses made in other markets, for example, the recent amendments to the German Mortgage Bank Act. Nevertheless, the concept of Irish covered bonds still represents an improved version of the German Pfandbrief. Ireland's rules for investor protection are the most stringent in the market—with strict supervision of the Central Bank of Ireland and an Independent Cover Asset Monitor approved by the regulator, controls on assets eligible for cover pools and no possibility of risk from duration mismatching.

The Irish steering committee decided against adopting a policy such as that used by Luxembourg's Lettres de Gage with regard to "eligible assets." They felt that allowing loans made in any OECD country as collateral for their bonds would compromise the credit quality of their Irish Asset Covered Securities. Instead, Ireland has limited the asset pool to the EEA, along with G7 countries and Switzerland.

A maximum of 10% of the cover pool can be commercial property loans and substitution assets cannot exceed 20%. To limit cash flow mismatching risk, the Irish bonds exhibit tight matching requirements. For example, the nominal value of the cover assets must at all times exceed the value of the corresponding securities. The aggregate interest from the assets must also exceed that of the covered bond and the currency of the cover assets must be similar to the related bonds. In addition to this, the duration of the cover assets must be greater than the duration of the bonds.

Critically, it is only in Ireland where the regulator has further stipulated that "the weighted average duration of the cover assets should not exceed the weighted average duration of the Irish covered bonds by a period greater than three years."

There is a loan-to-value limit imposed of 60% for residential mortgages and 100% for public-sector loans and hedging contracts against interest rate risk are permitted in the collateral pool.

The Irish product provides an interesting enhancement to the range of high quality products available in this sector. The legal framework combines all the traditional elements of covered bonds from existing European markets with innovative augmentations that serve to strengthen credit quality further.

SUMMARY

Covered bonds offer high safety while at the same time granting the investor an enhanced yield in comparison to government bonds. The sheer size of the Pfandbrief market with its market-making obligations has the potential to offer good liquidity and it is gradually breaking away from its reputation as a German "closed shop." However, it still has some way to go to catch up with the very markets that it purports to challenge, the aforementioned government markets, in terms of professionalism and ability to provide a credible liquid marketplace. Mortgage banks have now been given the opportunities to operate beyond European borders and truly market their product globally. Failure to take advantage of this situation could prove extremely detrimental to their standing. One issuer in particular, DEPFA, has already tapped into the United States with a Pfandbrief issue denominated in U.S. dollars. This offers US investors a high-quality investment alternative to US agencies and triple-A ABS and can give them much sort after diversification.

New and sophisticated covered bond laws, offering significant improvements to the original Pfandbrief model, have been introduced in France, Spain, Luxembourg, and now Ireland. Germany has responded with its amendments to the Mortgage Bank Act 2002 and furthermore with the introduction of the Pfandbrief Act 2005.

The development of these other markets comes at a time when the Pfandbrief market is experiencing a difficult period, featuring several prominent downgrades and the near closure of Allgemeine HypothekenBank Rheinboden Aktiengesellschaft (now known as COREALCREDIT BANK AG) in the summer of 2005. Their introduction is, for the first time, representing increased competition for the German market, albeit still some way off posing a serious threat.

The legislation changes throughout the European covered bond markets, also bring another possibility a step closer—a European Pfandbrief.

REFERENCES

Association of German Mortgage Banks. (2000a). *The Pfandbrief: A European Perspective*. London: Euromoney Publications.

Association of German Mortgage Banks. (2000b). *The Pfandbrief: Facts and Figures*. Cologne.

Barclays Bank. (2000). The Luxembourg Pfandbrief. *European Covered Bond Series*, London.

Barclays Bank. (2001). Asset covered securities: The Irish Pfandbrief. *EU Covered Bond Series*, no. 4, London.

Deutsche Bank. (2000). The Luxembourg mortgage law. *Global Market Research Series*. Frankfurt.

Euromoney. (2001). Covered bonds struggle to compete. *Euromoney*. July.

Fitch IBCA. (2000). *German Pfandbriefe and Analogous Funding Instruments Elsewhere in Europe*. May.

Langerbein, M., and Schulte, M. (2001). The European covered bond family and the Luxembourg Pfandbrief. *Fixed Income Market Review*. Luxembourg.

Morgan Stanley Dean Witter. (1999). Obligations Foncières: Shaping up well. *Fixed Income Research Series*. London, October.

CHAPTER 27

Commercial Paper

MOORAD CHOUDHRY, PhD
Head of Treasury, KBC Financial Products, London

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

STEVEN V. MANN, PhD
Professor of Finance, Moore School of Business, University of South Carolina

Abstract: A corporation that needs long-term funds can raise those funds in either the bond or equity markets. Alternatively, if a corporation needs short-term funds, it may attempt to acquire funds via bank borrowing. One close substitute to bank borrowing for larger corporations with strong credit ratings is commercial paper. Commercial paper is a short-term promissory note issued in the open market as an obligation of the issuing entity. Commercial paper is sold at a discount and pays face value at maturity. The discount represents interest to the investor in the period to maturity. Although some issues are in registered form, commercial paper is typically issued in bearer form.

Keywords: direct paper, dealer paper, rollover risk, yield on a bank discount basis, asset-backed commercial paper, special purpose corporation, conduit, single seller, multi-seller, liquidity enhancement

The commercial paper market was developed in the United States in the latter days of the nineteenth century and was once the province of larger corporations with superior credit ratings. However, in recent years, many lower-credit-rated corporations have issued commercial paper by obtaining credit enhancements or other collateral to allow them to enter the market as issuers. Issuers of commercial paper are not limited to U.S. corporations; non-U.S. corporations and sovereign issuers also issue commercial paper. Commercial paper was first issued in the United Kingdom in 1986 and was subsequently issued in other European countries.

Although the original purpose of commercial paper was to provide short-term funds for seasonal and working capital needs, it has been issued for other purposes, most prominently for "bridge financing." For example, suppose that a corporation desires long-term funds to build a plant or acquire equipment. Rather than raising long-term funds immediately, the issuer may choose to postpone the offering until more favorable capital market conditions prevail. The funds raised by issuing commercial paper are employed until longer-term securities are issued. Commercial paper is also used as bridge financing to finance corporate takeovers.

In this chapter, we describe the characteristics of commercial paper and its investment characteristics.

CHARACTERISTICS OF COMMERCIAL PAPER

The maturity of commercial paper is typically less than 270 days; a typical issue matures in less than 45 days. Naturally, there are reasons for this. First, the Securities and Exchange Act of 1933 requires that securities be registered with the Securities and Exchange Commission (SEC). Special provisions in the 1933 act exempt commercial paper from these registration requirements so long as the maturity does not exceed 270 days. To avoid the costs associated with registering issues with the SEC, issuers rarely issue commercial paper with a maturity exceeding 270 days. In Europe, commercial paper maturities range between 2–365 days. To pay off holders of maturing paper, issuers generally "rollover" outstanding issues; that is, they issue new paper to pay off maturing paper.

Another consideration in determining the maturity is whether the paper would be eligible collateral by a bank if it wanted to borrow from the Federal Reserve Bank's discount window. In order to be eligible, the paper's maturity may not exceed 90 days. Because eligible paper trades at a lower cost than paper that is ineligible, issuers prefer to sell paper whose maturity does not exceed 90 days.

The combination of its short maturity and low credit risk make commercial paper an ideal investment vehicle for short-term funds. Most investors in commercial paper are institutional investors. Money market mutual funds are the largest single investor of commercial paper. Pension funds, commercial bank trust departments, state and local governments, and nonfinancial corporations seeking short-term investments comprise most of the balance.

The market for commercial paper is a wholesale market and transactions are typically sizeable. The minimum round-lot transaction is $100,000. Some issuers will sell commercial paper in denominations of $25,000. Commercial paper comprises one of the largest sectors of money market approaching $2 trillion outstanding at the end of 2006 according to the Federal Reserve.

Direct Paper versus Dealer Paper

Commercial paper is classified as either direct paper or dealer paper. *Direct paper* is sold by an issuing firm directly to investors without using a securities dealer as an intermediary. The vast majority of the issuers of direct paper are financial firms. Because financial firms require a continuous source of funds in order to provide loans to customers, they find it cost effective to have a sales force to sell their commercial paper directly to investors. Direct issuers post rates at which they are willing to sell commercial paper with financial information vendors such as Bloomberg, Reuters, and Telerate.

Although commercial paper is a short-term security, it is issued within a longer term program, usually for three to five years for European firms: U.S. commercial paper programs are often open-ended. For example, a company might establish a five-year commercial paper program with a limit of $100 million. Once the program is established, the company can issue commercial paper up to this amount. The program is continuous and new paper can be issued at any time, daily if required.

In the case of dealer placed commercial paper, the issuer uses the services of a securities firm to sell its paper. Commercial paper sold in this manner is referred to *as dealer paper.* Competitive pressures have forced dramatic reductions in the underwriting fees charged by dealer firms.

Historically, the dealer market has been dominated by large investment banking firms because the Glass-Steagall Act prohibited commercial banks from underwriting commercial paper. In June 1987, however, the Federal Reserve granted subsidiaries of bank holding companies the power to underwrite commercial paper. Commercial banks began immediately making inroads into the dealer market that was once the exclusive province of investment banking firms. This process was further accelerated when the Gramm-Leach-Bliley Act was signed into law in November 1999. The reforms enacted in the Gramm-Leach-Bliley Act repealed the Glass-Steagall Act that mandated artificial barriers between commercial banks, investment banks, and insurance companies. Now each is free to expand into the others' businesses.

The Secondary Market

Although commercial paper is one of the largest sectors of the money market, there is relatively little trading in the secondary market. The reason is that most investors in commercial paper follow a "buy-and-hold" strategy. This is to be expected because investors purchase commercial paper that matches their specific maturity requirements. Any secondary market trading is usually concentrated among institutional investors in a few large, highly rated issues. If investors wish to sell their commercial paper, they can usually sell it back to the original seller either dealer or issuer.

COMMERCIAL PAPER CREDIT RATINGS

All investors in commercial paper are exposed to credit risk. Credit risk is the possibility the investor will not receive the timely payment of interest and principal at maturity. While some institutional investors do their own credit analysis, most investors assess a commercial paper's credit risk using ratings by a nationally recognized statistical rating organizations (NRSROs). Table 27.1 presents the commercial paper ratings from Fitch, Moody's, and Standard & Poor's.

Table 27.1 Ratings of Commercial Paper

	Fitch	Moody's	S&P
Superior	F1+/F1	P1	A1+/A1
Satisfactory	F2	P2	A2
Adequate	F3	P3	A3
Speculative	F4	NP	B, C
Defaulted	F5	NP	D

The risk that the investor faces is that the borrower will be unable to issue new paper at maturity. This risk is referred to as *rollover risk.* As a safeguard against rollover risk, commercial paper issuers secure backup lines of credit sometimes called "liquidity enhancement." Most commercial issuers maintain 100% backing because the NRSROs that rate commercial paper usually require a bank line of credit as a precondition for a rating. However, some large issues carry less than 100% backing. Backup lines of credit typically contain a "material adverse change" provision that allows the bank to cancel the credit line if the financial condition of the issuing firm deteriorates substantially (see Stojanovic and Vaughan, 1998). Historically, defaults on commercial paper have been relatively rare.

The commercial paper market is divided into tiers according to credit risk ratings. The "top top tier" consists of paper rated A1+/P1/F1+. "Top tier" is paper rated A1/P1, F1. Next, "split tier" issues are rated either A1/P2 or A2/P1. The "second tier" issues are rated A2/P2.

Yields on Commercial Paper

The yields offered on commercial paper track those of other money market instruments. Like Treasury bills, commercial paper is a discount instrument. In other words, it is sold at a price less than its maturity value. The difference between the maturity value and the price paid is the interest earned by the investor, although some commercial paper is issued as an interest-bearing instrument.

As an example, consider some 30-day commercial paper issued with a yield on a bank discount basis of 5.24%. Assume that the relevant day-count convention is actual/360. Given the yield on a bank discount basis, the price is found by first solving for the dollar discount as follows:

$$\text{Dollar discount} = \text{Discount yield} \times \text{Face value} \times (\text{number of days until maturity}/360)$$

The price is then found as follows:

$$\text{Price} = \text{face value} - \text{discount}$$

Assuming a face value of \$100, the discount is equal to

$$\text{Discount} = 0.0524 \times \$100 \times 30/360 = \$0.4367.$$

Therefore,

$$\text{Price} = \$100 - \$0.4367 = \$99.5633.$$

The yields offered on commercial paper are highly correlated with those of other money market instruments. Moreover, the yields on commercial paper are higher than Treasury bill yields, other things being equal. There are three reasons for this relationship. First, the investor in commercial paper is exposed to credit risk. Second, interest earned from investing in Treasury bills is exempt from state and local income taxes. As a result, commercial paper has to offer a higher yield to offset this tax advantage offered by Treasury bills. Finally, commercial paper is far less liquid than Treasury bills. The liquidity premium demanded is probably small, however, because commercial paper investors typically follow a buy-and-hold strategy and therefore they are less concerned with liquidity.

The yields offered on commercial paper track those of other money market instruments. Generally, CP trades below LIBOR because of a liquidity premium, although lower-tier paper sometimes trades above LIBOR, depending on the appetite for corporate credit at the time.

ASSET-BACKED COMMERCIAL PAPER

Asset-backed commercial paper (hereafter, ABCP) is commercial paper issued by either corporations or large financial institutions through a bankruptcy-remote special purpose corporation.

ABCP is usually issued to finance the purchase of receivables and other similar assets, Some examples of assets underlying these securities include trade receivables (that is, business-to-business receivables), credit card receivables, equipment loans, automobile loans, health care receivables, tax liens, consumer loans, and manufacturing-housing loans. According to FitchRatings (Fitch, 1999), historically trade receivables have been securitized most often. The reason being is that trade receivables have maturities approximating that of the commercial paper. Recently, the list of assets has expanded to include rated asset-backed, mortgage-backed, and corporate debt securities as ABCP issuers have attempted to take advantage of arbitrage opportunities in bond markets. There are three types of securities arbitrage programs in existence at the time of this writing: limited purpose investment companies, market value ABC paper programs, and credit arbitrage ABC paper programs. For a discussion of this process, see Dierdorff (1999).

The issuance of ABCP may be desirable for one or more of the following reasons: (1) it offers lower-cost funding compared with traditional bank loan or bond financing; (2) it is a mechanism by which assets such as loans can be removed from the balance sheet; and (3) it increases a borrower's funding options.

According to Moody's (see Adelson, 1993) an investor in ABCP is exposed to three major risks. First, the investor is exposed to credit risk because some portion of the receivables being financed through the issue of ABCP will default, resulting in losses. Obviously, there will always be defaults so the risk faced by investors is that the losses will be in excess of the credit enhancement. Second, liquidity risk which is the risk that collections on the receivables will not occur quickly enough to make principal and interest payments to investors. Finally, there is structural risk that involves the possibility that the ABCP conduit may become embroiled in a bankruptcy proceeding, which disrupts payments on maturing commercial paper.

Legal Structure

An ABCP issue starts with one seller or multiple sellers' portfolio of receivables generated by a number of obligors (e.g., credit card borrowers). A corporation using structured financing seeks a rating on the commercial paper it issues that is higher than its own corporate rating. This is accomplished by using the underlying loans or receivables as collateral for the commercial paper rather than the issuer's general credit standing. Typically, the corporation (that is, the seller of the collateral) retains some interest in the collateral. Because the corporate entity retains an interest, the NRSROs want to be assured that a bankruptcy of that corporate entity will not allow the issuer's creditors access to the collateral. Specifically, there is a concern that a bankruptcy court could redirect the collateral's cash flows or the collateral itself from the ABCP investors to the creditors of the corporate entity if it became bankrupt.

To allay these concerns, a bankruptcy-remote special entity (SPE) is formed. The issuer of the ABCP is then, the SPE Legal opinion is needed stating that in the event or the bankruptcy of the seller of the collateral, counsel does not believe that a bankruptcy court will consolidate the collateral sold with the seller's assets.

The SPE, is set up as a wholly owned subsidiary of the seller of the collateral. Despite this fact, it is established in such a way that it is treated as a third-party entity relative to the seller of the collateral. The collateral is sold to the SPE which it turn resells the collateral to a conduit (that is, trust). The conduit holds the collateral on the investors' behalf. It is the SPE that holds the interest retained by the seller of the collateral.

The other key party in this process is the conduit's administrative agent. The administrative agent is usually a large commercial bank that oversees all the operations of the conduit. The SPE usually grants the administrative agent power of attorney to take all actions on their behalf with regard to the ABCP issuance. The administrative agent receives fees for the performance of these duties.

Basic Types of ABCP Conduits

ABCP conduits are categorized on two critical dimensions. One dimension involves their level of program-wide credit support either fully or partially supported. The other dimension is as either a single-seller or a multi-seller program. In this section, we will discuss each type.

Fully versus Partially Supported

In a fully supported program, all of the credit and liquidity risk of an ABCP conduit is assumed by a third-party guarantor usually in the form of a letter of credit from a highly rated commercial bank. The ABCP investor's risk depends on the financial strength of the third-party guarantor rather than the performance of the underlying assets in the conduit. Thus, investors can expect to receive payment for maturing commercial paper regardless of the level of defaults the conduit experiences. Accordingly, in determining a credit rating, the NRSROs will focus exclusively on the financial strength of the third-party guarantor.

Partially supported programs exposes the ABCP investors directly to credit and liquidity risk to the extent that losses in the conduit exceed program-wide and pool-specific credit enhancements. The conduit has two supporting facilities. The program-wide credit enhancement facility covers losses attributable to the default of the underlying assets up to a specified amount. Correspondingly, the program-wide liquidity facility provides funds to the conduit to ensure the timely payment of maturing paper for reasons other than defaults (e.g., market disruptions). Since investors are exposed to defaults of the underlying assets, the NRSROs make their expected performance under various scenarios a central focus of the ratings process.

Single-Seller versus Multi-Seller Programs

The other key dimension used to categorize ABCP conduits is as either single-seller or multiseller. Single-seller conduits securitize assets purchased from a single seller (e.g., a single originator). Conversely, multiseller conduits pool assets purchased from several disparate sellers and the ABCP issued is backed by the portfolio of these assets.

Credit and Liquidity Enhancement

In a multiseller partially supported ABCP conduit, there are two levels of credit enhancement. The first line of defense is pool-specific credit enhancement that provides protection from the defaults on assets from a particular seller. Pool-specific credit enhancement may include overcollateralization, third-party credit support, or excess spread. The second line of defense is program-wide credit enhancement that provides protection after the pool-specific credit enhancement is depleted. Program-wide credit enhancement is usually supplied by a third-party in the form of an irrevocable loan facility, letter of credit, surety bond from a monoline insurance company, or cash invested in permitted securities (see Fitch, 1999).

Liquidity enhancement is also structured in two levels—pool-specific or program-wide. Liquidity enhancement usually takes the one of two forms. One form of liquidity support is a loan agreement in which the liquidity facility agrees to extend loans to the conduit if maturing paper cannot be rolled over due to say, a disruption in the commercial paper market due to a financial crisis. Note that the liquidity facility is not responsible for interjecting needed funds into the conduit due to defaults in the asset portfolio. The other form of liquidity support is an asset purchase agreement in which the liquidity facility agrees to purchase non-defaulted assets if funds are needed.

Figure 27.1 presents a flow chart illustrating the basic structure of a partially supported, multiseller ABCP program. Note the administrative agent invests no cash into the deal but instead provides a flow of services, as a result, the administrative agent's connection to the conduit is represented with a dashed line.

Extendable Note Commercial Paper

Extendable commercial paper is a newer development in the asset-backed commercial paper market and a number of conduits have been established or restructured to enable them issuance. The first extendable ABC paper issue was in 2002, from a number of vehicles, including ABN Amro "Tulip" conduit and AIG's Orchard Park and

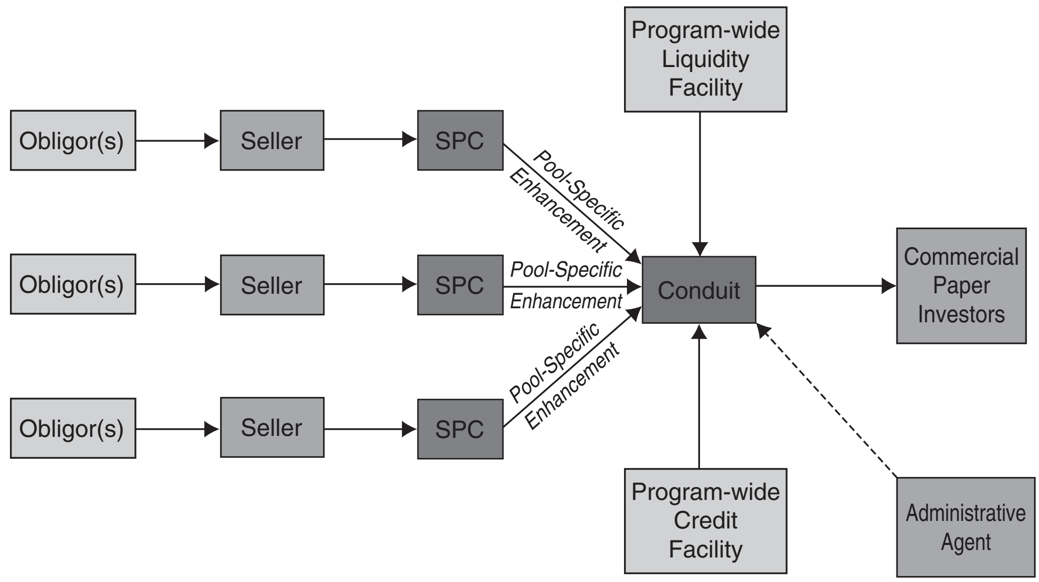


Figure 27.1 Basic Structure of a Partially Supported, Multiseller ABCP Program

Bluegrass conduits. In many cases extendable commercial paper serves as a backup or substitute to the conventional bank liquidity line on a conduit. In this section we describe a generic extendable ABC paper vehicle.

Extendable Notes

Before we describe an extendable note ABC paper structure, we should consider the form that the notes themselves take. Extendable notes are short-term liabilities issued by commercial paper conduits in the normal way, but with certain structural features that enable them to function more as liquidity reserve facilities rather than typical commercial paper liability.

Extendable notes are issued in the following form: secured liquidity notes (SLNs) and collateralized callable notes (CCNs). SLNs are also referred to variously as *liquidity notes, structured liquidity notes* and *extendable commercial paper notes*. An SLN issued by a conduit is a secured note issued with a formal maturity date of up to 397 days from original issuance. (As such, its maturity exceeds the 270 or 364 days maximum maturity of U.S. dollar or euro paper, respectively). The key aspect of the SLN however, is that its *expected* maturity date is shorter than the formal maturity date. The last expected maturity date of the note will be a function of the underlying assets in the vehicle, and the nature of the cash flows associated with these assets. Generally, the issuer is free to set the expected maturity date in line with its requirements, up to a maximum term in line with the formal maturity date.

On the expected maturity date of the SLN, the issuer will repay the note principal and interest, usually through a rollover issue of new SLNs. At that point, the note is no different from a normal issue of ABC paper. However, if for any reason a new issue of SLNs cannot be placed, then the SLN will not be repaid and instead it will be extended until its final maturity date. This is in effect similar to a liquidity facility; if the SLN cannot be rolled over, underlying assets must be sold to cover repayment on the formal maturity date. So for instance, if an SLN is issued with expected maturity of 90 days, and on the 90th day new SLNs cannot be issued, the SLN remains outstanding from the 91st to the 397th day. During the 307-day period after the expected repay date, underlying assets are sold or amortized, and the proceeds are used to repay the SLNs on or before the 397th day.

The advantage of the SLN facility over a traditional bank liquidity line is that the credit rating agencies assess the cash flow from the underlying assets (needed to repay the SLNs) for the end of the extension period. Hence, no bank liquidity would be required until this period, which would reduce the liquidity fee. Investors also view the extension of SLNs to be an unexpected occurrence, and would treat the initial issue to be normal ABCP in terms of required return.

Therefore, an SLN is essentially an ABC paper issue with an extension feature at the option of the issuer. The most common occurrence is for SLNs to be issued with 90- or 180-day maturities, with a legal final maturity date of 397 days.

A CCN is a collateralized callable note issued with a final maturity date again of maximum 397 days. The CCN has a call option that can be exercised by the issuer on a date prior to the final maturity date. The expected call date will depend on the nature of the cash flows of the underlying assets, but will be for a period inside the 397-day maximum. On the call date, the CCN will be called by the issuance of new CCNs. Again, this is similar to conventional ABC paper. If new CCNs cannot be issued, then the CCN will not be called and it remains outstanding to its final maturity date. Unlike with an SLN, there is a yield penalty: if the issue is not called when expected, its yield is increased (by anything from 10 to 25 basis points) for the remaining term. If the CCN is not called, underlying assets must be realized to repay the proceeds on final maturity.

Investor Perspective

In economic terms, CCNs are identical to SLNs, although investors may view CCNs as more favorable because there is no extension risk associated with them. Also, from the point of view of a rating agency, a callable note that is not called can be considered a not abnormal occurrence, while the extension of a SLN could be construed as a serious negative occurrence.

Where the market has a reasonable idea of the likelihood of an extended note facility actually being used, it is better able to determine how much of a return premium should be demanded by investors. For instance, the view among investors is that the ABN Amro "Tulip" and the Citibank "Dakota" vehicles are highly unlikely to exercise their extension facilities; hence, this paper is treated more or less as conventional ABCP.

Issuer Benefits

By structuring, or restructuring, a conduit with an extendable note facility, issuers can reduce their overall cost of funding. It also gives issuers more flexibility with managing their liquidity requirements and allow for unexpected occurrences.

The main advantages of an extendable note facility are:

- The freedom of having a one-year liquidity facility at lower cost than a normal liquidity line.
- The flexibility to issue to any term within the 397-day period.
- Favorable credit rating agency treatment, who view the extendable notes as 397-day liabilities, thus any liquidity back-up need not kick in until then.
- If backed with a traditional liquidity facility, or (in synthetic ABC paper programs) a guaranteed total-return swap (TRS) contract, the extended note facility is viewed very favorably by investors and traded as conventional ABCP.

The credit rating agencies consider its liquidity management capabilities as an essential component when making their rating assessment of a bank. Typically, a rating agency will analyze the following factors in assigning a bank's rating:

- Diversity of funding sources.
- Structure and maturity of liabilities.
- Balance sheet flexibility.
- Ability to access the markets for funding in time of correction or illiquidity.

The addition of an extendable note facility to a bank's ABC paper funding vehicles should strengthen the above points from the perspective of the ratings agencies. In fact, a number of banks have set up extendable note ABC paper vehicles or restructured existing vehicles to issue both straight and extendable ABC paper.

Conduit Structuring

It is possible to structure a commercial paper vehicle to issue straight and extendable commercial paper from inception or modify an existing vehicle for subsequent extendable note issuance. In the case of existing conduits that are set up to issue extendable paper, the restructuring can be effected by allowing extendable notes to be issued that are backed with:

- A facility to liquidate or amortize underlying assets within the extension period; market value risk of assets not being able to cover liabilities can be hedged through overcollateralization, or a swap arrangement that pays out on any underperformance.
- Setting up a TRS with a highly rated counterparty or guaranteed by another bank, that supports the extendable notes on final maturity; a traditional bank liquidity facility that is drawn on to repay notes on final maturity.

A traditional bank liquidity facility is the most expensive option, as it carries with it a standing fee that is payable irrespective of whether the line is ever drawn on.

For existing vehicles, legal documentation describing the conduit structure (the Issue and Paying Agency agreement and the Placement agreement or "Private Placement Memorandum") would need to be redrafted and executed. The redrafted documents would describe the new facility to issue both extendable and straight ABCP.

Figure 27.2 illustrates the structure diagram for a multiseller, multi-SPV combined ABC paper and extendable note program.

The ABCP Market Outside the United States

There are also well-developed ABCP markets in Europe and Australia. The assets underlying these European ABCP, are similar to those in the United States, namely, trade receivables, consumer loans, credit card receivables, equipment leases, etc. Moreover, there are an increasing number of programs designed to engage in arbitrage in the fixed income market by financing the purchase of asset-backed and mortgage-backed securities with ABCP. Another expanding area is using structured finance to finance cross-border trade receivables for multinational corporations.

Foreign Currency Denominated Commercial Paper

Synthetic foreign currency denominated commercial paper allows investors to earn non-U.S. interest rates without exposure to non-U.S. counterparties or political risk. Two examples are Goldman Sach's Universal Commercial Paper or Merrill Lynch's Multicurrency Commercial Paper. The process works as follows. First, a U.S. borrower issues commercial paper in a currency other than U.S. dollars, say British pounds, while simultaneously entering into a currency swap with a dealer. The commercial paper issuer faces no foreign exchange risk because the currency swap effectively allows the issuer to borrow U.S. dollars at British interest rates. Investors can then invest in commercial paper issued by a U.S. counterparty denominated in British pounds.

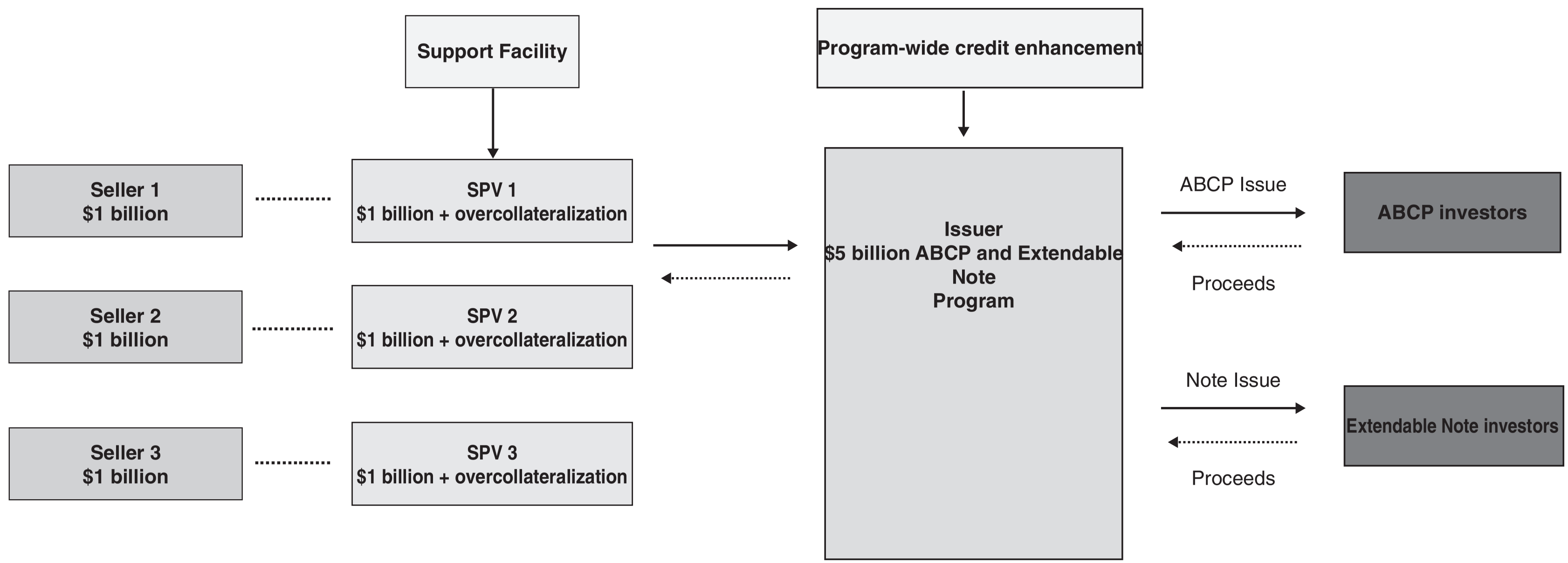


Figure 27.2 Combined ABCP/Extendable Note Program

SUMMARY

This chapter examines commercial paper, which is a vehicle for corporations to access short-term funding. The characteristics of commercial paper are discussed, including how it is issued and the secondary market. Investors in commercial paper are exposed to credit risk and most investors rely on the rating agencies to assess this risk. Commercial paper is discount instrument and pays interest at maturity. The process for issuing asset-backed commercial paper is also discussed. Finally, an overview of commercial paper denominated in currencies other than U.S. dollars is presented.

REFERENCES

Adelson, M. H. (1993). Asset-backed commercial paper: Understanding the risks. New York: Moody's Investor Services, April.

Coen, M. R., Lee, W., and Maas, B. (2000). ABCP market overview: ABCP enters the new millennium. New York: Moody's Investors Service.

Dierdorff, M. D. (1999). ABCP market overview: Spotlight on changes in program credit enhancement and growth and evolution of securities arbitrage programs. New York: Moody's Investors Service.

Fabozzi, F. J., Mann, S. V., and Choudhry, M. (2002). *Global Money Markets*. Hoboken, NJ: John Wiley & Sons.

FitchRating. (1999). Understanding asset-backed commercial paper. New York: Fitch.

Marc, R. S., and Strahan, P. E. (1999). Are banks still important for financing large businesses? Federal Reserve Bank of New York, Current Issues in Economics and Finance, August: 1–6.

Stojanovic, D., and Vaughan, M. D. (1998). Who's minding the shop? Federal Reserve Bank of St. Louis, The Regional Economist, April: 1–8.

CHAPTER 28

Money Market Calculations

STEVEN V. MANN, PhD
Professor of Finance, Moore School of Management, University of South Carolina

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: The money market is the market for short-term debt instruments. Short-term debt is traditionally defined as having original maturities of one year or less. Some of these instruments are interest bearing while others are discount instruments. Moreover, many of these securities calculate interest based on a 360-day year while others use a 365-day year. There are some essential money market calculations including day count conventions and basic formulas of price/yield one needs to know to understand how the money market works.

Keywords: day count convention, day count basis, actual/actual, actual/360, 30/360, yield on a bank discount basis, CD equivalent yield, bond-equivalent yield, effective annual yield

The purpose of this chapter is to introduce some of the fundamental calculations used every day in the money market. First, we will introduce day count conventions used in markets around the world. In addition, we will discuss the basic formulas for price/yield for both discount and interest-bearing securities.

DAY COUNT CONVENTIONS

To those unfamiliar with the workings of financial markets, it may come as a shock that there is no widespread agreement as to how many days there are in a year. The procedures used for calculating the number of days between two dates (e.g., the number of days between the settlement date and the maturity date) are called *day count conventions.* Day count conventions vary across different types of securities and across countries. In this section, we will introduce the day count conventions relevant to the money markets.

Day Count Basis

The *day count basis* specifies the convention used to determine the number of days in a month and a year. According to the *Securities Industry Association Standard Securities Calculation Methods* book, Volume 2, the notation used to identify the day count basis

(Number of days in a month)/(Number of days in a year)

Although there are numerous day count conventions used in the fixed income markets around the world, there are three basic types. All day count conventions used worldwide are variations of these three types. The first type specifies that each month has the actual number of calendar days in that month and each year has the actual number of calendar days in that year or in a coupon period (e.g., actual/actual). The second type specifies that each month has the actual number of calendar days in that month but restricts the number of days in each year to a certain number of days regardless of the actual number of days in that year (e.g., actual/360). Finally, the third types restricts both the number of days in a month and in a year to a certain number of days regardless of the actual number of days in that month/year (e.g., 30/360). Below we will define and illustrate the three types of day count conventions.

Actual/Actual

Treasury notes, bonds, and separate trading of registered interest and principal of securities (STRIPS) use an actual/actual (in period) day count convention. When calculating the number of days between two dates, the Actual/Actual day count convention uses the actual number of calendar days as the name implies. Let's illustrate the actual/actual day count convention with a 4.75% coupon, 30-year U.S. Treasury bond with a maturity date of February 15, 2037. Interest starts accruing on February 15, 2007 (the issuance date) and the first coupon date is August 15, 2007. Suppose this bond is traded with a settlement date of June 4, 2007. How many days are there between February 15, 2007, and June 4, 2007, using the actual/actual day count convention?

To answer this question, we simply count the actual number of days between these two dates or 109 actual days between February 15, 2007, and June 4, 2007. In the same manner, we can also determine the actual number of calendars days in the full coupon period. A full six-month coupon period can only have 181, 182, 183, or 184 calendar days. For example, the actual number of days between February 15, 2007, and August 15, 2007, is 181 days.

Actual/360

Actual/360 is the second type of day count convention. Specifically, Actual/360 specifies that each month has the same number of days as indicated by the calendar. However, each year is assumed to have 360 days regardless of the actual number of days in a year. Actual/360 is the day count convention used in U.S. money markets and most money markets around the world. Let's illustrate the actual/360 day count with a 26-week U.S. Treasury bill that matures on November 29, 2007. Suppose this Treasury bill is purchased with a settlement date on June 4, 2007 at a price of 97.640. How many days does this bill have until maturity using the actual/360 day count convention? The answer is 178 days.

When computing the number of days between two dates, Actual/360 and actual/actual will give the same answer. What, then, is the importance of the 360-day year in the actual/360 day count? The difference is apparent when we want to compare, say, the yield on 26-week Treasury bill with a coupon Treasury that has six months remaining to maturity. U.S. Treasury bills, like many money market instruments, are discount instruments. As such, their yields are quoted on a bank discount basis which determine the bill's price. The quoted yield on a bank discount basis for a Treasury bill is not directly comparable to the yield on a coupon Treasury using an actual/actual day count for two reasons. First, the Treasury bill's yield is based on a face-value investment rather than on the price. Second, the Treasury bill yield is annualized according to a 360-day year while a coupon Treasury's yield is annualized using the actual number of days in a calendar year (365 or 366). These factors make it difficult to compare Treasury bill yields with yields on Treasury notes and bonds. We demonstrate how these yields can be adjusted to make them comparable shortly.

Another variant of this second day count type is the actual/365. Actual/365 specifies that each month has the same number of days as indicated by the calendar and each year is assumed to have 365 days regardless of the actual number of days in a year. Actual/365 does not consider the extra day in a leap year. This day count convention is used in the UK money markets.

30/360

The 30/360 day count is the most prominent example of the third type of day count convention which restricts both the number of days in a month and in a year to a certain number of days regardless of the actual number of days in that month/year. With the 30/360 day count all months are assumed to have 30 days and all years are assumed to have 360 days. The number of days between two dates using a 30/360 day will usually differ from the actual number of days between two dates.

To determine the number of days between two dates, we will adopt the following notation:

$Y1$ = year of the earlier date
$M1$ = month of the earlier date
$D1$ = day of the earlier date
$Y2$ = year of the later date
$M2$ = month of the later date
$D2$ = day of the later date

Since the 30/360 day count assumes that all months have 30 days, some adjustments must be made for months having 31 days and February which has 28 days (29 days in a leap year). The following adjustments accomplish this task.

1. If the bond follows the end-of-month rule and $D2$ is the last day of February (the 28th in a non–leap year and the 29th in a leap year) and $D1$ is the last day of February, change $D2$ to 30.

2. If the bond follows the end-of-month rule and $D1$ is the last day of February, change $D1$ to 30.
3. If $D2$ is 31 and $D1$ is 30 or 31, change $D2$ to 30.
4. If $D1$ is 31, change $D1$ to 30.

Once these adjustments are made, the formula for calculating the number of days between two dates is as follows:

$$\text{Number of days} = [(Y2 - Y1) \times 360] + [(M2 - M1) \times 30] + (D2 - D1)$$

To illustrate the 30/360 day count convention, let's use a 4.625% coupon bond which matures on December 15, 2009, issued by Fannie Mae. Suppose the bond is purchased with a settlement date of June 4, 2007. The first coupon date is June 15, 2007, and the first interest accrual date is December 19, 2006. How many days have elapsed in the first coupon period from December 19, 2006, until the settlement date of June 4, 2007, using the 30/360 day count convention?

Referring back to the 30/360 day count rule, we see that adjustments 1 through 4 do not apply in this example so no adjustments to D1 and D2 are required. Accordingly, in this example:

$Y1 = 2006$
$M1 = 12$
$D1 = 19$
$Y2 = 2007$
$M2 = 6$
$D2 = 4$

Inserting these numbers into the formula, we find that the number of days between these two dates is 165, which is calculated as follows:

$$\text{Number of days} = [(2007 - 2006) \times 360] + [(6 - 12) \times 30] + (4 - 19) = 360 + -180 + -15 = 165$$

The actual number of days between these two dates is 165.

DISCOUNT INSTRUMENTS

Many money market instruments are discount securities (e.g. U.S. Treasury bills, agency discount notes, and commercial paper). Unlike bonds that pay coupon interest, discount securities are like zero-coupon bonds in that they are sold at a discount from their face value and are redeemed for full face value at maturity. Further, most discount securities use an actual/360 day count convention. In this section, we discuss how yields on discount securities are quoted, how discount securities are priced, and how the yields on discount securities can be adjusted so that they can be compared to the yields on interest-bearing securities.

Yield on a Bank Discount Basis

The convention for quoting bids and offers is different for discount securities from that of coupon-paying bonds. Prices of discount securities are quoted in a special way. Bids and offers of these securities are quoted on *a bank discount basis*, not on a price basis. The *yield on a bank discount basis* is computed as follows:

$$Y_d = \frac{D}{F} \times \frac{360}{t}$$

where

Y_d = annualized yield on a bank discount basis (expressed as a decimal)
D = dollar discount, which is equal to the difference between the face value and the price
F = face value
t = actual number of days remaining to maturity

As an example, suppose a Treasury bill with 91 days to maturity and a face value of \$100 trading at a price of \$98.5846. The dollar discount, D, is computed as follows:

$$D = \$100 - \$98.5846 = \$1.4054$$

Therefore, the annualized yield on a bank discount basis (expressed as a decimal)

$$Y_d = \frac{\$1.4054}{\$100} \times \frac{360}{91} = 0.0556$$

Given the yield on a bank discount basis, the price of a Treasury bill is found by first solving the formula for the dollar discount (D) as follows:

$$D = Y_d \times F \times (t/360)$$

The price is then

$$\text{price} = F - D$$

As an example, suppose a 91-day bill with a face value of \$100 has a yield on bank discount basis of 5.56%, D is equal to

$$D = 0.0556 \times \$100 \times 91/360 = \$1.4054$$

Therefore,

$$\text{price} = \$100 - \$1.4054 = \$98.5946$$

As noted earlier, the quoted yield on a bank discount basis is not a meaningful measure of the potential return from holding a discount instrument for two reasons. First, the measure is based on a face-value investment rather than on the actual dollar amount invested. Second, the yield is annualized according to a 360-day rather than a 365-day year, making it difficult to compare discount yields with the yields on Treasury notes and bonds that pay interest on an actual/actual basis. The use of 360 days for a year is a common money market convention. Despite its shortcomings as a measure of return, this is the method that dealers have adopted to quote discount notes like Treasury bills. Many dealer quote sheets and some other reporting services provide two other yield measures that attempt to make the quoted yield comparable to that for a coupon bond and interest-bearing money market instrument—the CD equivalent yield and the bond equivalent yield.

CD Equivalent Yield

The *CD equivalent yield* (also called the money market equivalent yield) makes the quoted yield on a bank discount basis more comparable to yield quotations on other money market instruments that pay interest on a 360-day basis. It does this by taking into consideration the price of the discount security (that is, the amount invested) rather than its face value. The formula for the CD equivalent yield is

$$\text{CD equivalent yield} = \frac{360\,Y_d}{360 - t(Y_d)}$$

To illustrate the calculation of the CD equivalent, suppose a 91-day Treasury bill has a yield on a bank discount basis is 5.56%. The CD equivalent yield is computed as follows:

$$\begin{aligned}\text{CD equivalent yield} &= \frac{360(0.0556)}{360 - 91(0.0556)}\\ &= 0.05639 = 5.639\%\end{aligned}$$

Bond-Equivalent Yield

The measure that seeks to make a discount instrument like a Treasury bill or an agency discount note comparable to coupon Treasuries is the *bond equivalent yield* as discussed earlier in the chapter. This yield measure makes the quoted yield on a bank discount basis more comparable to yields on Treasury notes and bonds that use an actual/actual day count convention. The calculations depend on whether the short-term discount instrument has 182 days or less to maturity or more than 182 days to maturity.

Discount Instruments with Less Than 182 Days to Maturity

To convert the yield on a bank discount to a bond-equivalent yield for a bill with less than 182 days to maturity, we use the following formula:

$$\text{Bond-equivalent yield} = \frac{T(Y_d)}{360 - t(Y_d)}$$

where T is the actual number of days in the calendar year (that is, 365 or 366). As an example, using a Treasury bill with 91 days to maturity yielding 5.56% on a bank discount basis, the bond-equivalent yield is calculated as follows:

$$\begin{aligned}\text{Bond-equivalent yield} &= \frac{365(0.0556)}{360 - 91(0.0556)}\\ &= 0.0572 = 5.72\%\end{aligned}$$

Note the formula for the bond-equivalent yield presented above assumes that the current maturity of the discount instrument in question is 182 days or less.

Discount Instruments with More Than 182 Days to Maturity

When a discount instrument (e.g., a 52-week Fannie Mae Benchmark bill) has a current maturity of more than 182 days, converting a yield on a bank discount basis into a bond-equivalent yield is more involved. Specifically, the calculation must reflect the fact that a Benchmark bill is a discount instrument while a coupon Treasury delivers coupon payments semiannually and the semiannual coupon payment can be reinvested. In order to make this adjustment, we assume that interest is paid after six months at a rate equal to the discount instrument's bond-equivalent yield and that this interest is reinvested at this rate.

INTEREST AT MATURITY INSTRUMENTS

In contrast to discount instruments, some money market instruments pay interest at maturity on a simple interest basis. Notable examples include federal funds, repos, and certificates of deposit. Interest accrues for these instruments using an actual/360 day count convention. We define the following terms:

F = face value of the instrument
I = amount of interest paid at maturity
t = actual number of days until maturity
Y_{360} = yield on a simple interest basis assuming a 360 day year

The following formula is used to calculate the dollar interest on a certificate of deposit:

$$I = F \times Y_{360} \times (t/360)$$

As an illustration, suppose a bank offers a rate of 4% on a 180-day certificate of deposit with a face value of \$1 million. If an investor buys this CD and holds it to maturity, how much interest is earned? The interest at maturity is \$20,000 and is determined as follows:

$$I = \$1{,}000{,}000 \times 0.04 \times (180/360) = \$20{,}000$$

Converting a CD Yield into a Simple Yield on a 365-Day Basis

It is often helpful to convert a CD yield that pays simple interest on an actual/360 into a simple yield on an actual/365 basis. The transformation is straightforward and is accomplished using the following formula:

$$Y_{365} = Y_{360}(365/360)$$

To illustrate, let's return to the 180-day certificate of deposit yielding 4% on a simple interest basis. We pose the question: What is this investor earning on an actual/365 basis? The answer is 4.056% and is calculated as follows:

$$Y_{365} = 0.04(365/360) = 0.04056$$

Converting a Periodic Interest Rate into an Effective Annual Yield

Suppose that \$100 is invested for one year at an annual interest rate of interest of 4%. At the end of the year, the interest received is \$4. Suppose, instead, that \$100 is invested for one year at an annual rate, but the interest is paid semiannually at 2% (one-half the annual interest rate). The interest at the end of the year is found by first calculating the future value of \$100 one year hence:

$$\$100(1.02)^2 = \$104.04$$

Interest is therefore \$4.04 on a \$100 investment. The interest rate or yield on the \$100 invested is 4.04%. The 4.04% is called the *effective annual yield.*

Investors in certificates of deposit will at once recognize the difference between the annual interest rate and effective annual yield. Typically, both of these interest rates are quoted for a certificate of deposit, the higher interest rate being the effective annual yield.

To obtain the effective annual yield corresponding to a given periodic rate, the following formula is used:

$$\text{Effective annual yield} = (1 + \text{Periodic interest rate})^m - 1$$

where m is equal to the number of payments per year.

To illustrate, suppose the periodic yield is 2% and the number of payments per year is two. Therefore,

$$\begin{aligned}\text{Effective annual yield} &= (1.02)^2 - 1 \\ &= 0.0404 \text{ or } 4.04\%\end{aligned}$$

We can also determine the periodic interest rate that will produce a given effective annual yield. For example, suppose we need to know what semiannual interest rate would produce an effective annual yield of 5.25%. The following formula can be used:

$$\text{Periodic interest rate} = (1 + \text{Effective annual yield})^{1/m} - 1$$

Using this formula to determine the semiannual interest rate to produce an effective annual yield of 5.25%, we find

$$\begin{aligned}\text{Periodic interest rate} &= (1.0525)^{1/2} - 1 \\ &= 0.0259 \text{ or } 2.59\%\end{aligned}$$

SUMMARY

This chapter introduces some of the fundamental calculations used in money markets around the world. The chapter started with the procedures used for calculating the number of days between two dates called day count conventions. Some money market instruments are discount instruments. The basic formulas for yield and price are discussed for each. Furthermore, some of these securities calculate interest using a 360-day while others use a 365-day year. The method for converting a CD yield into a simple yield on a 365-day basis is presented. The last section of the chapter details the conversion of a periodic interest rate into an effective annual yield.

REFERENCES

Fabozzi, F. J., and Mann, S. V. (2001). *Introduction to Fixed Income Analytics*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F., Mann, S. V., and Choudhry, M. (2002). *The Global Money Markets*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., and Thurston, T. (1986). State taxes and reserve requirements as major determinants of yield spreads among money market instruments. *Journal of Financial and Quantitative Analysis*, December: 427–436.

Mayle, J. (1994). *Standard Securities Calculation Methods*, vols. 1 and 2. New York: Securities Industry Association.

Stigum, M., and Robinson, F. (1996). *Money Market & Bond Calculations*. Chicago: Irwin.

CHAPTER 29

Convertible Bonds

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

STEVEN V. MANN, PhD
Professor of Finance, Moore School of Business, University of South Carolina

FILIPPO STEFANINI
Deputy Chief Investment Officer, Aletti Gestielle Alternative SGR Professor of Risk Management, Faculty of Engineering at Bergamo University in Italy

Abstract: Corporations finance themselves by selling claims to the expected future cash flows generated by their assets. The two basic claims issued are debt claims in the form of bonds and equity claims in the form of common stocks. Between these two endpoints, there exists a continuum of securities that possess features of both bonds and stocks. Among the most prominent of these hybrid securities is convertible bonds. A convertible bond is a combination of an option-free bond and the option to convert the bond into a given number of shares of the issuer's common stock. Convertible bonds may also be callable or putable or both. Depending on the performance of the underlying company, a convertible bond may behave more like a common stock or more like an option-free bond.

Keywords: conversion ratio, conversion premium, conversion price, hard put, soft put, conversion value, straight value, premium payback period, premium over straight value, fixed income equivalent, common stock equivalent, mandatory convertible, reverse convertible, convertible bond arbitrage

A convertible bond is a security that gives the investor the option to convert into a specified number of shares of the issuer's common stock. Convertible bonds issued today typically possess more than one embedded option in that they can be callable and putable. Accordingly, valuing convertible bonds is more challenging because the bond's value will depend on: (1) how interest rate changes impact the bond's expected future cash flows via call and/or put options, (2) how creditworthiness of the underlying company impacts expected future cash flows, (3) how the stock price changes impact the value of the conversion feature, and (4) how volatile the stock price is.

In its most basic form, there are two equivalent ways to describe a convertible bond. First, a convertible bond represents the combination of an option-free bond and call option on the common stock. However, unlike the exercise price of a call option, which is fixed, the value of the bond is surrendered to obtain a predetermined number of shares of stock. Second, a convertible bond is a combination of common stock and a put option, which gives the bondholder the right to sell the stock back to the issuer with an exercise price equal to the market value of the convertible. If the investor chooses not to convert, she effectively exercises the put and thereby keeps receiving the bond's cash flows.

Closely related to a convertible bond is an exchangeable bond. An exchangeable bond gives the bondholder the right, but not the obligation, to exchange the bond for the common stock of a firm other than the security issuer. For example, in April 2007, UBS AG issued 6% six-month notes that were exchangeable into a fixed number of shares of Honda Motor Corporation. For the remainder of the chapter, we will use the term "convertible bond" to refer to both convertible and exchangeable bonds.

In this chapter, we describe the defining characteristics of convertible bonds and provide a sketch of the convertible bond market. The traditional approach to analyzing convertibles is examined. Variants of the traditional convertible bond structure are introduced. Finally, the issue of why *convertible bond arbitrage* is a popular hedge fund strategy is discussed.

CONVERTIBLE BOND MARKET

The U.S. convertible bond market is by far the largest convertible bond market. Most U.S. convertible bonds are issued as private placements under Securities and Exchange Commission (SEC) Rule 144A. These issues can be sold to only qualified investors. Conversely, in Japan, the convertible bond market is comprised of a large number of very small domestic issues that are listed and traded on the Tokyo stock exchange.

There are marked differences in the size and creditworthiness of issuers of convertible in the United States and Europe. Most European convertible bonds are issued by large corporations with investment-grade credit ratings. In contrast, the majority of U.S. convertible bonds are issued by smaller corporations that are rated below investment grade.

GENERAL CHARACTERISTICS OF CONVERTIBLE BONDS

In this section, we will introduce the defining characteristics of a convertible bond using an example. Consider a convertible bond issued by United Auto Group in April 2006. This convertible bond carries a 3.5% coupon rate and the right to convert extends until the maturity date of April 1, 2026. As noted, the conversion provision grants the security holder the right, but not the obligation, to convert the bond into a predetermined number of shares of the issuer's common stock. The predetermined number of shares is called the *conversion ratio.* This ratio is always adjusted proportionally for stock splits and stock dividends. For the United Auto Group convertible, the conversion ratio is 42.2052 shares. Accordingly, the security holder at a time of her choosing may surrender the $1,000 maturity value bond for 42.2052 shares of United Auto Group common stock.

Armed with the conversion ratio, it is straightforward to determine the price per share the convertible bondholder pays when purchasing the share via the conversion mechanism. This is called the *conversion price* and is found by dividing the bond's price by the conversion ratio. If the United Auto Group bond is converted, the investor will receive 42.2052 shares of its common stock. Accordingly, at issuance, the shares are purchased at $23.69 per share ($1,000/42.2052). Purchasing the common stock with a convertible security requires that the investor pay a premium over the current share price. Investors accede to this because of the embedded optionality. The premium is often measured in percentage terms and is called the *conversion premium.* When the United Auto Group bonds were issued, the stock price was $18.95 and the conversion price was $23.69, so the initial conversion premium was 25%

Virtually all convertible bonds are callable such that the call feature gives the investor the right to buy the bond back at a given price (that is, the call price) before maturity. The United Auto Group bond has a 5-year call protection period such that the first call date is April 6, 2011, and gives the issuer the option to buy the bonds back before maturity. The call price is 100.

Many convertible bonds also possess a put feature. The put feature gives the bondholder the right but not the obligation to sell the bond back to the issuer at par before the maturity date. The United Auto Group bond is putable at par starting on April 6, 2011, five years after issuance. Put features may be classified as either "hard" or "soft" and differ as to the form of payment to the bondholder when the put is exercised. A *hard put* requires the convertible security to be redeemed for cash. Conversely, when a *soft put* is exercised, the issuer is allowed to choose the form of payment, which may be cash, common stock, subordinated debt, or some combination of the three.

ANALYZING CONVERTIBLE SECURITIES: THE TRADITIONAL APPROACH

The traditional approach to the analysis and valuation of convertible bonds predates models designed to value bonds with embedded options. As such, traditional analysis does not take into consideration the value of any of the convertible bond's embedded options directly.

Conversion Value versus Straight Value

The traditional approach to the valuation of convertible bonds begins with the determination of two values—

conversion and straight. A bond's *straight value* is found by valuing the bond as if the conversion feature does not exist. The *conversion value* is the security's value if it converted immediately. Specifically,

$$\text{conversion value} = \text{conversion ratio} \times \text{market price of common stock}$$

At any point before maturity, a convertible bond must be worth at least as much as the greater of the conversion value or the straight value. This is an arbitrage-enforced result. To see this, suppose the conversion value is greater than the straight value and the convertible bond's price is equal to its straight value. To exploit this arbitrage opportunity, an investor would buy the convertible and immediately convert. These actions enable the investor to capture the difference between the conversion value and the straight value less transaction costs. Suppose the opposite is true: The straight value is greater than the conversion value and the bond trades at its conversion value. If this occurs, the investor will be holding a bond that is undervalued relative to an otherwise the same straight bond.

Market Conversion Price

When an investor takes a position in a convertible bond, he or she is buying the upside potential driven by the common stock with the downside protection of the straight bond. Accordingly, investors are willing to pay a premium over the current share price to purchase the common stock using the convertible. The price paid per share if the convertible is purchased and then converted is called the *market conversion price* or the conversion parity price. The market conversion price is computed as follows:

$$\text{market conversion price} = \frac{\text{market price of convertible security}}{\text{conversion ratio}}$$

Analysts view the market conversion price as a "break-even price" because if the share price rises to this level, it just equals the price at which the investor purchased the shares using the convertible security.

The next phase in our analysis is to recast the premium paid for buying the shares using the convertible as a call option. This is true because the conversion feature allows for upside share price appreciation with a limited downside. To do this, we calculate the market conversion premium per share as follows:

$$\text{market conversion premium per share} = \text{market conversion price} - \text{current share price}$$

The market conversion premium per share can be viewed as the value of the call option on common stock underlying the convertible. An important difference between the two positions is the downside risk exposure. The downside of a long call position is limited to the price paid for the option. Moreover, the downside risk exposure is the convertible's straight value serves as a floor of the convertible security's value. The floor, unfortunately, is like the floor of an elevator in that it can go up or down. This is true because the straight bond's value is a function of the level of interest rates, credit risk, and the like.

The market conversion premium per share can also be expressed as a percentage of the current share price. Specifically, the *market conversion premium ratio* is computed as follows:

$$\text{market conversion premium ratio} = \frac{\text{market conversion premium per share}}{\text{market price of common stock}}$$

As an illustration, consider a convertible issued by Ford Motor Company in December 2006. The issue pays a 4.5% coupon semiannually and matures in December 2036. This bond has a conversion ratio of 108.6957 shares and the bond's current market value is 1,113.80. The market conversion price is

$$\text{market conversion price} = \frac{1{,}113.80}{108.6957} = 10.247$$

Accordingly, the investor is paying \$10.247 a share for Ford Motor Company common stock. Suppose the current market share price of Ford is \$8.01. The market conversion premium per share is computed as follows:

$$\begin{aligned}\text{market conversion premium per share} &= 10.247 - 8.01\\ &= 2.24\end{aligned}$$

This number tells us that if investors buy Ford common stock via the convertible, they pay a premium of \$2.24 a share as opposed to buying the stock at the prevailing market price. Finally, the market conversion premium ratio is computed as follows:

$$\begin{aligned}\text{market conversion premium ratio} &= \frac{2.24}{8.01} = 0.2793\\ &= 27.93\%\end{aligned}$$

This percentage is interpreted as the investor is paying a premium of 27.93% to purchase the Ford common stock through the convertible.

Measuring the Convertible Bond's Income Advantage

Assuming that the issuer does not default on its debt, convertible bonds usually generate more in coupon interest than dividend income received from a number of common shares equal to the conversion ratio. This income advantage counterbalances the premium paid for common stock purchased via the convertible bond. When analysts assess relative value, they often compute a measure called the *premium payback period* (also called the break-even time). The premium payback period measures how long it takes to pay for the market conversion premium per share with the convertible's income advantage. The premium payback period is computed with the following:

$$\text{premium payback period} = \frac{\text{market conversion premium per share}}{\text{favorable income differential per share}}$$

where the favorable income differential per share is computed as follows:

$$\frac{\text{coupon interest} - (\text{conversion ratio} \times \text{common stock dividend per share})}{\text{conversion ratio}}$$

The ratio has two parts. The first part (numerator) is the favorable income differential which is simply the coupon interest paid by the convertible less the dividend income forgone by not converting. The second part (denominator) is just the conversion ratio and puts the income advantage on a per share basis. Accordingly, the premium payback period answers the question: How long must one hold the convertible bond with its favorable income differential until the premium per share for buying the common stock via the convertible is recovered? We hasten to add that this measure does not account for future dividend changes or the time value of money.

We will use the Ford Motor Company convertible bond to illustrate this measure. Our first task is to find the favorable income differential per share. We need to calculate the following:

$$\begin{aligned}\text{coupon interest from Ford bond} &= 0.045 \times \$1,000\\ &= \$45.00\\ \text{conversion ratio} \times \text{dividend per share} &= 108.6957 \times 0.05\\ &= \$5.435\end{aligned}$$

Accordingly, we input these numbers into the expression for

$$\begin{aligned}\text{favorable income differential per share} &= \frac{\$45.00 - \$5.435}{108.6957}\\ &= \$0.36\end{aligned}$$

and then compute the

$$\text{premium payback period} = \frac{\$2.24}{\$0.36} = 6.2 \text{ years.}$$

This number tells us it will take approximately 6.2 years for the higher income of the convertible bond versus holding the common stock directly to recover the market conversion premium per share. The preceding ignores any dividend changes and the time value of money.

Measuring the Convertible Bond's Downside Risk

Traditional convertible analysis erroneously views a convertible bond's straight value as the floor for the bond's value. Following this line of reasoning, the distance between the current market price and the straight value can be viewed as a measure of the investor's downside risk exposure. Formally, the downside risk is measured as a percentage of the straight value (that is, the floor), referred to as the *premium over straight value*. It is calculated using the following formula:

$$\begin{aligned}&\text{premium over straight value}\\ &\quad = \frac{\text{market price of convertible bond}}{\text{straight value}} - 1\end{aligned}$$

All else being equal, the greater the premium over the straight value, the greater the investor's exposure to downside risk.

The flaw in this measure is that the straight value is mistakenly viewed as a fixed and an immoveable barrier. The straight value depends on the level of yields and will move inversely to changes in those yields. The "floor" is a moving target.

We illustrate this measure using the Ford Motor Company convertible bond, whose current market value is $1,112.80:

$$\begin{aligned}\text{premium over straight value} &= \frac{\$1,112.80}{\$922.40} - 1 = 0.2064\\ &= 20.64\%\end{aligned}$$

CONVERTIBLE BONDS AS AN INVESTMENT

A convertible bond is a hybrid financial instrument that combines elements of a position in a fixed income security and a position in the underlying common stock. The relative importance of each component is driven primarily by the financial performance of the underlying company that is ultimately reflected in the stock price. The relationship between the convertible bond price and the underlying stock price can be described as a continuum (see Figure 29.1). At one end of the continuum, the stock price is relatively low, such that the straight value of the convertible is considerably higher than the conversion value. When this occurs, convertibles have a low sensitivity to the underlying stock price because the conversion option is deep out-of-the-money and will trade like a high-yield straight bond. Convertibles in such circumstances are termed *fixed-income equivalent* or busted convertible. At the opposite end of the continuum, the stock price is relatively high, such that the conversion value is considerably higher than the straight value. The convertible bond will be highly responsive to changes in the stock price and possess a low conversion premium. When this occurs, the convertible bond will trade much like a common stock. The convertible under these conditions is said to be a *common stock equivalent.* At points in between the two endpoints, the convertible trades like a hybrid security possessing the characteristics of both a bond and a stock.

OTHER TYPES OF CONVERTIBLES

There are two prominent variants of the traditional convertible bond—mandatory convertibles and reverse convertibles. Each is discussed in turn in this section.

Mandatory Convertibles

Mandatory convertibles are equity-linked hybrid securities that convert automatically at maturity into shares of the issuer's common stock. This automatic conversion differs from convertible bonds where conversion is optional.

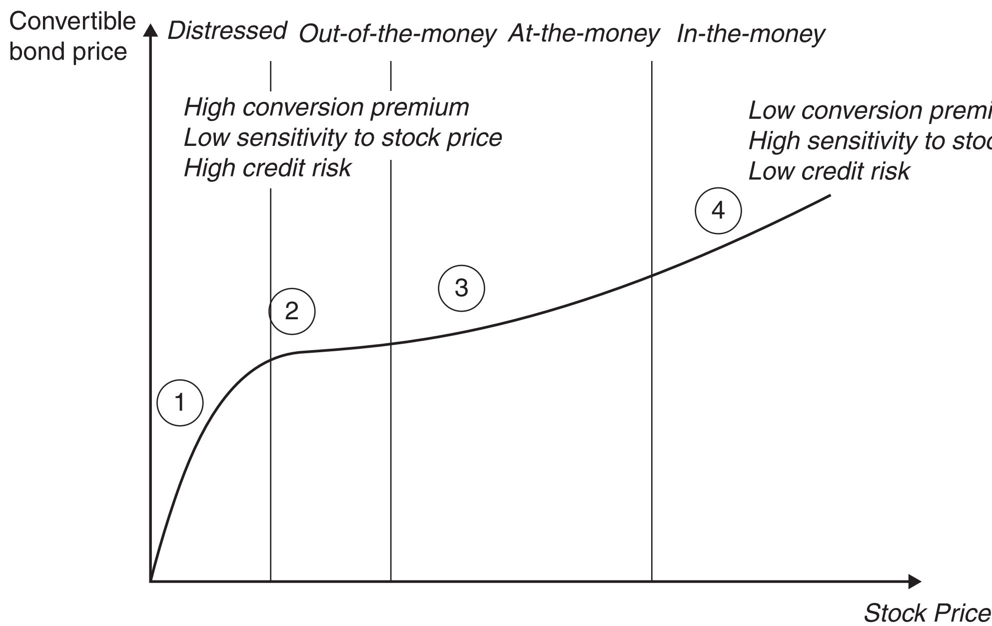


Figure 29.1 Relationship Between the Convertible Bond Price and the Underlying Stock Price

Mandatory convertibles offer higher coupon payments relative to the dividend income from holding the common stock directly. To glean these benefits, investors in mandatory convertibles pay a premium for the shares to be acquired at maturity. Moreover, these securities provide investors with limited upside participation in the underlying common stock. Mandatory convertibles are known by other trade names, including debt exchangeable for common stock (DECS).

Mandatory convertibles are converted at maturity into a number of shares determined by the underlying share price as presented in Figure 29.2. There are three possible outcomes. First, if the share price at maturity is below the lower exercise price P_1, the investor receives a fixed number of shares. Second, if the share price at maturity falls between the two exercise prices, the investor receives a variable number of shares such that their value remains constant (V_N) between the two exercise prices. Third, if the share price at maturity is above the higher exercise price P_2, the investor receives a fixed number of shares that is lower than in the first outcome.

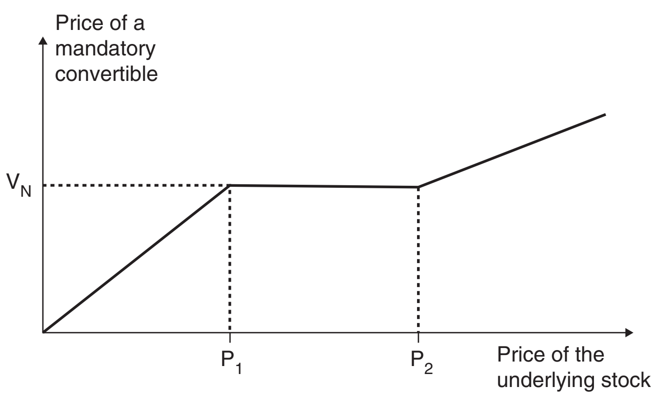


Figure 29.2 Relationship Between Price of Underlying Stock and Price of Mandatory Convertible

As an illustration, Deutsche Telekom AG issued a mandatory convertible on February 24, 2003, through its Dutch financial company Deutsche Telekom International Finance B.V. The amount issued was €2.2885 billion. The bond paid a 6.5% coupon annually and matured on June 1, 2006. Conversion was mandatory at maturity, but investors had the option to convert from July 1, 2003, until April 30, 2006. The lower exercise price was €11.80, and the higher exercise price was €14.632. The minimum/maximum conversion ratios were 3,417.17/4,237.29.

One appealing feature of a regular convertible bond is the downside protection of the bond component. If the underlying common stock price performance is anemic and the conversion feature has no value, investors still have the bond. Due to the automatic conversion at expiration, a mandatory convertible has no bond floor and offers no downside protection.

Reverse Convertibles

The major difference between a regular convertible and reverse convertible turns on who owns the conversion option. A *reverse convertible* can be thought of as the combination of a long position in an option-free bond and a short position in a put option. The issuer of the reverse convertible owns the put option and has the right but not the obligation to exercise its option to sell. If the price of the issuer's common stock is below the exercise price at the exercise date, the bondholder receives a fixed number of shares of stock. The investor is obligated in effect to purchase shares above its market value. Conversely, if the stock price is above the exercise price at expiration, the bondholder receives the bond's maturity value.

CONVERTIBLE BOND ARBITRAGE

As explained in Chapter 48 of Volume II, convertibles are ideal securities for arbitrage because the convertible itself, namely, the underlying stock and the associated derivatives, are value expressions of the same company and any discrepancy or mispricing would give rise to arbitrage opportunities for hedge fund managers to exploit. As we have seen, the valuation of convertible bonds is driven by four primary factors: (1) interest rates, (2) credit spreads, (3) stock prices, and (4) volatility of stock prices. Convertible bond arbitrage involves taking a leveraged position (long or short) in the convertible bond to gain exposure to a mispriced factor while simultaneously hedging interest rates and small changes in stock prices. Now, suppose that a hedge fund manager has a view that a convertible bond is undervalued. How would the hedge fund manager take advantage of this view? The natural response is to take a long position in the convertible bond. This strategy is exposed to at least three significant risks. First, the hedge fund manager could be wrong about his/her valuation assessment. Since this is how hedge fund managers add value, this is a risk they are willing to bear.

Second, a long position in the convertible is exposed to adverse movements in stock prices. To neutralize this risk, the hedge fund manager establishes a short position in an appropriate amount of the underlying common stock. The appropriate amount depends on the sensitivity of the convertible bond's value and changes in the underlying stock price. We call this price sensitivity "delta," and it tells us for a \$1 change in the conversion value what is the associated change in the convertible bond's value. Suppose that delta is 0.60. This measure tells that for a \$1 change in the conversion value, the convertible bond's value changes by approximately \$0.60. To hedge the exposure to adverse share price movements, the hedge fund manager would short a number of shares per bond equal to 0.60 multiplied by the conversion ratio. The combined value of the long convertible bond position and the short common stock position should be invariant to small change in the underlying stock price.

The third risk exposure is the adverse movements in interest rates. This risk can be measured, for example, by the effective dollar duration of the convertible position. Effective dollar duration tells us the dollar price change in the value of a bond position given a 100 basis point shift in yield. (For a discussion of effective duration, see Chapter 13 of Volume III.) To hedge this risk, managers take a short position in Treasury securities or a short position in interest rates futures.

The fourth risk exposure is the drop in the stock volatility that can happen in lackluster equity markets. The fifth risk exposure is the lack of liquidity in the market and the consequent widening of bid-ask spreads.

SUMMARY

A convertible bond is a security that gives the investor the option to convert into a specified number of shares of the issuer's common stock. In this chapter, we described the basic structure of a convertible bond. The traditional approach to the analysis of convertible bonds was presented. This approach does not attempt to value the options embedded in the convertible directly. Other types of convertibles—mandatory and reversible—are introduced. To the savvy hedge fund manager, convertible bonds represent a "target-rich" environment for finding and exploiting mispriced securities.

REFERENCES

Arak, M., and Martin, A. (2005). Convertible bonds: How much equity, how much debt? *Financial Analysts Journal* 61, 2: 44–50.

Bhattacharya, M. (2005). Convertible securities and their valuation. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Securities* (pp. 1393–1442). New York: McGraw-Hill.

Calamos, J. P. (1998). *Convertible Securities: The Latest Instruments, Portfolio Strategies, and Valuation Analysis*. New York: McGraw-Hill.

Connolly, K. B. (1998). *Pricing Convertible Bonds*. England: John Wiley & Sons.

de La Grandville, O. (2001). Bond Pricing and Portfolio Analysis, 1st edition. Cambridge, MA: MIT Press.

Dialynas, C. P., and Ritchie, J. C. (2005). Convertible securities and their investment characteristics. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Securities* (pp. 1371–1392). New York: McGraw-Hill.

Ho, T., and Pfeffer, D. (1996). Convertible bonds: Model, value attribution, and analytics. *Financial Analysts Journal* 51, 1: 35–44.

Mann, S., Moore, W., and Ramanlal, P. (1999). Timing of convertible debt issues. *Journal of Business Research* 45, 2: 101–105.

Moore, W. (2001). Real Options and Option Embedded Securities, 1st edition. Hoboken, NJ: John Wiley & Sons.

Moore, W., and Korkeamaki, T. (2004). Convertible bond design and capital investment: The role of call provisions. *Journal of Finance* 59, 1: 391–405.

Reverre, S. (2001), *The Complete Arbitrage Deskbook*, 1st edition. New York: McGraw-Hill.

Stefanini, F. (2006). *Investment Strategies of Hedge Funds.* Chichester, UK: John Wiley & Sons.

CHAPTER 30

Syndicated Loans

STEVEN MILLER
Managing Director, Standard & Poor's LCD

Abstract: Leveraged loans are the primary financing vehicle for leveraged buyouts, recapitalizations and other transactions that rely heavily on debt. Loans are underwritten and arranged by one more investment banks and syndicated to a group of bank and non-bank lenders. Unlike bonds, however, loans are private agreements between an issuer and a lender group that are not registered with, nor regulated by, the Securities & Exchange Commission. Therefore loans are embodied in idiosyncratic documents that require the issuer to meet specific covenant tests and, typically, pledge collateral for the benefit of lenders in the event of default.

Keywords: loans, market-flex language, leveraged loans, LBO, second-lien loans, syndicated loans, high-yield, private equity, leveraged finance, nonperforming loans

A *syndicated loan* is one that is provided by a group of lenders and is structured, arranged, and administered by one or several commercial or investment banks known as arrangers.

Starting with the large leveraged buyout (LBO) loans of the mid-1980s, the syndicated loan market has become the dominant way for issuers to tap banks and other institutional capital providers for loans. The reason is simple: Syndicated loans are less expensive and more efficient to administer than traditional bilateral, or individual, credit lines.

This chapter is a primer on the leveraged loan market detailing how loans are underwritten, arranged, syndicated and traded. As well, various aspects of the loan agreement are discussed, including covenants, collateral, coupon, amortization, events of default, assignment terms, and fees.

OVERVIEW OF THE SYNDICATED LOAN

At the most basic level, arrangers serve the time-honored investment-banking role of raising investor dollars for an issuer in need of capital. The issuer pays the arranger a fee for this service, and, naturally, this fee increases with the complexity and riskiness of the loan. As a result, the most profitable loans are those to leveraged borrowers—issuers whose credit ratings are noninvestment–grade and who are paying spreads (premiums above the London Interbank Offered Rate [LIBOR] or another base rate) sufficient to attract the interest of non-bank term loan investors, typically LIBOR + 200 or higher, though this threshold moves up and down depending on market conditions.

Indeed, large, high-quality companies pay little or no fee for a plain vanilla loan, typically an unsecured revolving credit instrument that is used to provide support for short-term commercial paper borrowings or for working capital. In many cases, moreover, these borrowers will effectively syndicate a loan themselves, using the arranger simply to craft documents and administer the process. For leveraged issuers, the story is a very different one for the arranger, and, by "different," we mean much more lucrative. A new leveraged loan can carry an arranger fee of up to 2.5% of the total loan commitment, depending on the complexity of the transaction. Merger and acquisition (M&A) and recapitalization loans will likely carry high fees, as will exit financings and restructuring deals. Seasoned leveraged issuers, by contrast, pay radically lower fees for refinancings and add-on transactions.

Because investment-grade loans are infrequently used and, therefore, offer drastically lower yields, the ancillary business is as important a factor as the credit product in arranging such deals, especially because many acquisition-related financings for investment-grade companies are large in relation to the pool of potential investors, which would consist solely of banks.

The "retail" market for a syndicated loan consists of banks and, in the case of leveraged transactions, finance companies and institutional investors. As in other capital markets, an arranger will get a market "read" by informally polling potential investors. After this market read, the arrangers will launch the deal at a spread and fee that it thinks will clear the market. Until 1998, this would have been it. Once the pricing was set, it was set, except in the most extreme cases. If the loan were undersubscribed, the arrangers could very well be left above their desired hold level. Since the Russian debt crisis roiled the market in 1998, however, arrangers have adopted *market-flex language*, which allows them to change the pricing of the loan based on investor demand—in some cases within a predetermined range—as well as shift amounts between various tranches of a loan, as a standard feature of loan commitment letters. Market-flex language, in a single stroke, pushed the loan market, at least the leveraged segment of it, across the Rubicon, to a full-fledged capital market.

Initially, arrangers invoked flex language to make loans more attractive to investors by hiking the spread or lowering the price. Over time, however, market-flex became a tool either to increase or decrease pricing of a loan, based on investor reaction.

As a result of market flex, a loan syndication today functions as a book-building exercise, in the parlance of the bond market. A loan is originally launched to market at a target spread or with a range of spreads referred to as "price talk" (that is, a target spread of, say, LIBOR + 250 to LIBOR + 275). Investors then will make commitments that in many cases are tiered by the spread. For example, an account may put in for $25 million at LIBOR + 275 or $15 million at LIBOR + 250. At the end of the process, the arranger will total up the commitments and then make a call on where to price the paper. Following the example above, if the paper is vastly oversubscribed at LIBOR + 250, the arranger may slice the spread further. Conversely, if it is undersubscribed even at LIBOR + 275, then the arranger will be forced to raise the spread to bring more money to the table.

Types of Syndications

There are three types of syndications: an underwritten deal, a "best-efforts" syndication, and a "club deal."

An underwritten deal is one for which the arrangers guarantee the entire commitment, then syndicate the loan. If the arrangers are unable to fully subscribe the loan, they are forced to absorb the difference, which they may later try to sell to investors. This is easy, of course, if market conditions, or the credit's fundamentals, improve. If not, the arranger may be forced to sell at a discount and, potentially, even take a loss on the paper. Or the arranger may just be left above its desired hold level of the credit. So, why do arrangers underwrite loans? First, offering an underwritten loan can be a competitive tool to win mandates. Second, underwritten loans usually require more lucrative fees because the agent is on the hook if potential lenders balk. Of course, with flex-language now common, underwriting a deal does not carry the same risk it once did when the pricing was set in stone prior to syndication.

A best-efforts syndication is one for which the arranger group commits to underwrite less than the entire amount of the loan, leaving the credit to the vicissitudes of the market. If the loan is undersubscribed, the credit may not close—or may need major surgery to clear the market. Traditionally, best-efforts syndications were used for risky borrowers or for complex transactions. Since the late 1990s, however, the rapid acceptance of market-flex language has made best-efforts loans the rule even for investment-grade transactions.

A club deal is a smaller loan (usually $25 million to $100 million, but as high as $150 million) that is premarketed to a group of relationship lenders. The arranger is generally a first among equals, and each lender gets a full cut, or nearly a full cut, of the fees.

THE SYNDICATION PROCESS

The Information Memo or "Bank Book"

Before awarding a mandate, an issuer might solicit bids from arrangers. The banks will outline their syndication strategy and qualifications, as well as their view on the way the loan will price in market. Once the mandate is awarded, the syndication process starts. The arranger will prepare an information memo (IM) describing the terms of the transactions. The IM typically will include an executive summary, investment considerations, a list of terms and conditions, an industry overview, and a financial model.

Because loans are not securities, this will be a confidential offering made only to qualified banks and accredited investors. If the issuer is non-investment grade and seeking capital from non-bank investors, the arranger will often prepare a "public" version of the IM. This version will be stripped of all confidential material such as management financial projections so that it can be viewed by accounts that operate on the public side of the wall or that want to preserve their ability to buy bonds or stock or other public securities of the particular issuer (see the Public Versus Private section below). Naturally, investors that view materially non-public information of a company are disqualified from buying the company's public securities for some period of time.

As the IM (or "bank book," in traditional market parlance) is being prepared, the syndicate desk will solicit informal feedback from potential investors on what their appetite for the deal will be and at what price they are willing to invest. Once this intelligence has been gathered, the agent will formally market the deal to potential investors. The information provided would include:

- The executive summary will include a description of the issuer, an overview of the transaction and rationale, sources and uses, and key statistics on the financials.
- Investment considerations will be, basically, management's sales "pitch" for the deal.
- The list of terms and conditions will be a preliminary term sheet describing the pricing, structure, collateral, covenants, and other terms of the credit (covenants are usually negotiated in detail after the arranger receives investor feedback).
- The industry overview will be a description of the company's industry and competitive position relative to its industry peers.
- The financial model will be a detailed model of the issuer's historical, pro forma, and projected financials including management's high, low, and base case for the issuer.

Most new acquisition-related loans are kicked off at a bank meeting at which potential lenders hear management and the sponsor group (if there is one) describe what the terms of the loan are and what transaction it backs. Management will provide its vision for the transaction and, most important, tell why and how the lenders will be repaid on or ahead of schedule. In addition, investors will be briefed regarding the multiple exit strategies, including second ways out via asset sales. (If it is a small deal or a refinancing instead of a formal meeting, there may be a series of calls or one-on-one meetings with potential investors.)

Once the loan is closed, the final terms are then documented in detailed credit and security agreements. Subsequently, liens are perfected and collateral is attached.

Loans, by their nature, are flexible documents that can be revised and amended from time to time. These amendments require different levels of approval (see the discussion of voting rights later in this chapter). Amendments can range from something as simple as a covenant waiver to something as complex as a change in the collateral package or allowing the issuer to stretch out its payments or make an acquisition.

The Loan Investor Market

There are three primary-investor consistencies: banks, finance companies, and institutional investors.

"Banks," in this case, can be either a commercial bank, a savings and loan institution or a securities firm that usually provide investment-grade loans. These are typically large revolving credits that back commercial paper or general corporate purposes or, in some cases, acquisitions. For leveraged loans, banks typically provide unfunded revolving credits, letter of credits (LOCs), and—although they are becoming increasingly less common—amortizing term loans, under a syndicated loan agreement.

Finance companies have consistently represented less than 10% of the leveraged loan market, and tend to play in smaller deals—$25 million to $200 million. These investors often seek asset-based loans that carry wide spreads and that often feature time-intensive collateral monitoring.

Institutional investors in the loan market are principally structured vehicles known as collateralized loan obligations (CLO) and loan participation mutual funds (known as "prime funds" because they were originally pitched to investors as a money market–like fund that would approximate the prime rate). In addition, hedge funds, *high-yield* bond funds, pension funds, insurance companies, and other proprietary investors do participate opportunistically in loans. Typically, however, they invest principally in wide-margin loans (referred to by some players as "high-octane" loans), with spreads of LIBOR + 500 or higher.

Collateralized loan obligations (CLOs) are special-purpose vehicles set up to hold and manage pools of leveraged loans. The special-purpose vehicle is financed with several tranches of debt (typically a AAA rated tranche, a AA tranche, a BBB tranche, and a mezzanine tranche) that have rights to the collateral and payment stream in descending order. In addition, there is an equity tranche, but the equity tranche is usually not rated. CLOs are created as arbitrage vehicles that generate equity returns through leverage, by issuing debt 10 to 11 times their equity contribution. There are also market-value CLOs that are less leveraged—typical 3 to 5 times—and allow managers more flexibility than more tightly structured arbitrage deals. CLOs are usually rated by two of the three major ratings agencies and impose a series of covenant tests on collateral managers, including minimum rating, industry diversification, and maximum default basket.

Prime funds are how retail investors can access the loan market. They are mutual funds that invest in leveraged loans. Prime funds were first introduced in the late 1980s. Most of the original prime funds were continuously offered funds with quarterly tender periods. Managers then rolled true closed-end, exchange-traded funds in the early 1990s. It was not until the early 2000's that fund complexes introduced open-ended funds that were redeemable each day.

PUBLIC VERSUS PRIVATE

In the old days, the line between public and private information in the loan market was a simple one. Loans were strictly on the private side of the wall and any information transmitted between the issuer and the lender group remained confidential.

In the late 1980s that line began to blur as a result of two market innovations. The first was more active secondary trading that sprung up to support (1) the entry of non-bank investors in the market, such as insurance companies and loan mutual funds and (2) to help banks sell rapidly expanding portfolios of distressed and highly leveraged loans that they no longer wanted to hold. This meant that parties that were insiders on loans might now exchange confidential information with traders and potential investors who were not (or not yet) a party to the loan. The second innovation that weakened the public-private divide was trade journalism focused on the loan market.

Despite these two factors, the public versus private line was well understood and rarely controversial for at least a decade. This changed in the early 2000s as a result of (1) the explosive growth of non-bank investors groups, which included a growing number of institutions that operated on the public side of the wall, including a growing number of mutual funds, hedge funds and even CLO boutiques (2) the growth of the credit default swaps market, in which insiders like banks often sold or bought protection from institutions that were not privy to inside information and (3) a more aggressive effort by the press to report on the loan market.

Some background is in order.

The vast majority of loans are unambiguously private financing arrangements between issuers and their lenders. Even for issuers with public equity or debt that file with the Securities & Exchange Commission (SEC), the credit agreement only becomes public when it is filed, often long after closing, as an exhibit to an annual report (10K), a quarterly report (10Q), a current report (8K) or some other document (proxy statement, securities registration, etc.).

Beyond the credit agreement, there is a raft of ongoing correspondence between issuers and lenders that is made under confidentiality agreements including quarterly or monthly financial disclosures, covenant compliance information, amendment and waiver requests and financial projections as well as plans for acquisitions or dispositions. Much of this information may be material to the financial health of the issuer and may be out of the public domain until the issuer formally puts out a press release or files an 8K or some other document with the SEC.

In recent years, this information has leaked into the public domain either via off-line conversations or the press. It has also come to light through mark-to-market pricing services, which often report significant movement in a loan price without any corresponding news. This is usually an indication that the banks have received negative or positive information that is not yet public.

By 2006, there was growing concern among issuers, lenders and regulators that this migration of once private information into public hands might breach confidentiality agreements between lenders and issuers and, more important, could lead to illegal trading. How has the market contended with these issues?

- **Traders.** In order to insulate themselves from violating regulations, some dealers and buyside firms have set up their trading desks on the public side of the wall. Consequently, traders, salespeople and analysts do not receive private information even if somewhere else in the institution the private data are available. This is the same technique that investment banks have used from time immemorial to separate their private investment banking activities from their public trading and sales activities.
- **Underwriters.** As mentioned above, in most primary syndications arrangers will prepare a public version of information memoranda that is scrubbed of private information like projections. These IM's will be distributed to accounts that are on the public side of the wall. As well, underwriters will ask public accounts to attend a public version of the bank meeting and distribute to these accounts only scrubbed financial information.
- **Buy-side accounts.** On the buy side there are firms that operate on either side of the public-private fence. Accounts that operate on the private side receive all confidential materials and agree to not trade in public securities of the issuers for which they get private information. These groups are often part of wider investment complexes that do have public funds and portfolios but, via Chinese walls, are sealed from these parts of the firms. There are also accounts that are public. These firms take only public IMs and public materials and, therefore, retain the option to trade in the public securities markets even when an issuer for which they own a loan is involved. This can be tricky to pull off in practice because in the case of an amendment the lender could be called on to approve or decline in the absence of any real information. Or the account could either (1) designate one person who is on the private side of the wall to sign off on amendments (2) empower its trustee or the loan arranger to do so. But it's a complex proposition.
- **Vendors.** Vendors of loan data, news and prices also face many challenges in managing the flow of public and private information. In generally, the vendors operate under the freedom of the press provision of the first amendment and report on information in a way that anyone can simultaneously receive it; for a price of course. Therefore, the information is essentially made

public in a way that doesn't deliberately disadvantage any party—whether it's a news story discussing the progress of an amendment or an acquisition. Or it's a price change reported by a mark-to-market service. This, of course, doesn't deal with the underlying issue that someone who is a party to confidential information is making it available via the press or prices to a broader audience.

Another way in which participant deal with the public versus private issue is to ask counterparties to sign "big boy" letters acknowledging that there may be information they are not privy to and they are agreeing to make the trade in any case. They are, effectively, big boys and will accept the risks.

The introduction of loan credit default swaps into the fray (see below) adds another wrinkle to this topic because a whole new group of public investors could come into play if that market catches fire.

CREDIT RISK: AN OVERVIEW

Pricing a loan requires arrangers to evaluate the risk inherent in a loan and to gauge investor appetite for that risk. The principal credit risk factors that banks and institutional investors contend with in buying loans are default risk and loss-given-default risk. Among the primary ways that accounts judge these risks are: ratings, credit statistics, industry sector trends, management strength and sponsor. All of these, together, tell a story about the deal.

Below we provide a brief description of the major risk factors.

Default Risk

Default risk is simply the likelihood of a borrower's being unable to pay interest or principal on time. It is based on the issuer's financial condition, industry segment, and conditions in that industry and economic variables and intangibles, such as company management. Default risk will, in most cases, be most visibly expressed by a public rating from Standard & Poor's or another ratings agency. These ratings range from AAA for the most creditworthy loans to CCC for the least.

The market is divided, roughly, into two segments: investment grade (loans rated BBB– or higher) and leveraged (borrowers rated BB+ or lower). Default risk, of course, varies widely within each of these broad segments. Since the mid-1990s, public loan ratings have become a de facto requirement for issuers that wish to tap the leveraged loan market, which, as noted above, is now dominated by institutional investors. Unlike banks, which typically have large credit departments and adhere to internal rating scales, fund managers rely on agency ratings to bracket risk and explain the overall risk of their portfolios to their own investors.

Loss-Given-Default Risk

Loss-given-default risk measures how severe a loss the lender would incur in the event of default. Investors assess this risk based on the collateral (if any) backing the loan and the amount of other debt and equity subordinated to the loan. Lenders will also look to covenants to provide a way of coming back to the table early—that is, before other creditors—and renegotiating the terms of a loan if the issuer fails to meet financial targets.

Investment-grade loans are, in most cases, senior unsecured instruments with loosely drawn covenants that apply only at incurrence, that is, only if an issuer makes an acquisition or issues debt. As a result, loss given default may be no different from risk incurred by other senior unsecured creditors. Leveraged loans, by contrast, are, in virtually all cases, senior secured instruments with tightly drawn maintenance covenants, that is, covenants that are measured at the end of each quarter whether or not the issuer takes any action. Loan holders, therefore, almost always are first in line among prepetition creditors and, in many cases are able to renegotiate with the issuer before the loan becomes severely impaired. It is no surprise, then, that loan investors historically fare much better than other creditors on a loss-given-default basis.

Credit statistics are used by investors to help calibrate both default and loss-given-default risk. These statistics include a broad array of financial data, including credit ratios measuring leverage (debt to capitalization and debt to earnings before interest, taxes, depreciation, and amortization [EBITDA]) and coverage (EBITDA to interest, EBITDA to debt service, operating cash flow to fixed charges). Of course, the ratios investors use to judge credit risk vary by industry.

In addition to looking at trailing and pro forma ratios, investors look at management's projections and the assumptions behind these projections to see if the issuer's game plan will allow it to pay its debt comfortably. There are ratios that are most geared to assessing default risk. These include leverage and coverage. Then there are ratios that are suited for evaluating loss-given-default risk. These include collateral coverage, or the value of the collateral underlying the loan relative to the size of the loan and the ratio of senior secured loan to junior debt in the capital structure.

Logically, the likely severity of loss-given-default for a loan increases with the size of the loan just as a percent of the overall debt structure also does. After all, if an issuer defaults on $100 million of debt, of which $10 million is in the form of senior secured loans, the loans are more likely to be fully covered in bankruptcy than if the loan totals $90 million.

Industry is a factor, because sectors, naturally, go in and out of favor. For that reason, having a loan in a desirable sector, like telecom in the late 1990s or healthcare in the early 2000s, can really help a syndication along. Also, defensive loans (like consumer products) can be more appealing in a time of economic uncertainty, whereas cyclical borrowers (like chemicals or autos) can be more appealing during an economic upswing.

Sponsorship is a factor, too. Needless to say, many leveraged companies are owned by one or more private equity firms. These entities, such as Kohlberg Kravis & Roberts or Carlyle Group, invest in companies that have leveraged capital structures. To the extent that the sponsor group has a strong following among loan investors, a loan will be easier to syndicate and, therefore, can be priced lower. In contrast, if the sponsor group does not have a loyal set of relationship lenders, the deal may need to be priced higher to clear the market. Among banks, investment factors may include whether or not the bank is party to the sponsor's equity fund. Among institutional investors, weighting is given to an individual deal sponsor's track record in fixing its own impaired deals by stepping up with additional equity or replacing a management team that is failing.

SYNDICATING A LOAN BY FACILITY

Most loans are structured and syndicated to accommodate the two primary syndicated lender constituencies: banks (domestic and foreign) and institutional investors (primarily structured finance vehicles, mutual funds and insurance companies). As such, leveraged loans consist of:

- Pro rata debt consists of the revolving credit and amortizing term loan (TLa), which are packaged together and, usually, syndicated to banks. In some loans, however, institutional investors take pieces of the TLa and, less often, the revolving credit, as a way to secure a larger institutional term loan allocation. Why are these tranches called "pro rata?" Because arrangers historically syndicated revolving credit and TLa's on a pro rata basis to banks and finance companies.
- Institutional debt consists of term loans structured specifically for institutional investors, though there are also some banks that buy institutional term loans. These tranches include first-lien loans, second-lien loans as well as prefunded letters of credit. Traditionally, institutional tranches were referred to a TLb's because they were bullet payments and lined up behind TLa's.

Finance companies also play in the leveraged loan market, and buy both pro rata and institutional tranches. With institutional investors playing an ever-larger role, however, by 2006 many executions were structured as simply revolving credit/institutional term loans, with the TLa falling by the wayside.

PRICING A LOAN IN THE PRIMARY MARKET

Pricing loans for the institutional market is a straightforward exercise based on simple risk/return consideration and market technicals. Pricing a loan for the bank market, however, is more complex. Indeed, banks often invest in loans for more than pure spread income. Rather, banks are driven by the overall profitability of the issuer relationship, including noncredit revenue sources.

Pricing Loans for Bank Investors

Since the early 1990s, almost all large commercial banks have adopted portfolio-management techniques that measure the returns of loans and other credit products relative to risk. By doing so, banks have learned that loans are rarely compelling investments on a stand-alone basis. Therefore, banks are reluctant to allocate capital to issuers unless the total relationship generates attractive returns—whether those returns are measured by risk-adjusted return on capital, by return on economic capital, or by some other metric.

If a bank is going to put a loan on its balance sheet, then it takes a hard look not only at the loan's pricing, but also at other sources of revenue from the relationship, including noncredit businesses—like cash-management services and pension-fund management—and economics from other capital markets activities, like bonds, equities, or M&A advisory work.

This process has had a breathtaking result on the leveraged loan market—to the point that it is an anachronism to continue to call it a "bank" loan market.

What this means is that the spread offered to pro rata investors is important, but even more important, in most cases, is the amount of other, fee-driven business a bank can capture by taking a piece of a loan. For this reason, issuers are careful to award pieces of bond- and equity-underwriting engagements and other fee-generating business to banks that are part of its loan syndicate.

Pricing Loans for Institutional Players

For institutional investors, the investment decision process is far more straightforward, because, as mentioned above, they are focused not on a basket of revenue, but only on loan-specific revenue.

In pricing loans to institutional investors, it's a matter of (1) the spread of the loan relative to credit quality and (2) market-based factors. This second category can be divided into liquidity and market technicals (that is, supply/demand).

Liquidity is the tricky part, but, as in all markets, all else being equal, more liquid instruments command thinner spreads than less liquid ones. In the old day—before institutional investors were the dominant investors and banks were less focused on portfolio management—the size of a loan didn't much matter. Loans sat on the books of banks and stayed there. But now that institutional investors and banks put a premium on the ability to package loans and sell them, liquidity has become important. As a result, smaller executions—generally those of $200 million or less—tend to be priced at a premium to the larger loans. Those in the middle, $200 million to $2 billion, were, through 2006 at least, the market's "sweet spot," within investor capacity and sizable enough to generate secondary interest. Those exceeding $2 billion would often command a spread premium to compensate investors for stepping up for larger pieces. These numbers represent a rough guide, although they do, naturally, move around, depending on the supply/demand dynamics of the market.

Market technicals, or supply relative to demand, is a matter of simple economics. If there are a lot of dollars chasing little product, then, naturally, issuers will be able to command lower spreads. If, however, the opposite is true, then spreads will need to increase for loans to clear the market.

MARK-TO-MARKET'S EFFECT

Beginning in 2000, the SEC directed bank loan mutual fund managers to use available mark-to-market data (bid/ask levels reported by secondary traders and compiled by mark-to-market services) rather than fair value (estimated prices), to determine the value of broadly syndicated loans for portfolio-valuation purposes. In broad terms, this policy has made the market more transparent, improved price discovery and, in doing so, made the market far more efficient and dynamic than it was in the past. In the primary, for instance, leveraged loan spreads are now determined not only by rating and leverage profile, but also by trading levels of an issuer's previous loans and, often, bonds. Issuers and investors can also look at the trading levels of comparable loans for market-clearing levels. What's more, and market sentiment tied to supply and demand. As a result, new-issue spreads rise and fall far more rapidly than in the past, when spreads were more or less the same for every leveraged transaction.

TYPES OF SYNDICATED LOAN FACILITIES

There are four main types of syndicated loan facilities: a revolving credit (within which are options for swingline loans, multicurrency-borrowing, competitive-bid options, term-out, and evergreen extensions); a term loan; an LOC; and an acquisition or equipment line (a delayed-draw term loan).

A revolving credit line allows borrowers to draw down, repay, and reborrow. The facility acts much like a corporate credit card, except that borrowers are charged an annual commitment fee on unused amounts, which drives up the overall cost of borrowing (the facility fee). Revolvers to non-investment-grade issuers are often tied to borrowing-base lending formulas. This limits borrowings to a certain percentage of collateral, most often receivables and inventory. Revolving credits often run for 364 days. These revolving credits—called, not surprisingly, 364-day facilities—are generally limited to the investment-grade market.

The reason for what seems like an odd term is that regulatory capital guidelines mandate that, after one year of extending credit under a revolving facility, banks must then increase their capital reserves to take into account the unused amounts. Therefore, banks can offer issuers 364-day facilities at a lower unused fee than a multiyear revolving credit.

There are a number of options that can be offered within a revolving credit line:

- A swingline is a small, overnight borrowing line, typically provided by the agent.
- A multicurrency line may allow the borrower to borrow in several currencies.
- A competitive-bid option (CBO) allows borrowers to solicit the best bids from its syndicate group. The agent will conduct what amounts to an auction to raise funds for the borrower, and the best bids are accepted. CBOs, typically, are available only to large, investment-grade borrowers.
- A term-out will allow the borrower to convert borrowings into a term loan at a given conversion date. This, again, is usually a feature of investment-grade loans. Under the option, borrowers may take what is outstanding under the facility and pay it off according to a predetermined repayment schedule. Often the spreads ratchet up if the term-out option is exercised.
- An evergreen is an option for the borrower—with consent of the syndicate group—to extend the facility each year for an additional year.

A term loan is simply an installment loan, such as a loan one would use to buy a car. The borrower may draw on the loan during a short commitment period and repays it based on either a scheduled series of repayments or a one-time lump-sum payment at maturity (bullet payment). There are two principal types of term loans. The first is an amortizing term loan (A-term loans, or TLa), which is a term loan with a progressive repayment schedule that typically runs six years or less. These loans are normally syndicated to banks along with revolving credits as part of a larger syndication. Starting in 2000, A-term loans became increasingly rare, as issuers bypassed the less accommodating bank market and tapped institutional investors for all or most of their funded loans.

The other type of term loan is an institutional loan (B-term, C-term, or D-term loans), which is a term loan facility carved out for nonbank, institutional investors. These loans came into broad usage during the mid-1990s as the institutional loan investor base grew. Until 2001, these loans were, in almost all cases, priced higher than amortizing term loans, because they had longer maturities and back-end–loaded repayment schedules. The tide turned, however, in late 2001, and through 2006 the spread on a growing percentage of these facilities into parity with (in some cases even lower than) revolvers and A-term loans. This is especially true when institutional demand runs high. This institutional category also includes second-lien loans and covenant-lite loans, which are described below.

LOCs differ, but, simply put, they are guarantees provided by the bank group to pay off debt or obligations if the borrower cannot.

Acquisition/equipment lines (delayed-draw term loans) are credits that may be drawn down for a given period to purchase specified assets or equipment or to make acquisitions. The issuer pays a fee during the commitment period (a ticking fee). The lines are then repaid over a specified period (the term-out period). Repaid amounts may not be reborrowed.

SECOND-LIEN LOANS

Although they are really just another type of syndicated loan facility, *second-lien loans* are sufficiently complex to warrant a separate section in this primer. After a brief flirtation with second-lien loans in the mid-1990s, these facilities fell out of favor after the Russian debt crisis caused investors to adopted a more cautious tone. But after default rates fell precipitously in 2003, arrangers rolled out second-lien facilities to help finance issuers struggling with liquidity problems. By 2006, the market had accepted second-lien loans to finance a wide array of transactions, including acquisitions and recapitalizations. Arrangers tap nontraditional accounts—hedge funds, distress investors, and high-yield accounts—as well as traditional CLO and prime fund accounts to finance second-lien loans.

As their name implies, the claims on collateral of second-lien loans are behind those of first-lien loans. Second-lien loans also typically have less restrictive covenant packages, in which maintenance covenant levels are set wide of the first-lien loans. As a result, second-lien loans are priced at a premium to first-lien loans. This premium typically starts at 200 basis points (bp) when the collateral coverage goes far beyond the claims of both the first- and second-lien loans to more than 1,000 bps for less generous collateral.

There are, lawyers explain, two main ways in which the collateral of second-lien loans can be documented. Either the second-lien loan can be part of a single security agreement with first-lien loans, or they can be part of an altogether separate agreement. In the case of a single agreement, the agreement would apportion the collateral, with value going first, obviously, to the first-lien claims and next to the second-lien claims. Alternatively, there can be two entirely separate agreements. Here's a brief summary:

- In a single security agreement, the second-lien lenders are in the same creditor class as the first-lien lenders from the standpoint of a bankruptcy, according to lawyers who specialize in these loans. As a result, for adequate protection to be paid the collateral must cover both the claims of the first- and second-lien lenders. If it does not, the judge may choose to not pay adequate protection or to divide it pro rata among the first- and second-lien creditors. In addition, the second-lien lenders may have a vote as secured lenders equal to those of the first-lien lenders. One downside for second-lien lenders is that these facilities are often smaller than the first-lien loans and, therefore, when a vote comes up first-lien lenders can outvote second-lien lenders to promote their own interests.
- In the case of two separate security agreements, divided by a standstill agreement, the first- and second-lien lenders are likely to be divided into two separate creditor classes. As a result, second-lien lenders do not have a voice in the first-lien creditor committees. As well, first-lien lenders can receive adequate protection payments even if collateral covers their claims, but does not cover the claims of the second-lien lenders. This may not be the case if the loans are documented together and the first- and second-lien lenders are deemed a unified class by the bankruptcy court.

COVENANT-LITE LOANS

Like second-lien loans, covenant-lite loans are really just another type of syndicated loan facility. But they also are sufficiently different to warrant a more detailed discussion in this chapter.

At the most basic level, covenant-lite loans are loans that have bond-like financial incurrence covenants rather than traditional maintenance covenants that are normally part and parcel of a loan agreement. What's the difference?

Incurrence covenants generally require that if an issuer takes an action (paying a dividend, making an acquisition, issuing more debt), it would need to still be in compliance. So, for instance, an issuer that has an incurrence test that limits its debt to five times cash flow would only be able to take on more debt if, on a pro forma basis, it was still within this constraint. If, not then it would be in breech of the covenant and in technical default on the loan. If, on the other hand, an issuer found itself above this five times threshold simply because its earnings had deteriorated, it does not violate the covenant.

Maintenance covenants are far more restrictive. This is because they require an issuer to meet certain financial tests every quarter whether or not it takes an action. So, in the case above had the 5 times leverage maximum been a maintenance rather than incurrence test, the issuer would need to pass it each quarter and would be in violation if either its earnings eroded or its debt level increased. For lenders, clearly, maintenance tests are preferable because it allows them to take action earlier if an issuer experiences financial distress.

Conversely, issuers prefer incurrence covenants precisely because they are less stringent. Covenant-lite loans, therefore, thrive only in the hottest markets when the supply/demand equation is tilted persuasively in favor of issuers.

LENDER TITLES

In the formative days of the syndicated loan market (the late 1980s), there was usually one agent that syndicated each loan. "Lead manager" and "manager" titles were doled out in exchange for large commitments. As league tables (which rank underwriters by their transaction volume in different capital market segments each year) gained influence as a marketing tool, "co-agent" titles were often used in attracting large commitments or in cases where these institutions truly had a role in underwriting and syndicating the loan.

During the 1990s, the use of league tables and, consequently, title inflation exploded. Indeed, the co-agent title has become largely ceremonial today, routinely awarded for what amounts to no more than large retail commitments. In most syndications, there is one lead arranger. This institution is considered to be on the "left" (a reference to its position in a tombstone ad). There are also likely to be other banks in the arranger group, which may also have a hand in underwriting and syndicating a credit. These institutions are said to be on the "right."

The different titles used by significant participants in the syndications process are administrative agent, syndication agent, documentation agent, agent, co-agent or managing agent, and lead arranger or book runner and they are described below:

- **Administrative agent.** The bank that handles all interest and principal payments and monitors the loan.
- **Syndication agent.** The bank that handles, in purest form, the syndication of the loan. Often, however, the syndication agent has a less specific role.
- **Documentation agent.** The bank that handles the documents and chooses the law firm.
- **Agent.** Title used to indicate the lead bank when there is no other conclusive title available, as is often the case for smaller loans.
- **Co-agent or managing agent.** Largely a meaningless title used mostly as an award for large commitments.
- **Lead arranger or book runner.** A league table designation used to indicate the "top dog" in a syndication.

SECONDARY SALES

Secondary sales occur after the loan is closed and allocated, when investors are free to trade the paper. Loan sales are structured as either assignments or participations, with investors usually trading through dealer desks at the large underwriting banks. Dealer-to-dealer trading is almost always conducted through a "street" broker.

When a loan is paying current interest, it is referred to as a performing loan. These loans are typically traded in the par market, where prices are typically close to 100 cents on the dollar. If the loan defaults or the borrower's credit condition deteriorates significantly, the secondary price of the loan will naturally decline. When loans are trading at 80 cents on the dollar or less, they are referred to as distressed loans and usually traded off the distressed debt desks of dealers.

Assignments

In an assignment, the assignee becomes a direct signatory to the loan and receives interest and principal payments directly from the administrative agent.

Assignments typically require the consent of the borrower and agent, although consent may be withheld only if a reasonable objection is made. In many loan agreements, the issuer loses its right to consent in the event of default.

The loan document usually sets a minimum assignment amount, usually $5 million for pro rata commitments. In the late 1990s, however, administrative agents started to break out specific assignment minimums for institutional tranches. In most cases, institutional assignment minimums were reduced to $1 million in an effort to boost liquidity. There were also some cases where assignment fees were reduced or even eliminated for institutional assignments, but these lower assignment fees remained rare into 2006, and the vast majority was set at the traditional $3,500.

One market convention that became firmly established in the late 1990s was assignment-fee waivers by arrangers for trades crossed through its secondary trading desk. This was a way to encourage investors to trade with the arranger rather than with another dealer. This is a significant incentive to trade with arranger—or a deterrent to not trade away, depending on your perspective—because a $3,500 fee amounts to between 7 bps to 35 bps of a $1 million to $5 million trade.

Primary Assignments

The term "primary assignments" is something of an oxymoron. It applies to primary commitments made by offshore accounts (principally CLOs and hedge funds). These vehicles, for a variety of tax reasons, suffer tax consequence from buying loans in the primary. The agent will therefore hold the loan on its books for some short period after the loan closes and then sell it to these investors via an assignment. These are called primary assignments and are effectively primary purchases.

Participations

A participation is an agreement between an existing lender and a participant. As the name implies, it means the buyer is taking a participating interest in the existing lender's commitment.

The lender remains the official holder of the loan, with the participant owning the rights to the amount purchased. Consents, fees, or minimums are almost never required. The participant has the right to vote only on material changes in the loan document (rate, term, and collateral). Nonmaterial changes do not require approval of participants. A participation can be a riskier way of purchasing a loan, because, in the event of a lender becoming insolvent or defaulting, the participant does not have a direct claim on the loan. In this case, the participant then becomes a creditor of the lender and often must wait for claims to be sorted out to collect on its participation.

DERIVATIVES—LOAN CREDIT DEFAULT SWAPS

Traditionally, accounts bought and sold loans in the cash market through assignments and participations. Aside from that, there was little synthetic activity outside over-the-counter total rate of return swaps. By 2006, however, a nascent market for synthetically trading loans was budding.

Loan credit default swaps (LCDS) are standard derivatives that have secured loans as reference instruments. In June 2006, The International Settlement and Dealers Association (ISDA) issued a standard trade confirmation for LCDS contracts.

Like all credit default swaps (CDS), LCDS is basically an insurance contract. The seller is paid a spread in exchange for agreeing to buy at par, or a pre-negotiated price, a loan in the event that loan defaults. LCDS enables participants to synthetically buy a loan by going short the CDS or sell the loan by going long the CDS. Theoretically, then, a loan holder can hedge a position either directly (by buying CDS protection on that specific name) or indirectly (by buying protection on a comparable name or basket of names).

Moreover, unlike the cash markets, which are long-only markets for obvious reasons, the CDS market provides a way for investors to short a loan. To do so, the investor would buy protection on a loan that it doesn't hold. If the loan subsequently defaults, the buyer of protection should be able to purchase the loan in the secondary market at a discount and then deliver it at par to the counterparty from which it bought LCDS contract. For instance, say an account buys five-year protection for a given loan, for which it pays 250 bps a year. Then in year two the loan goes into default and the market price falls to 80 percent of par. The buyer of the protection can then buy the loan at 80 and deliver to the counterpart at 100, a 20-point pickup. Or instead of physical delivery, some buyers of protection may prefer cash settlement in which the difference between the current market price and the delivery price is determined by polling dealers or using a third-party pricing service. Cash settlement could also be employed if there's not enough paper to physically settle all LCDS contracts on a particular loan.

As of this writing the LCDS market was still in its infancy and therefore additional context is yet to come. In addition to these specific loan contracts—or single-name LCDS—the market was also developing methods to synthetically trade a basket, or index, of loans synthetically. This is similar to the IBOXX for high-yield bonds. Investors can trade into an index of loans at a price. For instance, it can buy the index at a price today and a month later sell it at the current price to close the position. Or, conversely, it can sell the index at a price and then cover that sale later buy buying the index. As this implies, index-based derivatives are a way to take a bet on the market or hedge a portfolio against market risk.

PRICING TERMS

Rates

Bank loans usually offer borrowers different interest-rate options. Several of these options allow borrowers to lock in a given rate for one month to one year. Pricing on many loans is tied to performance grids, which adjust pricing by one or more financial criteria. Pricing is typically tied to ratings in investment-grade loans and to financial ratios in leveraged loans. Communications loans are invariably tied to the borrower's debt-to-cash-flow ratio.

Syndication pricing options include prime, LIBOR, CD, and other fixed-rate options:

- The prime is a floating-rate option. Borrowed funds are priced at a spread over the reference bank's prime lending rate. The rate is reset daily, and borrowers may be repaid at any time without penalty. This is typically an overnight option, because the prime option is more costly to the borrower than LIBOR or CDs.
- The LIBOR (or Eurodollars) option is so called because, with this option, the interest on borrowings is set at a spread over LIBOR for a period of one month to one year. The corresponding LIBOR rate is used to set pricing. Borrowings cannot be prepaid without penalty.
- The CD option works precisely like the LIBOR option, except that the base rate is certificates of deposit, sold by a bank to institutional investors.
- Other fixed-rate options are less common but work like the LIBOR and CD options. These include federal funds (the overnight rate charged by the Federal Reserve to member banks) and cost of funds (the bank's own funding rate).

Fees

The fees associated with syndicated loans are the upfront fee, the commitment fee, the facility fee, the administrative agent fee, the letter of credit (LOC) fee, and the cancellation or prepayment fee.

An up-front fee, which is the same as an original-issue discount in the bond market, is a fee paid by the issuer. It is often tiered, with the lead arranger receiving a larger amount in consideration of its structuring and/or underwriting the loan. Co-underwriters will receive a lower fee, and then the general syndicate will likely have fees tied to their commitment. Most often, fees are paid on a lender's final allocation. For example, a loan has two fee tiers: 100 bps (or 1%) for \$25 million commitments and 50 bps for \$15 million commitments. A lender committing to the \$25 million tier will be paid on its final allocation rather than on initial commitment, which means that, in this example, the loan is oversubscribed and lenders committing \$25 million would be allocated \$20 million and the lenders would receive a fee of \$200,000 (or 1% of \$20 million). Sometimes upfront fees will be structured as a percentage of final allocation plus a flat fee. This happens most often for larger fee tiers, to encourage potential lenders to step up for larger commitments. The flat fee is paid regardless of the lender's final allocation. Fees are usually paid to banks, mutual funds, and other non-offshore investors as an upfront payment. CLOs and other offshore vehicles are typically brought in after the loan closes as a "primary" assignment, and they simply buy the loan at a discount equal to the fee offered in the primary assignment, for tax purposes.

A commitment fee is a fee paid to lenders on undrawn amounts, under a revolving credit or a term loan prior to draw-down. On term loans, this fee is usually referred to as a "ticking" fee.

A facility fee, which is paid on a facility's entire committed amount, regardless of usage, is often charged instead of a commitment fee on revolving credits to investment-grade borrowers, because these facilities typically have collateralized bond obligations (CBOs) that allow a borrower to solicit the best bid from its syndicate group for a given borrowing. The lenders that do not lend under the CBO are still paid for their commitment.

A *usage fee* is a fee paid when the utilization of a revolving credit falls below a certain minimum. These fees are applied mainly to investment-grade loans and generally call for fees based on the utilization under a revolving credit. In some cases, the fees are for high use and, in some cases, for low use. Often, either the facility fee or the spread will be adjusted higher or lower based on a preset usage level.

A prepayment fee is a feature generally associated with institutional term loans. This fee is seen mainly in weak markets as an inducement to institutional investors. Typical prepayment fees will be set on a sliding scale; for instance, 2% in year one and 1% in year two. The fee may be applied to all repayments under a loan or "soft" repayments, those made from a refinancing or at the discretion of the issuer (as opposed to hard repayments made from excess cash flow or asset sales).

An administrative agent fee is the annual fee typically paid to administer the loan (including to distribute interest payments to the syndication group, to update lender lists, and to manage borrowings). For secured loans (particularly those backed by receivables and inventory), the agent often collects a collateral monitoring fee, to ensure that the promised collateral is in place.

A letter of credit fee can be any one of several types. The most common—a fee for standby or financial letters of credit—guarantees that lenders will support various corporate activities. Because these LOCs are considered "borrowed funds" under capital guidelines, the fee is typically the same as the LIBOR margin. Fees for commercial LOCs (those supporting inventory or trade) are usually lower, because in these cases actual collateral is submitted). The LOC is usually issued by a fronting bank (usually the agent) and syndicated to the lender group on a pro rata basis. The group receives the LOC fee on their respective shares, while the fronting bank receives an issuing (or fronting, or facing) fee for issuing and administering the LOC. This fee is almost always 12.5 bps to 25 bps (0.125% to 0.25%) of the LOC commitment.

Voting Rights

Amendments or changes to a loan agreement must be approved by a certain percentage of lenders. Most loan agreements have three levels of approval: required-lender level, full vote, and supermajority:

- The required-lenders level, usually just a simple majority, is used for approval of nonmaterial amendments and waivers or changes affecting one facility within a deal.
- A full vote of all lenders, including participants, is required to approve material changes such as RATS (rate, amortization, term, and security; or collateral) rights, but, as described below, there are occasions when changes in amortization and collateral may be approved by a lower percentage of lenders (a supermajority).
- A supermajority is typically 67% to 80% of lenders and is sometimes required for certain material changes such as changes in amortization (in-term repayments) and release of collateral. Used periodically in the mid-1990s, these provisions fell out of favor by the late 1990s.

COVENANTS

Loan agreements have a series of restrictions that dictate, to varying degrees, how borrowers can operate and carry themselves financially. For instance, one covenant may require the borrower to maintain its existing fiscal-year end. Another may prohibit it from taking on new debt. Most agreements also have financial compliance covenants, for example, that a borrower must maintain a prescribed level of equity, which, if not maintained, gives banks the right to terminate the agreement or push the borrower into default. The size of the covenant package increases in proportion to a borrower's financial risk. Agreements to investment-grade companies are usually thin and simple. Agreements to leveraged borrowers are often much more onerous.

The three primary types of loan covenants are affirmative, negative, and financial.

Affirmative covenants state what action the borrower must take to be in compliance with the loan, such as that it must maintain insurance. These covenants are usually boilerplate and require a borrower to pay the bank interest and fees, maintain insurance, pay taxes, and so forth.

Negative covenants limit the borrower's activities in some way, such as regarding new investments. Negative covenants, which are highly structured and customized to a borrower's specific condition, can limit the type and amount of investments, new debt, liens, asset sales, acquisitions, and guarantees.

Financial covenants enforce minimum financial performance measures against the borrower, such as that he must maintain a higher level of current assets than of current liabilities. The presence of these maintenance covenants—so called because the issuer must maintain quarterly compliance or suffer a technical default on the loan agreement—is a critical difference between loans and bonds. Bonds and covenant-lite loans (see above), by contrast, usually contain incurrence covenants that restrict the borrower's ability to issue new debt, make acquisitions, or take other action that would breach the covenant. For instance, a bond indenture may require the issuer to not incur any new debt if that new debt would push it over a specified ratio of debt to EBITDA. But, if the company's cash flow deteriorates to the point where its debt to EBITDA ratio exceeds the same limit, a covenant violation would not be triggered. This is because the ratio would have climbed organically rather than through some action by the issuer.

As a borrower's risk increases, financial covenants in the loan agreement become more tightly wound and extensive. In general, there are five types of financial covenants—coverage, leverage, current ratio, tangible net worth, and maximum capital expenditures:

- A coverage covenant requires the borrower to maintain a minimum level of cash flow or earnings, relative to specified expenses, most often interest, debt service (interest and repayments), fixed charges (debt service, capital expenditures, and/or rent).
- A leverage covenant sets a maximum level of debt, relative to either equity or cash flow, with the debt-to-cash-flow level being far more common.
- A current-ratio covenant requires that the borrower maintain a minimum ratio of current assets (cash, marketable securities, accounts receivable, and inventories) to current liabilities (accounts payable, short-term debt

of less than one year), but sometimes a "quick ratio," in which inventories are excluded from the numerate, is substituted.

- A tangible-net-worth covenant (TNW) requires that the borrower have a minimum level of TNW (net worth less intangible assets, such as goodwill, intellectual assets, excess value paid for acquired companies), often with a build-up provision, which increases the minimum by a percentage of net income or equity issuance.
- A maximum-capital-expenditures covenant requires that the borrower limit capital expenditures (purchases of property, plant, and equipment) to a certain amount, which may be increased by some percentage of cash flow or equity issuance, but often allowing the borrower to carry forward unused amounts from one year to the next.

MANDATORY PREPAYMENTS

Leveraged loans usually require a borrower to prepay with proceeds of excess cash flow, asset sales, debt issuance, or equity issuance.

Excess cash flow is typically defined as cash flow after all cash expenses, required dividends, debt repayments, capital expenditures and changes in working capital. The typical percentage required is 50% to 75%.

Asset sales are defined as net proceeds of asset sales, normally excluding receivables or inventories. The typical percentage required is 100%.

Debt issuance is defined as net proceeds from debt issuance. The typical percentage required is 100%.

Equity issuance is defined as the net proceeds of equity issuance. The typical percentage required is 25% to 50%.

Often, repayments from excess cash flow and equity issuance are waived if the issuer meets a preset financial hurdle, most often structured as a debt/EBITDA test.

COLLATERAL

In the leveraged market, collateral usually includes all the tangible and intangible assets of the borrower and, in some cases, specific assets that back a loan.

Virtually all leveraged loans and some of the more shaky investment-grade credits are backed by pledges of collateral. In the asset-based market, for instance, that typically takes the form of inventories and receivables, with the amount of the loan tied to a formula based off of these assets. The common rule is that an issuer can borrow against 50% of inventory and 80% of receivables. Naturally, there are loans backed by certain equipment, real estate, and other property.

In the leveraged market, there are some loans—since the early 1990s, very few—that are backed by capital stock of operating units. In this structure, the assets of the issuer tend to be at the operating-company level and are unencumbered by liens, but the holding company pledges the stock of the operating companies to the lenders. This effectively gives lenders control of these units if the company defaults. The risk to lenders in this situation, simply put, is that a bankruptcy court collapses the holding company with the operating companies and effectively renders the stock worthless. In these cases, which happened on a few occasions to lenders to retail companies in the early 1990s, loan holders become unsecured lenders of the company and are put back on the same level with other senior unsecured creditors.

Springing Liens/Collateral Release

Some loans have provisions that borrowers that sit on the cusp of investment-grade and noninvestment-grade must either attach collateral or release it if the issuer's rating changes.

A BBB or BBB– issuer may be able to convince lenders to provide unsecured financing, but lenders may demand springing liens in the event the issuer's credit quality deteriorates. Often, an issuer's rating being lowered to BB+ or exceeding its predetermined leverage level will trigger this provision. Likewise, lenders may demand collateral from a strong, noninvestment-grade issuer, but will offer to release under certain circumstances, such as if the issuer loses its investment-grade rating.

Change of Control

Invariably, one of the events of default in a credit agreement is a change of issuer control.

For both investment-grade and leveraged issuers, an event of default in a credit agreement will be triggered by a merger, an acquisition of the issuer, some substantial purchase of the issuer's equity by a third party, or a change in the majority of the board of directors. For sponsor-backed leveraged issuers, the sponsor's lowering its stake below a preset amount can also trip this clause.

Asset-Based Lending

Most of the information above refers to "cash flow" loans, loans that may be secured by collateral, but are repaid by cash flow. Asset-based lending is a distinct segment of the loan market. These loans are secured by specific assets and usually governed by a borrowing formula (or a "borrowing base"). The most common type of asset based loans are receivables and/or inventory lines. These are revolving credits that have a maximum borrowing limit, say $100 million, but also have a cap based on the value of an issuer's pledged receivables and inventories. Usually, the receivables are pledged and the issuer may borrow against 80%, give or take. Inventories are also often pledged to secure borrowings. However, because they are obviously less liquid than receivables, lenders are less generous in their formula. Indeed, the borrowing base for inventories is typically in the 50% to 65% range. Moreover, the borrowing base may be further divided into subcategories—for instance, 50% of work-in-process inventory and 65% of finished goods inventory.

In many receivables-based facilities, issuers are required to place receivables in a "lock box." That means that the bank lends against the receivable, takes possession of it and then collects it to pay down the loan.

In addition, asset-based lending is often done based on specific equipment, real estate, car fleets, and an unlimited number of other assets.

Loan Math: The Art of Spread Calculation

Calculating loan yields or spreads is not straightforward. Unlike most bonds, which have long no-call periods and high-call premiums, most loans are prepayable at any time typically without prepayment fees. And, even in cases where prepayment fees apply, they are rarely more than 2% in year one and 1% in year two. Therefore, affixing a spread-to-maturity or a spread-to-worst on loans is little more than a theoretical calculation.

This is because an issuer's behavior is unpredictable. It may repay a loan early because a more compelling financial opportunity presents itself or because the issuer is acquired or because it is making an acquisition and needs a new financing. Traders and investors will often speak of loan spreads, therefore, as a spread to a theoretical call. Loans, on average, between 1997 and 2004 had a 15-month average life. So, if you buy a loan with a spread of 250 bps at a price of 101, you might assume your spread-to-expected-life as the 250 bps less the amortized 100 bps premium or LIBOR + 170. Conversely, if you bought the same loan at 99, the spread-to-expect life would be LIBOR + 330.

SUMMARY

This chapter gives readers a detailed primer on the leveraged loan market, including an overview of how loans are underwritten, arranged, syndicated and traded. It also details the various aspect of the loan agreement, including coupon, covenants, amortization, security and collateral, facility types, events of default and assignment terms.

REFERENCES

Bavaria, S. (2006). Leveraged loans white hot liquidity, white hot risk. *CreditWeek*, May 3.

Berman, D. (2007). Sketchy loans abound. *Wall Street Journal*, March 27.

Kerr, S. (2005). New report shows "silent" second lien lenders not so silent. *S&P CreditWeek*, August 23.

Miller, S., Donnelly, C., Polenberg, R., and Fuller, M. (2006). Doctor's bill. *The Deal*, October 15.

Miller, S., Donnelly, C., Polenberg, R., and Lauritsch, D. (2006). Second helping. *The Deal*, March 12.

Miller, S., Donnelly, C., Polenberg, R., and Lauritsch, D. (2006). Hot credits. *The Deal*, July 23.

Taylor, A. (2006). The LSTA 2005 trade data study. *LSTA Loan Market Chronicle*, April.

Tesher, D. (2001). Cash flow CDOs: Continued growth despite economic risks. *S&P CreditWeek*, August 7.

Tett, G. (2007). Cov-lite loans come to Europe. *Financial Times*, March 20.

Williams, M. (2005). The rise of second lien loans in the U.S. *FinancierWorldwide*, January.

(2007). Junk in a mirror. *Grant's Interest Rate Observer*, January 12.

CHAPTER 31

Emerging Markets Debt

MARIA MEDNIKOV LOUCKS, CFA
Senior Managing Director, Black River Asset Management

JOHN A. PENICOOK, CFA
Managing Director, UBS Global Asset Management

UWE SCHILLHORN, CFA
Executive Director, UBS Global Asset Management

Abstract: In the broadest sense, the group of emerging countries includes all nations not considered industrialized or already "developed." Since the latter group has only two dozen or so members, the emerging country universe encompasses most of the world's population and geography. However, because many emerging market countries have no investable debt securities, only a subset of these countries comprises the emerging markets debt universe. Hence, the more precise terminology is emerging market, rather than emerging country. While convention and market terminology lump all of these countries into one market, there are profound, fundamental differences among them. Many Latin American countries have a history of poor macroeconomic management and suffer from deep social inequality, but their recent economic performances have largely improved. Eastern Europe is recovering from decades of central planning, but some countries have prewar histories of success with capitalism. Opening these markets up to the rest of the world has the potential of producing large growth rates. Africa is generally income poor but commodity rich. Finally, a number of southeastern Asian countries have very high savings rates, resulting in exportation rather than importation of capital. As more emerging countries develop sovereign bond markets, inter-regional and intercountry differences will expand diversification opportunities, improving the risk/return profile of the asset class.

Keywords: emerging markets debt, Brady Plan, rolling interest guarantee, emerging market local debt, quasi-sovereign, credit default swaps, sovereign credit analysis

Emerging markets debt (EMD) warrants consideration in diversified portfolios based upon its normal return potential, risk characteristics and portfolio diversification benefits. Fundamental investment analysis of this market requires an understanding of sovereign credit risk and the compositional complexities of the emerging market bonds themselves. The normal return potential of the market, in conjunction with its low correlation to other bond and equity markets, offers the opportunity to improve a portfolio's risk/reward profile.

Emerging market issuers rely on international investors for capital. Emerging markets cannot finance their fiscal deficits domestically because domestic capital markets are poorly developed and local investors are unable or unwilling to lend to the government. Although emerging market issuers differ greatly in terms of credit risk, dependence on foreign capital is the most basic characteristic of the asset class. After the Asian crisis in 1997, investors realized that even investment-grade sovereign issuers can run into problems when access to foreign capital is constrained.

The growth of emerging market economies and the greater reliance of emerging markets on bond financing lead to an increase in importance of developing countries' debt securities in the international marketplace. The market capitalization of emerging markets sovereign debt indices (external and local currency debt) totaled $1,000 billion in March 2006, roughly 10% of the market capitalization of the developed markets sovereign bond index.

In this chapter, we describe the various EMD instruments and an overview of sovereign credit analysis.

EMERGING MARKET DEBT INSTRUMENTS

Sovereign EMD instruments can be divided into two broad segments: external debt, and local currency debt. The characteristics and histories of these market segments are examined in this section. This section also discusses emerging market corporate bonds, emerging market credit derivatives, and popular index alternatives.

Emerging Market External Debt

EM external debt is denominated in a developed market currency, typically U.S. dollars. Over the years, Euro-denominated issuance has grown. As of September 2005, 20% of the external debt bonds outstanding were denominated in euros. Despite this growth, euro-denominated bonds have lower issuance size and less liquidity than comparable U.S. dollar (USD)-denominated bonds. There is more institutional investment in USD-denominated bonds, while a larger portion of Euro-denominated bonds is held by retail accounts.

EM external debt trading occurs primarily in New York and London. Most issues settle via the ordinary Euroclear mechanism; costly local custody arrangements are unnecessary. USD-denominated, EM external debt has two important characteristics in common with U.S. corporate bonds: direct currency risk is not a major consideration and U.S. interest rate risk has an impact on EM external debt. However, while U.S. corporate bonds are subject to corporate default risk, EM external bonds are impacted by sovereign default risk.

Besides being denominated in U.S. dollars, EM external debt is governed under international law (either New York or U.K. law). There are emerging market bonds that are denominated in U.S. dollars, but governed under the local law of the issuing country. These bonds will be discussed later.

Although investor interest in EMD securities goes back to the 1800s, the acceptance of EMD into modern institutional portfolios began in the 1980s with the Brady Plan. In the 1980s, the series of bank loan defaults by many developing countries forced U.S. and some foreign commercial banks to write down the value of their loans. Narrowly defined, the *Brady Plan* refers to an innovative debt renegotiation format, whereby defaulted sovereign bank loans were written down and converted into bonds; the bonds themselves also have unique structures. Mexico was the first Brady Plan participant in 1989. (Former U.S. Treasury Secretary Nicholas Brady was credited with this approach.) More broadly, the Brady Plan encompasses the entire set of economic policy prescriptions that developing countries adopted in order to receive additional international aid. This aid allowed them to meet their responsibilities under the Brady Plan.

The Brady Plan differed from previous approaches in a number of respects. For the first time, underlying structural problems of the debtor countries were addressed (such as protected markets and controlled prices). Typically, the principal amount of the defaulted loans was effectively reduced by 35% to 50%; sometimes interest and interest arrears were also reduced. This principal forgiveness had the effect of both raising the loans' value in the secondary market and lowering the borrowers' debt burden. Further, the commercial banks' loans to private and sovereign entities were transformed into sovereign bonds, thus enhancing their appeal to investors.

The features of Brady bonds vary. Most were issued with a final maturity between 10 and 30 years and have semiannual coupons; many Brady bonds have amortizing principal payments. Coupons may be fixed, floating, step-up, or a hybrid combination. Unique features such as principal collateral, rolling interest guarantees, and value recovery rights were added to Brady bonds in order to improve creditworthiness and attract investors. Collateral for Brady bonds is invested in high-quality securities. Collateralized principal is invested in U.S. zero-coupon bonds; collateralized interest is invested in AA money market securities. Typically, the two next coupon payments are collateralized and the guarantee rolls when a sovereign makes the current coupon payment (hence the name "rolling interest guarantee").

Because of their interest and principal guarantees, market participants established new conventions for calculating Brady bond yields, spreads, and durations. The correct analytical procedure is to value the collateral by discounting the collateral cash flows at the appropriate spot interest rate and to subtract this collateral value from the bond's market price; the remainder is the price of the sovereign cash flows. Given the sovereign cash flows and their derived price, the bond yields and spreads can then be calculated. Statistics calculated by removing the collateral value and looking solely at the sovereign cash flows are referred to as "stripped" or "sovereign."

In addition to yield calculations, market participants have adapted traditional price sensitivity measures to the special features of Brady bonds. Interest rate duration estimates a bond's price responsiveness to changes in U.S. interest rates—all cash flows are revalued given changes in the U.S. yield curve. A bond is less sensitive to changes in U.S. interest rates if the bond's coupons are floating (that is, reset at a spread above Treasury yields). The investor is also concerned with isolating the bond's price response to a change in creditworthiness. Since only a portion of the bond's cash flows are exposed to sovereign credit risk (in some cases as little as 50%), a change in stripped spread will result in the repricing of only a subset of the cash flows (the sovereign cash flows). Thus, in addition to the standard interest rate duration measure, "spread duration" measures the bond's price responsiveness to movements in the stripped spread. If an overall widening of credit spreads is expected, the investor now has the tool to estimate which bonds will be more or less adversely affected.

Over time, the proportion of Brady bonds as a percentage of the sovereign debt universe has declined considerably. As of December 2006, Brady bonds accounted for only a small fraction of the USD-denominated sovereign bonds outstanding. The decline in importance of Brady bonds has occurred because new EM external debt issuance happens in the form of Eurobonds. Eurobonds are internationally issued securities denominated in hard currencies. Most Eurobonds have a fixed coupon and a bullet maturity.

In addition, many countries have done exchanges where they bought back Brady bonds and issued Eurobonds. Exchanging Brady bonds for Eurobonds is attractive if a country can issue Eurobonds for lower yields than existing Brady bonds (net present value savings). In some cases, countries have chosen to exchange Brady bonds for Eurobonds in order to receive cash-flow savings through lower coupons/amortizations or to release the Treasury collateral backing certain Bradys.

Eurobonds were serviced during the 1980s' bank loan crisis. A possible motivation behind such an admirable repayment history may have been that these obligations were small compared to bank debt, so that defaulting on them was much less economical. A second possible motivation may derive from the unique nature of bonds relative to loans. Debt restructuring negotiations of loans are easier because loans involve a small, easily identified, and relatively homogeneous group of creditors (that is, banks). It is difficult for a bank to not restructure its loan to a country and to free-ride on other banks' willingness to do so. By contrast, bondholders are a large and diverse group with no incentives to stay in good terms with the country. This makes broad approval of a bond restructuring more difficult.

Both Eurobonds and Brady bonds are held by a diverse group of creditors. There used to be a market perception that distressed sovereign issuers could try to selectively default on Bradys while they continue to service Eurobonds, thus maintaining some type of reputation in the market.

As it turned out, the growth of the Eurobond market eliminated the potential of preferential treatment. Since Eurobonds are now a larger portion of a sovereign's total debt, restructurings in Ecuador (2000) and Argentina (2005) included Eurobonds in order to meaningfully decrease debt burden. In the case of Ecuador, restructuring took approximately a year and only a few investors did not participate.

The Argentine default was considerably larger and the government of Argentina took a much harder negotiating stance with creditors; the government's restructuring proposal came three years after the default. A sizeable part of Argentine bondholders did not participate in the restructuring of Argentine debt in the first half of 2005. Legal battles between holdout creditors and the government of Argentina are likely to continue for many years to come. The Argentine restructuring illustrates the uncertainties of sovereign restructuring because there is no legal framework for default resolution (in contrast to the application of bankruptcy law to corporates). It also illustrates that lengthy legal battles and temporary lack of market access will not prevent an insolvent country from defaulting.

In the aftermath of the Argentine debt default, Uruguay and the Dominican Republic were able to effectively reschedule their external bonds with broad creditor consent. Creditors that did not consent had their obligations honored fully. Every sovereign default is different; it is difficult to predict how future distressed sovereigns will act.

Emerging Market Local Debt

Many developing countries have functioning and relatively liquid domestic debt markets. Local issues are issued under local law. The bulk of local issues are denominated in local currencies, but a large portion is denominated in major currencies (U.S. dollar, euro, yen) or linked to a major currency. Besides evaluating direct currency risk, international investors need to be compensated for the lack of protection offered by local laws, potential settlement difficulties and taxation issues.

The development of EM local debt markets has been encouraged by the growth of local pension fund industries and increased foreign interest. Thanks to a young population and changes in pension fund regulations, the public and private pension funds in developing countries have grown tremendously creating a stable holder of government debt.

Foreign interest in EM local debt has grown tremendously since the 1990s. After a series of devastating crises (Mexico devaluation in 1994, Asia devaluations in 1997, Russia default in 1998, Brazil devaluation in 1999, and Argentina default in 2001), many EM countries abandoned fixed exchange rates and built up their foreign exchange reserves. Many EM countries benefited from the commodity boom and greatly improved their external and fiscal ratios. This improvement led to improved credit ratings and increased investor interest and lower yields on external debt. International investors turned to local currency debt for the higher yields and because floating exchange rates and more stable economics made the massive devaluations of the past less likely.

Historically, EM local debt had very short maturity, so changes in investor sentiment quickly escalated into liquidity crises as the country had to raise interest rates rapidly to encourage rollover. In addition, a large portion of the debt was dollar-denominated or dollar-linked (similar in currency risk to external debt), so country's debt grew quickly during a currency devaluation.

Emerging markets viewed increased demand for local debt from local pension funds and foreign investors as an opportunity to improve the composition of their debt by increasing maturities and converting USD-linked debt to local currency debt. Mexico exemplifies a country at the forefront of improving its debt structure. In 2006, Mexico issued a fixed-rate 30-year peso-denominated bond. The Mexican government also issued warrants that allow holders to exchange USD-denominated debt for peso-denominated debt. Mexico's debt structure improved tremendously from 1994 when the country faced a massive liquidity crisis due in part to the rollover of short-term dollar-denominated bonds.

Since the universe of EM local debt is very large and complicated (most local bond markets have their own peculiarities), this section offers only a basic review of different types of EM local debt including USD-denominated local debt, fixed-rate bonds, floating-rate bonds, and inflation-linked bonds.

USD-denominated and USD-linked local law bonds are still a significant portion of the EM local debt universe. Historically, these instruments were issued because locals did not want to take currency risk for longer maturity bonds. However, the biggest issuer of USD-denominated local law bonds as of the end of 2006 was Argentina. Argentina issues only local law instruments because they are currently being sued in international courts following a default on external debt in 2001; any proceeds from an international law bond issue by Argentina face the risk of attachment. Since Argentine USD-denominated local law bonds are settled internationally and traded out of New York, many holders do not see a practical difference between these bonds and external debt. The difference between local law and international law bonds is limited in normal times, but comes to the forefront if there is a default.

In the 1990s, inflation declined substantially throughout EMs. But due to a long history of high inflation and substantial currency depreciations, many countries are still not able to issue long-dated fixed-rate bonds. A large part of public debt is therefore short term or floating rate.

To significantly extend the maturity structure of their domestic debt, some countries are required to issue inflation-linked bonds. In 2005, Brazil issued an inflation linked bond maturing in 2045. Inflation often increases after strong currency depreciations. Therefore, inflation-linked bonds can be expected to provide some implicit compensation for currency devaluations in the long run.

Emerging Market Corporate Debt

The risk analysis of an EMD corporation hinges on its ownership type and its sensitivity to domestic economy. EMD corporations may be owned by the sovereign or an established multinational or have local ownership. Sovereign-owned corporations are often referred to as *quasi-sovereign*. A corporation may sell its product domestically (e.g., a cable operator) or it may earn hard currency by exporting its product (e.g., an oil company).

Rating agencies have historically limited a corporation's debt rating to its country's sovereign credit rating because corporate debt manifests specific corporate business risk in addition to the sovereign risk of its government. In effect, a sovereign ceiling limited a corporation's credit ratings. The theory behind the sovereign ceiling is that the sovereign entity ultimately controls the corporation's access to foreign currency and its tax burden. Essentially, the corporation depends on a benevolent legal and institutional framework from the sovereign government and, therefore, is never a better credit risk than the sovereign itself.

Standard and Poor's (S&P) and Moody's have weakened the sovereign ceiling by allowing certain corporates to receive ratings above their sovereign ceiling. S&P sometimes rates corporations above their respective sovereign ceiling if they operate in highly dollarized economies, if they are geographically diversified or if they have offshore parent support or structural enhancements. Moody's allows corporations to be rated above the sovereign ceiling if there are external support mechanisms (that is, support from a multinational parent), if there is a low chance of a moratorium in the event of a sovereign default, and if the borrower has access to foreign exchange.

In the debt crises of the 1980s, Latin countries imposed a blanket debt moratorium on all foreign currency borrowers, many of which were corporations and banks. In recent sovereign defaults (Ecuador, Pakistan, Russia, and Ukraine), corporate access to foreign currency was not restricted by the government, but there were few corporate foreign currency borrowers. When Argentina defaulted in 2001, the government did not impose a blanket debt moratorium. However, most of the Argentine corporates defaulted on their external debt obligations.

With the exception of a few foreign-owned exporters, there remain strong arguments in support of the sovereign ceiling. Some argue that a particular international corporation, like a government-owned oil company, may be so vital to the country's access to foreign currency, that the corporation's credit reputation may supersede the country's ability to access international capital markets. However, while a nationally vital corporation may receive government assistance, it does not follow that its bondholders in general will prosper. Thus, one would expect most corporate issues to offer higher yields than their sovereign counterparts.

The credit improvement of EM sovereigns has led to less sovereign issuance (countries have smaller fiscal gaps to fill) and lower spreads for sovereign external debt. In their search for yield, investors are not only looking at USD-denominated international law EM corporate bonds, but also local currency–denominated EM corporate bonds.

Emerging Market Credit Derivatives

Derivatives instruments that combine or eliminate different risks of EM securities represent a large and fascinating universe. Instruments can be created that combine the default risk of one country with the currency risk of another. The most basic and commonly used emerging markets debt derivatives are *credit default swaps* (CDSs).

The terminology used in trading CDS contracts is similar to the terminology used when someone buys an insurance contract. Each CDS contract specifies an issuer and a specific length of time (the term of the protection). One party "buys protection" against a credit event (e.g., a default); another party receives periodic payments for "selling protection." The seller of protection is long the default risk of a specific issuer for a specific length of time, similar to the holder of a bond. Conversely, the buyer of protection is short the default risk of a specific issuer, similar to someone who sells a bond short.

Sovereign CDS is quoted as a spread over the London Interbank Offered Rate (LIBOR) for a specific tenor (typically out to 10 years) creating a CDS curve for each sovereign. If an investor sells $10 million of protection for Brazil for 5 years at 200 basis points, the investor will be paid $10 million × 0.02 = $200,000 a year for 5 years. Like the holder of a Brazilian bond, the seller of Brazilian protection is long Brazilian default risk. However, the seller of Brazilian protection only has to post a small initial margin to establish a notional exposure of $10 million.

If Brazil defaults before the CDS contract expires, the buyer of Brazilian protection will receive the "insurance payment." The buyer of protection delivers $10 million notional of Brazilian securities to the seller of protection in exchange for $10 million. If Brazilian securities trade at a 70% discount post-default, $10 million notional of Brazilian securities will be worth $3 million at market prices, so the buyer of protection receives a $7 million insurance payment if Brazil defaults ($10 million payment from the protection seller in exchange for bonds valued at $3 million).

It is easy to replicate a sovereign bond by combining a sale of protection with an equal notional amount of cash. If an investor sells $10 million of Brazilian 5-year protection at 2% and holds $10 million in high-quality short-term securities earning LIBOR, the investor will create exposure that mimics a 5-year Brazilian floating-rate bond with a coupon of LIBOR+2%. An investor that creates a synthetic bond by selling CDS can calculate a price for the specific synthetic security by valuing the securities cash flows using the current CDS curve. As with a regular bond, the synthetic security's price will decline (increase) as spreads increase (decline). Fortunately, the CDSW screen on Bloomberg has become the market standard for valuing CDS.

The specific legal terminology used in CDS contracts are established by the International Swaps and Derivatives Association (ISDA). Before entering into CDS transactions, market counterparties typically sign an ISDA master agreement stating that they agree to the basic definitions. Settlement confirms outline specifics of individual trades. As with all legal documents, the devil is in the details. The ISDA master documents and the settlement confirms should clearly state what constitutes a credit event, what type of instruments are deliverable onto a CDS contract, and settlement logistics if a credit event occurs.

Emerging Market Bond Indices

The J. P. Morgan Emerging Markets Bond Indices are the most commonly used by international investors. External debt investors can use one of the indices in the Emerging Markets Bond Index (EMBI) family, while investors into local markets can use the Emerging Local Markets Index (ELMI) or one of the indices in the Emerging Markets Global Bond Index (EM GBI) family. Increasingly, EMD mandates will want both external and local markets exposure with investors combining an external debt index and a local market index as a benchmark.

Bonds in the EMBI family contain either U.S. dollar- or euro-denominated sovereign bonds that were issued internationally (under either New York or U.K. law). Depending on the actual index, the EMBI indices use either ratings, country income per capita, or a debt restructuring criterion as inclusion criteria. Popular Emerging Markets Equity indices also use income per capita to classify countries for inclusion. The EMBI indices include nonrated and defaulted issuers.

Two widely used EMBI indices are the EMBI Global or the EMBI+. Once a country meets the criteria to be included in the EMBI Global or the EMBI+, a particular bond issue must meet certain liquidity requirements. The liquidity requirements used by EMD indices are stringent in comparison to those used by other bond indices. In order to be included in the EMBI Global, a bond must have at least $500 million face amount outstanding, at least 2.5 years to maturity, verifiable prices, and verifiable cash flows. In comparison, the minimum face outstanding for the Lehman Investment Grade Corporate Index and the Merrill Lynch High-Yield Index was $150 million and $100 million, respectively. The liquid nature of the JPM EM indices facilitates the trading of index swaps and allows investors to quickly implement top-down strategy changes.

EM external debt indices have poor issuer diversification when compared to U.S. High-Yield and U.S. Investment-Grade Credit indices. While most U.S. High-Yield and U.S. Investment-Grade Credit indices have hundreds of issuers, the EMBI Global and the EMBI+ have only around 40 issuers. In addition, EM external debt indices are heavily weighted toward several large countries (Brazil, Mexico, and Russia) and to the Latin region in general. Investors have an option of using a more equal-weighted version of these indices (referred to as the EMBI Global Diversified and the EMBI+ Diversified). Given the small universe of issuers, there is no way to completely remove concerns about diversification. However, since an EMD portfolio is usually a small piece of an institutional investor's portfolio, issuer diversification should be less of a concern.

Indices in the EM GBI family have only around 18 sovereign issuers. They are quite concentrated in some eastern European countries, Malaysia, Mexico, and South Africa. They include fixed-rate bonds and zero-coupon bonds. There is no other liquidity criterion than the availability of daily prices.

SOVEREIGN CREDIT ANALYSIS

A country's bond spreads (spread over U.S. Treasuries for sovereign external debt) are related to its willingness and capacity to repay its debt. The latter depends directly on the amount of obligations coming due at a point in time and the foreign exchange resources and refinancing opportunities available at that time. Both economic and political factors should be considered when analyzing the resources available for a sovereign.

Economic Considerations

Many economic measures are relevant to assessing the credit risk of a developing country. One manner of organizing economic and financial considerations is to compartmentalize measures into three categories: structural, solvency, and serviceability. In addition to making the analysis more manageable by removing redundancies, this categorization produces a term structure of credit risk, akin to the well-known notion of the term structure of interest rates. An understanding of individual country politics, as well as the role of various international agencies, is also an essential part of sovereign credit analysis.

Structural

Measures belonging to this category describe the long-term fundamental health of the country. They include economic variables such as reliance on a particular commodity for export earnings, welfare indicators such as per capita gross disposable product (GNP), and social/economic measures such as income distribution. These variables generally are not directly linked to default, but countries with poor structural fundamentals are likely to develop economic problems. Further, given two countries which are similar in other respects, the one with the inferior structural measures will have a lower capacity to tolerate adverse economic shocks.

Solvency

In contrast to the structural variables, the solvency class contains intermediate-term measures of a country's economic health. In particular, these variables should reflect the country's ability, over time, to meet its central government debt obligations. Both local and external debt ratios are included in this category. Countries with inferior solvency measures, all else being equal, have higher default risk because international debt service competes with local economic constituencies for resources.

Serviceability

The factors in this category are of short-term, if not immediate, concern. They reflect the country's foreign exchange reserve position relative to its obligations (and are therefore usually presented in ratio form). Some examples include debt service (percent of exports) and short-term debt (percent of reserves). Despite good or improving fundamentals and strong solvency measures, a developing country may be forced into a crisis if its reserves are (or will become) deficient, or if alternative reserve sources, such as the International Monetary Fund (IMF), are circumscribed. Experience suggests that serviceability, or liquidity, is a paramount concern (Mahoney 1999).

Political Considerations

Peculiar to analyzing developing country investments are certain critical political issues such as international aid and policy instability. The United States and multilateral agencies such as the World Bank and IMF have invested a great deal of political and financial capital in the recovery of developing countries and their return to the global marketplace. Therefore, an event that would ordinarily raise the likelihood of default may actually induce international organizations to assist the emerging country and reduce the probability of default. Alternatively, the movement to representative government and open markets is a recent phenomenon, and in many developing countries, there are few institutions in place to serve as anchors to these policies. The resignation or death of one key policy maker may be enough to alter economic policy. In sum, political factors can cut both ways: the politics of individual countries are often fragile, but international politics often acted as counterbalances in the past.

Despite notable attempts, multilateral agencies have been unable to influence a sovereign's relationship with its bondholders. Multilateral agencies attempted to establish a process for sovereign bankruptcy, but there was no political support to corral the interests of various bondholder groups. Despite concerns that litigious bondholders would prevent a sovereign from recovering from default, the Argentina's take-it-or-leave-it approach to bondholders in its 2005 restructuring illustrates that bondholders have limited rights against a sovereign.

While nascent representative governments may suffer from institutional instability, it is important to recognize that these countries have undergone profound political change in a short time. Most countries have moved from military rule to competitive, multiparty democracies within the decade. For example, in 1982, approximately 80% of the emerging market countries' populations were under communist or military rule; now approximately 97% are governed by democratic rule.

Willingness to Pay

Some argue that sovereign risk analysis is doomed to failure because, notwithstanding the ability to pay, a country may be unwilling to make good on its debt obligations. Distinguishing sovereign risk from corporate or municipal credit risk on this basis alone exposes a deficient understanding of default risk. Borrowers default when their competing economic interests override the damage done by default. Default is never a casual decision. Corporations and municipalities are faced with the same decision as sovereign borrowers: at what point are you willing to capitulate and damage your reputation?

Sovereign Credit Perspective

The major risk in emerging economies is often not the government's debt load on the economy, but access to foreign exchange. Because of previous poor policy management, weak banking systems, and ineffective leadership, many emerging countries are forced to borrow in foreign currency (usually U.S. dollars). Developing countries access foreign currency through foreign direct investment, exports, portfolio investment and official loans, all of which depend upon sound economic management and stable political leadership. This access to dollars, which is a serviceability issue, can largely be a matter of investor confidence in policy makers and is a unique risk to this market. Total external debt (public and private foreign currency denominated debt) relative to GDP for emerging countries is not significantly different from that of developed countries, but, in some cases, these developing countries have difficulty accessing foreign currency through exports or through foreign direct investment.

This additional risk aside, three macro trends may lead investors to be optimistic that emerging countries will continue their economic development process and eventually become better credit risks. First, the retreat of communism and the Soviet state signal an end to dismal economic incentives for much of the world. Second, the movement to more democratic forms of government should, in the long run, stimulate a more competitive marketplace of ideas and policies. Finally, the high rate of integration (trade, tourism, information technology, and so on) and the rapid pace of technological change make economic isolation more costly and less acceptable to the populace.

The current economic position of emerging countries is in some ways not radically different from their developed counterparts. What differentiates them is that EM borrowers have less institutional stability, less demonstrated commitment to free market principles and less reliable access to foreign exchange. These problems lead primarily to a weaker serviceability measure, but do not necessarily imply structural infirmity or insolvency.

SUMMARY

Emerging market debt has come a long way since defaulted bank loans were restructured into Brady bonds in the early 1990s. Most countries have moved to flexible exchange rates and built up their foreign exchange reserves. In addition, the boom in commodity prices has led to considerable improvement in economic statistics and upgrades by rating agencies. Many emerging markets have improved their institutional stability by implementing laws to keep central banks independent and lower fiscal deficits. Even in cases of considerable political change (Mexico, 2000, and Brazil, 2002, being two examples), emerging market institutions proved robust and helped maintain investor confidence. With the improvement in credit quality, new investors have opted for exposure to emerging markets debt. The increase of longer-term buy-and-hold investors has broadened the investor base and decreased volatility.

The asset class expanded to include bonds issued under local law (denominated in major currencies and emerging market local currencies) as well as external debt governed by international laws. Emerging market debt mandates increasingly want external and local markets exposure and combine an external debt index and a local market index to create their benchmark. Emerging market debt investors are also increasing their exposure to emerging market corporate debt and actively use credit default swaps to gain or hedge market exposure.

The development of the emerging market asset class has not been smooth, with major periods of volatility including the Mexican devaluation (1994), Russian devaluation/default (1998), and Argentine default (2001). Investors need to be familiar with sovereign credit analysis including both economic and political considerations. Economic risks include long-term structural problems, medium-term measures of indebtedness, and short-term measures of a country's reserves relative to its obligations. Political risks include the stability of local institutions and the potential for international agencies to help in times of crisis.

REFERENCES

Chambers, J. (2000). The rise and fall of sovereign ratings: 2000. *Standard & Poor's CreditWeek* (December): 27.

Conybeare, J. A. C. (1990). On the repudiation of sovereign debt: Sources of stability and risk. *Columbia Journal of World Business* (Spring/Summer).

Dym, S. (1994). Identifying and measuring the risks of developing country bonds. *Journal of Portfolio Management* (Winter): 61–66.

Dym, S. (1992). Global and local components of foreign bond risk. *Financial Analysts Journal* (March/April): 83–91.

Eichengreen, B., and Portes, R. (1988). Setting defaults in the era of bond finance. *Centre for Economic Policy Research*, Discussion Paper No. 272 (September).

Grief, A. (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of Political Economy* 102 (October): 912–950.

Hausmann, R., and Gavin, M. (1999). Preventing crisis and contagion: fiscal and financial dimensions. *Inter-American Development Bank and Warburg Dillon Read* (March).

Lee, S. Y. and Venezia, M. E. (2000). A primer on Brady bonds. *Emerging Markets Fixed Income*. Salomon Brothers (March).

Mahoney, C. (1999). What have we learned? Explaining the world financial crisis. *Moody's Investor Service* (March).

Mathieson, D. J., and Schinasi, G. J. (2001). *International Capital Markets Developments, Prospects, and Key Policy Issues*. Washington, DC: International Monetary Fund (July).

Policy Development and Review Department (1999). *Involving the Private Sector in Forestalling and Resolving Financial Crises*. Washington, DC: International Monetary Fund (April).

Purcell, J. F. H., and Kaufman, J. A. (1993). The risks of sovereign lending: Lessons from history. *Emerging Market Research*. Salomon Brothers (September).

Tonge, D. (2001). The risk-return properties of euro-denominated emerging market debt. *Emerging Market Research*. Salomon Brothers (May).

Truglia, V. (1995). Sovereign risk: Bank deposits vs. bonds. *Moody's Investor Service* (October).

Winkler, M. (1933). *Foreign Bonds: An Autopsy*. Philadelphia: Roland Swain Co.

CHAPTER 32

Introduction to Mortgage-Backed Securities

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

ANAND K. BHATTACHARYA, PhD
Managing Director, Countrywide Securities Corporation

WILLIAM S. BERLINER
Executive Vice President, Countrywide Securities Corporation

Abstract: Mortgage-backed securities (MBSs) are constructed by aggregating large numbers of similar mortgage loans in mortgage pools. There are two mechanisms for securitizing MBS pools; they can be issued by governmental agencies and quasi-agencies (that is, the government-sponsored enterprises, or GSEs) as agency pools, or structured with separate credit enhancement as "private-label" securities. MBS trading conventions reflect the nature of mortgage lending. Loans begin as an application that is "locked" at some point prior to the loan's closing date, while it is being underwritten and processed. Trading in many securities is done on a forward basis, where trading is executed to settle at some date in the future. This trading convention also implicitly creates a financing vehicle, where securities can be financed in an efficient and inexpensive fashion. CMOs are created by carving up MBS principal and interest cash flows in order to target the needs of specific investor clienteles. As MBS pools are closed universes of cash flows, structuring involves transferring prepayment and (for private-label structures) credit risk within the deal, with the goal of maximizing the deal's proceeds while better meeting the objectives and preferences of various investor constituencies.

Keywords: mortgage-backed securities (MBSs), private-label securities, agency securities, senior pass-throughs, credit enhancements, guaranty fee, structured securities, tranching, collateralized mortgage obligations (CMOs), mortgage strips, base servicing, excess servicing, government sponsored enterprises, agencies, Fannie Mae, Freddie Mac, Ginnie Mae, weighted average coupon (WAC), subordination, shifting-interest structures, overcollateralization, forward market, dollar roll.

This chapter provides an overview of *mortgage-backed securities* (*MBSs*) and the markets in which they trade. It discusses the mechanics of issuing different forms of MBS, along with many of the market practices, conventions, and terms associated with the MBS markets.

CREATING DIFFERENT TYPES OF MBS

The fundamental unit in the MBS market is the pool. At its lowest common denominator, mortgage-backed pools are aggregations of large numbers of mortgage loans with similar (but not identical) characteristics. Loans with a commonality of attributes such as note rate (that is, the interest rate paid by the borrower on the loan), term to maturity, credit quality, loan balance, and product type are combined using a variety of legal mechanisms to create relatively fungible investment vehicles. With the creation of MBS, mortgage loans were transformed from a heterogeneous group of disparate assets into sizeable and homogenous securities that trade in a liquid market.

The transformation of groups of mortgage loans with common attributes into tradable and liquid MBS occurs using one of two mechanisms. Loans that meet the guidelines of the agencies (that is, Fannie Mae, Freddie Mac, and Ginnie Mae) in terms of credit quality, underwriting standards, and balance are assigned an insurance premium, called a *guaranty fee*, by the agency in question and securitized as an agency pool. Loans that either do not qualify for agency treatment, or for which agency pooling execution is not efficient, are securitized in nonagency or "private label" transactions. These types of securities do not have an agency guaranty, and must therefore be issued under the registration entity or "shelf" of the issuer. As noted later in this chapter, the insurance (or "credit enhancement") for the loans is in the form of either a private guaranty or, more commonly, structured in the deal through so-called "subordinate" classes.

The senior portions of these deals are very similar in profile to agency pools, and are often referred to as *private-label* or *senior pass-throughs*. The term "pass-through" indicates that principal and interest is passed on to the investor pro rata with their holding. Using this definition, the senior portion of a private label deal is technically not a pass-through, because principal is redistributed within the structure; however, the term is nonetheless utilized to describe the senior cash flows before they are restructured.

Once a pool (in either agency or private-label form) is created, it can be sold to investors in the form of a pass-through, in which principal and interest is paid to investors based on their pro rata share of the pool. However, the cash flows of pools can also be carved up to meet the requirements of different types of investors. The creation of so-called "structured securities" involves dividing (or "tranching") the underlying pools' cash flows into securities that have varying average lives and durations, different degrees of prepayment protection or exposure, and (in the case of private label deals) different degrees of credit risk. These types of securities are broadly referred to as *collateralized mortgage obligations* (*CMOs*). The flexibility inherent in tranching, along with the broad range of loan instruments, allows the MBS market to reflect a large degree of market segmentation. In turn, this allows a wide range of investor types with different investment objectives and risk tolerances to invest in the MBS market, supplying the funds that ultimately are recycled into new mortgage lending.

It will be helpful at this juncture to briefly discuss and contrast the processes of creating and structuring agency CMOs and private label transactions. To create an agency CMO deal, the underwriter buys agency MBS pools in the primary or secondary markets and places them in a trust-like entity. The different tranches are then created from the principal and interest cash flows generated by the MBS pools (or "collateral"). In contrast, private label transactions are created by placing large numbers of loans directly in a securitization vehicle, from which the structured transaction is subsequently created by the issuer. (This accounts for why these transactions are sometimes referred to as "whole loan CMOs." CMOs are also referred to as real estate mortgage investment conduits, or REMICs. The terminology refers to a provision in the Tax Reform Act of 1986 in order to remedy some double taxation inefficiencies inherent in earlier collateralized structures. While the term "REMIC" is essentially a tax election, often the terms "CMO" and "REMIC" are used interchangeably.) While the agency transaction is an arbitrage of sorts, the private-label securitization serves as the process by which loans are directly distributed into the capital markets.

A different subset of the MBS sector is the market for *mortgage strips*, or more precisely the market for principal-only and interest-only securities. Since mortgages are comprised of both principal and interest, the two components can be separated and sold independently. The holder of the principal-only security (or PO) receives only principal (scheduled and unscheduled) paid on the underlying loans. The holder of the interest-only security (or IO) receives the interest generated by the underlying loans. Although IOs are quoted with a principal balance, this balance is notional in nature; it is used only as a point of reference for settling the transaction and calculating monthly interest cash flows generated by the security. The most common mortgage strips are created simply by putting agency pools into a trust and splitting principal and interest cash flows into IOs and POs. (Note that IOs in this context should not be confused with interest-only loans; the two concepts are totally different, even though they do share some of the same nomenclature.)

The market for mortgage developed to allow MBS investors a means of trading directly on prepayment speeds and expectations. POs typically have long positive durations and rise in value when rates decline, while IOs have negative durations, behaving in a fashion similar to bond puts when rates rise. However, the critical driver of performance strips is prepayment expectations. POs perform well if prepayment speeds are fast, in the same way that returns would be enhanced if a zero-coupon bond were called prior to maturity at par. By contrast, IOs perform well if prepayment speeds are slow; they can be viewed as an annuity where the value increases the longer it remains outstanding.

While IOs and POs are most commonly created in trust form, both types of bonds can also be created as part of a CMO deal. Structured IOs and POs have a similar appeal to investors as strips, and are evaluated in a similar fashion. They are created as part of the process of structuring certain bonds with a targeted coupon or dollar price. If an investor seeks a bond with a lower dollar price, for example, the coupon on the bond must be reduced; this can be accomplished by stripping some coupon off the tranche in question and selling it as a structured IO.

Agency MBS Creation

While both agency adjustable rate pools and private label securities have existed for many years, the agency fixed rate market remains the most widely quoted and liquid benchmark in the MBS market. Therefore, a discussion of pooling practices and the securitization process logically begins with the formation of fixed rate agency pools. In this section, we will first address the basics of agency fixed rate pools, which dictate to a large extent how such pools are created. Subsequently, we will discuss the creation of adjustable rate mortgage (ARM) pools, which have many similarities to fixed rate products but are pooled quite differently.

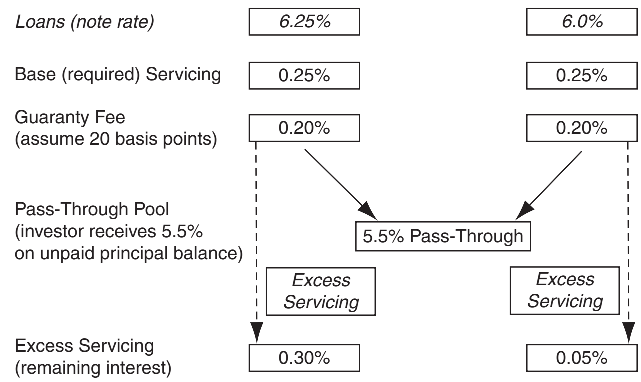


Figure 32.1 Cash Flow Allocation for a 5.5% Agency Pass-Through Pool for Loans with Different Note Rates

Fixed Rate Agency Pooling

Agency fixed rate MBS are traded according to their coupons, which are normally securitized in 50 basis point increments. There are liquid markets in both even coupons and half-coupons (e.g., 6.0% and 6.5%), although quarter- and eighth-coupon pools are sometimes originated. Loans, by contrast, are normally originated in increments of 12.5 basis points (or one-eighth of a percent). As part of the transformation process, certain cash flows from the loan interest stream are allocated for servicing and credit support payments. These apportionments are as follows:

- Guaranty fees (or "g-fees") are, as described earlier, fees paid to the agencies to insure the loan. Since these fees essentially represent the price of credit risk insurance, g-fees vary across loan types. In the conventional universe, g-fees vary depending on the perceived riskiness of the individual loans (based on credit metrics such as credit score, loan-to-value (LTV) ratio, and documentation). However, high-volume lenders may be able to negotiate generally lower guaranty fees. For Ginnie Mae pools, the guaranty fee is almost always six basis points, reflecting the loan-level guarantees provided by the Federal Housing Administration and Veterans Administration.
- Required servicing or *base servicing* refers to a portion of the loan's note rate that must be held by the servicer of the loan. This entity collects payments from mortgagors, makes tax and insurance payments for the borrowers, and remits payments to investors. The amount of base servicing required differs depending on the agency and program. At this writing, base servicing is 25 basis points in the fixed rate pass-through market.
- Excess servicing is the amount of the loan's note rate in excess of the desired coupon remaining after the g-fees and base servicing are subtracted.

Both base and excess servicing (sometimes described as mortgage servicing rights or MSRs) can be capitalized and held by the servicer after the loan is funded. However, secondary markets exist for trading servicing, either in the form of raw mortgage servicing rights or interest-only securities created from excess servicing.

A simple schematic showing how two loans with different fixed note rates can be securitized into a 5.5% agency pass-through pool is shown in Figure 32.1. For both loans, the amount of base servicing and guaranty fee is the same, with the difference being the amount of excess servicing created by the issuer. This diagram ignores some of the complexities of pooling, however, which will be addressed later in this section.

General pooling practices in the fixed rate market mandate that the note rate of the loans must be greater than the pool's coupon. However, loans with a wide range of note rates can be securitized in pools. For example, *Freddie Mac* and *Fannie Mae* allow the note rate to be as much as 250 basis points higher than the coupon rate. *Ginnie Mae* pooling rules depend on the program used to securitize the loans. The Ginnie I program, where the majority of loans are pooled, requires that the note rate be 50 basis points over the coupon rate. The multi-issuer Ginnie II program allows the note rate to be up to 150 basis points higher than the coupon.

Pooling economics normally dictate that the note rate of the loan is between 25 and 75 basis points higher than the coupon rate since retaining large amounts of excess servicing is generally uneconomical. In addition, guaranty fees can be capitalized, or "bought down," and paid as an up-front fee to the GSE at the loan's funding. This typically occurs when the lender wishes to create pools with relatively high coupon rates (e.g., pool a 6.25% loan into a 6.0% pool) based on market conditions, a practice known as "pooling up." (Naturally enough, pooling this loan into a 5.5% pool would be called "pooling down.") Because of the base servicing requirement, however, at this writing the spread between a loan's note rate and pooling coupon cannot be less than 25 basis points.

Once large numbers of loans are funded, lenders will group loans with the same coupon in order to form pools. To create a pool, the lender effectively transfers loans earmarked for a particular coupon to the agency and receives the same face value of MBS in exchange. The MBS received may consist of a pool collateralized by only its loans, or it may be part of a multi-issuer pool. After receiving the security, the lender can either sell the pool into the secondary market or (in the case of a depository) hold it in its investment portfolio.

The GSEs also buy loans for cash proceeds through what is called, appropriately enough, the cash window. This is often used for loan programs with unusual specifications such as certain documentation styles or loan-to-value ratios, as well as by smaller lenders that engage in piecemeal sales. Loans purchased through the cash window can either be securitized in multi-issuer pools or retained in the GSEs' portfolios.

Adjustable Rate Agency Pooling

As noted earlier in this section, pooling practices in the agency ARM market are currently somewhat different. As in the fixed rate market, a standard amount of base servicing is held on each loan, and guaranty fees are assigned and paid on the loans based on each loan's perceived riskiness. (Base servicing in the ARM market has historically been 37.5 basis points, but at this writing some lenders have begun to hold only 12.5 basis points of base servicing.) The lender's current production, with loans having a range of note rates, is then pooled, with the pool's coupon being an average of the net note rates in the pool weighted by the loans' balances. This is referred to as having a *weighted average coupon* or *WAC*. Using this methodology means that:

- No excess servicing is held in order to decrease the net note rates of the loans to a targeted level.
- G-fees are generally not bought down, since buydown pricing is not efficient in the ARM sector.
- Pools will contain loans with note rates below the coupon rate.

There are important implications of this different pooling methodology. ARM pools typically are originated with uneven coupons taken to three decimal places (e.g., a pool might have an initial coupon of 5.092%). In addition, coupon rates on ARM pools (and, in fact, any security with a WAC coupon) change slightly over time, as individual loans in the pool are paid off. The result of these factors is that agency ARMs trade on a pool-specific basis, rather than by specific coupons as in the fixed rate universe. (There have been a number of initiatives designed to create ARM securities that can trade in forward markets more like those of the fixed rate universe, although none have yet been adopted in a broad fashion.)

Private-Label Securitization

While the creation of private-label deals is conceptually similar to agency pooling practices, the lack of involvement by the agencies necessitates significant differences. Since there is no guaranty fee, alternative forms of credit enhancement must be utilized as noted previously. Private credit enhancement is most commonly created in the form of subordination, which means that a portion of the deal is subordinate or "junior" in priority of cash flows, and is the first to absorb nonrecoverable losses in order to protect the remaining (or "senior") tranches. A common technique is to divide the subordinated part of the deal into different tranches, each with different ratings (which typically range from double-A to unrated first-loss pieces) and degrees of exposure to credit losses. (For example, the nonrated "first loss" tranche is the first to absorb losses; if this tranche is exhausted, the losses are then allocated to the tranche second-lowest in initial priority). Subordinate tranches trade at significantly higher yields than the seniors to compensate investors for the incremental riskiness and greater likelihood of credit-related losses.

Figure 32.2 shows an example of a senior/subordinate deal structured in this fashion. The amount of subordination required for a deal and the relative sizes of the different subordinate tranches (often referred to as the "splits") are dictated by the rating agencies, based upon their assessment of the likelihood of losses for the subject collateral. Prior to being structured, the senior portion of the deal in the example has cash flows that are very similar (but not identical) to agency pools, as noted previously. These private label pass-throughs are sometimes sold directly in unstructured form, although it is more common to see them restructured into tranches.

Deals with *subordination* (also called senior/sub deals) typically have an additional feature designed to insure the adequacy of credit enhancement levels. All unscheduled principal payments (that is, prepayments) are initially directed to the senior tranches, and the subordinates are locked out from prepayments (although they do receive scheduled principal payments, or amortizations). This feature causes the subordination (as a percentage of the deal) to grow over time, and increases the degree of protection for the senior sector. The subordinates eventually begin to receive some unscheduled principal payments (although the actual schedule depends on the type of collateral), and ultimately receive prepayments pro rata with their size. The technique is referred to as "shifting interest," and deals with this type of subordination are commonly called *shifting-interest structures*.

Other variations of the senior/subordinate structure are used in the MBS markets, especially for subprime and other loans that have a greater degree of default risk. Some deals are structured such that there is more loan collateral than bonds in a deal, lending additional credit support to the senior bonds (in addition to some subordinate classes). This structuring technique is referred to as *overcollateralization*, and deals structured in this fashion are referred to as OC structures.

As with agency ARM pools, private label deals typically securitize a wide range of note rates, due in part to the desire of issuers to capitalize on economies of scale by issuing large deals. However, the market for fixed rate securities is generally not receptive to WAC coupons. In order to create a fixed coupon rate, the loan collateral must be modified before the credit enhancement is structured. This technique is somewhat different from that utilized in creating agency pools. Both the range of note rates included in a deal, as well as the desire to include loans with note rates below the deal's coupon (once base servicing and fees are taken into account), necessitate the creation of WAC IOs and POs, securities unique to the private label market.

The decision with respect to which coupon is to be produced is a function of market conditions, including investor's interest rate and prepayment outlook. Once the

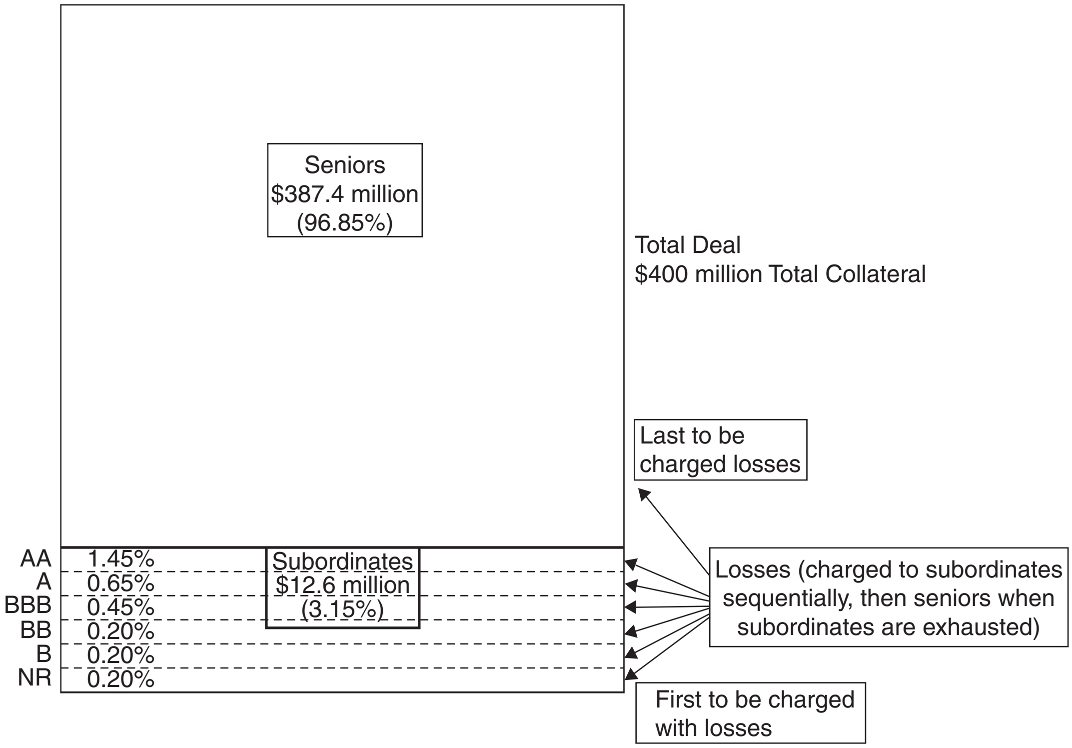


Figure 32.2 Diagram of a Senior/Subordinate Structure

coupon is designated, the loans are divided into "discount" and "premium" loan groups. This calculation subtracts the base servicing and fees from each loan's note rate to create the net note rate. The net note rate is then compared to the deal's designated coupon. Discount loans are those loans that have a net note rate lower than that of the deal's coupon; premium loans are those where the net note rate is above the deal coupon.

At this point, the two loan groups are each structured to give them the deal's coupon. The discount loans are "grossed up" to the deal's coupon rate by creating, for each note rate, a small amount of PO. (By creating some PO for each strata, the available interest is allocated over a smaller amount of principal, effectively raising its net note rate to that of the deal coupon.) The amount of PO created for each rate stratum is computed based on the PO percentage, which is calculated as follows:

$$\text{PO percentage} = [\text{Deal coupon} - \text{Note rate}] \div \text{Deal coupon}$$

The PO percentage for each note rate stratum is then multiplied by its face value, and the sum of the POs created for all discount note rates is the size of the WAC PO.

The loans in the premium loan group are stripped of some of the interest in order to reduce their net note rates to that of the deal coupon. The interest strip is assigned a notional value equal to the face value of the stratum. As an example, assume that $20 million face value of loans has a 6.5% note rate, and the designated deal coupon is 6.0%. Assuming 25 basis points of base servicing and no fees gives it a 6.25% net note rate. Therefore, 25 basis points of interest is stripped from these loans, creating $20 million notional value of a strip with a coupon of 0.25%. The notional value of the WAC IO is simply the combined notional value of all loans having premium net note rates, and its coupon is the average of the strip coupons weighted by their notional balances. (Note that in some cases the strip cash flows generated by the premium loans are held by the originator in the form of excess servicing, rather than securitized into a WAC IO.)

The breakdown and grouping of loans backing a hypothetical private label deal, and the structuring of the loans into a pool with one fixed coupon rate, is shown in Table 32.1. The table shows the calculations for a package of loans with various note rate strata for a deal with a 5.75% coupon, assuming 25 basis points of base servicing and 0.9% trustee fee (which are both standard assumptions at this writing). All loans with note rates of 6.125% and higher are considered premium loans, since they will have a net note rate higher than the 5.75% cutoff; loans with note rates below 6.125% are classified as discount loans. Notice that changing the deal's coupon changes the sizes of the WAC IO and PO, as well as the WAC IO's coupon. In the example, lowering the deal coupon to 5.5% pushes $82 million face value of loans, with note rates of 5.75% and 5.875%, into the premium sector, increasing the WAC IOs notional face value from $333.5 million to $415.5 million. The face value of the WAC PO declines, however, from approximately $7.21 million to $2.72 million. Therefore, the "market conditions" influencing the choice of coupon include the preferences of investors for premium or discount coupons, as well as the relative demand for IOs and/or POs.

Table 32.1 Example of Loan Stratification and Coupon Creation for a Hypothetical Private-Label Deal (assuming 25 basis points base servicing, 0.9 basis points trustee fee, and a 5.75% security coupon)

	Note Rate	Net Note Rate[a]	Balance in Cohort	Difference—Net Note Rate and Coupon	Net Contribution to WAC[b]	PO %[c]	PO% × Balance	Face Value Added to WACIO
Discount Loans	5.000%	4.741%	500,000	−0.0101	0.0000	17.5%	87,739	0
	5.125%	4.866%.	2,600,000	−0.0088	0.0000	15.4%	399,722	0
	5.250%	4.991%	5,000,000	−0.0076	0.0000	13.2%	660,000	0
	5.375%	5.116%	8,000,000	−0.0063	0.0000	11.0%	882,087	0
	5.500%	5.241%	16,400,000	−0.0051	0.0000	8.9%	1,451,757	0
	5.625%	5.366%	21,000,000	−0.0038	0.0000	6.7%	1,402,435	0
	5.750%	5.491%	31,000,000	−0.0026	0.0000	4.5%	1,396,348	0
	5.875%	5.616%	37,000,000	−0.0013	0.0000	2.3%	862,261	0
	6.000%	5.741%	45,000,000	−0.0001	0.0000	0.2%	70,435	0
Premium Loans	6.125%	5.866%	55,000,000	0.0012	0.0012	0.0%	0	55,000,000
	6.250%	5.991%	70,000,000	0.0024	0.0024	0.0%	0	70,000,000
	6.375%	6.116%	41,000,000	0.0037	0.0037	0.0%	0	41,000,000
	6.500%	6.241%	42,000,000	0.0049	0.0049	0.0%	0	42,000,000
	6.625%	6.366%	37,000,000	0.0062	0.0062	0.0%	0	37,000,000
	6.750%	6.491%	30,500,000	0.0074	0.0074	0.0%	0	30,500,000
	6.875%	6.616%	22,000,000	0.0087	0.0087	0.0%	0	22,000,000
	7.000%	6.741%	21,000,000	0.0099	0.0099	0.0%	0	21,000,000
	7.125%	6.866%	8,000,000	0.0112	0.0112	0.0%	0	8,000,000
	7.250%	6.991%	4,000,000	0.0124	0.0124	0.0%	0	4,000,000
	7.375%	7.116%	3,000,000	0.0137	0.0137	0.0%	0	3,000,000
						Total	7,212,783	333,500,000
	Total Deal Size		500,000,000					
	WAC IO Size[d]		333,500,000					
	WAC PO Size[e]		7,212,783					

[a]Note rate less base servicing and trustee fee.
[b]For premium loans, the net contribution is defined as: Net note rate—Security coupon. It is 0 for discount loans.
[c]For discount loans, the PO percentage is defined as: (Security coupon—Net note rate)/Security coupon.
[d]The face value of the WAC IO is the sum of the face value of the premium loans.
[e]The face value of the WAC PO is the sum of the face value of the discount loans times the PO percentage.

MBS TRADING

The structure of the MBS markets has long reflected the practices of both originators and borrowers in the primary mortgage market. This discussion is facilitated by a brief overview of the timeline of a mortgage loan, illustrated in Figure 32.3. A loan begins as an application, which may either be associated with a designated rate (making the loan "locked" or "committed") or carried as a floating rate obligation (to be locked at a point prior to funding). Borrowers that lock their loan at the time of application pay slightly more for their loans (in terms of either a rate differential or slightly higher fees) to account for the cost of hedging the loan for the period between application and funding. Most importantly, there is a lag between the points in time when borrowers apply for their loans and the loans are funded that lenders must take into account in managing their book of business, or pipeline. This lag reflects the time necessary for lenders to underwrite the loan and process the paperwork, which includes appraisals, title searches and insurance, geological and flood surveys, and credit analysis. In addition, purchase transactions often require additional time to process and register the underlying real estate transaction.

The lag between application and funding, which varies depending on the type of loans and market conditions, allows lenders to sell their expected production for

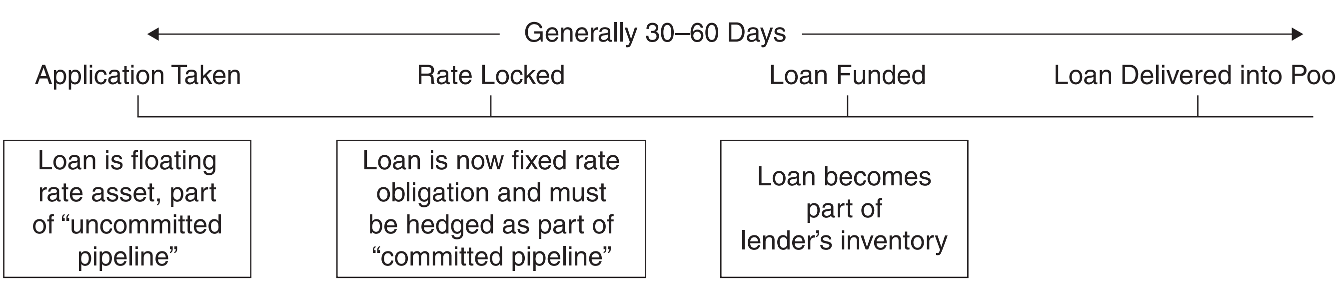


Figure 32.3 Timeline of Loan from Application to Agency Pooling

settlement in the future. However, it also forces lenders to manage and hedge their production pipeline in order to control the variability of their proceeds and maximize their profitability. While hedging a loan pipeline is similar in concept to hedging other portfolios, it requires lenders to continuously be appraised of the rate of new applications (which adds to the position) as well as so-called "fall out," which occurs as some borrowers allow their loan applications to lapse. A fairly consistent amount of loans will fall out under all circumstances, reflecting transactions that fail to close for a variety of reasons. However, fallout of committed loans can change sharply if lending rates fluctuate. For example, a drop in rates typically causes an increase in the number of loans that fall out as applicants let their existing application lapse and apply for new loans. In the same fashion as negative convexity occurs with mortgage loans, changing fallout rates complicate the process of hedging by making changes in the pipeline's value nonlinear with respect to interest rates.

The need for lenders to sell their expected production for future settlement has resulted in the MBS market being structured as a so-called *forward market*. In a forward market, a trade is agreed upon between two parties at a price for settlement (that is, the exchange of the item being traded for the agreed-upon proceeds) at some future date.

MBS Market Structure

The MBS market has evolved a number of conventions unique to the needs of both mortgage originators and investors. For example, settlements occur each month according to a predetermined calendar which specifies the delivery date for a variety of products over the course of each month. (The calendar is developed by the Bond Market Association (BMA) and published roughly six months in advance.) Prices are typically quoted for three settlement months (e.g., a quote sheet in March would post prices for April, May, and June settlement). However, trades can be executed farther in the future, subject to accounting and counterparty risk considerations.

Transactions in fixed-rate pass-through securities can be effected in one of three ways:

1. **A preidentified pool or pools can be traded.** In this type of "specified pool" trade, the pool number and "original balance" of the pool (that is, the amount of the pool as if it were a brand-new pool, before the effects of paydowns) are identified at the time the transaction is consummated.
2. **A so-called to-be-announced (TBA) trade.** In this case, the security is identified (e.g., Fannie Mae 6.0s) and a price is set; but the actual pools identities are not provided by the seller until just before settlement. (This process is referred to as pool allocation.) The attributes of the pools that are eligible for delivery into TBA trades is specified by the BMA in order to effect a degree of standardization.
3. **A "stipulated" trade.** This is a variation on a TBA trade, but the underlying characteristics of the pool are specified more precisely than in a standard TBA trade. In some cases, the pools in a stipulated trade are not deliverable, under the BMA rules, into TBA pools. In other instances, the pools can be delivered, but are viewed as having incremental value to investors and trade at a premium to TBAs.

The TBA market only exists, at this writing, in the fixed rate market for agency pools. As noted previously, there is currently no equivalent to the TBA in the ARM market for conventional ARMs because of the wide variety of product types and specifications in the ARM market. (There has been a TBA market in the Ginnie ARM product, but trading in that sector became fairly illiquid in the late 1990s.) ARM products trade almost entirely as either specified or stipulated (or "stipped," as it is sometimes called) pools, although they generally settle based on "good-day" delivery specified by the BMA calendar. Both agency and private label deals are settled at the end of the month; secondary trading typically occurs for settlement three business days after the trade is executed, for so-called "corporate settlement."

Financing and the Dollar Roll Market

An interesting attribute of forward markets that has appeal to MBS participants is the fact that they implicitly create a built-in financing vehicle. The forward market mechanism allows trading in the same securities for settlement in different months. As noted, originators generally sell their production for forward settlement in order to monetize and hedge their pipelines. However, there is also demand for MBS pools for settlement in the early or "front" months. For example, some types of investors (such as depository institutions) generally put securities on their books rather than forward obligations, which may not receive favorable accounting treatment. In addition, dealers acquiring agency pools as collateral for agency CMO deals must take delivery of the pools before their structured transaction settles. Therefore, MBS trading involves pricing the same securities for different settlement dates. In addition, dealers make active markets in TBAs for different settlements, simultaneously buying positions for one settlement month and selling the identical position for another. This type of transaction is known as a *dollar roll* or simply a "roll."

Simply put, valuing dollar rolls involves weighing the benefits and costs, over a holding period, of either:

1. Buying the security for the earlier (or "front") month, and owning (and financing) it for the period ending with the latter (or "back" month) settlement date.
2. Buying the security for the back month's settlement.

In the first case, where the security is bought for the front month, the investor receives coupon payments and reinvested interest for the holding period, along with principal payments (both amortizations and prepaid principal). The investor must also finance the position, typically through the repurchase market. In theory, the back month price is such that the investor is indifferent between the two alternatives. In practice, the price difference (or "drop") between the two settlement dates is often greater than that implied by the break-even calculation, which means that the investor buying the position for back-month settlement is effectively financing the security at an implied repo rate lower than that available in the repurchase market.

CASH FLOW STRUCTURING

As noted previously, the cash flows generated by agency pools and senior private-label pass-throughs are very similar in nature. Both securities can be structured to take advantage of demand for a variety of securities by different segments of the fixed income investment community. Various investor clienteles have different investment objectives and risk tolerances, and thus tend to invest in securities with different cash flow and performance attributes. Some different market segments include:

- Banks and other depository institutions, which generally seek short securities where they can earn a spread over their funding costs.
- Life insurance companies and pension funds, which typically invest in bonds with longer maturities and durations in order to immunize long-dated expected liabilities.
- Investment managers, who typically manage fixed income assets versus performance indexes.
- Hedge funds, which typically seek investment vehicles that offer the potential for very high-leveraged returns.

The nature of mortgage cash flows makes mortgage loans and mortgage-backed securities ideal vehicles for creating a variety of bonds. Their long-term principal and interest cash flows allow structurers to create securities of varying average lives and durations in order to meet the needs of different classes of investors. In addition, different structures allow different risks (both prepayment and, for private-label deals, credit) to be transferred within the structure, and creates a rich environment for the wide variety of structures and structuring techniques available. For a discussion of MBS structuring, see Chapters 5 through 8 of Fabozzi, Bhattacharya, and Berliner (2007).

However, mortgage structures are closed universes by nature, in that all balances and cash flows generated by the collateral within the structure must be taken into account. For example, a structure where the coupon of one bond is stripped below that of the collateral must allocate the incremental interest cash flows elsewhere in the structure. This shifting of interest cash flows can be done in a number of different forms (see Fabozzi, Bhattacharya, and Berliner (2007)). Another example might be a bond that pays principal to investors based on a schedule. This stabilizes the "scheduled" bond's average life and duration, but cash flow uncertainty is transferred to other bonds in the structure, giving their cash flows greater variability.

Therefore, the process of MBS structuring requires examining and valuing the trade-offs necessary to create a variety of bonds designed to meet the needs of multiple investor clienteles. To create a more desirable bond within a structure, for example, the underwriter must be able to sell the enhanced bond (or combination of bonds) at a better valuation than the original tranche, in order to offset the concession that must be given to attract investors to the bond with less appealing attributes. Understanding the trade-offs involved in structuring therefore requires an understanding of how the different structuring techniques work, and how they impact other bonds within the structure.

SUMMARY

The market for mortgage-backed securities is the largest cash securities market in the world, and is almost half again as large as the Treasury market. The development of the MBS pool, which facilitates the aggregation of many thousands of unique assets into fungible securities, has been a critical factor in the growth of the MBS market to its current size. While a large portion of the MBS market has credit enhancement from government or quasi-governmental agencies, large volumes of securities are issued without such guarantees. These so-called private-label securitizations, issued without the credit support of government agencies or enterprises, typically use subordination as a mechanism for creating large amounts of triple A senior securities. The cash flows of both agency and senior private-label pass-throughs can be restructured or "tranched" to tailor securities more closely to different investors' preferences, as well as to redistribute risk and yield within the structure.

REFERENCES

Davidson, A., Ho, T., and Lim, Y. (1994). *Collateralized Mortgage Obligations: Analysis, Valuation and Portfolio Strategy*. New York: McGraw-Hill.

Fabozzi, F. J. (ed.) (2006). *The Handbook of Mortgage-Backed Securities*, 6th edition. New York: McGraw-Hill.

Fabozzi, F. J., Bhattacharya, A. K., and Berliner, W. S. (2007). *Mortgage-Backed Securities: Products, Structuring, and Analytical Techniques*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., and Dunlevy, J. (2001). *Real Estate Backed Securities*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., and Kalotay, A. (2003). *Ginnie Mae and the Secondary Mortgage Market: An Integral Part of the American Economic Engine*. Government National Mortgage Association, March 2003.

Fabozzi, F. J., and Modigliani, F. (1992). *Mortgage and Mortgage-Backed Securities Markets*. Boston: Harvard Business School Press.

Fabozzi, F. J., and Ramsey, C. (1999). *Collateralized Mortgage Obligations*, 3rd edition. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., Ramsey, C., and Marz, M. (eds.) (2000). *The Handbook of Nonagency Mortgage-Backed Securities*. Hoboken, NJ: John Wiley & Sons.

CHAPTER 33

Structuring Collateralized Mortgage Obligations and Interest-Only/ Principal-Only Securities

ANDREW DAVIDSON
President and Founder, Andrew Davidson & Co., Inc.

ANTHONY SANDERS, PhD
John W. Galbraith Chair and Professor of Finance, Ohio State University

LAN-LING WOLFF
Fredell & Co. Structured Finance Ltd.

ANNE CHING
Senior Consultant, Andrew Davidson & Co., Inc.

Abstract: The mortgage market has found a way to restructure mortgage cash flows to meet the needs and views of a variety of investors. The basic mortgage pass-throughs all have very similar cash-flow structures and performance characteristics. Discounts and premiums differ to some extent, but the overall investment patterns are quite similar. While the investment characteristics of these loans are similar, the needs of the investors vary significantly. The collateralized-mortgage obligation (CMO) has become the vehicle to transform mortgage cash flow into a variety of investment instruments. The driving force behind the creation of CMOs is arbitrage. CMOs will be created when the underwriter sees the ability to buy mortgage collateral, structure a CMO, and sell the CMO bonds for more than the price of the underlying collateral plus expenses. Because of the dynamic nature of arbitrage opportunities, the types of CMOs created will reflect current market conditions and can change significantly. If the end result of the cash flow structure were simply a rearrangement of cash flows, it would be difficult to create added value. CMOs are successful, nevertheless, for two main reasons. First, investors have varying needs and are willing to pay extra for a bond that meets their specific needs. Second, investors misanalyze bonds. Many investors rely on tools such as yield spread and average-life analysis. These tools are insufficient to analyze mortgage-backed securities and CMO bonds.

Keywords: collateralized mortgage obligation (CMO), principal-pay types, sequential bonds, pro rata bonds, planned amortization classes (PACs), interest-pay types, floating-rate bond, inverse-floating-rate bonds, sequential bonds, IOettes, scheduled bonds, sequential PACs, interest-only bonds, principal-only bonds, senior/subordinated structures

A *collateralized mortgage obligation (CMO)* can be defined as a bond secured by mortgage cash flows. The mortgage cash flows are distributed to the bond based on a set of prespecified rules. The rules determine the order of principal allocation and the coupon level. The specific choice of CMOs and structured mortgage products created is determined by interaction of market demand with credit, legal, tax, and accounting requirements. The primary ingredient in CMO creation is the availability of mortgage cash flows. The cash flows provide the raw material for the CMOs. Every CMO must address the amount and availability of cash flow. In this chapter, we explain the different types of bond classes that can be created in an asset securitization.

NON-CASH-FLOW ASPECTS OF CMOs

While our focus is on the cash flow aspects of CMOs, a brief discussion of the other aspects of CMO creation is warranted. An important component of the CMO is the assurance that the investor will get the promised cash flows. The market has developed several methods for achieving this goal. Generally, the mortgages or the agency-backed mortgage pools are placed in a trust and the CMO bonds are issued out of that trust. Various legal structures can be used to create a bankruptcy-remote entity to hold the mortgages and issue the bonds. The investor looks to the trust and cash flows of the mortgages to provide the bond's principal and interest payments. These payments are assured through either a rating agency assurance (that is, a triple A rating) or through the guarantee of a government-sponsored enterprise (GSE), either Fannie Mae or Freddie Mac. These mortgages back private-label CMOs issued by Wall Street firms, large mortgage originators, or mortgage conduits and are rated by rating agencies.

The tax treatment of CMOs is generally covered under the provisions of the Real Estate Mortgage Investment Conduit (REMIC) rules. CMOs, at times, are referred to as REMICs. In order to be a REMIC, the bonds must have a certain structure and must elect REMIC status. REMIC election drives the tax treatment of the bonds. The regular interests of the REMIC are generally taxed as ordinary bonds, whereas the residual interest bears the tax consequences of the CMO structure. Originally, the residual interest was intended to receive excess interest not distributed to regular interests. However, most residuals no longer have cash flow attached to them and are distributed primarily based on their tax consequences.

The accounting treatment for most CMOs is straightforward. However, CMO bonds sold at a premium or discount must be evaluated on a level yield basis. That is, income is determined by the yield of the bond rather than its coupon. When prepayments change, the expected cash flows of the security changes. As a result, the income stream must be adjusted accordingly. This is a complex area, especially for some CMO residuals, interest-only (IO) and principal-only (PO) securities. The rules for treatment of these bonds are subject to change. Please consult with your tax and accounting advisors before purchasing CMOs.

Once the legal, tax, and accounting issues are resolved, the investment characteristics of a CMO will be driven by the cash flows of the underlying collateral and the structure of the CMO deal. In order to understand CMOs, it is necessary to understand the rules by which the mortgage cash flows are distributed to the bonds.

COLLATERALIZED MORTGAGE OBLIGATIONS AS RULES

The CMO can be thought of as a set of rules. The rules tell the trustee how to divide the payments that it receives on the mortgages. The rules tell the trustee in what order to pay the bondholders and how much to pay them. The rules generally can be split between principal-payment rules and interest-payment rules. Market participants have developed standard definitions for CMO types; these types specify the nature of the rules used to distribute cash flows. These standard types include principal-pay types and interest-pay types. Each bond has both a principal-pay type and interest-pay type. Table 33.1 shows the standard CMO definitions.

Each CMO represents a combination of these bond types and, hence, of the mortgage rules. In the following examples, we show how these rules are applied and how complex CMO structures can grow out of these simple rules. In this chapter, we concentrate on several of the most common principal and interest rules.

The starting point for the creation of the CMO is the mortgage collateral. For the following examples, we use newly originated agency collateral with a net coupon of 8% and a gross coupon of 8.6%. Assume a 30-year maturity and an age of 5 months. CMOs are generally priced as structured using the Public Securities Association (PSA) model. We use 175 PSA as our base speed and look at the effects on the structure of prepayments at 100 PSA and at 400 PSA. Figures 33.1a and 33.1b show the principal balance outstanding and the cash flows of the mortgages. Note that the cash flows consist of principal and interest payments. The principal payments represent both the scheduled principal payments and the unscheduled payments (prepayments). The interest cash flows consist of the net interest payment to the investor and the payment to the servicer and guarantor. The change in balances and cash flows for speeds of 100 PSA and 400 PSA are shown in Figures 33.2 and 33.3. The cash flows of mortgages are the raw material for the CMO. The cash flows of the CMO bonds must come from the mortgage cash flows. As the cash flows of the underlying pass-through change, the cash flows of the CMO bonds must also change.

PRINCIPAL-PAY TYPES

Principal-pay rules determine how principal payments are split among CMO tranches. These rules can be applied alone or in combination with one another. For *sequential*

Table 33.1 Standard Definitions for REMIC and CMO Bonds

Agency Acronym	Category of Class	Principal-Pay Types
AD	Accretion directed	Pays principal from specified accretions of accrual bonds. ADs may, in addition, receive principal from the collateral paydowns.
AFC	Available funds	May receive as principal, in addition to other amounts, interest paid on the underlying assets of the series trust to the extent that the interest exceeds certain required interest distributions on this class (or related class—Freddie Mac).
CALL	Call	Freddie Mac only; Holders have the right to direct the issuer to redeem the related callable class or classes.
CALLABLE/CC	Callable	Receive payments based on distributions on underlying callable securities (directly or indirectly, at the direction of the holder of the related call class—Freddie Mac).
GMC	Guaranteed maturity class	Freddie Mac only: Final payment date is earlier than the latest date by which those classes could be retired by payments on their underlying assets. Typically, holders of a guaranteed maturity class receive payments up to their final payment date from payments made on a related underlying REMIC class. On its final payment date, however, the holders of an outstanding GMC will be entitled to receive the entire outstanding principal balance of their certificates, plus interest at the applicable class coupon accrued during the related accrual period, even if the related underlying REMIC class has not retired.
NPR	No payment residual	Receives no payments of principal.
NSJ	Nonsticky jump	Principal pay down is changed by the occurrence of one or more triggering events. The first time the trigger condition is met, the bond changes to its new priority for receiving principal and remains in its new priority for the life of the bond.
NTL	Notional	No principal balance and bears interest on its notional principal balance. The notional principal balance is used to determine interest distributions on an interest-only class that is not entitled to principal.
PAC	Planned amortization class	Pays principal based on a predetermined schedule established for a group of PAC bonds. The principal redemption schedule of the PAC group is derived by amortizing the collateral based on two collateral prepayment speeds. These two speeds are endpoints for the "structuring PAC range." A PAC group is therefore defined as PAC bonds having the same structuring range. A group can be a single bond class.
PT	Pass-through	Receives principal payments in direct relation to actual or scheduled payments on the underlying securities, but is not a strip class.
SC	Structured collateral	Receives principal payments based on the actual distributions on underlying securities representing regular interests in a REMIC trust.
SCH	Scheduled	Pays principal to a set redemption schedule(s), but does not fit the definition, of a PAC or TAC.
SEG	Segment	Combined, in whole or part, with one or more classes (or portion of classes) to form a segment group or aggregate group for purposes of allocating certain principal distribution amounts.
SEQ	Sequential pay	Starts to pay principal when classes with an earlier priority have paid to a zero balance. SEQ bonds enjoy uninterrupted payment of principal until paid to a zero balance. SEQ bonds may share principal paydown on a pro rata basis with another class.
SJ	Sticky jump	Principal paydown is changed by the occurrence of one or more triggering events. The first time the trigger condition is met, the bond changes to its new priority for receiving principal and remains in its new priority for the life of the bond.
SPP	Shifting-payment percentage	Freddie Mac only: Receives principal attributable to prepayments on the underlying mortgages in a different manner than principal attributable to scheduled payments or shifting proportions over time.
STP	Strip	Receives a constant proportion, or strip, of the principal payments on the underlying securities or other assets of the series trust.
SUP	Support (or companion)	Receives principal payments after scheduled payments have been paid to some or all PAC, TAC, or SCH bonds for each payment date.
TAC	Target (or targeted amortization class)	Pay principal based on a predetermined schedule, derived by amortizing the collateral based on a single prepayment speed.
XAC	Index allocation	Principal payment allocation is based on the value of an index.

(*continued*)

Table 33.1 (*Continued*)

Agency Acronym	Category of Class	Interest-Pay Types
AFC	Available funds	Receives, as interest, certain interest or principal payments on the underlying assets of the related trust. These payments may be insufficient on any distribution date to cover, fully, the accrued and unpaid interest of this class at its specified interest rate for the related interest accrual period. In this case, the unpaid interest amount may be carried over to subsequent distribution dates (and any unpaid interest amount may itself accrue interest) until payments are sufficient to cover all unpaid interest amounts. It is possible that these insufficiencies will remain unpaid and, If so, they will not be covered by issuer's guaranty.
ARB	Ascending-rate bond	Have predetermined coupon rates that take effect one or more times on dates set forth at issuance.
DLY	Delay	A floating rate, inverse-floating rate, or weighted-average coupon class for which there is a delay of 15 or more days from the end of its accrual period to the related payment date.
DRB	Descending-rate bond	Have predetermined class coupons that decrease one or more times on dates determined before issuance.
EXE	Excess	Entitled to collateral principal and interest paid that exceeds the amount of principal and interest obligated to all bonds in the deal.
FIX	Fixed	Coupons are fixed throughout the life of the bond.
FLT	Floater (or floating rate)	Coupons reset periodically based on an index and may have a cap or floor. The coupon varies directly with changes in the index.
IDC/DIF	Index differential	Has an interest rate that reset periodically computed in part on the basis of the difference (or other specified relationship—Freddie Mac, Fannie Mae) between two designated indexes (e.g., LIBOR and the 10-year Treasury index).
INV	Inverse floater	Coupons reset periodically (like floaters) based on an index and may have a cap or floor. The coupon varies inversely with changes in the index.
IO	Interest only	Receive some or all of the interest portion of the underlying collateral and little or no principal. A notional amount is the amount of principal used as a reference to calculate the amount of interest due. A nominal amount is actual principal that will be paid to the bond. It is referred to as "nominal" since it is extremely small compared to other classes.
NPR (also above)	No payment residual	Receives no payment of interest.
PEC	Payment exchange certificates	Freddie Mac only: Class coupons vary, in whole or in part, based on payments of interest made to or from one or more related classes.
PO	Principal only	Receives no interest.
PZ	Partial accrual	Accretes interest (which is added to the outstanding principal balance) and receives interest distributions in the same period. These bonds have a stated coupon, which is equal to the sum of the accretion coupon and interest distribution coupon.
W/WAC	Weighted-average coupon	Represent a blended interest rate (effective weighted-average interest rate—Fannie Mae), which may change in any period. Bonds may be comprised of nondetachable components, some of which have different coupons.
Z	Accrual	Accrete interest that is added to the outstanding principal balance. This accretion may continue until the bond begins paying principal or until some other event has occurred.
CPT	Component	Comprised of nondetachable components. The principal pay type or sequence of principal pay of each component may vary.
IMD	Increased minimum denomination	Ginnie Mae only: To be offered and sold in higher minimum denominations than those of other classes.
LIQ	Liquidity	Intended to qualify as a liquid asset for savings institutions. LIQ bonds are any agency-issued bonds that have a stated maturity of 5 years (or less), or any non-agency-issued bonds that have a stated maturity of 3 years (or less), in each case from issue date.
RTL	Retail	Designated to be sold to retail investors.
R, RS, RL	Residual	Designated for tax purposes as the residual interest in a REMIC.
RDM	Redeemable	Fannie Mae only: Certificate that is redeemable directly or indirectly as specified in the prospectus.
SP	Special	Ginnie Mae only: Having characteristics other than those identified above.
TBD	To be defined	Does not fit into any of the current definitions

Sources: Fannie Mae, Freddie Mac, and Ginnie Mae.

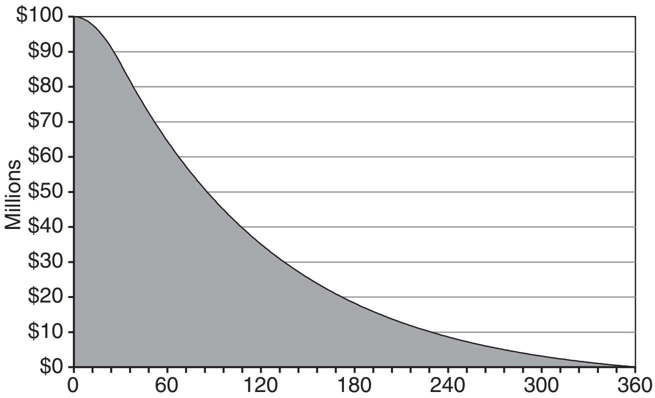


Figure 33.1a Balance at 175 PSA

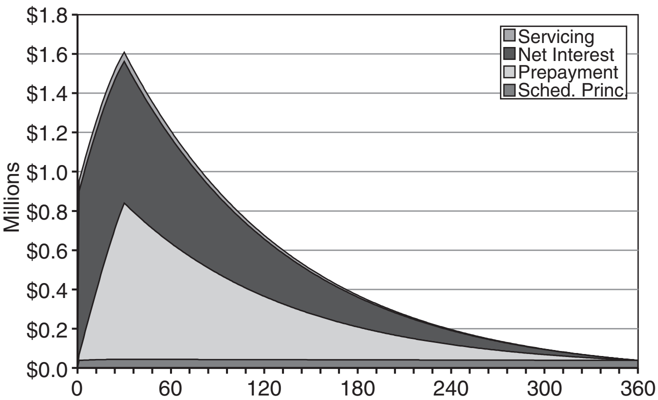


Figure 33.1b Cash Flow at 175 PSA

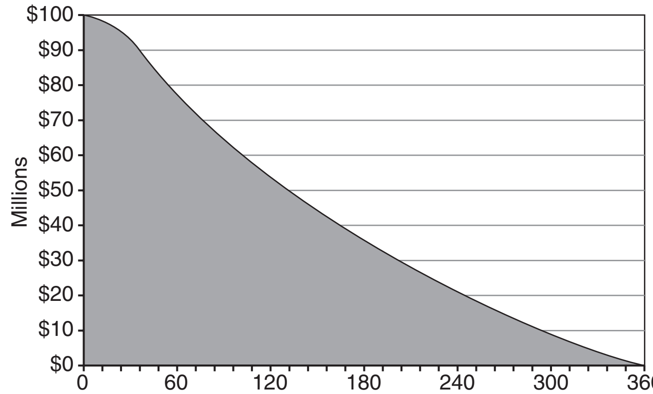


Figure 33.2a Balance at 100 PSA

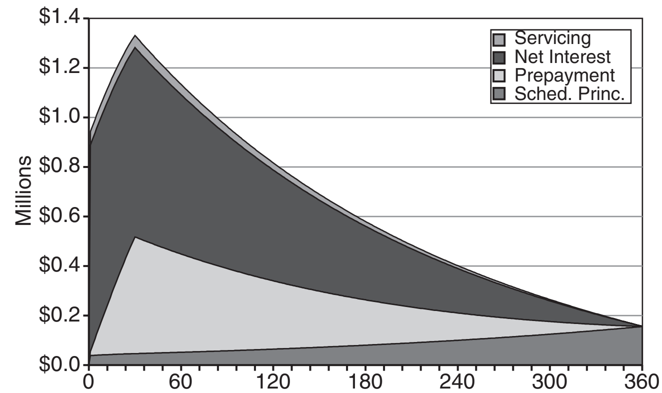


Figure 33.2b Cash Flow at 100 PSA

bonds, one bond is completely paid down before principal payment begins on the next. *Pro rata bonds* pay down simultaneously according to a fixed allocation. *Planned amortization classes (PACs)* are part of a group of bonds classified as scheduled bonds. These bonds receive priority within the structure for certain principal payments. *Support bonds* are created in conjunction with the scheduled bond and absorb the remaining principal payments.

INTEREST-PAY TYPES

Interest-pay rules determine the amount of interest received by the CMO bondholders each period. Interest-payment rules are linked with principal-pay rules to produce a wide variety of bond types. One rule of CMO creation is that the combined interest on the CMO bonds must be less than the available interest from the collateral. Fixed interest payments are the most common type. The bondholder receives an interest amount, which is a constant percentage of the outstanding balance. In an *accrual bond* (or Z bond), the bondholder does not receive interest payments for some time period. During this time, the interest payments are converted to principal, and the balance of the investment increases. The coupon for a *floating-rate bond* changes based on an underlying index. Floating-rate bonds typically pay a margin above an index (frequently the London Interbank Offered Rate [LIBOR]) and have an interest rate cap. *Inverse-floating-rate bonds* are usually produced in conjunction with floating-rate bonds. Their coupon moves inversely with the index, usually at some multiple of the index. They typically have a cap and a floor. *Principal-only* and *interest-only* payment types provide for bonds with principal payments and no interest payment or interest payments with no principal.

SEQUENTIAL BONDS

The first CMO bonds were sequential CMOs. They were created to turn mortgage-backed securities (MBS) into more corporate bond-like investments. Sequential bonds tend to narrow the time over which principal payments are received, creating a more bullet-like structure. Figure 33.4 shows the cash flow of a typical sequential CMO. In this example, classes A, B, C, and Z are sequential bonds. Class A receives all of the principal payments first. Once class A is completely paid off, then class B begins to receive principal payments. Once class B is paid off, then C begins principal payments, and so on until class Z is paid off. Note that each bond receives principal payments over a relatively narrow time period.

The interest payment on each of these tranches is fixed and equal to the net coupon of the underlying MBS. Each bond receives a monthly interest payment equal to the

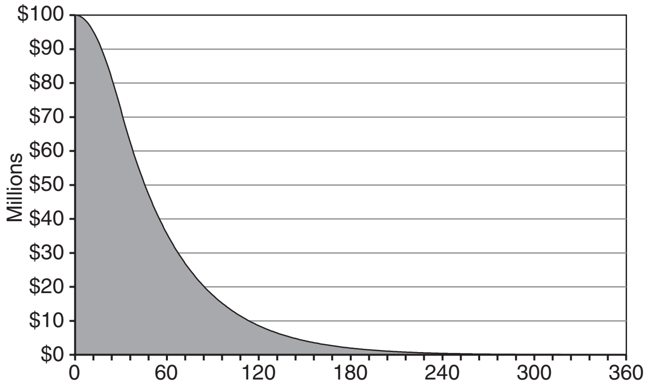


Figure 33.3a Balance at 400 PSA

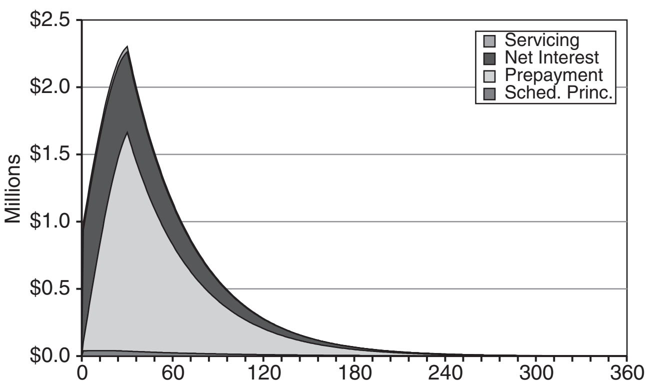


Figure 33.3b Cash Flow at 400 PSA

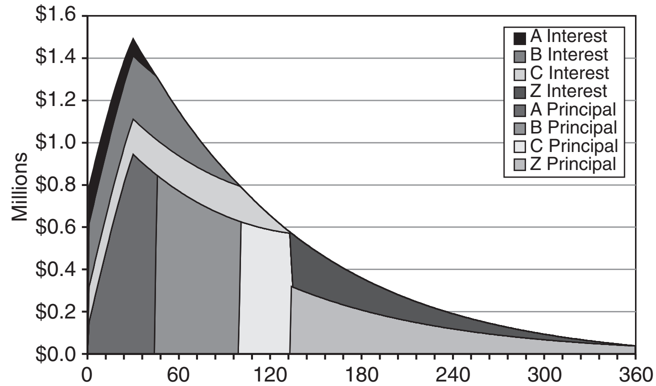


Figure 33.4 Cash Flows of a CMO

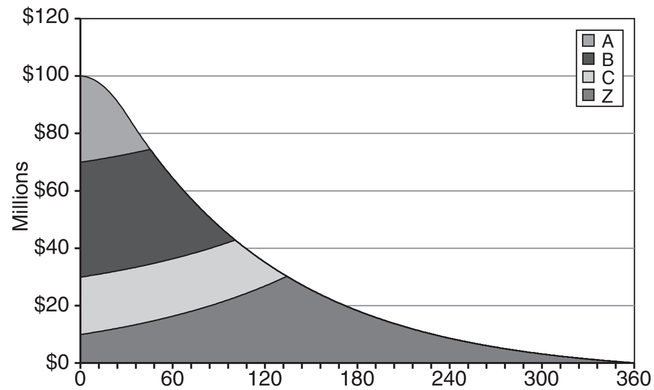


Figure 33.5 Balance Outstanding Over Time

coupon divided by 12 times the outstanding balance of that tranche. Thus, tranches A, B, and C all receive interest payments beginning in month one. Class Z is an accrual bond. Rather than receiving its share of interest, its interest payment is converted to principal and is used to increase the balance of the Z tranche. The interest payment that should have gone to Z is used to make principal payments to tranches A, B, and C. This can also be seen in Figure 33.4. Figure 33.5 shows the balance outstanding of each tranche over time. Note that the balance of tranche A begins to decline immediately. The balance of tranches B and C are constant until the prior tranches are paid off. Tranche Z shows an increasing balance until all of the earlier tranches are paid off. The net effect is to shorten the average life of tranches A, B, and C. Typically, shorter tranches are priced at lower yields. By increasing the amount of principal received by the shorter tranches, CMO structures are able to increase the value of the CMO arbitrage.

As prepayments increase, the payments on the bonds will be received earlier. While some analysts have argued that sequential bonds offer protection from prepayment risk, it is difficult to make general statements about the amount of risk in a sequential bond. Rather than rely on a general prescription of which bonds are safe and which are risky, it is better to perform an analysis of the specific tranche that is a potential acquisition.

PRO RATA BONDS

Much of the complexity of CMO structures arises from layering different types of principal-payment rules. Pro rata bonds provide one means to affect this layering. Pro rata bonds are two or more bonds that receive cash flows according to exactly the same rules. Cash flows available to these bonds are divided proportionally. Figure 33.6 shows an example of pro rata bonds. The figure shows only the principal payments of the bonds. Tranches B1 and B2 receive a pro rata share of the cash flows that went to class B in the earlier example. Here, B1 receives 40% of the principal while B2 receives 60% of the principal.

Pro rata bonds are created to allow for different interest-payment rules for the same principal payment and will change the risk characteristics of the bonds to make them attractive to different investors. Given this pro rata structure, there are several choices of interest-pay rules possible for B1 and B2. One possibility is that they will both be fixed-rate bonds, but with different coupons. Say B1 has a coupon of 6%. The coupon of B2 cannot exceed 9.33%, since the weighted-average coupon cannot exceed 8%. Through this mechanism it is possible to create bonds that have coupons that are higher and lower than the collateral coupon.

Another example of a pro rata bond is an IO strip. It is possible to create an IO off of any bond. For some time, REMIC rules required that all regular interests have a principal component. At that time, IOs were created with a tiny

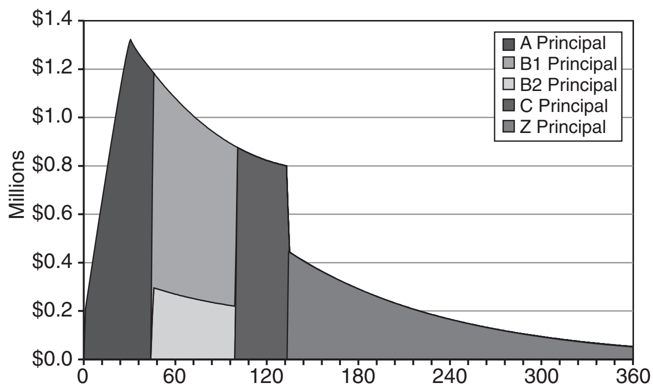


Figure 33.6 Principal Payments of a Pro Rata Bond

piece of principal and generally had coupons of 1,200% due to federal wire requirements. The limitation on principal has now been removed so that a pure interest strip can be created off of any bond. Some people call IOs that are created off of CMO bonds *IOettes* to distinguish them from IO strips created using all the interest payments of an MBS. In a CMO, IO strips are used to lower the coupon of the CMO tranche. Due to prepayment risk, investors prefer to buy bonds at a slight discount, so their yield will be less affected by changing prepayments.

Other forms of pro rata bonds are floaters and inverse floaters. Just as it is possible to create two fixed-rate bonds, where the combined coupon equals the coupon of the collateral, it is possible to create a *floating-rate bond* and an *inverse-floating-rate bond* whose coupon equals the collateral coupon.

Suppose bond B2 is a floating-rate bond with a coupon equal to LIBOR + 50 basis points: If LIBOR is currently 4%, then the coupon on B2 is 4.5%. The coupon on B1 would then be 13.25%. If LIBOR rises to 5%, the coupon on B2 becomes 5.5%, while the coupon on B1 must fall to 11.75%. As LIBOR rises, the coupon on B2 floats with LIBOR, while the coupon on B1 moves inversely to LIBOR. The coupon on B1 changes by 1.5 times the amount of the change in LIBOR. This inverse floater is said to have a slope of 1.5. Because the interest must come from the fixed-interest payment of the collateral, the coupon on these floaters must be capped. The floating-rate bond cannot exceed 13.33%, while the inverse floater cannot exceed 19.25%. Figure 33.7 is a graph of the possible coupon combinations of B1 and B2. The coupon on the inverse floater is usually described by a formula. In this case, the formula would be 19.25% – 1.5 × LIBOR.

SCHEDULED BONDS

While sequential bonds may offer some allocation of prepayment risk, investors seeking more protection from prepayment risk have turned to scheduled bonds for greater certainty of cash flow. Several types of scheduled bonds exist. Here, we concentrate on planned amortization classes (PACs). PACs are designed to produce constant cash flows within a range of prepayment rates. Unlike sequential bonds, PAC bonds provide a true allocation of risk. PAC bonds clearly have more stable cash flows than comparable non-PAC bonds. The additional stability of the PAC bonds comes at a cost. In order to create a more stable PAC bond, it is necessary to create a less stable support bond. Support bonds bear the risk of cash-flow changes. Although PAC bonds are somewhat protected from prepayment risk, they are not completely risk free. If prepayments are fast enough or slow enough, the cash flows of the PAC bonds will change. Furthermore, there can be great differences in the performance of PAC bonds. As with all CMO bonds, it is better to evaluate the cash-flow characteristics of the bond you are considering as an investment, than to rely on the type of the bond to indicate its riskiness. Some PAC bonds can be variable, and some support bonds can be very stable.

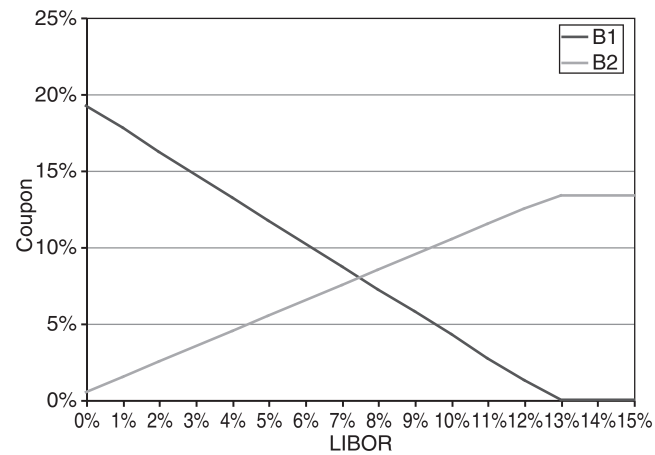


Figure 33.7 Floater and Inverse Floater Coupons

PACs are created by calculating the cash flow available from the collateral using two different prepayment speeds: a fast speed, 300 PSA, for example, and a slow speed, such as 100 PSA.

Figure 33.8 shows the principal cash flows of our collateral using 100 PSA and 300 PSA. The cash flow available each period under each scenario is the cash flow that can be used to construct the PAC bond. Under the 100 PSA assumption, there is less available in the early years of the CMO and more available in the later years. Under the 300

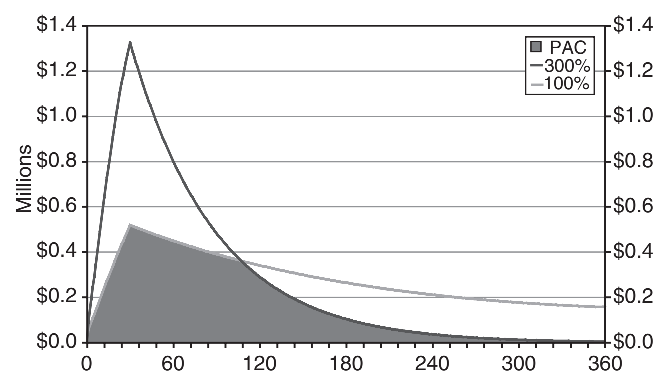


Figure 33.8 Determining PAC Bond Schedule

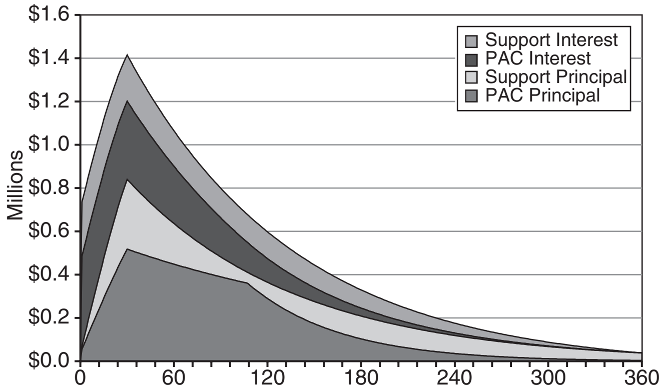


Figure 33.9 Cash Flows of a CMO at 175 PSA

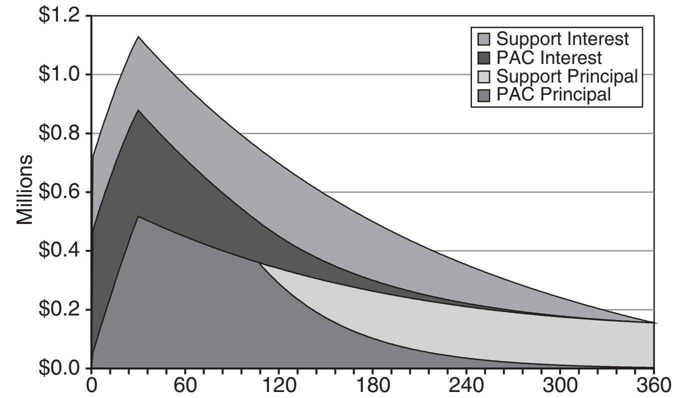


Figure 33.11 Cash Flows at 100 PSA

PSA assumption, there is more cash flow available during the early years of the CMO, and less available in the later years.

Figure 33.9 shows the cash flows of a CMO consisting of two classes, a PAC bond and a support bond, assuming a prepayment speed of 175 PSA. The PAC bond was constructed with a PAC bond of 100 PSA to 300 PSA. The principal payment pattern of the PAC bond is exactly equal to the schedule created using the two PAC band speeds. The cash flows are neither sequential nor pro rata. The support bond pays down simultaneously with the PAC bond, but the ratio of payments is determined by the PAC schedule and varies depending on prepayment rates. Figure 33.10 shows the paydown of the balance of the two classes at 175 PSA.

As prepayment rates change, the cash flow characteristics shift. At a speed of 100 PSA, far less cash flow is available in the early years of the CMO. Figure 33.11 shows the cash flows are 100 PSA. In the early years, all the principal cash flows go to the PAC bond. Principal payments on the support bond are deferred. Under this scenario, the cash flows of the support bond extend. The support bond does, however, receive more interest payments. Since 100 PSA is within the PAC bonds, the PAC bond still receives cash flow according to the original PAC schedule.

If prepayments are outside of the PAC bands, then the PAC schedule cannot be met. Figure 33.12 shows the cash flows of the CMO assuming prepayments at 400 PSA. Here the prepayment speed is outside the PAC bands. In this case, it is impossible to keep the PAC schedule. The cash flows to the support bond are accelerated. The support bond is fully paid off by month 91 and all remaining principal payments go to the PAC bond, significantly shortening its life. Even though the PAC schedule cannot be met, the PAC bond is still more stable than the support bond.

PAC bands are expressed as a range of prepayment speeds. In our example, we use 100 PSA to 300 PSA, which means that if prepayments occur at any single constant speed between 100 PSA and 300 PSA, the PAC schedule will be met. It does not mean that the PAC schedule will be met if prepayments on the collateral stay between 100 PSA and 300 PSA. Even if prepayments vary within the PAC bands, it is possible that the schedule will not be met. For example, if prepayments stay near 300 PSA for several years and then fall to near 100 PSA, the PAC bond will probably extend outside the PAC band. During the years at 300 PSA, the support bond will be nearly paid off. Then when prepayments slow, there is no way to cushion the extension of the PAC bond. Once again, analysis of the

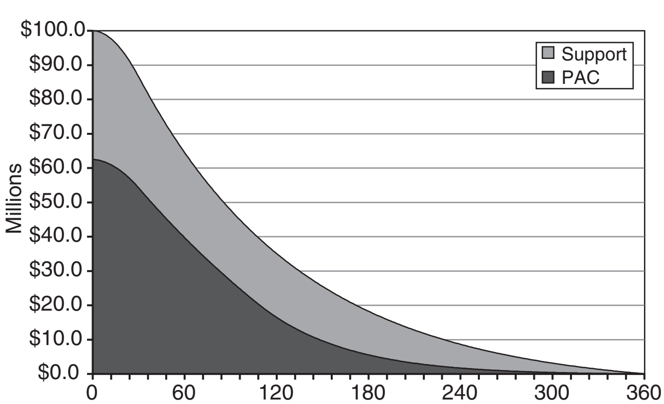


Figure 33.10 Principal Balance at 175 PSA

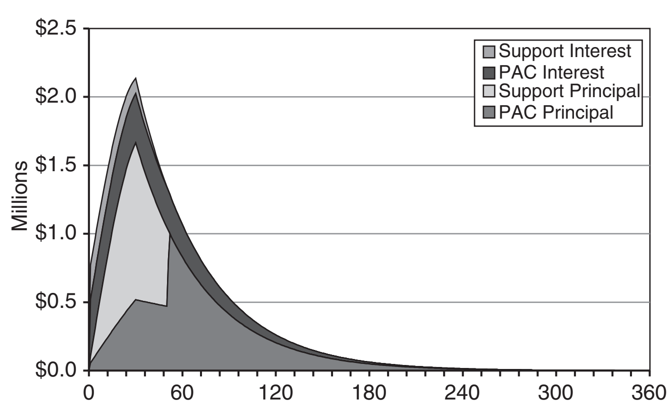


Figure 33.12 Cash Flows at 400 PSA

bond's cash flows are a more useful measure of the value of the bond than can be determined from its name alone.

SEQUENTIAL PACs AND OTHER COMBINATIONS

Just as sequential bonds were created to allow investors to specify a maturity range for their investments, PACs can be divided sequentially to provide more narrow paydown structures. Figure 33.13 shows the cash flows of a CMO in which the PAC has been split into several different bonds. These sequential PACs narrow the range of years over which principal payments occur. Investors with short horizons choose the earlier PACs, whereas investors with longer horizons choose the longer PACs. While these bonds were all structured using the 100 PSA to 300 PSA PAC band, the actual range of speeds over which their schedules will be met may differ. In particular, the early bonds can withstand faster speeds than the top of the PAC band, without varying from their scheduled payments.

Sequential PACs represent another example of how the CMO structuring rules can be combined to create more complex structures in order to meet a wider variety of investor requirements. It is possible to take any bond and further structure it. For example, the sequential PACs could be split using a pro rata structure to create high- and low-coupon PACs.

Another common PAC structure strips an IO piece to lower the coupon of the CMO classes and splits the collateral into a PAC and support bond. Another PAC class is created within the existing PAC class. The more stable PACs are called PAC Is and the less stable ones are called PAC IIs. These bonds are divided sequentially into PACs with various average lives. These sequential PACs are divided pro rata in order to strip down their coupons so that the bonds trade at or below par. The remaining IO strips are sold individually or together as IOettes. The structuring continues with the support piece, which is divided sequentially. The sequential support pieces are split pro rata to create a variety of fixed-rate, floating-rate, and inverse-floating-rate bonds. Using the few simple structures we saw previously, structures with 50 or more classes can be created.

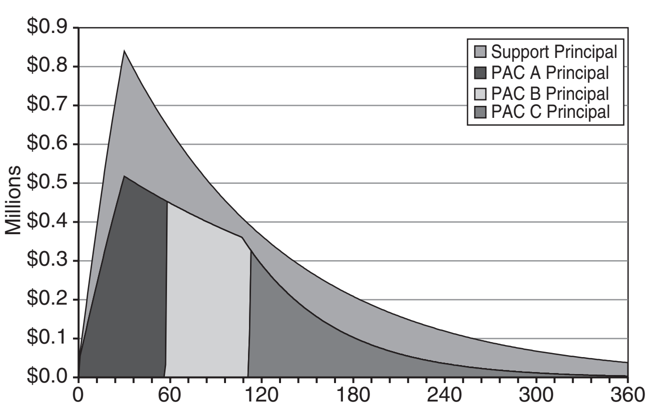


Figure 33.13 Cash Flows of a CMO at 175 PSA

INTEREST ONLY AND PRINCIPAL ONLY

IO and PO bonds can be created within CMO structures as a type of pro rata bond as described earlier. They can also be created independently by stripping MBS. Both FNMA and FHLMC offer programs under which MBS can be split into IOs and POs. IOs and POs tend to be very volatile. By splitting principal and interest, the effects of prepayments on value are amplified.

PO bonds receive all of the principal payments. Therefore, the total amount of cash flow to be received by the PO investor is known from the start. On our \$100 million of collateral, the PO investor will receive \$100 million of cash flow. The uncertainty is over the timing of the cash flows. At faster prepayment speeds, the cash flow is received over a relatively short period. The cash-flow pattern can vary greatly. Figure 33.14 shows the cash flows of the PO at various prepayment speeds.

Since POs receive no interest, they are priced at a discount. Due to discounting effects, the value of the PO increases as the cash flow is received sooner. That is, other things being equal, you would rather receive your money sooner than later. The value of a PO will then be affected by the discount rate and the timing of prepayments. As interest rates fall, the discount rates fall and prepayment rates increase. Both factors serve to increase the value of the PO. POs thus become very bullish instruments, strongly benefiting from falling rates. The performance characteristics of POs, however, are very dependent on the characteristics of the underlying collateral. POs from premium collateral have very different performance characteristics than POs from discount collateral due to their very different prepayment characteristics. Furthermore, POs are very volatile instruments because slight changes in prepayment rates can have a significant impact on value.

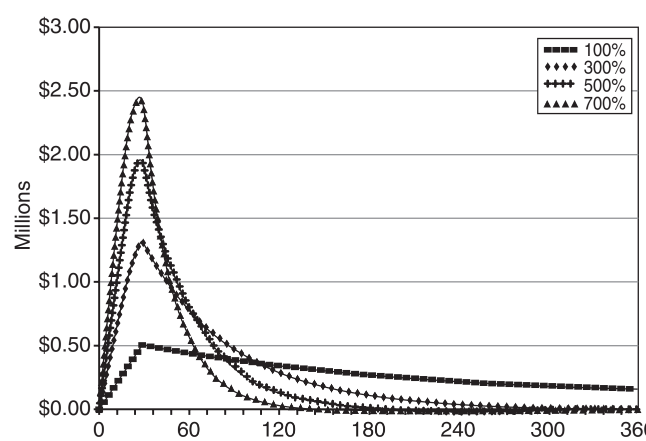


Figure 33.14 Cash Flows of a PO at Various Prepayment Speeds

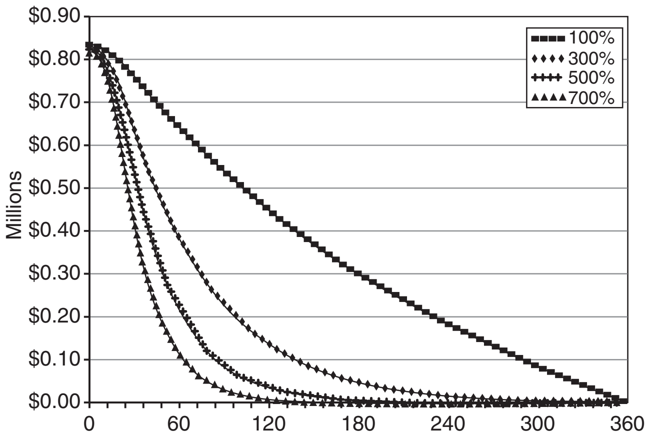


Figure 33.15 Cash Flows of an IO at Various Prepayment Speeds

IO securities have no assured cash flows. The amount of interest received depends on the balance outstanding. As prepayments increase, the amount of cash flow received by the IO decreases. As prepayments decrease, the amount of cash flow increases. The change in cash flow can be significant. Figure 33.15 shows the cash flow of the IO under several prepayment assumptions. Cash flow under the 700%PSA assumption is just a fraction of the cash flow under the 100%PSA assumption.

IO securities have a feature that is unique in the fixed-income world. IOs tend to increase in value as rates rise and decrease in value as rates fall, because as rates rise, prepayments tend to slowdown. Slower prepayments lead to greater cash flows to the IO investor. The increase in cash flow more than compensates for the higher discount rate.

Investors should be cautious about using IOs to hedge other fixed income instruments. Although the general direction of movement of an IO is opposite to other fixed-income instruments, it is difficult to assess precise hedge ratios. IOs are extremely sensitive to prepayment expectations, and it is difficult to establish precise relationships between interest rates and prepayment rates. Highly sensitive instruments such as IOs and POs require sophisticated analysis tempered with a good deal of judgment. The difficulty in assessing these types of instruments is an indication of the limitation of current valuation tools.

SENIOR/SUBORDINATED STRUCTURES

So far, we have concentrated on CMO structures in which the underlying collateral or the CMO itself is guaranteed by the agencies (GNMA, FNMA, and FHLMC). For collateral that does not meet the agency standards, a different type of structure is required. With agency collateral, the investor bears no default risk. In the case of a default, the investor receives a prepayment equal to the full principal amount of the loan. If the loan is not guaranteed by the agencies, other forms of credit enhancement are required to attract investors. Credit enhancement can be either external or internal. External credit enhancement is provided by a mortgage insurance company. The insurance operates at the pool level and provides investors with the assurance that they will not suffer from mortgage delinquencies and defaults.

Internal credit enhancement operates by relying on the overall credit quality of the mortgages to produce different classes of bonds with different exposure to credit loss. Generally, a senior class is produced, which is protected from credit losses, along with a junior piece, which absorbs the losses. Senior/subordinated structures are somewhat akin to PAC bonds. However, instead of offering protection against prepayment risk, the senior class is protected from default risk.

The construction of senior/subordinated deals can become quite complex. New structures are continually being developed to make the execution more efficient. In some structures, a junior class is set up so that it is large enough to absorb worst-case losses. The guidelines for the size of the subordinated structures are set by the rating agencies (Standard & Poor's, Moody's, or Fitch). Underwriters and issuers generally seek at least a double A rating on the senior class. The junior piece generally remains outstanding until the balance of the senior piece has declined sufficiently, so that the risk of loss is minimal. Additional credit protection may come from a reserve account funded with cash or with excess interest that is not going to either the senior or junior class.

The senior/subordinated structures may be further structured using any of the tools described earlier. These CMOs tend to have fewer classes than the agency-backed CMOs.

SUMMARY

CMOs allow mortgage cash flows to be restructured to create securities with a wide variety of investment performance characteristics. Complex CMOs are generally the result of application of relatively simple cash-flow allocation rules. While the rules may be simple, the resulting securities may be quite complex. Knowing the structure of a CMO may provide some insight into the risks of the bond. However, analysis of the actual cash flows under a variety of interest-rate and prepayment scenarios will produce more reliable results. Very often, the performance characteristics of two same-type bonds can differ dramatically.

REFERENCES

Ambrose, B., Lacour-Little, M., and Sanders, A. B. (2004). The effect of conforming loan status on mortgage yield spreads: A loan level analysis. *Real Estate Economics* 32, 4: 541–569.

An, X., Deng, Y., and Sanders, A. B. (2007). Subordination levels in structured financing. (2007). In A. Boot and A. Thakor (eds.), *Handbook of Financial Intermediation*. North Holland, forthcoming.

Bruskin, E., Sanders, A. B., and Sykes, D. (2000). The nonagency mortgage market: Background and overview. In F. Fabozzi, C. Ramsey, and M. Marz (eds.), *The Handbook of Nonagency Mortgage-Backed Securities* (pp. 5–38). Hoboken, NJ: John Wiley & Sons.

Buser, S., Hendershott, P., and Sanders, A. B. (1990). On the determinants of the value of call options on default-free bonds. *Journal of Business* 63, 1: 33–50.

Davidson, A., Herskovitz, M., and Van Drunen, L. (1988). The refinancing threshold pricing model: An economic approach to valuing MBS. *Journal of Real Estate Finance and Economics* 1, 2: 117–130.

Davidson, A., Ho, T., and Lim, Y. (1994). *Collateralized Mortgage Obligations: Analysis, Valuation and Portfolio Strategy*. New York: McGraw-Hill.

Davidson, A., Sanders, A. B., Wolff, L., and Ching, A. (2003). *Securitization: Structuring and Investment Analysis*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F., Sanders, A. B., Yuen, D., and Ramsey, C. (2005). Nonagency CMOs. In F. Fabozzi (ed.), *The Handbook of Fixed Income Securities* (pp. 579–588). New York: McGraw-Hill.

McConnell, J., and Singh, M. (1994). Rational prepayments and the valuation of collateralized mortgage obligations. *Journal of Finance* 49, 3: 891–921.

Mohebbi, C, Li, G., and White, T. (2006). Stripped mortgage-backed securities. In F. Fabozzi (ed.), *The Handbook of Mortgage-Backed Securities* (pp. 465–480). New York: McGraw-Hill.

Richard, S., and Roll, R. (1989). Prepayments on fixed-rate mortgage-backed securities. *Journal of Portfolio Management* 15, 3: 73–82.

Rothberg, J., Nothaft, F., and S. Gabriel. (1989). On the determinants of yield spreads between mortgage passthrough and Treasury securities. *Journal of Real Estate Finance and Economics* 2, 4: 301–315.

CHAPTER 34

Commercial Mortgage-Backed Securities

JAMES MANZI, CFA
Consultant

DIANA BEREZINA
Associate Director, Fitch Ratings

MARK ADELSON
Principal, Adelson & Jacob Consulting, LLC

Abstract: Commercial mortgage-backed securities (CMBSs) are a type of fixed income investment backed by commercial loans. They can be appealing to investors because they generally offer high credit quality, a reasonable degree of credit stability, cash-flow stability, and low-spread volatility. Choosing a particular CMBS investment involves careful consideration of the characteristics of the underlying commercial loans, bond structure, risk appetite, and typical deal features.

Keywords: commercial mortgage-backed securities (CMBSs), commercial loans, CMBS bond structure, CMBS deal features, prepayment penalties, defeasance, yield maintenance

This chapter is intended to provide an overview of the *commercial-mortgage backed securities (CMBSs)* market and the tools to choose between different CMBS investments based on relative value considerations, risk characteristics, and nuances in bond structure.

INVESTMENT CHARACTERISTICS OF TRADITIONAL (CASH) CMBS

CMBSs are backed by commercial mortgage loans. That is, the underlying mortgage loans are secured by commercial, rather than residential, real estate. In contrast to residential mortgage loans, most commercial mortgage loans in the United States do not allow for unrestricted prepayments by the borrowers. Accordingly, a key distinction between CMBS and residential mortgage-backed securities (RMBSs) is that CMBS embody little or no prepayment risk.

A typical CMBS is a "pass-through" security that represents partial ownership of an underlying pool of commercial mortgage loans. An investor who owns the CMBS is entitled to receive collections of interest and principal on the loans. In CMBS jargon, the payments on the loans are "passed through" to the investors. However, a small portion of the interest collections is not passed through. Instead, it is used to cover expenses of the deal. Thus, a CMBS has a "pass-through rate," which is the *net* rate at which investors receive interest on the balance of the mortgage loans backing the security.

Unlike Treasury securities or regular corporate bonds, a typical CMBS is an amortizing security. That is, a typical CMBS returns principal to investors incrementally during its life. Monthly distributions to investors ordinarily include both interest and principal. Accordingly, for purposes of pricing CMBSs, market participants frequently use a security's weighted-average life (WAL) instead of its final maturity.

Every CMBS has at least one servicer. A servicer is a company that collects payments from borrowers and handles the administrative task of aggregating the collected funds and transmitting the funds to a deal's trustee for distribution to investors. Naturally, the servicer receives a fee for its services. In most CMBSs, the fees to the servicer consume all or nearly all of the difference between the interest rate on the mortgage loans and the pass-through rate on the security. Many CMBSs have more than one servicer. In such a case, a primary servicer handles routine servicing functions and a "special servicer" takes over on loans that become seriously delinquent.

Unlike most residential MBSs, most CMBSs do not carry guarantees from the U.S. government or from government-sponsored enerprises (GSEs). Accordingly, a typical CMBS transaction uses "credit tranching" as a form of credit enhancement to counterbalance the risk of defaults and losses on the underlying loans. Credit tranching creates multiple classes (or tranches) of securities, each of which has a different seniority relative to the others. Senior classes receive protection from junior classes that bear amplified exposure to credit risk. Rating agencies measure the credit strength of a transaction's different tranches and assign ratings accordingly.

Finally, a major distinction between CMBS and RMBS deals is the role of the buyers of the junior (non-investment-grade) bond classes. In a CMBS, no deal is done without first finding the buyers for the junior classes. The potential buyers first review the proposed pool, and may exclude loans that they do not like. This provides an extra layer of security for the senior buyers, particularly because the buyers of the junior classes tend to be real estate experts. For their extra credit work, the junior buyers generally seek yields in the range of 10% to 15%, or sometimes more, depending on the quality of the underlying collateral.

BASIC COMMERCIAL LOAN AND CMBS DEAL STRUCTURES

Typical Loan Features

The most typical commercial mortgage is a nonrecourse, fixed-rate loan with a 7- to 10-year balloon payment, although shorter maturity loans, such as 5-year balloons, have gained in popularity as well. A typical loan provides for partial amortization prior to the balloon date on a schedule corresponding to full amortization over a period of 25 to 30 years. However, in periods of increased competition among lenders, it is not uncommon to see loans offered with interest-only (IO) periods, sometimes for the entire term of the loan (to the balloon date). For example, almost three-fourths of the loans securitized to create CMBS during 2006 had at least a partial IO period. Since these loans do not amortize during the IO period, their inclusion in CMBS pools increases the risk that a borrower will not be able to make the balloon payment. Negatively amortizing loans are rare (except with construction loans). Additionally, there have been loans where the amortization rate is accelerated as well as some loans with payment schedules designed to match lease payments.

Cash Flow

Cash flows arising from a typical commercial mortgage include monthly interest, principal, and possibly *prepayment penalties*, and default or extension penalties. CMBS deals have mechanisms to allocate all such cash flows to the respective bond classes. Loans in CMBS deals have ranged in size from just about $1 million to several hundred million dollars. As for the deals themselves, almost all recently issued, fixed-rate deals have combined large loans and small loans and are typically referred to as "conduit/fusion deals." Transaction sizes have varied between a few hundred million dollars to several billion dollars, and generally include several hundred separate loans/properties.

Most commercial mortgages are fixed rate, although there are floating-rate mortgages as well. Generally, fixed- and floating-rate mortgages are not mixed in the same pool. To the extent that there is a great disparity in the interest rates among the loans in a pool, the weighted-average coupon (WAC) can vary considerably over time. The difference in coupon at the inception of the deal can arise due to the loans having been originated over time as interest rates have changed, or due to varying degrees of risk of the loans. Over the life of the deal, even more dispersion can occur. This is true even if the amortization or principal occurs as expected. It gets more complicated if there are unanticipated principal paydowns due to either prepayments, defaults, or extensions.

Table 34.1 Typical Conduit/Fusion Deal (JPMCC 2006-LDP6)

Class	Size ($MM)	Rating	Credit Support (%)	Average Life (yrs)	Coupon	Principal Window (Months from Issue)	Notes
A1	59.9	Aaa	30.00	3.01	5.160	1–58	Super-duper senior
A2	155.9	Aaa	30.00	4.90	5.379	59–61	Super-duper senior
A3B	55.7	Aaa	30.00	7.14	5.559	81–96	Super-duper senior
A3FL	100.0	Aaa	30.00	6.71	L+16	81–81	Floating rate
A4	819.3	Aaa	30.00	9.85	5.475	110–120	Super-duper senior
ASB	103.7	Aaa	30.00	7.08	5.490	58–110	Amortization bond
A1A	205.0	Aaa	30.00	8.28	5.471	1–120	Multifamily carve-out (FNMA and Freddie Mac only)
AM	214.2	Aaa	20.00	9.96	5.525	120–120	Mezzanine super senior
AJ	163.3	Aaa	12.38	9.96	5.565	120–120	Junior triple-A
B	48.2	Aa2	10.13	10.02	5.702	> 120	
C	18.7	Aa3	9.25	10.04	5.722	> 120	
D	34.8	A2	7.63	10.04	5.775	> 120	
E	21.4	A3	6.63	10.04	5.775	> 120	
F	29.5	Baa1	5.25	10.04	5.775	> 120	
G	21.4	Baa2	4.25	10.04	5.775	> 120	
H	21.4	Baa3	3.25	10.04	5.775	> 120	
J	10.7	Ba1	2.75	10.04	5.155	> 120	"B" piece
K	10.7	Ba2	2.25	10.04	5.155	> 120	
L	5.4	Ba3	2.00	10.04	5.155	> 120	
M	5.4	B1	1.75	10.04	5.155	> 120	
N	5.4	B2	1.50	10.04	5.155	> 120	
P	8.0	B3	1.13	10.11	5.155	> 120	
NR	24.1	NR	0.00	12.53	5.155	> 120	
X1 (Comp)	2,142.1*	Aaa	N/A	8.58	0.040		IO class
X2 (PAC)	2,096.7*	Aaa	N/A	5.53	0.255		IO class

*Signifies notional balance.

CMBS Bond Structure

CMBS use subordination for credit enhancement. Table 34.1 shows the typical structure of a conduit/fusion transaction.

The senior-subordinate structure creates senior and junior interests in the underlying asset pool. In Table 34.1 each tranche, from AM down to NR provides protection to (that is, is subordinate to) each other tranche listed above it and receives protection from (that is, is senior to) each other tranche listed below it. The structure requires that all principal payments, both scheduled and from recoveries on defaulted loans, be used to retire the most senior outstanding bond. In addition, the structure allocates losses to the most junior outstanding classes. In a typical deal, most of the bond classes have fixed coupons, but the IO classes (X1 and X2) and a few of the bonds near the bottom of the capital structure are WAC bonds. Essentially, a WAC bond pays a varying coupon over time, based on the weighted-average interest rate of the loans in the pool. As the balances on the loans change, their relative weightings change, which results in a changing coupon on the WAC bond. The creation of WAC bonds generally is necessary (for a greater number/amount of bonds on the capital structure) when rates increase sharply during the loan aggregation phase of a CMBS deal.

Super-Senior and Super-Duper Senior Bonds

Due to the generally positive credit performance of CMBSs dating from the early 1990s to the mid-2000s, the rating agencies steadily lowered subordination levels. However, some investors believed that the credit enhancement levels had dropped too low. In response to investor worries about falling subordination levels in CMBS conduit/fusion deals, dealers started to break up the triple-A-rated class into super-senior, "mezzanine," and "junior" parts. In the structure shown in Table 34.1, classes A1, A2, A3B, A3FL, A4, ASB, and A1A have 30% credit support from subordination and are called "super-duper seniors." Class AM is the mezzanine triple-A-rated tranche. It has 20% credit enhancement from subordination and is called the "super-senior" class. Class AJ is the "junior," or "mezzanine," triple-A-rated tranche and has approximately 12.4% credit enhancement. With this structure, investors who are worried about future CMBS credit performance (e.g., the effects of another prolonged real estate recession like that of the early 1990s) can buy bonds with more protection. Those who are comfortable with the current triple-A subordination levels and the added extension risk (relative to the super-senior AM and the super-duper senior class), can invest in the mezzanine or junior triple-A classes, and receive a small amount of incremental compensation in the form of a slightly wider spread on their bonds.

CMBS Relative Value

Drivers of CMBS Spreads

CMBS spread levels are highly dependent on the balance of supply and demand within the sector, as well as spreads on competing investments, such as residential MBS and corporate bonds. Bonds from seasoned deals may trade with varying spreads depending on the credit performance of the underlying collateral. Information on the CMBS market has become readily available to investors, contributing greatly to the market's strong liquidity. Organizations such as Trepp, Intex, and Bloomberg collect extensive data on the secondary market while publications such as *Commercial Mortgage Alert* provide updates on issuance as well as information on the pipeline of upcoming deals (as far as three months ahead). All other things being equal, when the market is flooded with new deals or when the pipeline projects heavy issuance, spreads tend to widen. However, at times when there is a lull in issuance and investors do not have many deals to choose from, spreads generally tighten. Movements in spreads of other investment products, namely corporate bonds and RMBSs, also tend to drive CMBS spreads. This is true predominantly for the top of the capital structure (classes rated single-A or higher) as investors may cross over between these products in search of higher yield.

Deal Selection

With similar vintage deals generally pricing within a few basis points of each other at the top of the capital structure and offering similar credit enhancement levels, how do investors decide which deal or tranche to buy? Thanks to the increasing transparency of the CMBS market, investors who are willing to carefully scrutinize deals have access to extensive information, such as rating agency presale reports. Naturally, investors who buy riskier tranches (that is, lower in the capital structure) typically spend far more time analyzing a collateral pool than those who invest in the senior triple-A-rated tranches that have 30% credit support. Some deal features to focus on are reviewed below.

Stressed Loan-to-Value Ratio and Debt Service Coverage Ratio Each rating agency calculates a "stressed" loan-to-value ratio (LTV) and stressed debt service coverage ratio (DSCR) for each deal that it rates using its own definition of sustainable cash flow and consistent capitalization rates. This is useful in discerning how aggressive or conservative the loan originator's underwriting may have been even if the reported (underwritten) LTVs are all around 70% to 75%, and the reported DSCR is around 1.3 to 1.5 times.

Interest-Only Loans An IO loan presents greater risk than an amortizing loan because the full original principal amount of the loan is due on the balloon date—there is no amortization during the life of the loan to reduce the balloon risk. In a deal, the greater the share of IO loans, the greater the exposure to balloon risk.

Top 10 Many investors are weary of "lumpy" deals, where several big loans make up a significant portion of the pool. The inherent risk of such deals is that the deterioration of just a few large loans can jeopardize the entire deal.

Property Mix As the real estate market undergoes cycles, different property types tend to encounter difficult periods. For example, during the housing boom of the mid-2000s, the multifamily sector suffered high delinquency rates and their respective portion of CMBS collateral pools declined. Market size can also be important, as deals with concentrations of collateral in economically booming large metropolitan statistical areas (MSAs) are likely to perform better than those backed by properties in economically troubled secondary or tertiary markets.

RISK CHARACTERISTICS OF COMMERCIAL REAL ESTATE COLLATERAL

Property Types

Since the inception of the market in the early mid-1990s, the "big 5" property types—office, retail, multifamily, industrial, and hotels—made up an overwhelming majority of CMBS pools on a weighted-average basis. While the performance of the individual property types is interrelated and subject to "macro" risk variables such as interest rates and inflation, each sector also has certain idiosyncratic risks. For example, job growth, outsourcing, and the growth of "telecommuting" are factors usually associated with office properties, while high levels of consumer spending and disposable income, in part due to home price appreciation, were linked with strength in the retail and lodging sectors during the early mid-2000s. Typically, to judge the health of a particular market and property type combination, CMBS professionals look at recent trends in supply (completions/construction), demand (absorption), effective rent (property income), and vacancy (occupancy).

In addition, trends in capitalization rates, and the recent levels of sales/transactions (both number and dollar value) may be used in conjunction with the previously mentioned "fundamentals" to measure property values and whether they are forecasted to appreciate or depreciate.

Diversification

One of the main advantages to investing in conduit/fusion CMBSs is the diversity of the collateral pools. As noted above, pools backing an average deal may contain several hundred individual loans, some of which may be multiproperty loans. Additionally, there are typically seven or eight different property types represented in a transaction, though office, multifamily, and retail tend to make up the lion's share of the collateral. While this "concentration" may seem worrisome at first, several rating agency loss and default studies have shown that loss severities in the core property types (multifamily, retail, industrial, and office) tend to be lower than those of noncore property types. In addition, there are several other types and measures of diversity related to CMBS transactions.

Geography

Diversification by geography protects against downturns in specific real estate markets, and the fact that different markets may be at different phases of their real estate cycles at the same time. A typical CMBS transaction contains loans from most of the 50 states, with a large majority of the collateral located in the top 50 MSAs by population. Thus, it should come as no surprise that roughly 25% of the domestic CMBS market is made up of loans secured by properties in New York and California.

While this apparent lack of diversity may seem to be a cause for concern, as with the concentration in a limited number of property types, based on historical performance, it is also likely a benefit to investors. Rating agency studies have shown that defaulted loans in smaller secondary and tertiary markets generally experience higher loss severities than defaulted loans in larger, primary markets. We believe that the amount of available land on which to build is one of the primary reasons for this phenomenon.

Measurement and Analysis of Diversification

In addition to geographic and property type diversity, the rating agencies (and the rest of the market) typically utilize Moody's Herfindahl score and the top 10 percentage to measure the concentration risk (by loan balance) in a collateral pool. Moody's Herfindahl score measures a pool's lumpiness and is calculated as:

$$\text{Herfindahl score} = 1/\sum_{i=1}^{n}(p_i/P)^2$$

where n is the number of assets, p_i is the principal balance of each asset, and P is the aggregate principal balance. A credit-neutral score is 100, while scores, on average, have ranged between 40 and 140 over the late 1990s to mid-2000s. Concentration measures are of particular interest to buyers at the lower end of the capital structure, as lower-rated tranches in transactions with lumpy collateral are more exposed to the default of a relatively smaller number of large loans.

Prepayments

One crucial difference between residential and commercial mortgages, which tends to give CMBSs better cash flow stability and positive convexity, is strict rules regarding prepayments. Most commercial mortgages prohibit or severely limit voluntary prepayments through a lockout period, combined with *defeasance* and *yield maintenance* provisions. In other words, if the loan does not contain a lockout feature, it will likely have features designed to discourage the borrower from making a prepayment and to compensate the lender in the event of a prepayment. Often, the prohibitions are for the majority of the life of a loan, with a small open period lasting three to six months before the balloon date. The short open period is designed to give the borrower the opportunity to refinance the balloon. By far, the most common feature discouraging prepayment in most conduit/fusion loans is defeasance. A small percentage of yield maintenance provisions are present in the remainder. The two prepayment deterrents contain important differences regarding the amount and timing of the cash flows.

Defeasance

Under the defeasance approach, the borrower is required to purchase U.S. Treasury securities whose cash flows equal or exceed the remaining payments of the mortgage loan. In this case, the cash flow to securities backing the loan remains identical to what it would have been without the defeasance. As stated before, defeasance has become the most popular form of prepayment protection because of the virtual elimination of credit risk that it affords the investor, as well as the simplicity of the structure. Also attractive to lenders is the fact that the cost of defeasance to the borrower is, on average, quite high—which strongly discourages prepayments.

Yield Maintenance

The concept of yield maintenance is to make the lender indifferent to the prepayment of a loan with a cash premium equal to the future value of the loan's cash flows. Unlike defeasance, yield maintenance provisions require a one-time, lump-sum cash payment, rather than replication of the cash flows of the mortgage loan. As such, issues may arise over the correct discount rate to be applied in calculating the lump-sum payment. Additionally, each deal's structure dictates the allocation rules regarding the penalty cash flows. The allocation of these penalties among the bond classes can vary considerably between deals. The specifics of the allocation can have a significant impact on the performance of the different bond classes. In a multiclass deal, while the penalty still serves as a deterrent to the borrower prepaying, it may not be sufficient to fully compensate all of the bonds within the transaction.

STRUCTURAL NUANCES OF CMBSs

ARD Loans

In the mid-1990s a variation to the balloon loan was developed, known as an ARD (anticipated repayment date) loan. The ARD is similar in many respects to a balloon date, except for one important difference. Failure to fully repay principal on a loan's balloon date is an event of default. In contrast, failure to retire a loan on its ARD is not an event of default. The borrower could keep on paying scheduled principal and interest after the ARD. However, in order to motivate the borrower to pay off the loan on the ARD, the loan's interest rate would rise sharply and all excess cash flow (above the debt service, insurance, taxes, funding of reserves, etc.) would be applied to pay down principal (a situation known as "hyperamortization"). From a credit perspective, the ARD feature alleviated the pressure caused by the required balloon payment, with some protection against interest-related balloon extension. During the mid-2000s, some of these protections against nonpayment at the ARD had been relaxed, but a full discussion is beyond the scope of this chapter.

Appraisal Reduction

When a loan meets certain criteria for being troubled, such as being delinquent for 120 days or experiencing foreclosure of its collateral (that is, becoming real estate owned or REO), it may trigger an appraisal reduction event within a CMBS transaction. In such a case, the principal balance of the first loss class(es) is written down in anticipation of a future loss, effectively reallocating the interest cash flow to the seniormost tranche. Thus, the senior bondholders are better protected against a scenario where a troubled property undergoes a long, drawn-out, workout before the loan can be fully resolved.

Appraisal reductions are interesting from the standpoint of derivatives. An appraisal reduction is an actual write-down of a security. The use of appraisal reductions in the CMBS sector arguably eliminates the need for implied write-downs as used in the standard documentation of credit default swaps (CDSs) on CMBS.

Servicing and Conflicts of Interest

The responsibilities of the servicer and special servicer in a CMBS deal are as follows. The servicer is responsible for supervising the regular cash flow aspects of the loan. It keeps track of the reserves, the insurance payments, the tax payments, and similar items. The servicer is also responsible for advancing principal and interest through foreclosure of a loan, for as long as it deems the advances recoverable. A loan is moved to the special servicer only when the borrower is in default, imminent default, or in violation of covenants. The special servicer is charged with the responsibility of working out the loan. Ideally, the special servicer can restore the loan to performing status. The special servicer has the authority to take the loan through the foreclosure process and is supposed to be guided by the principle of maximizing the present value of proceeds from the property. Sometimes, however, conflicts of interest can arise because the special servicer is often the owner of the junior (first-loss) classes.

Consider a potential balloon default as an example, where the special servicer can choose between extending the loan or foreclosing and selling the property. From a credit perspective, the senior class usually views an extension as an adverse event (unless there is little or no subordination left) because the real estate collateral could continue to deteriorate and thereby lessen the proceeds at a subsequent foreclosure. However, from a rate-of-return perspective, the senior bondholder could be better off with the extension in a falling interest rate environment. Conversely, an extension in a rising rate environment would negatively impact the performance of the senior bondholder. All else being equal, we think loan extensions are more likely in a rising-rate environment.

In contrast to a senior bondholder, a junior bondholder may prefer an extension. If the property value has deteriorated to the point where foreclosure proceeds would be less than the loan balance (plus unpaid interest), the junior class would surely suffer a loss. In this case the junior bondholder would prefer that the borrower be granted an extension, to keep the loan cash-flowing. If the property value in foreclosure is large enough to fully pay the junior class, however, the junior bondholder would likely align with the senior bondholder to push for foreclosure as quickly as possible. One would not expect the latter situation to occur often because if foreclosure proceeds would be sufficient to permit a full recovery for both senior and junior bondholders, the borrower might do better to sell the property and pay off the loan.

Interest Shortfalls and Recovery of Advances

As noted above, if a securitized commercial mortgage defaults during its term, the servicer is required to advance principal and interest through foreclosure, provided that it deems the advances to be recoverable. This enhances the timeliness of distributions to holders of the securities even when there are interruptions in the inflow of property income. The servicer is compensated with interest on these advances and is first in line to recover advances upon the liquidation of the property. In the case of a prolonged foreclosure, where the servicer continues to advance principal and interest, the proceeds from a sale may not be enough to fully compensate the servicer for his advances, plus interest. The result can be an interest shortfall to the bonds in a deal. Interest shortfalls on subordinate tranches are reasonably common. Interest shortfalls have occurred on CMBS tranches rated triple-A, though such events are very rare.

A/B Notes and Mezzanine Loans

A CMBS loan may be divided into senior and junior interests. Figure 34.1 illustrates a generic \$100 million property financed with (1) a \$60 million dollar investment grade

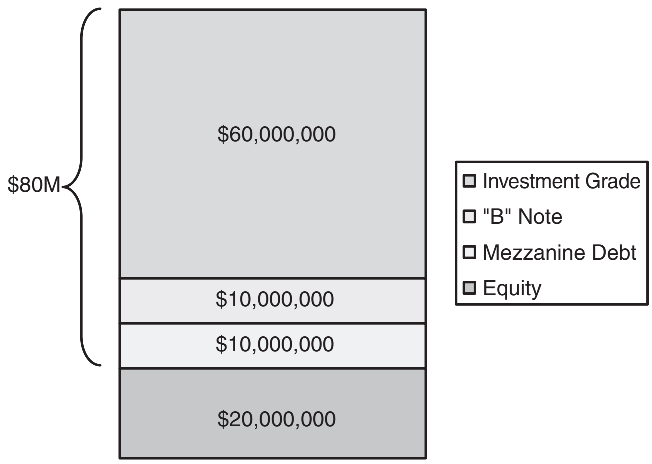


Figure 34.1 A-B Note Structures ($100 Million Total)

"A-note," (2) a $10 million "B-note," (3) $10 million in mezzanine debt, and (4) $20 million in hard equity. Typically, only the A-note will be included within a conduit/fusion deal. The B-note ordinarily would not be included in a CMBS trust. However, we have seen cases where both an A-note and its related B-note are securitized within the same CMBS transaction. During the mid-2000s, B-notes and mezzanine debt were popular collateral types for inclusion within commercial real estate (CRE) collateralized debt obligations (CDOs).

In the event of default, the B-note holder's right to principal and interest payments is subordinate to the rights of the holder of the A-note. Also, the A-note holder will generally have greater, if not exclusive, control over any bankruptcy proceedings dealing with the workout of a troubled loan. Typically, for the B-note holder to obtain greater control over the workout of a troubled loan, he will have to exercise the option usually granted the B-note holder of buying out the A-note holder's participation in the loan, at par plus accrued interest. Other important rights that the B-note holder may have include:

- The right to hire and fire the special servicer.
- The right to cure defaults in order to keep the senior lender from foreclosing (usually, there is a limit to the number of times this right can be exercised).
- Approval rights associated with the property budgets, leases, and property managers.

A mezzanine loan is not secured by a lien on the related property. Instead it is a pledge of stock in the special purpose entity that owns the property and that is the borrower on the A-note and on the B-note. In effect, the mezzanine lender is subordinate to the first mortgage (the A-note and the B-note) and senior to the hard equity holder. If the mezzanine loan gets into trouble, the holder cannot foreclose on the property directly. Rather, the mezzanine debt holder can foreclose on the equity interest of the first mortgage borrower, in effect taking over the borrowing entity, and therefore controlling the property in question.

Additional Debt

A common feature of older loans is a prohibition on additional debt against the real estate after the inception of the loan. This is important because adding extra debt can immediately raise the leverage against the property and increase the debt service burden. Many recent commercial mortgage loans permit borrowers to take additional debt. It is now quite common to see CMBS deals in which as much as 40% (or more!) of the underlying loan pool (by principal balance) allows additional debt. Although most such loans require mitigating factors—the loan must meet certain tests, such as maintaining a specified combined DSCR and combined LTV—this is still a worrisome trend.

SUMMARY

In this chapter, we covered general characteristics of CMBSs and the *commercial loans* underlying the transactions, as well as some of the nuances that an experienced CMBS market participant would consider before making an investment. It is clear that choosing a particular CMBS investment involves careful evaluation of the characteristics of the underlying commercial loans, the structure of the deal, the investor's risk appetite, and relative value, among many other important considerations!

REFERENCES

DeMichele, J. F., Adams, W. J., and Hewlett, D. C. (2002). Commercial mortgage-backed securities. In F. J. Fabozzi (ed.), *The Handbook of Financial Instruments* (pp. 399–422), Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (ed.) (2001). *Investing in Commercial Mortgage-Backed Securities*. New York: John Wiley & Sons.

Fabozzi, F. J., and Jacob, D. P. (eds.) (1999). *The Handbook of Commercial Mortgage-Backed Securities*, 2nd edition. New York: John Wiley & Sons.

Jacob, D. P., and Fabozzi, F. J. (2003). The impact of structuring on CMBS bond class performance. *Journal of Portfolio Management* (Special Real Estate Issue): 76–86.

Leffler, P., Malysa, J., Story, J., and Merrick, S. S. (2006). Cash-flow CDOs for CMBS investors. In F. J. Fabozzi (ed.), *The Handbook of Mortgage Backed Securities* (pp. 1209–1216), New York: McGraw Hill.

Obazee, P. O., and Hewlett, D. C. (2006a). CMBS collateral performance: Measures and valuations. In F. J. Fabozzi (ed.), *The Handbook of Mortgage Backed Securities* (pp. 1187–1198), New York: McGraw Hill.

Obazee, P. O. and Hewlett, D. C. (2006b). Value and sensitivity of CMBS IOs. In F. J. Fabozzi (ed.), *The Handbook of Mortgage Backed Securities* (pp. 1199–1208). New York: McGraw Hill.

Sanders, A. B. (2005). Commercial mortgage-backed securities. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Securities* (pp. 615–628). New York: McGraw Hill.

CHAPTER 35

Nonmortgage Asset-Backed Securities

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

LAURIE S. GOODMAN, PhD
Co-head of Global Fixed Income Research and Manager of U.S. Securitized Products Research, UBS

DOUGLAS J. LUCAS
Executive Director and Head of CDO Research, UBS

Abstract: Asset-backed securities are debt instruments that are backed by a pool of loans or receivables. There is considerable diversity in the types of assets that have been securitized. These assets can be classified as mortgage assets and nonmortgage assets. The former includes residential and commercial mortgage loans. The two largest types of nonmortgage assets that have been securitized are card receivables and auto loan receivables. Investors are attracted to asset-backed securities primarily because of their desirable investment and maturity characteristics.

Keywords: securitization, asset-backed securities (ABSs), credit card receivable-backed securities, monthly mortgage payment, early amortization trigger, auto loan-backed securities, prepayments, absolute prepayment speed, single monthly mortality rate, student loan asset-backed securities (SLABS), alternative loans, Small Business Administration (SBA) loan-backed securities, aircraft lease-backed securities, franchise loan-backed securities, loan-to-value ratio, fixed charge coverage ratio, rate reduction bonds (RRBs), stranded costs, stranded assets, competitive transition charge (CTC)

The process for the creation of *asset-backed securities (ABSs)*, referred to as securitization, is as follows. The owner of assets sells a pool of assets to a bankruptcy remote vehicle called a special purpose entity (SPE). The SPE obtains the proceeds to acquire the asset pool, referred to as the collateral, by issuing debt instruments. The cash flow of the asset pool is used to satisfy the obligations of the debt instruments issued by the SPE. The debt instruments issued by the SPE are referred to generically as *asset-backed securities*, asset-backed notes, asset-backed bonds, or asset-backed obligations.

ABSs issued in a single securitization can have different credit exposure and based on the credit priority, securities are described as senior notes and junior notes (subordinated notes). In the prospectus for a securitization, the securities are actually referred to as certificates: pass-through certificates or pay-through certificates. The distinction between these two types of certificates is the nature of the claim that the investor has on the cash flow generated by the asset pool. If the investor has a direct claim on all of the cash flow and the certificate holder has a proportionate share of the collateral's cash flow, the term "pass-through certificate" (or "beneficial interest certificate") is used. When there are rules that are used to allocate the collateral's cash flow among different classes of investors, the asset-backed securities are referred to as pay-through certificates.

The types of assets that have been securitized are generally classified as traditional assets and nontraditional or emerging assets. Market participants attribute different meaning as to what is meant by nontraditional assets. Some refer to nontraditional assets as assets other than the major types of assets that have been securitized at the time. In the early years of the ABS market, traditional assets included home equity loans, manufactured housing loans, credit card loans, and auto loans. The list of what is viewed as traditional assets has changed as securitization has become a more popular vehicle for issuers to raise funds. Others view nontraditional or emerging assets in a more limited way: those assets that are being securitized for the first time or for which there have been very few securitizations. For example, the recording artists David Bowie, James Brown, the Isley Brothers, and Rod Stewart have securitized their future music royalties, the first being by Bowie, who in 1997 issued $55 million of ABS backed by the current and future revenues of his first 25 music albums (287 songs) recorded prior to 1990. (These bonds, popularly referred to as "Bowie bonds," were purchased by Prudential Insurance Company and had a maturity of 10 years. When the bonds matured in 2007, the royalty rights reverted back to David Bowie.)

Another classification of ABS is based on whether the assets are mortgage-related assets or nonmortgage-related assets. The former includes residential mortgage loans such as home equity loans and manufactured housing; the latter includes a wide range consumer and business loans and receivables, as well as the securitization of whole businesses. In this chapter, we will discuss ABSs for which the collateral is a pool of traditional nonmortgage assets. More specifically, we will describe credit card receivable-backed securities, auto loan-backed securities, student loan-backed securities, Small Business Administration (SBA) loan-backed securities, aircraft lease-backed securities, franchise loan-backed securities, and rate reduction bonds. For a description of the structure of ABS in general, see Chapters 73 and 74 of Volume III.

CREDIT CARD RECEIVABLE-BACKED SECURITIES

A major sector of the ABS market is that of securities backed by credit card receivables. Credit cards are issued by banks (e.g., Visa and MasterCard), retailers (e.g., JCPenney and Sears), and travel and entertainment companies (e.g., American Express). Credit card deals are structured as either a discrete trust or master trust. With a master trust the issuer can sell several series from the same trust.

Cash Flow

For a pool of credit card receivables, the cash flow consists of finance charges collected, fees, and principal. Finance charges collected represent the periodic interest the credit card borrower is charged based on the unpaid balance after the grace period. Fees include late payment fees and any annual membership fees.

Interest to security holders is paid periodically (e.g., monthly, quarterly, or semiannually). The interest rate may be fixed or floating. The floating rate is uncapped.

A *credit card receivable-backed security* is a nonamortizing security. For a specified period of time, referred to as the lockout period or revolving period, the principal payments made by credit card borrowers comprising the pool are retained by the trustee and reinvested in additional receivables to maintain the size of the pool. The lockout period can vary from 18 months to 10 years. So, during the lockout period, the cash flow that is paid out to security holders is based on finance charges collected and fees.

After the lockout period, the principal is no longer reinvested but paid to investors. This period is referred to as the principal-amortization period, and the various types of structures are described later.

Performance of the Portfolio of Receivables

Several concepts must be understood in order to assess the performance of the portfolio of receivables and the ability of the issuer to meet its interest obligation and repay principal as scheduled.

The gross yield includes finance charges collected and fees. Charge offs represent the accounts charged off as uncollectible. Net portfolio yield is equal to gross portfolio yield minus charge-offs. The net portfolio yield is important because it is from this yield that the bondholders will be paid. So, for example, if the average yield (WAC) that must be paid to the various tranches in the structure is 5% and the net portfolio yield for the month is only 4.5%, there is the risk that the bondholder obligations will not be satisfied.

Delinquencies are the percentages of receivables that are past due for a specified number of months, usually 30, 60, and 90 days. They are considered an indicator of potential future charge-offs.

The monthly payment rate (MPR) expresses the monthly payment (which includes finance charges, fees, and any principal repayment) of a credit card receivable portfolio as a percentage of credit card debt outstanding in the previous month. For example, suppose a $500 million credit card receivable portfolio in January realized $50 million of payments in February. The MPR would then be 10% ($50 million divided by $500 million).

There are two reasons why the MPR is important. First, if the MPR reaches an extremely low level, there is a chance that there will be extension risk with respect to the principal payments on the bonds. Second, if the MPR is very low, then there is a chance that there will not be sufficient cash flows to pay off principal. This is one of the events that could trigger early amortization of the principal (described as follows).

At issuance, portfolio yield, charge-offs, delinquency, and MPR information are provided in the prospectus. Information about portfolio performance is thereafter available from various sources.

EARLY AMORTIZATION TRIGGERS

There are provisions in credit card receivable-backed securities that require early amortization of the principal if certain events occur. Such provisions, which are referred to as either early amortization or rapid amortization, are included to safeguard the credit quality of the issue. The only way that principal cash flows can be altered is by triggering the early amortization provision.

Typically, early amortization allows for the rapid return of principal in the event that the three-month average excess spread earned on the receivables falls to zero or less. When early amortization occurs, the credit card tranches are retired sequentially (that is, first the AAA bond, then the AA rated bond, and so on). This is accomplished by paying the principal payments made by the credit card borrowers to the investors instead of using them to purchase more receivables. The length of time until the return of principal is largely a function of the monthly payment rate. For example, suppose that a AAA tranche is 82% of the overall deal. If the monthly payment rate is 11%, then the AAA tranche would return principal over a 7.5-month period (82%/11%). An 18% monthly payment rate would return principal over a 4.5-month period (82%/18%).

Monthly information is available on each deal's trigger formula and base rate. The trigger formula is the formula that shows the condition under which the rapid amortization will be triggered. The base rate is the minimum payment rate that a trust must be able to maintain to avoid early amortization.

AUTO LOAN-BACKED SECURITIES

Auto loan-backed securities are issued by:

1. The financial subsidiaries of auto manufacturers (domestic and foreign).
2. Commercial banks.
3. Independent finance companies and small financial institutions specializing in auto loans.

In terms of credit, borrowers are classified as either prime, nonprime, or subprime. Each originator employs its own criteria for classifying borrowers into these three broad groups. Typically, prime borrowers are those that have had a strong credit history that is characterized by timely payment of all their debt obligations. The FICO score of prime borrowers is generally greater than 680. Nonprime borrowers have usually had a few delinquent payments. Nonprime borrowers, also called near-prime borrowers, typically have a FICO score ranging from the low 600s to the mid-600s. When a borrower has a credit history of missed or major problems with delinquent loan payments and the borrower may have previously filed for bankruptcy, the borrower is classified as subprime. The FICO score for subprime borrowers typically is less than the low 600s (Roever, 2005).

Cash Flows and Prepayments

The cash flow for auto loan-backed securities consists of regularly scheduled monthly loan payments (interest and scheduled principal repayments) and any prepayments. For securities backed by auto loans, prepayments result from (see Roever, 2005):

1. Sales and trade-ins requiring full payoff of the loan.
2. Repossession and subsequent resale of the automobile.
3. Loss or destruction of the vehicle.
4. Payoff of the loan with cash to save on the interest cost.
5. Refinancing of the loan at a lower interest cost.

While refinancings may be a major reason for prepayments of mortgage loans, they are of minor importance for automobile loans. Moreover, the interest rates for the automobile loans underlying some deals are substantially below market rates (subvented rates) since they are offered by manufacturers as part of a sales promotion.

Prepayments for auto loan-backed securities are measured in terms of the *absolute prepayment speed* (ABS). The ABS is the monthly prepayment expressed as a percentage of the original collateral amount. (Note that another measure of prepayments used for other asset classes that have been securitized is the *single monthly mortality rate* (SMM). The SMM is a monthly prepayment rate that expresses prepayments based on the prior month's balance).

Structures

When auto ABS were first issued, the typical structure was a grantor trust that issued passthrough certificates. A major drawback with using grantor trusts in creating efficient structures is the inability to time tranche securities. That is, while an issuer can use a grantor trust to create subordinate interests and thereby issue multiple bond classes, each with a different level of priority, it could not issue multiple bond classes with the same level of priority. Nor are issuers permitted to use interest rate derivatives within a grantor trust. This led to the extensive use of the pay-through structures by issuers. The most common pay-through structure used is the owner trust.

Moreover, because of the flexibility granted to issuers to manage the cash flows from the collateral when using pay-through structures such as the owner trust, issuers could include performance-related triggers. Because of the reduced credit risk resulting from the inclusion of these triggers, issuers could reduce the cost of credit enhancement.

There are two typical structures used in auto ABS pay-through structures. In both structures there are multiple sequential-pay senior classes and a subordinate class. One of the senior classes is a Rule 2a-7 of the Investment Company Act of 1940 eligible money market class. In one typical structure, the senior classes receives all principal until every senior class is paid off. Only after that time is the subordinate class paid any principal. In the other typical structure, once the money market class is paid off, the other senior classes and the subordinate class are paid principal concurrently. However, in this structure, the concurrent payments to the senior classes and subordinate classes require that a performance trigger be reached. If the performance trigger is breached, the principal distribution rules of the second structure will be the same as that for the first structure.

STUDENT LOAN ASSET-BACKED SECURITIES

Student loans are made to cover college cost (undergraduate, graduate, and professional programs such as medical school and law school) and tuition for a wide range of vocational and trade schools. Securities backed by student loans are popularly referred to as *SLABS* (*student loan asset-backed securities*).

The student loans that have been most commonly securitized are those that are made under the Federal Family Education Loan Program (FFELP). Under this program, the government makes loans to students via private lenders. The decision by private lenders to extend a loan to a student is not based on the applicant's ability to repay the loan. If a default of a loan occurs and the loan has been properly serviced, then the government will guarantee 97% of the principal and accrued interest (for loans originated in July 2006 or later).

Loans that are not part of a government guarantee program are called *alternative* or *private loans*. These loans are basically consumer loans, and the lender's decision to extend an alternative loan will be based on the ability of the applicant to repay the loan. Alternative loans are securitized in increasing amounts due to the rising cost of education.

The Student Loan Marketing Association (Sallie Mae) is a major issuer of SLABS, and its issues are viewed as the benchmark issues. Other entities that issue SLABS are either traditional for-profit issuers (e.g., the Key Corp Student Loan Trust) or nonprofit organizations (Michigan Higher Education Loan Authority and the Florida Educational Loan Marketing Corporation). The SLABS of the latter typically are issued as tax-exempt securities and therefore trade in the municipal market.

Collateral

There are different types of student loans under the FFELP, including subsidized and unsubsidized Stafford loans, Parent Loans for Undergraduate Students (PLUS), and Supplemental Loans to Students (SLS). These loans involve several periods with respect to the borrower's payments—deferment period, grace period, and loan repayment period. Typically, student loans work as follows. While a student is in school, no payments are made by the student on the loan. This is the in-school deferment period. Upon leaving school, the student is extended a grace period of usually six months when no payments on the loan must be made. After this period, payments are made on the loan by the borrower (repayment period).

Prepayments typically occur due to defaults or loan consolidation. Even if there is no loss of principal faced by the investor when defaults occur, the investor is still exposed to contraction risk. This is the risk that the investor must reinvest the proceeds at a lower spread and, in the case of a bond purchased at a premium, the premium will be lost. Consolidation of a loan occurs when the student who has taken out loans over several years combines them into a single loan. The proceeds from the consolidation are distributed to the original lender and, in turn, distributed to the bondholders. Loan consolidation allows student borrowers to achieve lower rates and longer terms. Student loan consolidation was very popular during the 2001–2005 period, and lead to prepayment rates during those years that were considerably higher than anticipated when the deals were priced.

Structures

Structures on student loan floaters have experienced more than the usual amount of change since 2000. The reason for this is quite simple.

The underlying collateral—student loans—is exclusively indexed to three-month Treasury bills, while a large percentage of securities are issued as London Interbank Offered Rate (LIBOR) floaters. This creates an inherent mismatch between the collateral and the securities.

Issuers have dealt with the mismatch in a variety of ways. Some issued Treasury bill floaters which eliminates the mismatch, others issued hedged or unhedged LIBOR floaters, while others switched back and forth between the two. More recently, some have issued both Treasury and LIBOR floaters in the same transaction. (Also in conjunction with the choice of index, issuers have incorporated a variety of basis swaps and/or have bought cap protection from third parties, while some have used internal structures to deal with the risk).

It is important to bear in mind that when an ABS structure contains a basis mismatch, it is not only the investor, but the issuer that bears a risk. Student loan deals (like deals in many other ABS classes) have excess spread; that is, roughly the difference between the net coupon on the collateral and the coupon on the bonds.

In mortgage-related ABS, the excess spread is much larger than in the student loan sector, and is used to absorb monthly losses. Because losses in federally guaranteed student loans are relatively small, the vast majority of the excess spread flows back to the issuer. Hence, the Treasury bill/LIBOR-basis risk is of major concern to issuers. When an issuer incorporates a swap in the deal, it not only reduces the risk to the investor (by eliminating the effect of an available funds cap) but reduces risk to the issuer by protecting a level of excess spread. When a cap is purchased, it is primarily for the benefit of the investor, because the cap only comes into play once the excess spread in the deal has been effectively reduced to zero.

The indices used on private and public student loan ABS transactions since the earliest deals in 1993 have changed over time (even though throughout this period, the index on the underlying loans was always three-month Treasury bills). From 1993 to 1995, most issuers, with the notable exception of Sallie Mae, used one-month LIBOR, which indicated strong investor preference for LIBOR floaters. By contrast, from Sallie Mae's first deal in late 1995, that issuer chose to issue Treasury bill floaters to minimize interest rate mismatch risk.

SBA LOAN-BACKED SECURITIES

The Small Business Administration (SBA) is an agency of the U.S. government empowered to guarantee loans made by approved SBA lenders to qualified borrowers. The loans are backed by the full faith and credit of the government. Most SBA loans are variable rate loans where the reference rate is the prime rate. The rate on the loan is reset monthly on the first of the month or quarterly on the first of January, April, July, and October. SBA regulations specify the maximum coupon allowable in the secondary market. Newly originated loans have maturities between five and 25 years.

The Small Business Secondary Market Improvement Act passed in 1984 permitted the pooling of SBA loans. When pooled, the underlying loans must have similar terms and features. The maturities typically used for pooling loans are 7, 10, 15, 20, and 25 years. Loans without caps are not pooled with loans that have caps.

Most variable rate SBA loans make monthly payments consisting of interest and principal repayment. The amount of the monthly payment for an individual loan is determined as follows. Given the coupon formula of the prime rate plus the loan's quoted margin, the interest rate is determined for each loan. Given the interest rate, a level payment amortization schedule is determined. This level payment is paid until the coupon rate is reset.

The monthly cash flow that the investor in an SBA-backed security receives consists of:

- The coupon interest based on the coupon rate set for the period.
- The scheduled principal repayment (that is, scheduled amortization).
- Prepayments.

Prepayments for *SBA loan-backed* securities are measured in terms of the conditional prepayment rate (CPR). Voluntary prepayments can be made by the borrower without any penalty. There are several factors contributing to the prepayment speed of a pool of SBA loans. A factor affecting prepayments is the maturity date of the loan. It has been found that the fastest speeds on SBA loans and pools occur for shorter maturities. The purpose of the loan also affects prepayments. There are loans for working capital purposes and loans to finance real estate construction or acquisition. It has been observed that SBA pools with maturities of 10 years or less made for working capital purposes tend to prepay at the fastest speed. In contrast, loans backed by real estate that have long maturities tend to prepay at a slow speed. All other factors constant, pools that have capped loans tend to prepay more slowly than pools of uncapped loans.

AIRCRAFT LEASE-BACKED SECURITIES

Aircraft financing has gone thorough an evolution over the past several years. It started with mainly bank financing, then moved to equipment trust certificates (ETCs), then to enhanced ETCs (EETCs), and finally to aircraft ABS. Today, both EETCs and *aircraft lease-backed securities* are widely used.

EETCs are corporate bonds that share some of the features of structured products, such as credit tranching and liquidity facilities. Aircraft ABS differ from EETCs in that they are not corporate bonds, and they are backed by leases to a number of airlines instead of being tied to a single airline. The rating of aircraft ABS is based primarily on the cash flow from their pool of aircraft leases or loans and the collateral value of that aircraft, not on the rating of lessee airlines.

One of the major characteristics that set aircraft ABS apart from other forms of aircraft financing is their diversification. ETCs and EETCs finance aircraft from a single airline. An aircraft ABS is usually backed by leases from a number of different airlines, located in a number of different countries and flying a variety of aircraft types. This diversification is a major attraction for investors. In essence, they are investing in a portfolio of airlines and aircraft types rather than a single airline—as in the case of an airline corporate bond. Diversification also is one of the main criteria that rating agencies look for in an aircraft securitization. The greater the diversification, the higher the credit rating, all else being equal.

Aircraft Leasing

Although there are various forms of financing that might appear in an aircraft ABS deal—including operating leases, financing leases, loans or mortgages—to date, the vast majority of the collateral in aircraft deals has been operating leases. In fact, all of the largest deals have been issued by aircraft leasing companies. This does not mean that a diversified finance company or an airline itself might not at some point bring a lease-backed or other aircraft ABS deal. It just means that so far, aircraft ABS have been mainly the province of leasing companies. Airlines, on the other hand, are active issuers of EETCs.

Aircraft leasing differs from general equipment leasing in that the useful life of an aircraft is much longer than most pieces of industrial or commercial equipment. In a typical equipment lease deal, cash flow from a particular lease on a particular piece of equipment only contributes to the ABS deal for the life of the lease. There is no assumption that the lease will be renewed. In aircraft leasing, the equipment usually has an original useful life of 20+ years, but leases run for only around 4 to 5 years. This means that the aircraft will have to be re-leased on expiration of the original leases. Hence, in the rating agencies' review, there is a great deal of focus on risks associated with re-leasing the aircraft.

The risk of being able to put the plane back out on an attractive lease can be broken down into three components: (1) the time it takes to re-lease the craft; (2) the lease rate; and (3) the lease term. Factors that can affect releasing include the general health of the economy, the health of the airline industry, obsolescence, and type of aircraft.

Servicing

Servicing is important in many ABS sectors, but it is crucial in a lease-backed aircraft deal, especially when the craft must be remarketed when their lease terms expire before the term of the aircraft ABS. It is the servicer's responsibility to re-lease the aircraft. To fulfill that function in a timely and efficient manner, the servicer must be both well-established and well-regarded by the industry.

As Moody's states, the servicer "should have a large and diverse presence in the global aircraft marketplace in terms of the number of aircraft controlled. Market share drives the ability of a servicer to meet aircraft market demand and deal with distressed airlines."

The servicer is also the key to maintaining value of the aircraft, through monitoring usage of the craft by lessees. If a lessee is not maintaining an aircraft properly, it is the servicer's responsibility to correct that situation. Because of servicers' vital role to the securitization, the rating agencies spend a great deal of effort ascertaining how well a servicer is likely to perform.

Defaults

In addition to the risk from needing to re-lease craft, rating agencies are also concerned about possible defaults. Because of protections under Section 1110 of the U.S. Bankruptcy Code, and international statutes that favor aircraft creditors, there is relatively little risk of losing an aircraft. There are, however, repossession costs, plus the loss of revenues during the time it takes to repossess and restore the aircraft to generating lease income.

The rating agencies will "stress" an aircraft financing by assuming a default rate and a period of time and cost for repossessing the aircraft. A major input into base default assumptions is the credit rating of airline lessees. For this part of the review, the ABS rating analyst does rely on the corporate rating of the airline.

While there is little risk of not recovering the aircraft in event of a default, the rating agencies do carefully review the legal and political risks that the aircraft may be exposed to, and evaluate the ease with which the aircraft can be repossessed in the event of a default, especially if any of the lessees are in developing countries.

Enhancement Levels

In aircraft ABS, as in every other ABS sector, the rating agencies attempt to set enhancement levels that are consistent across asset types. That is, the risk of not receiving interest or principal in a aircraft deal rated a particular credit level should be the same as in a credit card or home equity deal (or, for that matter, even for a corporate bond) of the same rating. The total enhancement ranges from 34% to 47%.

Since the early deals, there has been a change in enhancement levels. Early deals depended largely on the sale of aircraft to meet principal payments on the bonds. Since then, aircraft ABS has relied more on lease revenue. Because lease revenue is more robust than sales revenue, the enhancement levels have declined. To understand why a "sales" deal requires more enhancement than a "lease" deal, consider the following. If an aircraft is sold during a recession, the deal suffers that entire decline in market value. On the other hand, if a lease rate declines during a recession, the deal sustains only the loss on the re-lease rate.

FRANCHISE LOAN-BACKED SECURITIES

Franchise loan-backed securities are a hybrid between the commercial mortgage-backed securities (CMBS) and ABS markets. They are often backed by real estate, as in CMBS, but the deal structures are more akin to ABS. Also, franchise loans resemble SBA loans and CDOs more than they do consumer loan-backed ABS securities. Greater reliance is placed on examining each franchise loan within the pool than on using aggregate statistics. In a pool of 100 to 200 loans (typical franchise loan group sizing) each loan is significant. By contrast within the consumer sector, any individual loan from a pool of 10,000 loans (as in home equity deals) does not represent as large a percentage, thus is not considered quite as important.

Franchise loans are similar to SBA loans in average size, maturity, and end use. But whereas most SBA loans are floating rate loans indexed to the prime rate, most securitized franchise loans are fixed rate; if they are floating, they are likely to be LIBOR linked. Franchise loans are used to fund working capital, expansion, acquisitions, and renovation of existing franchise facilities.

The typical securitized deal borrower owns a large number of units, as opposed to being a small individual owner of a single franchise unit. However, individual loans are usually made on a single unit, secured either by the real estate, the building, or the equipment in the franchise.

The consolidation within the industry and the emergence of large operators of numerous franchise units has improved industry credit performance. A company owning 10 to 100 units is in a better position to weather a financial setback than is the owner of a single franchise location.

Loans can also be either fixed or floating rate, and are typically closed-end, fully amortizing with maturities of 7 to 20 years. If secured by equipment, maturities range from 7 to 10 years. If they are secured by real estate, maturities usually extend 15 to 20 years. Interest rates range from 8% to 11%, depending on maturity and risk parameters.

Security Characteristics

Because franchise loan collateral is relatively new to the ABS market, and deal size is small, most of these securitized packages have been issued as a 144a private placement (Rule 144a of the Securities Act of 1933 governing private resales of securities to institutions). Issuers also prefer the 144a execution for competitive reasons, because they are reluctant to publicly disclose details of their transactions.

Deals typically range from \$100 to \$300 million, and are customarily backed by 150 to 200 loans. Average loan size is around, \$500,000, while individuals loans may range from \$15,000 to \$2,000,000.

Most deals are structured as sequential-pay bonds with a senior/subordinate credit enhancement. Prepayments can occur if a franchise unit closes or is acquired by another franchisor. However, few prepayments have been experienced within securitized deals as of this writing, and most loans carry steep prepayment penalties that effectively discourage rate refinancing. Those penalties often equal 1% of the original balance of the loan.

Major Sectors

The vast majority of franchise operations consist of three types of retail establishments: restaurants, specialty retail stores (e.g., convenience stores, Blockbuster, 7–11s, Jiffy Lube, and Meineke Muffler), and retail gas stations (e.g., Texaco and Shell). The restaurant category has three major subsectors: quick-service restaurants (e.g., McDonald's, Burger King, Wendy's, and Pizza Hut), casual restaurants (e.g., T.G.I. Fridays, Red Lobster, and Don Pablo's), and family restaurants (e.g., Denny's, Perkins, and Friendly's).

A "concept" is simply another name for a particular franchise idea, since each franchise seeks to differentiate itself from its competitors. Hence, even though Burger King and Wendy's are both quick-service restaurants specializing in sandwiches, their menu and style of service are sufficiently different that each has its own business/marketing plan—or "concept." For example, Wendy's has long promoted the "fresh" market, because the firm mandated fresh (not frozen) beef patties in their hamburgers, and helped pioneer the industry's salad bars. Burger King is noted for its "flame-broiled" burgers, and doing it "your way."

In addition to segmenting the industry by functional types, it is also segmented by credit grades. For example, Fitch developed a credit tiering system based on expected recoveries of defaulted loans. Tier I concepts have a much lower expected default level than Tier II concepts, and so on. Many financial and operational variables go into these tiered ratings, including number of outlets nationwide (larger, successful concepts benefit from better exposure, national advertising, and the like); concept "seasoning" (especially if it has weathered a recession); and viability in today's competitive environment. (Yesterday's darlings may have become oversaturated, or unable to respond to changing tastes or trends by revamping and updating!)

Risk Considerations

There are several risk factors to be aware of when comparing franchise loan pools, and the following are some of the most important.

Number of Loans/Average Size

High concentrations of larger loans represent increased risk, just as in any other pool of securitized loans.

Loan-to-Value Ratio

The *loan-to-value* (LTV) ratio can be based on either real estate or business values. It is important to determine which is being used in a particular deal in order to make a valid comparison with other franchise issues. Note that when business value is used to compute LTV, it is common for a nationally recognized accounting firm to provide the valuation estimate.

Fixed Charge Coverage Ratio

The *fixed charge coverage ratio* (FCCR) is calculated as follows:

$$\text{FCCR} = \frac{\text{Adjusted free cash flow less occupancy costs}}{\text{Occupancy costs plus debt service}}$$

Typical FCCRs range from 1.00 to 3.00, and average around 1.5. A deal with most unit FCCRs below 1.5 would be viewed as having greater risk than average, while one with most FCCRs above 1.5 would be perceived as having less risk than average.

Diversification

As in all ABS sectors, a primary risk factor is the degree of diversification. In a franchise loan deal, important areas for diversification include franchise owner, concept, and location.

A typical franchise pool includes loans to 10 to 15 franchisees, each having taken out loans on 5 to 20 individual units. A large concentration of loans to any single franchise operator might increase deal risk. However, such concentration is sometimes allowed, and rating agencies will not penalize severely if that particular franchisee has a very strong record and the individual franchise units have strong financials. It might even be better to have a high concentration of high-quality loans than a more diverse pool of weaker credits.

Concept diversification is also important. Franchise loans extend for 10 to 20 years, and a profitable concept today may become unprofitable as the loans mature.

It is not as important that pooled loans include representation across several major sectors (such as more than one restaurant subsector, or loans from all three major groups). Many finance companies specialize in one or two segments of the industry, and know their area well. Thus, a deal from only one of the major sectors does not add any measurable risk as long as there is diversification by franchisee and concept.

Geographical diversification is also important, as it reduces risk associated with regional economic recessions.

Control of Collateral

A key factor in the event of borrower (franchisee) default is control of the collateral. If a franchise loan is secured by a fee simple mortgage, the lender controls disposition of collateral in a bankruptcy. However, if that collateral is a leasehold interest (especially if the lessor is a third party and not the franchisor), the lender may not be able to control disposition in the event of default.

RATE REDUCTION BONDS

The concept of *rate reduction bonds* (*RRBs*)—also known as *stranded costs* or *stranded assets*—grew out of the movement to deregulate the electric utility industry and bring about a competitive market environment for electric power. Deregulating the electric utility market was complicated by large amounts of “stranded assets” already on the books of many electric utilities. These stranded assets were commitments that had been undertaken by utilities at an earlier time with the understanding that they would be recoverable in utility rates to be approved by the states’ utility commissions. However, in a competitive environment for electricity, these assets would likely become uneconomic, and utilities would no longer be assured that they could charge a high enough rate to recover the costs. To compensate investors of these utilities, a special tariff was proposed. This tariff, which would be collected over a specified period of time, would allow the utility to recover its stranded costs.

This tariff, which is commonly known as the *competitive transition charge (CTC)*, is created through legislation. State legislatures allow utilities to levy a fee, which is collected from its customers. Although there is an incremental fee to the consumer, the presumed benefit is that the utility can charge a lower rate as a result of deregulation. This reduction in rates would more than offset the competitive transition charge. In order to facilitate the securitization of these fees, legislation typically designates the revenue stream from these fees as a statutory property right. These rights may be sold to an SPV, which may then issue securities backed by future cash flows from the tariff.

The result is a structured security similar in many ways to other ABS products, but different in one critical aspect: The underlying asset in a RRB deal is created by legislation, which is not the case for other ABS products.

In the first quarter of 2001 there was a good deal of concern regarding RRBs. The sector came under intense scrutiny as a result of the financial problems experienced by California’s major utilities. Yet despite the bankruptcy motion filed by Pacific Gas and Electric (PG&E) in 2001—a bellwether issuer of RRBs—rating agencies maintained their triple-A ratings on California’s existing RRB issues. This is not the first time the RRB sector had found itself in turmoil. Over much of 1998, the sector was roiled by a movement in California to overturn the existing legislation that had been created specifically for RRB securitization. This put existing RRB issues in jeopardy. However, the ultimate result—a voter initiative was defeated—proved to be positive for this product. The ability of this asset class to retain its rating despite a significant credit crisis at an underlying utility, as well as a serious challenge to the legislation that allows for the creation of these securities, speaks volumes for the soundness of the structures of RRB deals.

Structure

As noted above, state regulatory authorities and/or state legislatures must take the first step in creating RRB issues. State regulatory commissions decide how much, if any, of a specific utility’s stranded assets will be recaptured via securitization. They will also decide on an acceptable time frame and collection formula to be used to calculate the CTC. When this legislation is finalized, the utility is free to proceed with the securitization process.

The basic structure of an RRB issue is straightforward. The utility sells its rights to future CTC cash flows to an SPV created for the sole purpose of purchasing these assets and issuing debt to finance this purchase. In most cases, the utility itself will act as the servicer because it collects the CTC payment from its customer base along with the typical electric utility bill. Upon issuance, the utility receives the proceeds of the securitization (less the fees associated with issuing a deal), effectively reimbursing the utility for its stranded costs immediately.

RRBs usually have a "true-up" mechanism. This mechanism allows the utility to recalculate the CTC on a periodic basis over the term of the deal. Because the CTC is initially calculated based on projections of utility usage and the ability of the servicer to collect revenues, actual collection experience may differ from initial projections. In most cases, the utility can re-examine actual collections, and if the variance is large enough (generally a 2% difference), the utility will be allowed to revise the CTC charge. This true-up mechanism provides cash flow stability as well as credit enhancement to the bondholder.

Enhancement Levels

Credit enhancement levels required by the rating agencies for RRB deals are very low relative to other ABS asset classes. Although exact amounts and forms of credit enhancement may vary by deal, most transactions require little credit enhancement because the underlying asset (the CTC) is a statutory asset and is not directly affected by economic factors or other exogenous variables. Furthermore, the true-up mechanism virtually assures cash flow stability to the bondholder.

As an example, the AAA-rated bonds Detroit Edison Securitization Funding 1 issued in March 2001 were structured with 0.50% initial cash enhancement (funded at closing) and 0.50% overcollateralization (to be funded in equal semiannual increments over the terms of the transactions). This total of 1% credit enhancement is minuscule in comparison to credit cards, for example, which typically require credit enhancement at the AAA level in the 12% to 15% range for large bank issuers.

Unique Risks

RRBs are subject to risks that are very different from those associated with more traditional structured products (e.g., credit cards, home equity loans, and so on). For example, risks involving underwriting standards do not exist in the RRB sector because the underlying asset is an artificial construct. Underwriting standards are a critical factor in evaluating the credit of most other ABS. Also, factors that tend to affect the creditworthiness of many other ABS products—such as levels of consumer credit or the economic environment—generally do not have a direct effect on RRBs. Instead, other unique factors that must be considered when evaluating this sector. The most critical risks revolve around the legislative process and environment plus the long-term ability of the trust to collect future revenues to support the security's cash flows.

SUMMARY

In this chapter we described the characteristics of seven types of ABSs for which the asset pool consists of traditional nonmortgage assets: credit card receivable-backed securities, auto loan-backed securities, student loan-backed securities, SBA loan-backed securities, aircraft lease-backed securities, franchise loan-backed securities, and rate reduction bonds.

REFERENCES

Anderson, J. S., and Clark, K. L. (2000). Franchise loan securitization. In F. J. Fabozzi (ed.), *Investing in Asset-Backed Securities* (pp. 187–202). Hoboken, NJ: John Wiley & Sons.

Anderson, J. S., Shannon, S. H., and Clark, K. L. (2000). Equipment-financed ABS. In F. J. Fabozzi (ed.), *Investing in Asset-Backed Securities* (pp. 139–150). Hoboken, NJ: John Wiley & Sons.

Bakalar, N., and Nimberg, J. (2000). Securities backed by recreational vehicle loans. In F. J. Fabozzi (ed.), *Investing in Asset-Backed Securities* (pp. 101–138). Hoboken, NJ: John Wiley & Sons.

Batchvarov, A., Collins, J., and Davies, W. (2004). European mezzanine loan securitisations. In F. J. Fabozzi and M. Choudhry (eds.), *The Handbook of European Structured Financial Products* (pp. 252–262). Hoboken, NJ: John Wiley & Sons.

Davidson, A., Sanders, A., Wolff, L-L, and Ching, A. (2003). *Securitization: Structuring and Investment Analysis*. Hoboken, NJ: John Wiley & Sons.

Dennis, A. (2004). Securitisation of UK pubs. In F. J. Fabozzi and M. Choudhry (eds.), *The Handbook of European Structured Financial Products* (pp. 389–410). Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (2005). The structured finance market: An investor's perspective. *Financial Analysts Journal* 60, 3: 27–40.

Fabozzi, F. J., and Kothari, V. (2008). *Introduction to Securitization*. Hoboken, NJ: John Wiley & Sons.

Faulk, D. (1996). SBA loan-backed securities. In A. K. Bhattacharya and F. J. Fabozzi (eds.), *Asset-Backed Securities* (pp. 167–178). Hoboken, NJ: John Wiley & Sons.

Flanagan, C., and Tan, W. (2000). SBA Development Company participation certificates. In F. J. Fabozzi (ed.), *Investing in Asset-Backed Securities* (pp. 169–188). Hoboken, NJ: John Wiley & Sons.

Gintz, C. (2004). Stock securitisations. In F. J. Fabozzi and M. Choudhry (eds.), *The Handbook of European Structured Financial Products* (pp. 263–272). Hoboken, NJ: John Wiley & Sons.

Kothari, V. (2006). *Securitization: The Financial Instrument of the Future*, 3rd edition. Singapore: John Wiley & Sons.

Lucas, J., and Zimmerman, T. (1996). Equipment lease-backed securities. In A. K. Bhattacharya and F. J. Fabozzi (eds.), *Asset-Backed Securities* (pp. 147–166). Hoboken, NJ: John Wiley & Sons.

McElravey, J. (2005). Credit card asset–backed securities. In F. J. Fabozzi (ed.), *Handbook of Fixed Income Securities*, 7th edition (pp. 629–646). New York: McGraw-Hill.

McPherson, N. (2000). Healthcare receivable backed ABS. In F. J. Fabozzi (ed.), *Investing in Asset-Backed Securities* (pp. 211–222). Hoboken, NJ: John Wiley & Sons.

McPherson, N. (2000). Securities backed by auto loans and leases. In F. J. Fabozzi (ed.), *Investing in Asset-Backed Securities* (pp. 61–100). Hoboken, NJ: John Wiley & Sons.

Neimier, M. (2004). European credit card ABS. In F. J. Fabozzi and M. Choudhry (eds.), *The Handbook of European Structured Financial Products* (pp. 177–200). Hoboken, NJ: John Wiley & Sons.

Neimier, M. (2004). European auto and consumer loan ABS. In F. J. Fabozzi and M. Choudhry (eds.), *The Handbook of European Structured Financial Products* (pp. 201–222). Hoboken, NJ: John Wiley & Sons.

Philips, J., and Hsiang, O. (2000). Aircraft asset-backed securities. In F. J. Fabozzi (ed.), *Investing in Asset-Backed Securities* (pp. 151–168). Hoboken, NJ: John Wiley & Sons.

Pica, V. T., and Bhattacharya, A. K. (1996). In A. K. Bhattacharya and F. J. Fabozzi (eds.), *Asset-Backed Securities* (pp. 179–186). Hoboken, NJ: John Wiley & Sons.

Quisenberry, J. O. (1998). Securitization of non-traditional asset types: An investor's perspective. In F. J. Fabozzi (ed.), *Handbook of Structured Financial Products* (pp. 21–28). Hoboken, NJ: John Wiley & Sons.

Ramgarhia, A., Muminoglu, M., and Pankratov, O. (2004). Whole business securitisation. In F. J. Fabozzi and M. Choudhry (eds.), *The Handbook of European Structured Financial Products* (pp. 329–338). Hoboken, NJ: John Wiley & Sons.

Roever, W. (2005). Securities backed by automobile loans and leases. In F. J. Fabozzi (ed.), *Handbook of Fixed Income Securities*, 7th edition (pp. 647–668). New York: McGraw-Hill.

Schorin, C. N. (1998). Auto lease ABS. In F. J. Fabozzi (ed.), *Handbook of Structured Financial Products* (pp. 139–151). Hoboken, NJ: John Wiley & Sons.

Schorin, C. N. (1998). Credit card asset–backed securities. In F. J. Fabozzi (ed.), *Handbook of Structured Financial Products* (pp. 153–178). Hoboken, NJ: John Wiley & Sons.

Schorin, C. N., and Roach, K. (1998). Utility stranded costs securitization. In F. J. Fabozzi (ed.), *Handbook of Structured Financial Products* (pp. 263–270). Hoboken, NJ: John Wiley & Sons.

Thompson, A.V. (1998). Securitization of commercial and industrial loans. In F. J. Fabozzi (ed.), *Handbook of Structured Financial Products* (pp. 235–240). Hoboken, NJ: John Wiley & Sons.

Wagner, K., and Callahan, E. (1998). Student loan ABS. In F. J. Fabozzi (ed.), *Handbook of Structured Financial Products* (pp. 201–234), Hoboken, NJ: John Wiley & Sons.

Weaver, K., and Bakalar, N. (2000). In F. J. Fabozzi (ed.), *Investing in Asset-Backed Securities* (pp. 223–238). Hoboken, NJ: John Wiley & Sons.

Zimmerman, T. (2000). Student loan floaters. In F. J. Fabozzi (ed.), *Investing in Asset-Backed Securities* (pp. 203–210). Hoboken, NJ: John Wiley & Sons.

CHAPTER 36

Synthetic Asset-Backed Securities

MOORAD CHOUDHRY, PhD
Head of Treasury, KBC Financial Products, London

Abstract: The excess of demand over supply for specific tranches of structured finance transactions such as asset-backed securities and mortgage-backed securities has led to investors' accessing these assets via the credit default swap market. A credit default swap written on an asset-backed security possesses different contract mechanics to a standard credit default swap written on a corporate reference name. The main differences relate to the list of occurrences that constitute a credit event and the settlement mechanics whenever the underlying tranches of an asset-backed security or mortgage-backed security experience a paydown or other prepayment.

Keywords: asset-backed security (ABS), mortgage-backed securities (MBSs), cash settlement, credit default swaps (CDSs), credit event, pay-as-you-go, physical settlement, structured finance securities

Credit derivatives were first introduced during the 1990s, initially as a risk management tool for banks seeking to manage and transfer the credit exposure of their loan books. They have since developed into an asset class in their own right, in synthetic form, and in some cases are preferred to the cash version of an asset where the latter is in short supply or otherwise illiquid. The advent of a liquid and transparent market in credit derivatives has meant that investors are now looking at synthetic access to the *asset-backed security (ABS)* market. In this chapter we describe the form of *credit default swaps (CDSs)* written on *structured finance securities* such as ABSs and *mortgage-backed securities (MBSs)*, and their trading and settlement mechanics, which differ from CDSs written on plain vanilla bonds.

CREDIT DERIVATIVES AND ABS MARKETS

Credit derivatives markets have expanded rapidly since the first instruments were introduced in 1994. They have now been extended into the asset-backed and mortgage-backed markets, mainly due to the shortage of paper in the cash market. The standardization of CDS contracts and trading terminology also facilitated the expansion of credit derivatives into structured finance markets.

There are a number of detail differences between ABS credit risk and corporate credit risk. A single-name corporate CDS transacted under the standard 2003 International Swaps and Derivatives Association (ISDA) Credit Derivatives definitions (www.isda.org) will be based on clearly defined trigger events *("credit events")* and a transparent process of settlement, either physical or *cash settlement.* The settlement process for standard CDSs is shown as Figure 36.1. A CDS written on an ABS can present problematic issues with regard to both of these items. Corporate CDS trigger events are, following some initial problems with definitions, straightforward to describe. They include bankruptcy, failure to pay, debt restructuring, and, in some cases, ratings downgrade. Such occurrences can be identified easily in most cases. Also, the outstanding debt of a corporate entity can be expected to trade at roughly the same level in the event of issuer default, irrespective of coupon or maturity.

Structured finance securities such as ABSs and MBSs differ in both these respects. The key difference is that, unlike corporate bonds, most ABSs are issued by special purposes vehicles (SPVs), bankruptcy-remote legal entities created solely for the purpose of facilitating the bond

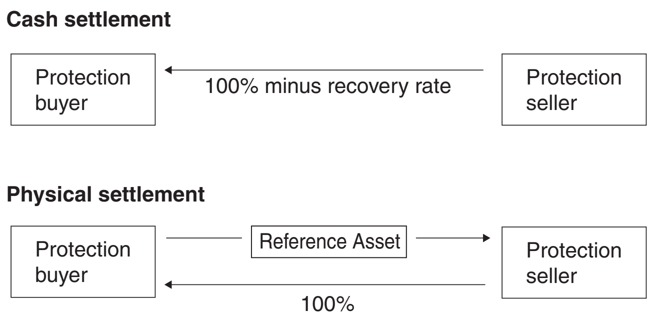


Figure 36.1 CDS Contract Settlement Mechanism, Cash or Physical Settlement

issue. Bankruptcy and restructuring rarely, if ever, apply to SPVs. Also, it may be less clear in the case of an SPV that there has been a failure to pay. Unlike corporate entities, credit ratings of SPVs are based essentially on the quality of the underlying assets. The repayment of these assets is not known with certainty, which is why ABS bonds are given long legal final maturities. Other issues that complicate the matter of CDSs on ABSs include the following:

- ABS structures with an element of uncertain cash flow patterns include the provision for the write-down of principal in the event of losses. This does not always constitute a "default" as the write-down can be reversed and made good later.
- Many ABS structures allow for a delay in interest payment, for example, during a time when the excess spread in the vehicle has been reduced. Again, this may not constitute default and may not necessarily lead instantly to a ratings downgrade, as the interest coverage may be expected to become sufficient again.
- The structure represents a distinct pool of assets, ring-fenced within the SPV. This contrasts with the general pool of assets represented in a corporate entity.
- It is quite possible for the more junior tranches of an ABS issue to be in default while the senior tranches are not, again representing the way the asset pool is performing.

The significant difference, therefore, between an ABS CDS and a single-name corporate CDS is that the former is written against a specific security, while the latter is written at an entity level on a corporate name. However, writing a contract on a specific security means that *physical settlement* on occurrence of a credit event is impractical. For this reason, physical settlement is not used. Cash settlement may also be problematic because of the difficulty with ascertaining the market value of the ABS tranche. A different type of CDS, the *pay-as-you-go* CDS (PAUG CDS) has been developed for this market.

PAY-AS-YOU-GO CDS

PAUG CDSs have been developed to meet the distinct requirements of synthetic investment in ABS issues. A PAUG CDS acts like a standard CDS, with provision for termination on occurrence of specified credit events. The protection buyer pays a fixed-basis-point fee to the protection seller, which is also standard. However, the PAUG contract also permits the following:

- Payment of an additional floating payment from protection seller to protection buyer in the event of principal write-down.
- Payment of a fixed payment from protection buyer to protection seller in the event of write-up.
- Provision of altering cash flows in the event of interest shortfall of ABS vehicle.

To illustrate, consider the case where the performance of the underlying asset pool in an ABS, due to underperformance or default, means that the principal amount of one or more of the overlying note tranches must be reduced. This action would normally be undertaken by the Trustee or servicer to the transaction. The protection seller would make a floating payment to the protection buyer to cover this written-down amount. The CDS itself would not terminate. If at a later date the principal balance is reinstated, for example, because the portfolio performance has improved, the protection buyer would then make a fixed payment to the seller. Figure 36.2 illustrates the mechanics of a PAUG CDS in the event of write-down.

A standard CDS would generally cover the following credit events:

- Failure to pay
- Credit rating downgrade to sub–investment grade
- Permanent write-down

A PAUG CDS would also cover the following without being terminated:

- Principal write-down
- Interest shortfall
- Failure to pay principal

By incorporating this flexibility, investors are better able to gain a realistic exposure to the ABS market, albeit synthetically.

MARKET CONSIDERATIONS

Generally, the protection sellers in the ABS CDS market include investors who would normally hold cash ABS bonds. This is marked when there is a shortage of paper in the cash market. The ability to short ABS tranches means that investors can also take a view on ABS credit; previously, this would not necessarily have been straightforward because of the illiquid nature of the ABS repo market. The differences between the corporate and ABS markets are mirrored in the synthetic market. Investors will be aware that corporate entities are dynamic corporations that are proactively able to avoid credit events, which is not the case with SPVs.

Synthetic ABS investors must therefore still be concerned primarily with the quality of the underlying collateral and the specific risk/return profile of the individual ABS tranche. Also, there is the issue of prepayment

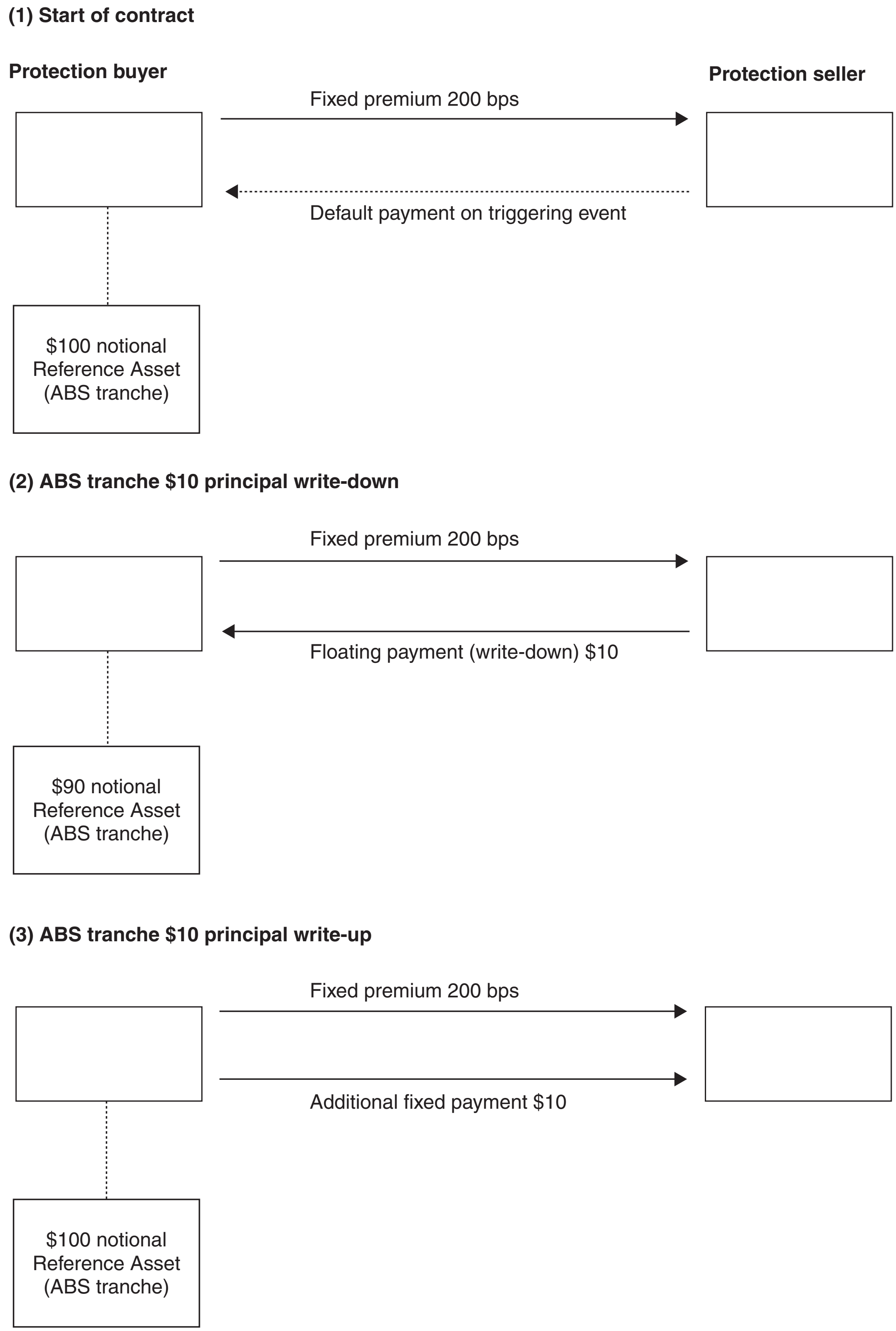


Figure 36.2 PAUG CDS Cash-Flow Mechanics in the Event of ABS Note Principal Write-down

uncertainty. Most corporate bonds have a bullet maturity or fixed redemption date. Nonredemption would constitute a credit event. ABS securities, however, amortize over time, with the redemption date not known with certainty. (For analysis purposes, the "average life" of the ABS note is used; this figure is an estimated repayment term based on an assumed level of prepayments). However, the nonredemption of a tranche in accordance with an average life estimate would not be deemed a credit event.

ABS tranches experience a declining notional balance over time as principal is repaid in stages in the underlying asset pool, due to prepayments and other factors. The outstanding notional value of ABS tranches therefore reduces over time; investors would observe this also occurring with ABS CDS contract notionals, as they mirror the behavior of the cash bond.

ABS CDS AND CASH BOND VALUATION

In theory, the basis between a PAUG CDS and its reference cash bond should be small because the contract mirrors the profile and behavior of the cash bond closely. (See Choudhry [2004b] for more detail on the CDS basis.)

In practice, a number of market and structural factors cause the cash and synthetic markets to trade at a negative or positive basis. These include the following:

- In the synthetic market, the investor is exposed additionally to counterparty risk, as it is the counterparty that is paying the coupon (CDS premium). The cash investor is exposed to the quality of the reference collateral only.
- The ABS CDS is an unfunded instrument and so carries no funding cost; this is an additional factor in relative value analysis.
- Supply-and-demand factors may be more prevalent in the synthetic market, as the availability of protection buyers may be limited. (Unlike ABS transaction originators in the cash market, there is no natural market for protection buyers in the ABS CDS market outside market makers.)

As a relatively new market as of this writing, the depth and transparency of the ABS CDS market may be limited for certain sectors. This should not be a problem once the market develops.

SUMMARY

The liquidity and transparency of the credit derivatives market, together with the adoption of standard terms and contractual documentation, has resulted in the market's being straightforward for investors to access. Where there is a shortage of required assets in the cash bond market, investors are able to access the same name in the credit derivative market. One asset class where this manifests itself is in the structured finance market, where specific tranches of ABS and MBS transactions are often in short supply or illiquid. Investors can access these tranches via a CDS written on the specific tranche. Such contracts differ in certain technical aspects from CDSs written on conventional bullet bonds. This includes the events that form a credit event, and the way that contract notionals are adjusted after a pay-down on the underlying reference tranche.

REFERENCES

Choudhry, M. (2004a). *Structured Credit Products: Credit Derivatives and Synthetic Securitisation*. Singapore: John Wiley & Sons

Choudhry, M. (2004b). The credit default swap basis: Analysing the relationship between cash and synthetic markets. *Journal of Derivatives Use, Trading and Regulation* 10, 1: 8–26

Das, S. (2006). *Credit Derivatives, CDOs, and Structured Credit Products*, 3rd edition. Singapore: John Wiley & Sons.

Deacon, J. (2004). *Global Securitisation and CDOs*. Chichester, UK: John Wiley & Sons.

Fabozzi, F. J., and Choudhry, M. (eds.) (2004). *The Handbook of European Structured Financial Products*. Chichester, UK: John Wiley & Sons.

CHAPTER 37

Catastrophe Bonds

WILLIAM L. MESSMORE
Vice President, Lehman Brothers

BETH STARR
Managing Director, Lehman Brothers

SUNITA GANAPATI

MARK RETIK

PAUL PULEO

Abstract: Catastrophe bonds represent a growing class of structured insurance risk products that offer returns that are linked to the occurrence of catastrophic events such as earthquakes and hurricanes. These securities can provide investors with diversification from corporate and asset-backed securities at comparable or wider spreads. Issued through special purpose vehicles, these bondlike securities are usually rated and offer an opportunity to participate directly in catastrophe risk with the benefit of an active secondary market. Investing in catastrophe risk can also improve the risk-return profile of a diversified portfolio of assets because this risk is generally uncorrelated with general credit and interest rate risk present in other securities markets.

Keywords: catastrophe bonds (cat bonds), sidecars, attachment point, trigger, extreme mortality securities, collateralized debt obligation (CDO), synthetics, industry loss warranties (ILWs), shelf issuance programs

The need for additional reinsurance capacity following Hurricane Andrew (1992) and the Northridge earthquake (1994), which in combination produced $27 billion in industry-wide insured losses, encouraged insurers to seek a new form of reinsurance protection. Driven by a particularly catastrophic 2005 U.S. wind season with Hurricanes Katrina, Rita, and Wilma, insurance companies faced capacity and pricing constraints in the broader reinsurance market and further turned to the capital markets to transfer risk. In exchange for a reinsurance premium (that is, interest on the securities), investors in *catastrophe bonds* (*cat bonds*) assume financial exposure to the risk that a catastrophe will strike and will be severe enough to exceed a certain trigger level. If such a catastrophe occurs, cat bond investors would receive a reduced yield and lose part or all of their principal; the insurer would receive a reinsurance claim payment. By transferring catastrophe risks to the capital markets in this manner, insurance companies are supplementing their use of traditional reinsurance and internal loss management mechanisms to reduce volatility in their financial statements and preserve overall liquidity.

In this chapter, we discuss catastrophes and the role that reinsurance has traditionally played in mitigating catastrophic losses. We describe developments in the capital markets that have led to catastrophe risk securitization and outline typical cat bond structures. We consider the third-party modeling analyses that accompany each cat bond and the related rating agency approaches. Finally, we discuss new risk-transfer products and market developments.

CATASTROPHE RISK MANAGEMENT

Traditional Reinsurance

Reinsurance gives an insurer the ability to transfer risk with the primary purpose of either smoothing its income stream or protecting its balance sheet. Catastrophe management is an essential component of a reinsurance program for large property insurers. Insurance companies structure their reinsurance coverage according to their internal risk tolerance, corporate ratings targets, and cost considerations. Traditional reinsurance market coverage comes on an unsecured basis and became particularly costly after the 2005 U.S. wind season due to reduced amounts of capital available. Rating agencies instituted more stringent ratings criteria (particularly for extreme or "tail" risks), and the third-party risk-modeling firms reassessed U.S. hurricane risk, resulting in more conservative risk quantifications. Some of the largest reinsurance companies, most notably Swiss Re and Hannover Re, have turned to the capital markets to buy protection on their underlying books of business.

Catastrophe Risk Layers

Catastrophe risk can be viewed as composed of layers of risk from events with decreasing probability of occurrence and increasing magnitude of losses. Historical and sophisticated modeling analyses indicate that catastrophic events occur at mostly unpredictable intervals and that less severe catastrophes occur with greater frequency. Risk management of catastrophe losses varies from one insurer to another. Figure 37.1 shows a probability distribution of insured losses and the sources of risk capital that an insurer may use to manage its catastrophe exposure.

Catastrophes resulting in gross insured losses of less than 5% of a major property insurer's statutory surplus occur frequently and are assumed to be part of the normal course of business. Losses from these events are absorbed by an insurer's operating cash flow, policyholders' surplus, or "working layer" reinsurance program. Events that cause losses between 5% and 10% of surplus are generally covered by purchasing traditional reinsurance contracts.

As insurers have increased their use of advanced catastrophe modeling to predict losses, they have tended to purchase coverage equal to their probable loss under various severe loss scenarios or, at a minimum, for losses in

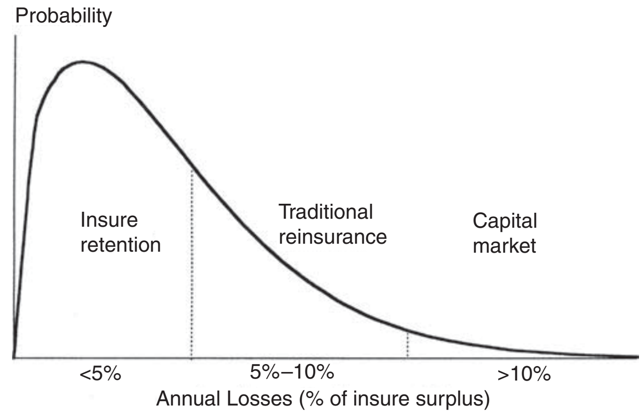


Figure 37.1 Illustration of Insurer Catastrophe Risk Management

excess of 10% of their capital. However, large insurers find that protecting their balance sheet against an infrequent but large catastrophe is often too expensive due to the concentration of risk and lack of capacity in the reinsurance industry for covering this type of risk. Reinsurers face the same constraints with respect to the overall exposure in retrocession market, which offers reinsurance to reinsurers. Therefore, insurers and reinsurers are seeking capital market solutions to bridge this gap in capacity and to create a more efficient risk transfer mechanism.

CAPITAL MARKET DEVELOPMENTS

As insurers explore alternative solutions for gaining additional reinsurance coverage, they have participated in several creative capital market–related developments, including government initiatives, cat bonds, sidecars, industry loss warranties, and synthetic cat bonds.

Government Initiatives

In response to reduced property insurance availability after Hurricane Andrew and the Northridge earthquake, the U.S. and state governments created various funds to provide additional insurance capacity. These include the Florida Hurricane Catastrophe Fund and the California Earthquake Authority, among others. These funds are set up to access the capital markets immediately after an event to provide additional funding either directly to homeowners or to insurance companies.

The Florida Hurricane Catastrophe Fund was designed to provide additional reinsurance capacity to primary insurers writing homeowner policies in Florida. The fund has expanded its risk capacity to cover $32 billion in losses, which would be funded largely by assessments on future Florida insurance premiums in the event of a major hurricane. Critics maintain that the fund artificially depresses policyholder premiums, and in the event of a particularly catastrophic event, could ultimately put the burden of claims repayment on Florida taxpayers instead of distributing the risk among the worldwide insurance industry. Several primary insurers have curtailed their participation in the Florida market.

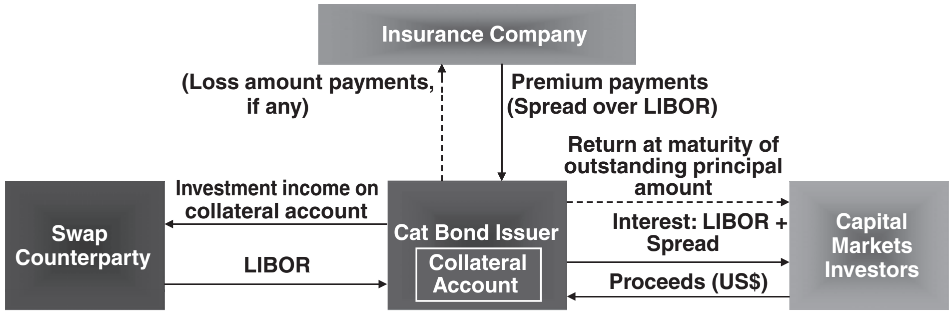


Figure 37.2 Catastrophe Bond Structure

The California Earthquake Authority (CEA) has over $8 billion of claims-paying capacity in California. The program was designed to include a combination of letter-of-credit facilities, reinsurance policies directly from reinsurers, assessments from participating insurance carriers, tax-exempt earnings on its reserves, and the capital markets. CEA policies are written by participating primary insurance companies as an add-on to a customer's existing residential policy.

These special funds are expected to provide incremental capacity to the property-casualty industry and potentially bridge part of the gap in reinsurance supply.

Securitization of Catastrophe Risk

Insurance risk securitization is the transfer or sale, in the form of an investment security, of part of the underwriting risks associated with a group of insurance policies. Insurance companies expect that insurance risk securitization will play a significant role in meeting the shortage in reinsurance capacity. Investing in cat bonds is akin to issuing a reinsurance contract where the investor covers the insurer for a fixed amount of losses over a specified value (the *attachment point* or *trigger*).

As with any capital market product, structures are still evolving. The cat bond universe covers a number of different peril types with customized loss trigger types. Dealers have been distributing securities that reference portfolios of insurance and reinsurance risks, known as *sidecars,* and managed *collateralized debt obligation (CDO)* technology has been utilized as well. In concert with the explosive popularity of derivative contracts in the credit markets, the cat market has seen several investors participate in synthetic catastrophe risk, using existing cat bonds as reference obligations, or writing derivative contracts that are linked to third-party estimates of a catastrophe's insured industry losses.

STRUCTURE OF CATASTROPHE BONDS

Cat bonds are issued for an expected maturity with the payment of coupon and retirement of principal dependent on the nonoccurrence of a catastrophic event with losses greater than a specified trigger during a defined risk or loss-occurrence period. As in other asset-backed transactions, the sponsor sets up a special purpose vehicle (SPV) that is bankruptcy remote. The vehicle is generally set up offshore for regulatory and tax reasons and issues securities that carry the risk of catastrophe losses over a specified level. It then issues a back-to-back reinsurance or derivative contract to the insurer, thus providing the loss protection.

The SPV invests cash raised from the issue in high-quality, liquid, fixed income instruments (typically AAA-rated securities). This portfolio is used to cover losses from events or to repay investors on maturity of the bond, and to provide a minimum rate of return (e.g., the London Interbank Offered Rate [LIBOR]). The return on the collateral account is guaranteed by a swap counterparty in the form of a total return swap on the assets. The catastrophe risk is transferred via a cash-collateralized reinsurance or derivative contract and, unlike traditional reinsurance contracts, does not carry any credit risk of the reinsurer. The coupon on the cat bond includes a spread over the minimum rate earned by the collateral account. The insurer pays the spread to the SPV, which passes through the total coupon payment to investors (see Figure 37.2).

The maturity of the security is based on the period during which a loss event can occur, called the risk period (or the loss occurrence period), and the time for computation of losses, called the development period. The development period may be up to two years, during which time the company works with a calculation agent to aggregate the final data surrounding the event, and inputs that data to determine whether a trigger event has taken place. The cat bond is usually structured to have a scheduled maturity date that can be extended for a maximum period equal to the development period (with reduced interest payments), thus exposing investors to some extension risk.

Cat bonds utilize a trigger structure that generally falls into one of four categories:

1. Parametric
2. Industry loss
3. Indemnity
4. Modeled loss

Parametric instruments rely on scientific readings surrounding an event: wind speeds, ground shake acceleration, or even measurements in flood recording stations. Typically, the data collected for the event are plugged into an index formula designed by the sponsor, and to the extent the output from that formula exceeds a certain predefined threshold ("attachment point"), the cat bond will suffer a principal loss. If the output exceeds a second threshold ("exhaustion point"), the cat bond will suffer a full principal loss.

For a cat bond using an industry loss trigger structure, principal losses occur if a third-party reporting agency's estimates of industry-wide insured loss for an event exceed a predetermined attachment point. Several sponsors have attempted to match their underlying books of business in their trigger formulas by applying weighted factors against the loss estimates, which are given on a state-by-state basis. A sponsor with concentrated obligations in the northeastern United States may build only those states into their trigger structure, and may further refine their coverage by making their loss trigger more sensitive to New York and less sensitive to Maryland, for example. Property Claim Services (PCS) is the most widely used third-party loss estimator for catastrophes in the United States.

From a risk perspective, the indemnity structure is less transparent than the parametric and industry loss structures, but it removes the basis risk between the insurer's underlying book of business and the cat bond coverage. In an indemnity structure, the insurance company receives a payout from the cat bond to the extent that the insurance company's book of business suffers losses above a certain attachment point. In a basic arrangement, the insurance company would retain the entire layer of losses up to the attachment point, and then retain a small pro-rata slice of the losses along with the cat bond; this pro-rata slice demonstrates to the cat bond investor that the insurance company retains an interest in its underwriting for the severe end of the risk spectrum. Protection on indemnity losses can be structured based on (1) losses from a single event (which has sensitivity to the severity of events) or (2) the aggregate annual losses from multiple events (which also has sensitivity to event frequency). Insurers may prefer indemnity structures because they are most similar to traditional reinsurance contracts, but the marketability of indemnity cat bonds relies on the insurance company's ability to demonstrate a strong and consistent underwriting history.

A variation on the indemnity trigger is the modeled loss trigger, where the sponsor designs and employs an escrowed loss model that inputs parametric data surrounding an event, and outputs an index value from the collected data. The modeled loss trigger is more transparent to investors than an indemnity structure because it limits the risk that underlying policy losses could occur that were not factored into the risk model used to evaluate an indemnity transaction. Therefore, in a modeled loss structure, the sponsor still retains some basis risk between their underlying book of business and the model output. However, an advantage to a modeled loss trigger is that the sponsor could receive a reinsurance payment immediately after data surrounding an event becomes available, instead of waiting for the underlying policy losses to develop and be aggregated.

RATING AGENCY AND MODELING CONSIDERATIONS

The rating agencies have developed criteria for rating catastrophe-linked securities and furnish ratings on most transactions. At present, the methodology used by each agency is similar—though each is being continually refined, reflecting the relative newness and prospect for growth of this asset class. The presentation in this section is based on discussions with analysts at Standard & Poor's (S&P) and Moody's, along with publications on catastrophe-linked security rating approaches promulgated by the agencies.

The agencies rate cat bonds to reflect loss to both principal and interest. Because this approach is also used to rate corporate credits and asset-backed security structures, it is possible to draw conclusions on relative creditworthiness between these securities and catastrophe-linked securities based on ratings.

In analyzing cat bonds, the rating agencies consider structural and insurance risks. The structural analysis is essentially the same as the analysis used to rate any structured security. This analysis focuses on the transaction's legal structure; the quality of collateral; the bankruptcy-remote status of the SPV issuer; the flow of funds; and the market, counterparty, and legal risks inherent in the transaction.

Although structural risk is an important element in the rating methodology, the key risk that the rating agencies analyze is insurance risk. Cat bonds have their principal and interest at risk in that their calculated index value could exceed a predetermined attachment point.

There are several independent modeling firms that specialize in catastrophe modeling, and their models are utilized throughout the property and casualty (P&C) insurance industry for risk management purposes. Typically, each cat bond transaction will have an independent risk analysis that the issuer publishes as a part of the offering materials. Modeling firms EQECAT, AIR, and RMS maintain various models for different perils, including U.S. Hurricane, U.S. Earthquake, European Windstorm, Japanese Typhoon, Japanese Earthquake, and so on. These models rely on historical data, prospective climatological analysis, and topographic information to predict frequency, severity, and location of potential future events. Once the catastrophic events have been simulated, the modeling firms overlay data on insured values in order to estimate the damage applicable to the specific transaction (if applicable) and create a loss exceedance curve. Each structure has an estimated annualized probability of attachment, loss, and exhaustion, which gives investors the ability to assess a deal's risk against its offered yield, as well as against other cat bonds exposed to the same peril or perils.

The rating agencies rely on the results of simulation-driven catastrophe models to assign their ratings. The agencies first validate the analytic integrity of the model and test the quality of the insurance company data used by the model, if any.

These "stress tests" are conducted through a due diligence process. This process typically involves assessing the appropriateness of the probability distributions employed by the model to simulate catastrophe frequency and intensity. Both the underlying density functions and parameters are considered. Occasionally, a rating agency will request a modification of the probability distribution to generate more conservative results (e.g., it might ask to recalculate the insured loss distribution using twice the assumed catastrophe frequency).

In addition, property damage vulnerability relationships are examined. Vulnerability functions are considered for each property characteristic (e.g., construction type, elevation, building usage, etc.) using engineering and actuarial analysis. In all cases, consistency with published industry and academic literature is tested. Some rating agencies retain the services of outside meteorological or seismic experts to assist in evaluating the model.

For indemnity-triggered structures, the insurance company data used by the model are reviewed by the rating agencies for accuracy. These data include both the book of insured properties and the policy provisions in place on each property. Conservative adjustments are made to account for incomplete data.

Finally, certain indirect factors are sometimes also factored into the rating analysis. These include demand surge (the effect of a catastrophe on local prices for building materials and wages), growth and change of mix in the insured book of business over the course of the security's term, and the insurance company's claims handling and loss management/settlement procedures.

An Uncorrelated Asset Class

Cat bonds offer investors the unique opportunity to invest exclusively in catastrophe risk and may provide potential diversification benefits. Although investors can invest in catastrophe risks by buying insurance and/or reinsurance company equity and debt, these investments are not perfect substitutes for the pure catastrophe exposure inherent in cat bonds. First, cat bonds do not carry the idiosyncratic or nondiversifable risks associated with an investment in securities of an insurance or reinsurance company. Cat bonds also allow investors to avoid principal-agent risks (such as the risk that equity holders may have incentives to restructure the debt or increase the overall riskiness of the company, to the disadvantage of bondholders) inherent in a corporate security.

Second, the occurrence and magnitude of natural hazards are expected to be largely uncorrelated with movements in the stock and bond markets. However, insurance and reinsurance company securities do involve a significant systematic risk. A study by Canter, Cole, and Sandor (1996) show that a portfolio of 10 prominent catastrophe reinsurance companies has a strong positive correlation (beta of 0.83) with stock market movements. As a result, buying reinsurance company equity does not bring significant diversification benefits. In this respect, cat bonds offer better diversification opportunities since they are expected to have near-zero betas.

Modern portfolio theory asserts that an uncorrelated asset would be an attractive addition to a well-diversified portfolio even at the risk-free rate of return. If cat bonds offer returns in excess of the risk-free rate and do not exhibit systematic risk, then investing in these securities can improve overall portfolio performance on a risk-adjusted basis. Investors who purchase cat bonds can potentially receive an attractive expected return and improve the diversification of their current portfolio.

A study by Froot, Murphy, Stern, and Usher (1995) based on pricing and claims on actual catastrophe reinsurance contracts brokered by the reinsurance intermediary, Guy Carpenter & Company Inc., draws three valuable conclusions. First, the correlation of catastrophe risk with stocks and bonds is statistically indistinguishable from zero. Second, assuming that returns on reinsurance contracts provide a reasonable proxy for expected returns on cat bonds, the study shows that investment in such a portfolio of catastrophe reinsurance contracts from 1970 to 1994 would have generated returns 200 basis points above the Treasury bill rate. Third, adding the portfolio of reinsurance contracts improves the efficiency of a diversified portfolio. Using a base portfolio of 70% domestic assets (70% stocks, 30% bonds) and 30% foreign assets (70% stocks, 30% bonds), the study shows that the reward-to-risk ratio (measured as the realized return minus the risk-free return divided by the standard deviation of the portfolio return) grows from 26% to 30% as the addition of catastrophe risk goes from 5% to 25%. Even though the past is no guarantee for future results, historical data provide strong evidence that catastrophe-linked securities offer portfolio opportunities to investors.

MARKET DEVELOPMENTS

Although cat bonds are the most widely issued and traded capital markets product for catastrophe risk transfer, other securities and derivatives have gained some popularity in the marketplace, including the following:

- **Sidecars**. Investor capital is collateralized against an entity writing a portfolio of reinsurance contracts, and the investor participates in the experience of the entity (that is, the investment return is linked to the premium income earned versus the claims paid on the underlying policies). In some sidecar structures, dealers have issued debt layers against the more remote end of the risk spectrum, which has the effect of adding leverage to the equity position in the vehicle. Most sidecars utilize the underwriting expertise of an established reinsurer.
- **Extreme mortality securities**. Similar to a cat bond in structure, an *extreme mortality security* protects a sponsor against a catastrophic increase in mortality rates over a short period of time. Insurance companies use these securities to transfer mortality risk associated with pandemics or (to some extent) terrorism, and sponsors tailor the trigger structures to match the geographic, age, and gender characteristics of their underlying books of life insurance policies.
- **Collateralized debt obligation**. In a cat CDO, dealers borrow technology from the asset-backed securities (ABS) market to create both actively and passively managed instruments. In an actively managed instrument, a risk manager assembles and manages a portfolio of cat bonds, other catastrophe instruments, and reinsurance contracts according to predetermined guidelines. In the passively managed structure, the basket of risks is fixed for the duration of the transaction. In both structures, the portfolio backs a capital structure that has both rated and unrated classes of securities.
- **Synthetics**. In the broader credit markets, credit default swaps (CDS) have taken enormous leaps in terms of trading volumes and liquidity, and the technology is being utilized increasingly in the cat space. Protection buyers and sellers can use existing cat bonds to get long

or short risk, using standard International Swaps and Derivatives Association (ISDA) documentation. Smaller reinsurance companies that want to offload risk but cannot achieve the economies of scale in launching a cat bond program may find benefits in referencing a different issuer's outstanding cat bond. Investors can also get short using *synthetics,* enabling them to make bearish trades on what they view as overpriced risk.

- **Industry loss warranties.** *Industry loss warranties (ILWs)* are another family of derivative contract that are increasingly traded by capital markets players. In an ILW, a protection buyer pays a premium to a counterparty in return for a payout to the extent that a catastrophe's estimated industry-wide losses exceed a certain trigger level in predetermined locations. Terms for an ILW are usually one season (e.g., June through November for the U.S. hurricane season). ILWs are useful risk management tools in that protection buyers may want to hedge certain geographies in the context of their entire risk portfolio, or investors may want to add additional risk in certain regions where they are underexposed. ILWs can be governed by standard ISDA documentation, which increases liquidity due to posting arrangements that exist between financial institutions.
- ***Shelf issuance programs.*** Similar to ABS issuers, many cat bond issuers have adopted shelf documentation technology for their risk-transfer programs. A shelf is an issuance platform that allows issuers to come to market with new securities quickly to take advantage of market conditions.

MARKET PARTICIPANTS

Participation in the cat bond market has seen a substantial transformation. The combination of a broader repricing of insurance risk after the devastating 2005 U.S. hurricane season and the emergence of multistrategy hedge funds as a large and active source of capital has brought in a number of new market participants. However, the largest and most consistent cat bond players continue to be dedicated catastrophe funds.

Most investors pursue diversification strategies when participating in the cat bond market. Since principal losses on cat bonds would tend to be binary, investors build diversity into their portfolios by adding securities that cover different peril types (e.g., earthquake versus hurricane) and that cover different geographical territories. Even though two cat bonds may have the same annualized loss probabilities in their respective risk analyses, a cat bond covering a less widely issued peril or geography generally will price at a lower spread than a more frequently issued risk profile.

Several noninsurance companies have issued cat bonds into the capital markets, bypassing the traditional primary insurers for their protection needs. Sponsors range from energy companies looking to protect against hurricane damage on their Gulf region oil production assets to entertainment companies protecting against localized earthquake damage. These cat bonds generally utilize customized trigger structures to protect against very specific geographies.

SUMMARY

The focus of this chapter is on catastrophe bonds, a financial security designed to transfer risk associated with natural catastrophic events, like hurricanes and earthquakes. Insurance and reinsurance companies issue catastrophe bonds as an alternative to the traditional reinsurance or retrocession markets to protect themselves from losses incurred during extreme events. Catastrophe bonds can offer attractive yields and give investors the opportunity to invest in an asset uncorrelated to the credit or rate markets. Discussion includes deal structure, the rating agencies' approach to this product, new product development, and how these securities can provide diversification benefits to investment portfolios.

ACKNOWLEDGMENTS

This chapter is an update and expansion of "Catastrophe-Linked Securities" by S. Ganapati, M. Retik, P. Puleo, and B. Starr which appears in *Investment Management for Insurers* (John Wiley & Sons, 1999).

REFERENCES

Canter, M. S., Cole, J. B., and Sandor, R. L. (1996). Insurance derivatives: A new asset class for the capital markets and a new hedging tool for the insurance industry. *Journal of Derivatives*, Winter: 89–104.

Coval, J. D., Jurek, J. W., and Stafford, E. (2007). Economic catastrophe bonds. Harvard Business School working paper, July.

Froot, K. A. (ed.) (1999). *The Financing of Catastrophe Risk*. Chicago and London: University of Chicago Press.

Froot, K. A. (2001). The market for catastrophe risk: A clinical examination. *Journal of Financial Economics* 60: 529–571.

Froot, K., Murphy, B., Stern, A., and Usher, S. (1995). The emerging asset class: Insurance risk. *Guy Carpenter & Company Inc.'s Review of Catastrophes Exposures and the Capital Markets*, July.

Ganapati, S., Retik, M., Puleo, P., and Starr, B. (1999). Catastrophe-linked securities. In D. F. Babbel and F. J. Fabozzi. (eds.), *Investment Management for Insurers* (pp. 209–234), New York: John Wiley & Sons.

Koutsaftis, V. (2000). The applications of insurance securitization: A new tool for risk managers; A new asset class for the capital markets and a new source of capital for the insurance industry. University of Chicago School of Business working paper.

Lakdawalla, D., and Zanjani, G. (2006). Catastrophe bonds, reinsurance, and the optimal collateralization of risk transfer. Federal Reserve of New York and Rand Corporation working paper.

Nell, M., and Richter, A. (2004). Improving risk allocation through indexed cat bonds. *Geneva Papers on Risk and Insurance* 29, 2: 183–201.

Woo, G. (2004). A catastrophe bond niche: Multiple event risk. Paper presented at the 2004 meeting of the NBER Insurance Project Group.

CHAPTER 38

Collateralized Debt Obligations

DOUGLAS J. LUCAS
Executive Director and Head of CDO Research, UBS

LAURIE S. GOODMAN, PhD
Co-head of Global Fixed Income Research Manager of U.S. Securitized Products Research, UBS

FRANK J. FABOZZ1, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract A collateralized debt obligation (CDO) is an asset-backed security backed by a diversified pool of one or more classes of debt (corporate loans, corporate bonds, emerging market bonds, asset-backed securities, residential mortgage-backed securities, commercial mortgage-backed, securities, and real estate investment trusts). The list of asset types included in a CDO portfolio is continually expanding. CDOs are categorized based on the motivation of the sponsor of the transaction: balance sheet, arbitrage, or origination. A synthetic CDO is so named because the CDO does not actually own the pool of assets on which it has the risk. Stated differently, a synthetic CDO absorbs the economic risks, but not the legal ownership, of its reference credit exposures. The nonsynthetic CDO is referred to as a "cash" structure. The building block for synthetic CDOs is a credit default swap, which allows the transfer of the economic risk of a pool of assets, but not the legal ownership, of underlying assets.

Keywords: collateralized debt obligations (CDOs), synthetic CDOs, arbitrage CDOs, balance sheet CDOs, origination CDOs, market value credit structures, cash flow credit structure, cash flow waterfall, coverage tests, overcollateralization tests, Interest coverage tests, par value test, overcollateralization trigger, pay-in-kind (PIK) feature, quality tests, diversity score, weighted average rating factor (WARF) recovery rates, diversification, loss distribution tests, full capital structure CDOs, single-tranche CDOs, standard tranches of credit default swap indices

In this chapter we provide the basics of *collateralized debt obligations (CDOs)*. We begin with an overview of cash CDOs. Then, we delve into more cash CDO details, including the cash flow credit structure, credit rating agencies' methodologies, interest rate hedging, and CDO call features. Finally, we discuss synthetic CDOs.

UNDERSTANDING CDOs

A CDO issues debt and equity and uses the money it raises to invest in a portfolio of financial assets such as corporate loans or mortgage-backed securities. It distributes the cash flows from its asset portfolio to the holders of its various liabilities in prescribed ways that take into account the relative seniority of those liabilities. This is just a starting definition; we will fill in the details of this definition over the next few pages.

Four Attributes of a CDO

Any CDO can be well described by focusing on its four important attributes: assets, liabilities, purposes, and credit structures. Like any company, a CDO has assets. With a CDO, these are financial assets such as corporate loans or mortgage-backed securities. And like any company, a CDO has liabilities. With a CDO, these run the gamut of preferred shares to AAA-rated senior debt. Beyond the seniority and subordination of CDO liabilities, CDOs have additional structural credit protections, which fall into the category of either cash flow or market value protections. Finally, every CDO has a purpose that it was created to fulfill, and these fall into the categories of *arbitrage*, balance sheet, or origination. In this section, we are going to look at the different types of assets CDOs hold, the different types of liabilities CDOs issue, the purposes for which CDOs are created, and the different credit structures CDOs employ.

Assets

CDOs own financial assets such as corporate loans or mortgage-backed securities. A CDO is primarily identified by its underlying assets.

Created in 1987, the first CDOs owned high-yield bond portfolios. In fact, before the term "CDO" was invented to encompass an ever-broadening array of assets, the term in use was "collateralized bond obligation" or "CBO." In 1989, corporate loans and real estate loans were used in CDOs for the first time, causing the term "collateralized loan obligation" or "CLO" to be coined. Generally, CLOs are comprised of performing high-yield loans, but a few CLOs, even as far back as 1988, targeted distressed and nonperforming loans. Some CLOs comprised of investment-grade loans have also been issued.

Loans and bonds issued by emerging market corporations and sovereign governments were first used as CDO collateral in 1994, thus "emerging-market CDO" or "EM CDO." In 1995, CDOs comprised of residential mortgage-backed securities (RMBS) were first issued. CDOs comprised of commercial mortgage-backed securities (CMBSs) and asset-backed securities (ABSs), or combinations of RMBS, CMBS, and, ABS followed, but they have never found a universally accepted name. We refer to them as a "structured finance CDO" or "SF CDO." However, Moody's champions the term "resecuritizations" and many others use "ABS CDO," even to refer to CDOs with CMBS and RMBS in their collateral portfolios.

Liabilities

Any company that has assets also has liabilities. In the case of a CDO, these liabilities have a detailed and strict ranking of seniority, going up the CDO's capital structure as equity or preferred shares, subordinated debt, mezzanine debt, and senior debt. These tranches of notes and equity are commonly labeled Class A, Class B, Class C, and so forth, going from top to bottom of the capital structure. They range from the most secured AAA-rated tranche with the greatest amount of subordination beneath it, to the most levered, unrated equity tranche. Table 38.1 shows a simplified tranche structure for a CLO.

Special purposes entities like CDOs are said to be "bankrupt remote." One aspect of the term is that they are new entities without previous business activities. They therefore cannot have any legal liability for sins of the past. Another aspect of their "remoteness from bankruptcy" is that the CDO will not be caught up in the bankruptcy of any other entity, such as the manager of the CDO's assets, or a party that sold assets to the CDO, or the banker that structured the CDO.

Another very important aspect of a CDO's bankruptcy remoteness is the absolute seniority and subordination of the CDO's debt tranches to one another. Even if it is a certainty that some holders of the CDO's debt will not receive their full principal and interest, cash flows from the CDO's assets are still distributed according to the original game plan dictated by seniority. The CDO cannot go into bankruptcy, either voluntarily or through the action of an aggrieved creditor. In fact, the need for bankruptcy is obviated because the distribution of the CDO's cash flows, even if the CDO is insolvent, has already been determined in detail at the origination of the CDO.

Within the stipulation, of strict seniority, there is great variety in the features of CDO debt tranches. The driving force for CDO structurers is to raise funds at the lowest possible cost. This is done so that the CDO's equity holder, who is at the bottom of the chain of seniority, can get the most residual cash flow.

Most CDO debt is floating rate off LIBOR (London Interbank Offered Rate), but sometimes a fixed rate

Table 38.1 Simple, Typical CLO Tranche Structure

Tranche	Percent of Capital Structure	Rating	Coupon
Class A	77.5	AAA	LIBOR + 26
Class B	9	A	LIBOR + 75
Class C	2.75	BBB	LIBOR + 180
Class D	2.75	BB	LIBOR + 475
Preferred shares	8	NR	Residual cash flow

tranche is structured. Avoiding an asset liability mismatch is one reason why floating rate, high-yield loans are more popular in CDOs than fixed rate, high-yield bonds. Sometimes a CDO employs short-term debt in its capital structure. When such debt is employed, the CDO must have a standby liquidity provider, ready to purchase the CDO's short-term debt should it fail to be resold or roll in the market. A CDO will only issue short-term debt if its cost, plus that of the liquidity provider's fee, is less than the cost of long-term debt.

Sometimes a financial guaranty insurer will wrap a CDO tranche. Usually this involves a AAA-rated insurer and the most senior CDO tranche. Again, a CDO would employ insurance if the cost of the tranche's insured coupon plus the cost of the insurance premium is less than the coupon the tranche would have to pay in the absence of insurance. To meet the needs of particular investors, sometimes the AAA tranche is divided into senior AAA and junior AAA tranches.

Some CDOs do not have all their assets in place when their liabilities are sold. Rather than receive cash that the CDO is not ready to invest, tranches might have a delay draw feature, where the CDO can call for funding within some specified time period. This eliminates the negative carry that the CDO would bear if it had to hold uninvested debt proceeds in cash. An. extreme form of funding flexibility is a revolving tranche, where the CDO can call for funds and return funds as its needs dictate.

Purposes

CDOs are created for one of three purposes:

Balance sheet. A holder of CDO-able assets desires to (1) its balance sheet, (2) reduce required regulatory capital, (3) reduce required economic capital, or (4) achieve cheaper funding costs. The holder of these assets sells them to the CDO. The classic example of this is a bank that has originated loans over months or years and now wants to remove them from its balance sheet. Unless the bank is very poorly rated, CDO debt would not be cheaper than the bank's own source of funds. But selling the loans to a CDO removes them from the bank's balance sheet and therefore lowers the bank's regulatory capital requirements. This is true even if market practice requires the bank to buy some of the equity of the newly created CDO.

Arbitrage. An asset manager wishes to gain assets under management and management fees. Investors wish to have the expertise of an asset manager. Assets are purchased in the marketplace from many different sellers and put into the CDO. CDOs are another means, along with mutual funds and hedge funds, for an asset management firm to provide its services to investors. The difference is that instead of all the investors sharing the fund's return in proportion to their investment, investor returns are also determined by the seniority of the CDO tranches they purchase.

Origination. Banks, insurance companies, and real estate investment trusts (REITs) wish to increase equity capital. Here the example is a large number of smaller-size banks issuing trust-preferred securities directly to the CDO simultaneous with the COO's issuance of its own liabilities. (Trust-preferred, securities are unsecured obligations that are generally ranked lowest in the order of repayment.) The bank capital notes would not be issued but for the creation of the CDO to purchase them.

Three purposes differentiate CDOs on the basis of how they acquire their assets and focus on the motivations of asset sellers, asset managers, and trust preferred securities issuers. From the point of view of CDO investors, however, all CDOs have a number of common purposes, which explain why many investors find CDO debt and equity attractive.

One purpose is the division and distribution of the risk of the CDO's assets to parties that have different risk appetites. Thus, a AAA investor can invest in speculative-grade assets on a loss-protected basis. Or a BB investor can invest in AAA assets on a levered basis.

For CDO equity investors, the CDO structure provides a leveraged return without some of the severe adverse consequences of borrowing via repo from a bank. CDO equity holders own stock in a company and are not liable for the losses of that company. Equity's exposure to the CDO asset portfolio is therefore capped at the cost of equity minus previous equity distributions. Instead of short-term bank financing, financing via the CDO is locked in for the long term at fixed spreads to LIBOR.

Credit Structures

Beyond the seniority and subordination of CDO liabilities, CDOs have additional structural credit protections, which fall into the category of either cash flow or market value protections.

The *market value credit structure* is less often used, but easier to explain, since it is analogous to an individual's margin account at a brokerage. Every asset in the CDO's portfolio has an advance rate limiting the amount that can be borrowed against that asset. Advance rates are necessarily less than 100% and vary according to the market value volatility of the asset. For example, the advance rate on a fixed rate B-rated bond would be far less than the advance rate on a floating rate AAA-rated bond, Both the rating and floating rate nature of the AAA bond indicate that its market value will fluctuate less than the B-rated bond. Therefore, the CDO can borrow more against it. The sum of advance rates times the market values of associated assets is the total amount the CDO can borrow.

The credit quality of a market value CDO derives from the ability of the CDO to liquidate its assets and repay debt tranches. Thus, the market value of the CDO's assets are generally measured every day, advance rates applied, and the permissible amount of debt calculated. If this comes out, for example, to \$100 million, but the CDO has \$110 million of debt, the CDO must do one of two things. It can sell a portion of its assets and repay a portion of its debt until the actual amount of debt is less than the permissible amount of debt. Or the CDO's equity holders can contribute more cash to the CDO. If no effective action is taken, the entire CDO portfolio is liquidated, all debt is repaid, and residual cash given to equity holders. The market value credit structure is analogous to an individual being faced with a collateral, call at his (or her) brokerage account. If he does not post additional collateral, his portfolio is at least partially liquidated.

The *cash flow credit structure* does not have market value tests. Instead, subordination is sized so that the after-default cash flow of assets is expected to cover debt tranche principal and interest with some degree of certainty. Obviously, the certainty that a AAA CLO tranche, with 23% subordination beneath it, will receive all its principal and interest is greater than the certainty a BB CLO tranche, with only 8% subordination beneath it, will receive all its principal and interest.

Most cash flow CDOs have overcollateralization and interest coverage tests. These tests determine whether collateral cash flow is distributed to equity and subordinate debt tranches or instead diverted to pay down senior debt tranche principal or used to purchase additional collateral assets. We will discuss these tests in detail later in this chapter, but their purpose is to provide additional credit enhancement to senior CDO debt tranches,

A CDO Structural Matrix

Table 38.2 shows the four CDO building blocks and a variety of options beneath each one. Any CDO can be well described by asking and answering the four questions implied by the exhibit:

- What are its assets?
- What are the attributes of its liabilities?
- What is its purpose?
- What is its credit structure?

This way of looking at CDOs encompasses all the different kinds of CDOs that have existed in the past and all the kinds of CDOs that are currently being produced. By adding "synthetic asset option" and "unfunded super senior" to the matrix, the matrix also encompasses synthetic CDOs, a type of CDO we discuss in detail later in this chapter.

Parties to a CDO

A number of parties and institutions contribute to the creation of a CDO.

CDO Issuer and Coissuer
A CDO is a distinct legal entity, usually incorporated in the Cayman Islands. Its liabilities are called CDOs, so one might hear the seemingly circular phrase "the CDO issues CDOs." Offshore incorporation enables the CDO to more easily sell its obligations to United States and international investors and escape taxation at the corporate entity level. When a CDO is located outside the U.S., it will typically also have a Delaware coissuer. This entity has a passive role, but its existence in the structure allows CDO obligations-to be more easily sold to U.S. insurance companies.

Asset Manager (Collateral Manager)
Asset managers (or collateral managers) select the initial portfolio of an arbitrage CDO and manage it according to prescribed guidelines contained in the CDO's indenture. Sometimes an asset manager is used in a balance sheet CDO of distressed assets to handle their workout or sale. A variety of firms offer CDO asset management services including hedge fund managers, mutual fund managers, and firms that specialize exclusively in CDO management.

Asset Sellers
Asset sellers supply the portfolio for a balance sheet CDO and typically retain its equity. In cash CDOs, the assets involved are usually smaller-sized loans extended to smaller-sized borrowers. In the United States, these are called "middle market" loans and in Europe these are called "small and medium enterprise" (SME) loans.

Investment Bankers and Structurers
Investment bankers and structurers work with the asset manager or asset seller to bring the CDO to fruition. They set up corporate entities, shepherd the CDO through the debt rating process, place the CDO's debt and equity with investors, and handle other organizational details. A big part of this job involves structuring the CDO's liabilities: their size and ratings, the cash diversion features of the structure, and, of course, debt tranche coupons. To obtain the cheapest funding cost for the CDO, the structurer must know when to use short-term debt or insured debt or senior/junior AAA notes, to name just a few structural options. Another part of the structurer's job is to negotiate an acceptable set of eligible assets for the CDO. These tasks obviously involve working with and balancing the

Table 38.2 CDO Structural Matrix

Assets	Liabilities	Purpose	Credit Structure
High-yield loans	Fixed/floating rate	Arbitrage	Cash flow
High-grade structured finance	PIK/non-PIK	Balance sheet	Market value
Mezzanine structured finance	Guaranteed/unenhanced	Origination	
Capital notes	Short term/long term		
High-yield bonds	Delayed draw/revolving		
Emerging market debt			
Synthetic assets	Unfunded super senior		

desires of the asset manager or seller, different debt and equity investors, and rating agencies.

Insurers/Guarantors

Monoline bond insurers or financial guarantors typically only guarantee the senior-most tranche in a CDO. Often, insurance is used when a CDO invests in newer asset types or is managed by a new CDO manager.

Rating Agencies

Rating agencies approve the legal and credit structure of the CDO, perform due diligence on the asset manager and the trustee, and rate the various seniorities of debt issued by the CDO. Usually two or three of the major rating agencies (Moody's, S&P, and Fitch) rate the CDO's debt. DBRS is a recent entrant in CDO ratings and A. M. Best has rated CDOs backed by insurance company trust preferred securities.

Trustees

Trustees hold the CDO's assets for the benefit of debt and equity holders, enforce the terms of the CDO indenture, monitor and report upon collateral performance, and disburse cash to debt and equity investors according to set rules. As such, their role also encompasses that of collateral custodian and CDO paying agent.

CASH FLOW CDOs

As explained earlier, arbitrage CDOs are categorized as either cash flow transactions or market value transactions. The objective of the asset manager in a cash flow transaction is to generate cash flow for CDO tranches without the active trading of collateral. Because the cash flows from the structure are designed to accomplish the objective for each tranche, restrictions are imposed on the asset manager. The asset manager is limited in his or her authority to buy and sell bonds. The conditions for disposing of issues held are specified and are usually driven by credit risk management. Also, in assembling the portfolio, the asset manager must meet certain requirements set forth by the rating agency or agencies that rate the deal.

In this section, we review cash flow transactions. Specifically, we look at the distribution of the cash flows, restrictions imposed on the asset manager to protect the noteholders, and the key factors considered by rating agencies in rating tranches of a cash flow transaction. We focus on establishing a basic understanding of cash flow CDO deals using examples. (For a discussion of deals based by other types of collateral, see Lucas, Goodman, and Fabozzi (2006).

Distribution of Cash Flows

In a cash flow transaction, the cash flows from income and principal are distributed according to rules set forth in the prospectus. The distribution of the cash flows is referred to as the "waterfall." We describe these rules below and will use a representative CDO to illustrate them.

The representative CDO deal we will use is a $300 million cash flow CDO with a "typical" cash flow structure. The deal consists of the following:

- $260 million (87% of the deal) Aaa/AAA (Moody's/S&P) floating rate tranche.
- $27 million ($17 million fixed rate + $10 million floating rate) Class B notes, rated A3 by Moody's.
- $5 million (fixed rate) Class C notes, rated Ba2 by Moody's.
- $8 million in equity (called "preference shares" in this deal).

The collateral for this deal consists primarily of investment-grade, CMBS, ABS, RETT, and RMBS; 90% of which must be rated at least "Baa3" by Moody's or BBB– by S&P. (At the time of purchase, the collateral corresponded, on average, to a Baa2 rating.) The asset manager is a well-respected money management firm.

Figure 38.1 illustrates the priority of interest distributions among different classes for our sample deal. Interest payments are allocated first to high priority deal expenses such as fees, taxes, and registration, as well as monies owed to the asset manager and hedge counterparties. After these are satisfied, investors are paid in a fairly straightforward manner, with the more senior bonds paid off first, followed by the subordinate bonds, and then the equity classes.

Note the important role in the waterfall played by what is referred to as the *coverage tests.* We explain these shortly.

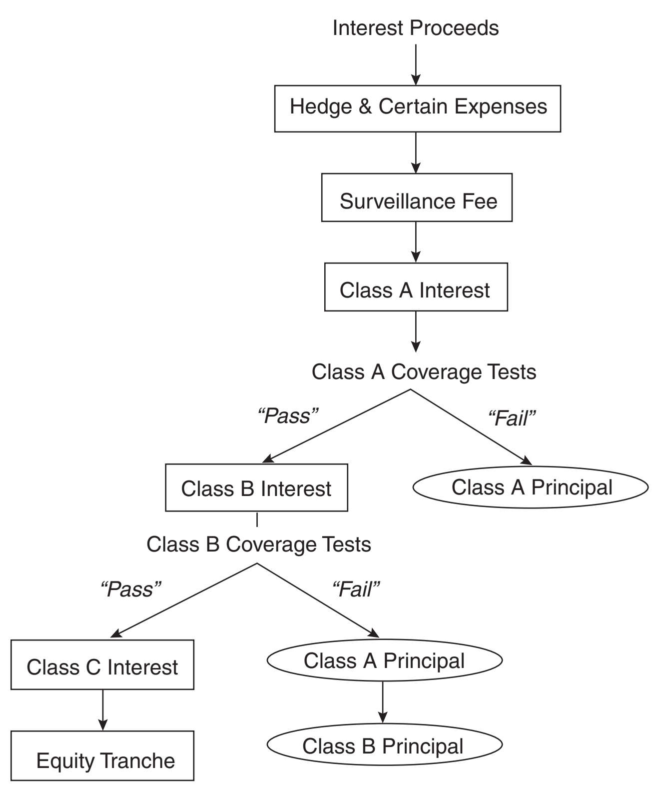


Figure 38.1 Interest Cash Flow "Waterfall"

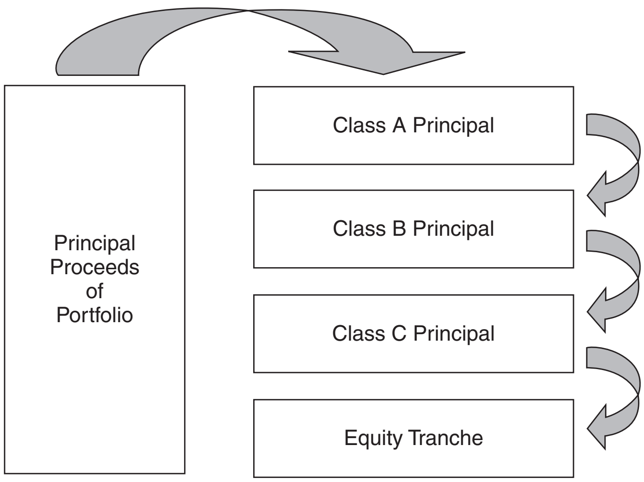


Figure 38.2 Principal Cash Flow Waterfall

They are important because, before any payments are made on Class B or Class C bonds, coverage tests are run to assure the deal is performing within guidelines. If that is not the case, consequences to the equity holders are severe. Note from Table 38.1 if the Class A coverage tests are violated, then excess interest on the portfolio goes to pay down principal on the Class A notes, and cash flows will be diverted from all other classes to do so. If the portfolio violates the Class B coverage tests, then interest will be diverted from Class C and the equity tranche to pay down first principal on Class A, or, if Class A is retired, Class B principal.

Figure 38.2 shows the simple principal cash flows for this deal. Principal is paid down purely in class order. Any remaining collateral principal from overcollateralization gets passed on to the equity piece.

Restrictions on Management: Safety Nets

Noteholders have two major protections provided in the form of tests. They are coverage tests and quality tests. We discuss each type in this section.

Coverage Tests

Coverage tests are designed to protect noteholders against a deterioration of the existing portfolio. There are actually two categories of tests—*overcollateralization tests* and *interest coverage tests.*

Overcollateralization Tests The overcollateralization or O/C ratio for a tranche is found by computing the ratio of the principal balance of the collateral portfolio over the principal balance of that tranche and all tranches senior to it. That is,

$$\text{O/C ratio for a tranche} = \frac{\text{Principal (par) value of collateral portfolio}}{\text{Principal of tranche} + \text{Principal of all tranches senior to it}}$$

The higher the ratio, the greater protection for the note holders. Note that the overcollateralization ratio is based on the principal or par value of the assets, (Hence, an overcollateralization test is also referred to as a *par value test.)* An overcollateralization ratio is computed for specified tranches subject to the overcollateralization test. The overcollateralization test for a tranche involves comparing the tranche's overcollateralization ratio to the tranche's required minimum ratio as specified in the CDO's guidelines. The required minimum ratio is referred to as the *overcollateralization trigger.* The overcollateralization test for a tranche is passed if the overcollateralization ratio is greater than or equal to its respective overcollateralization trigger. (Note that for market value CDOs, overcollateralization tests are based on market values rather than principal or par values.)

Consider our representative CDO. There are two rated tranches subject to the overcollateralization test—Classes A and B. Therefore, two overcollateralization ratios are computed for this deal. For each tranche, the overcollateralization test involves first computing tine overcollateralization ratio as follows:

$$\text{O/C ratio for Class A} = \frac{\text{Principal (par) value of collateral portfolio}}{\text{Class A prinicpal}}$$

$$\text{O/C ratio for Class B} = \frac{\text{Principal (par) value of collateral portfolio}}{\text{Class A prinicpal} + \text{Class B prinicpal}}$$

Once the overcollateralization ratio for a tranche is computed, it is then compared to the overcollateralization trigger for the tranche as specified in the guidelines. If the computed overcollateralization ratio is greater than or equal to the overcollateralization trigger for the tranche, then the test is passed with respect to that tranche.

For our representative deal, the overcollateralization trigger is 113% for Class A and 101% for Class B. Note that the lower the seniority, the lower the overcollateralization trigger. The Class A overcollateralization test is failed if the ratio falls below 113% and the Class B overcollateralization test is failed if the ratio falls below 101%.

Interest Coverage Test The interest coverage or I/C ratio for a tranche is the ratio of scheduled interest due on the underlying collateral portfolio to scheduled interest to be paid to that tranche and all tranches senior to it. That is,

$$\text{I/C ratio for a tranche} = \frac{\text{Scheduled interest due on underlying collateral portfolio}}{\begin{array}{c}\text{Scheduled interest}\\ \text{to that tranche}\end{array} + \begin{array}{c}\text{Scheduled interest to}\\ \text{all tranches senior}\end{array}}$$

The higher the interest coverage ratio, the greater the protection. An interest coverage ratio is computed for specified tranches subject to the interest coverage test. The interest coverage test for a tranche involves comparing the tranche's interest coverage ratio to the tranche's interest coverage trigger (that is, the required minimum ratio as specified in the guidelines). The interest coverage test for a tranche is passed if the computed interest coverage ratio is greater than or equal to its respective interest coverage trigger.

For our representative deal, Classes A and B are subject to the interest coverage test. The following two interest coverage ratios are therefore computed:

$$\text{I/C ratio for Class A} = \frac{\text{Scheduled interest due on underlying collateral portfolio}}{\text{Class A scheduled interest}}$$

$$\text{I/C ratio for Class B} = \frac{\text{Scheduled interest due on underlying collateral portfolio}}{\text{Class A scheduled interest} + \text{Class B scheduled interest}}$$

In the case of our representative deal, the Class A interest coverage trigger is 121%, while the Class B interest coverage trigger is 106%.

PIKing Occurs When Coverage Tests are Not Met We showed in Figure 38.1 that if the Class A coverage tests are violated, the excess interest on the portfolio goes to pay down principal on the Class A notes, and cash flows is diverted from the other classes to do so. In this case, what happens to the Class B notes?

They have a *pay-in-kind (PIK) feature*. This is a clearly disclosed structural feature in most CDOs where, instead of paying a current coupon, the par value of the bond is increased by the appropriate amount. So if a \$5 coupon is missed, the par value increases, say from \$100 to \$105. The next coupon is calculated based on the larger \$105 par amount. The PIK concept originated in the high-yield, market, and was employed for companies whose future cash flows were uncertain. The option to pay-in-kind was designed to help these issuers conserve scarce cash or even avoid default. It was imported to the CDO market as a structural feature to enhance the more senior classes.

The PIKability of subordinate tranches and the diversion of cash flows to cause early amortization of the Class A tranche naturally strengthens the Class A tranche. The Class A tranche can therefore either achieve a higher rating, or its size can be increased while still maintaining its original rating. CDO equity holders benefit from an overall lower cost of funds: They either have a lower coupon on the Class A tranche; or the Class A tranche, which enjoys the CDO's lowest funding cost, is larger. Either case lowers interest costs to the CDO and thus increases return to equity holders.

The effectiveness of PIK-ing in bolstering the credit quality of the Class A tranche depends upon the amount of collateral cash flow that exists in excess of Class A coupon. The higher the coupon on collateral, and the longer the tenor of collateral, then the more cash flow potentially available for diversion to pay down Class A principal. The effectiveness of PIKing (in bolstering the Class A tranche) also depends upon the looseness or tightness of the overcollateralization and interest coverage tests. The tighter the coverage tests are to the CDO's original par and coupon ratios, the sooner a deterioration in those ratios will cause cash flow to be diverted to repay Class A principal.

The effect of cash diversion to the Class A tranche in a high-yield-backed CDO can be dramatic. It is not unusual for subordinate tranches of a CDO to have been downgraded (and to be PIKing without any chance of ultimate payment) while the CDO's Aaa tranche maintains its credit quality and rating. That is due to the outlook for Class A receiving full principal and interest because of the diversion of cash to Class A principal.

In determining its optimal capital structure, CDO equity must weigh reduction in the overall cost of CDO debt against the potential for equity to receive less cash flow in severe default scenarios. Distribution of collateral cash flow among tranches in a CDO is a zero-sum game. And since equity receives residual cash flow after debt tranches are satisfied, PIK-ing and the diversion of cash flows to Class A principal affects it the most. First, the CDO's average cost of funds increases. Second, the CDO becomes more delevered. Finally, less cash reaches the equity tranche, and that which does is delayed.

Quality Tests

After the tranches of a CDO deal are rated, the rating agencies are concerned that the composition of the collateral portfolio may be adversely altered by the asset manager over time. Tests are imposed to prevent the asset manager from trading assets so as to result in a deterioration of the quality of the portfolio and are referred to as *quality tests*. These tests deal with maturity restrictions, the degree of diversification, and credit ratings of the assets in the collateral portfolio.

Credit Ratings

There are three key inputs to cash flow CDO ratings: collateral diversification, likelihood of default, and recovery rates. While each rating agency uses a slightly different methodology, they reach similar conclusions. For this analysis, we use a variation of Moody's methodology, as it is the most transparent and allows us to change inputs to show the import and impact of each.

Moody's uses the same objective process for developing liability structures regardless of the type of collateral, Moody's determines losses on each tranche under different default scenarios, and probability-weight those results. The resulting "expected loss" is then compared to the maximum permitted for any given rating. While that whole iterative process makes for a tedious analysis, it does help highlight why, for example, a deal backed by investment-grade corporate bonds will have a very high proportion of triple A tranches and a low proportion of equity compared to a deal backed by high-yield corporate bonds.

Collateral Diversification

Moody's methodology reduces the number of credits in the CDO portfolio to a smaller number of homogenous, uncorrelated credits. For example, for CDOs backed by corporate bonds, a *diversity score* is calculated by dividing the bonds into different industry classifications. Each industry group is assumed to have zero correlation with other industry groups. Two securities from different issuers within the same industry group are assumed, to have some correlation to each other. At the extreme, two securities from the same issuer are treated as having 100% correlation and thus providing zero diversification.

Reducing the portfolio to the number of independent securities allows the use of a binomial probability distribution. This is the distribution that allows one to figure out the probability of obtaining 9 "heads" in 10 flips of the coin. This distribution can also be applied to a weighted coin, where the probability of "heads" is substantially different than the probability of tails. Intuitively, each asset is a separate flip of the coin, and the outcomes ("heads" and "tails") corresponds to "no default" and "default." The use of this probability distribution makes it possible to define the likelihood of a given number of securities in the portfolio defaulting over the life of a deal.

One factor concerning investors in CDOs is the potential for the default on one bond to wipe out the equity. In fact, in addition to the general diversification methodology, there are single-name concentration rules that protect against too large a concentration within securities issued by any single entity. It is customary for issuer exposure to be no more than 2%. To allow asset managers some flexibility, a few exceptions are permitted. In one actual deal, for example, four positions could be as large as 3%, as long as no more than two of these exposures were in the same industry. If two of the exposures greater than 2% were in the same industry, additional restrictions apply.

Historical Defaults

Likelihood of default is provided by the *weighted average rating factor (WARF)*. This is a rough guide to the asset quality of a portfolio and is meant to incorporate the probability of default for each of the bonds backing a CDO. To see where this comes from, we need to look at actual default experience on corporate bonds.

Table 38.3 shows actual average cumulative default rates from 1 to 10 years based on Moody's data from 1983 to 2004. These data show that bonds with an initial rating of Baa3 experienced average default rates of 5.36% after 7 years, and 7.20% after 10 years. Compare that to the Bl default rate of 35.69% after 7 years and 47.43% after 10 years. Generally, as would be expected, bonds with lower ratings exhibit higher default patterns. Moreover, defaults rise exponentially, not linearly, as ratings decline.

However, it is difficult to use these data to construct a stylized default pattern, as some anomalies appear. For example, over some time periods, Aaa bonds default more frequently than do Aal bonds. And Aa2 bonds default more frequently than either Aa3 or Al bonds, while A2 bonds default more frequently than A3 bonds. Correspondingly, B2 bonds default less frequently than either Ba3 or Bl bonds.

Moody's smooths these data and constructs a WARF, shown in Table 38.4. Thus, a bond with a Baal rating has a Moody's score of 260, while one rated Baa3 would have a WARF score of 610. Note that these scores exhibit the same pattern as did actual default numbers: Scores are nonlinear and increase exponentially as ratings decline. These scores are also dollar-weighted across the portfolio to deliver a WARF for the portfolio.

The WARF for the portfolio translates directly into a cumulative probability of default. The cumulative probability of default will be larger the longer the portfolio is outstanding. A WARF score of 610 means that there is a 6.1% probability of default for each of the independent, uncorrelated assets defaulting in a 10-year period.

Table 38.3 Average Issuer-Weighted Cumulative Default Rates by Alphanumeric Rating, 1983–2004 Moody's

	Time Horizon (Years)									
Cohort Hating	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Aaa	0.00	0.00	0.00	0.06	0.18	0.24	0.32	0.40	0.40	0.40
Aa1	0.00	0.00	0.00	0.15	0.15	0.25	0.25	0.25	0.25	0.25
Aa2	0.00	0.00	0.04	0.13	0.28	0.34	0.40	0.48	0.57	0.67
Aa3	0.00	0.00	0.05	0.11	0.18	0.26	0.26	0.26	0.26	0.33
A1	0.00	0.00	0.19	0.30	0.38	0.47	0.50	0.58	0.67	0.84
A2	0.03	0.08	0.22	0.47	0.68	0.89	1.05	1.34	1.59	1.69
A3	0.03	0.21	0.37	0.50	0.65	0.86	1.19	1.38	1.55	1.69
Baa1	0.17	0.50	0.84	1.14	1.46	1.69	1.92	2.05	2.21	2.31
Raa2	0.12	0.40	0.81	1.52	2.11	2.74	3.39	3.98	4.62	5.49
Baa3	0.41	1.07	1.70	2.66	3.60	4.49	5.36	6.15	6.68	7.20
Ba1	0.66	2.07	3.55	5.23	6.76	8.67	9.70	10.85	11.61	12.38
Ba2	0.62	2.22	4.48	6.84	8.82	10.11	11.85	13.13	14.20	14.66
Ba3	2.23	6.10	10.62	15.03	19.14	23.05	26.56	30.00	33.35	36.24
B1	3.03	8.89	14.81	20.09	25.27	30.29	35.69	39.97	43.98	47.43
B2	5.93	13.73	20.58	26.58	31.24	34.54	37.39	39.60	42.19	44.48
B3	10.77	20.43	29.01	36.82	43.55	49.74	54.46	58.40	61.02	62.32
Caa-C	22.24	35.80	46.75	54.60	60.40	65.15	68.30	72.36	75.38	78.81
Investment grade	0.08	0.23	0.43	0.71	0.96	1.21	1.43	1.65	1.84	2.03
Speculative grade	5.26	10.84	16.06	20.63	24.54	28.00	31.04	33.63	35.87	37.66
All rated	1.79	3.66	5.38	6.89	8.13	9.17	10.04	10.75	11.35	11.83

Source: Exhibit 17 in David T. Hamilton, Praveen Vama, Sharon Ou, and Richard Cantor, *Default and Recovery Rates of Corporate Bond Issuers: 1920–2004*, Moody's Investors Service (January 2005), p. 17.

Table 38.4 Moody's Weighted Average Rating Factor

Rating	WARF	Rating	WARF	Rating	WARF
Aaa	1	Baa1	260	B1	2,220
Aa1	10	Baa2	360	B2	2,720
Aa2	20	Baa3	610	B3	3,490
Aa3	40	Ba1	940	Caa1	4,770
A1	70	Ba2	1,350	Caa2	6,500
A2	120	Ba3	1,780	Caa3	8,070
A3	180			Ca/C	10,000

Source: Moody's Investors Service.

(In general, the WARF score translates directly into the 10-year "idealized" cumulative default rate.) The same 610 WARF would correspond to a 4.97% probability of default after eight years, or a 5.57% probability of default after nine years.

When the desired rating on the CDO tranche is the same as the rating on the underlying collateral, Moody's uses the probability of default derived from the WARF score. For CDO ratings higher than the ratings on their underlying collateral, Moody's will use a higher default rate. The multiple applied to the idealized cumulative default rate is referred to as a stress factor. Thus, for example, in an investment-grade deal (Baa-rated collateral), Moody's uses a factor of 1.0 to rate a Baa tranche. If the rating on the CDO tranche is Aaa, Aa, or A, then Moody's uses a higher factor to stress the default rates.

Recovery Rates

Moody's recovery rates are dependent on the desired rating of the CDO tranche. To obtain the highest ratings (Aaa and Aa), Moody's generally assumes recovery rates of 30% on unsecured corporate bonds. To obtain an A or Baa rating, recovery assumptions are slightly higher, at 33% and 36%, respectively. It should be understood that actual average recovery rates are higher than these assumptions. A Moody's study covering the period 1981 to 2004 showed that the median, or midpoint, recovery rate for senior unsecured debt was $45.20 ($44.90 average or mean). For subordinated unsecured debt, the median recovery rate was $33.40 ($32.00 average). The bottom line is this: Moody's is again conservative, as it uses a recovery value consistent with subordinated unsecured debt on debt that is in most cases senior—and that builds in "extra" protection for the investors.

Putting It All Together

Moody's has an expected loss permissible for each CDO rating. That expected loss is derived as follows:

$$\text{Expected loss} = \sum_{i=1}^{n} (\text{Loss in default scenario } i) \times (\text{Probability of scenario } i \text{ occuring})$$

The following example, using an investment-grade corporate CDO, helps clarify this formula. Assume a typical CDO deal with 45 independent assets. Assume further that we are looking at a 10-year deal in which each asset has a probability of default of 5% corresponding to a WARF score of 500, which is well within the category of Baa-rated assets. Moreover, we assume a capital structure with 85% of the bonds Aaa-rated, 10% Baa-rated, and 5% equity. The recovery rate is assumed to be 30%.

To create an example that can be replicated with a simple spreadsheet, we assume all interim cash flows are distributed, and all defaults occur at the end of the life of the deal. Moody's actually runs each scenario through its CDO cash flow model in order to determine the loss to each bond in the CDO structure. Then Moody's assumes a number of different loss schedules and select the most detrimental.

We have simplified that whole analytical process to make it more transparent. Our methodology overstates losses to the bondholders, since we ignored all overcollateralization and interest coverage tests. As the portfolio deteriorated, those two tests kick in and would cut off cash flow to the equity tranche, redirecting cash flows to pay down the higher-rated tranches. We have also ignored the excess spread on these deals, which provides a very important cushion to the noteholders.

The probability of a scenario in which none of the 45 securities default is (probability of no default)^{45}, or $(0.95)^{45}$. This is equal to 9.94%. If there are zero defaults, there is obviously no loss. The probability of only one loss is found as follows:

$$[(\text{Probability of no default})^{44} \times (\text{Probability of 1 default}) \times 45] = (0.95)^{44} \times 0.95 \times 45 = 23.55\%$$

This frequency distribution for a selected number of defaults is shown in the column of Table 38.5, labeled "Probability."

With one default, the defaulted bond comprises 1/45 of the portfolio, or 2.22%. However, since a 30% recovery rate is assumed, that loss is lowered to 1.56% (2.22 x 0.7). Thus, the "Portfolio Loss" column of Table 38.5 shows that the loss with one default would be 1.56%. But the 5% equity in the deal acts as a buffer, and there would be no loss to the BBB bond. In order to impact the BBB bond, losses must total more than 5%.

Assume four defaults among the 45 assets. This means that 8.89% of the assets (4/45) are defaulting, and portfolio loss becomes 6.22% (8.89% × 0.7). The probability of this occurring is 11.37%. If that case does occur, the Baa bond would lose 12.22% of its value. That is, the equity would be eliminated, and the $10 Baa tranche ($10 per $100 par value), would be reduced by ($6.22 − $5.00), or $1.22, for a 12.22% reduction. Thus,

$$[(\text{Baa loss}) \times (\text{Probability of loss})] = 1.38\%$$

or

$$[(11.37\% \text{ probability of scenario}) \times (12.22\% \text{ loss if scenario materializes})]$$

Similarly, if there were five defaults (a 4.92% probability), the portfolio loss would be 7.78%. This corresponds to a loss of 27.78% on the Baa bond. The expected loss to the Baa bond in this scenario is (4.91 × 27.78), or 1.3629%. Note that if portfolio losses total more than 15%, the Baa

Table 38.5 Expected Loss on BBB Class, Investment-Grade CDO Deal (Given 45 Assets)

No. of securities: 45
Default probability: 5%
Loss given default: 70%

Portfolio loss for single default: 1.56% (1/45 × 70%)
Expected BBB loss: 3.9205%

No. of Defaults	Portfolio Loss (%)	Probability (%)	BBB Loss (%)	BBB Loss × Probability (%)
0	0.00	9.94	0.00	0.0000
1	1.56	23.55	0.00	0.0000
2	3.11	27.27	0.00	0.0000
3	4.67	20.57	0.00	0.0000
4	6.22	11.37	12.22	1.3895
5	7.78	4.91	27.78	1.3629
27	42.00	0.00	100.00	0.0000
28	43.56	0.00	100.00	0.0000
29	45.11	0.00	100.00	0.0000
30	46.67	0.00	100.00	0.0000
31	48.22	0.00	100.00	0.0000
32	49.78	0.00	100.00	0.0000
42	65.33	0.00	100.00	0.0000
43	66.89	0.00	100.00	0.0000
44	68.44	0.00	100.00	0.0000
45	70.00	0.00	100.00	0.0000

bond is eliminated, and only then does the Aaa bond, start incurring losses.

Adding expected losses in each of the scenarios across the binomial probability distribution, we find that the expected loss on this Baa CDO tranche is 3.92%. Realize again that this example is for illustrative purposes and will overstate losses to the bondholders. It ignores overcollateralization and interest coverage ratios and the excess spread in the deal.

Importance of Diversification We can now readily show the importance of diversification. No matter how many assets we have, if the probability of default on each is 5% and recovery is 30%, then the expected loss on the portfolio is 3.5%. However, this does not address the distribution of losses, which is certainly important to the bondholders.

In fact, the Baa bondholders are concerned about the likelihood of losses exceeding the amount of equity in the deal, while the Aaa bondholders are concerned about the likelihood of losses exceeding the amount of equity and Baa bonds. The greater the number of assets, the greater the likelihood that losses on those assets will cluster around 3.5% and the lower the likelihood that losses will exceed the 5% equity cushion and impact the Baa piece. On the flipside, the smaller the number of assets, the greater the likelihood that losses exceed the 5% equity cushion and will hit the Baa bonds.

Figure 38.3 shows probability distributions for losses on pools of 15, 30, and 45 securities. Note that the fewer the number of assets, the greater likelihood that losses will exceed a 5% equity cushion.

Table 38.6 supports the point that with fewer assets, expected losses to the Baa-rated tranche are much higher. Thus, for 15 assets, the loss to the Baa tranche is 9.15%; for 30 assets it is 5.62%. For 45 assets, the loss to the Baa tranche is 3.92%; and for 60 assets, it is 2.92%, Note also that the benefits of diversification diminish as more assets are added. The loss to the Baa tranche is 5.5% lower in

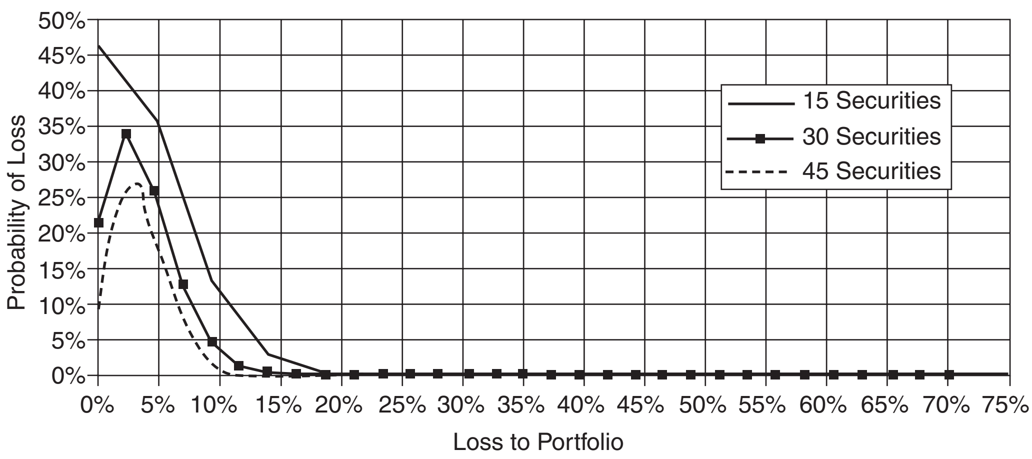


Figure 38.3 Benefits of Diversification

Table 38.6 Diversity and Expected Losses, %

No. of Securities	15	20	25	30	45	60
Aaa losses	0.0273	0.0091	0.0032	0.0012	0.0001	0.0000
Baa losses	9.1520	8.5074	6.8720	5.6216	3.9205	2.9262

moving from 15 to 30 assets. It only drops 1.7% in moving from 30 to 45 assets and only 1% from 45 to 60 assets.

What's "Too Much" Diversification? The above analysis suggests that greater diversification is always better, since it means less variation of collateral returns. However, a higher diversity score also means that it may be likely the asset manager pushed for, and achieved, less equity in the deal. In fact, with a diversity score of 60, the same losses on the Aaa and Baa bonds could have been achieved with less equity (on the order of 4.5% rather than the 5% required on a deal with a diversity score of 45).

Is there any such thing as too much of a diversification "good thing"? That depends on the asset manager. A large, broad-based asset manager may have considerable strength across all sectors and should be able to handle the analysis—and risks-—of a highly diverse portfolio. Even here, a very high diversity score can limit flexibility by requiring an asset manager with broad expertise to invest in an industry he does not like. Whether or not flexibility is being limited too much by a very high diversity score is dependent on the range of assets employed and the strengths of a particular asset manager.

Too much diversification, is even more a major problem for a smaller asset manager, where the portfolio may have selective strengths in fewer industries. This asset manager may be stretching to take on additional diversity to achieve a lower required equity. Investors should certainly be wary of deals in which very high diversity scores are achieved by managers straying from their fields of expertise.

Loss Distribution Tests

As can be seen from the discussion above, Moody's approach to rating CDOs involves (1) developing a diversity score; (2) calculating a weighted average rating factor; (3) using the binomial distribution to determine the probability of a specific number of defaults; and (4) calculating the impact of those defaults on bonds within the CDO structure. One element needed to calculate that impact is a distribution of defaults and losses across time. Let us look at this distribution of defaults and losses.

Moody's stresses CDOs via six different loss distributions, and a CDO tranche must pass each test. Moody's basic approach assumes 50% of the losses will occur at a single point in time, and that remaining losses are evenly distributed over a 5-year period. This single 50% loss is assumed to occur at a different point in each of the six tests.

Liability Structure

The structure of the liabilities will be primarily determined by the credit quality of the assets, the amount of diversification, and excess spread. That is, the combination of credit quality, diversification of assets, and excess spread dictate expected losses on each tranche. That is then compared to losses allowed to achieve a given rating. Realize that the structures have been optimized. If a structurer sees one of the tranches passing expected loss tests by a large margin, that means there is room to improve the arbitrage. That can be accomplished by leveraging the structure more (that is, reducing equity, reducing the amount of mezzanine bonds, or both).

Uses of Interest Rate Swaps and Caps in CDO Transactions

We have mentioned that a wide variety of collateral can be used to back CDO deals. Some of this collateral (high-yield bonds, investment-grade bonds) have fixed rate coupons, some (high-yield loans) have floating-rate coupons. SF collateral may be fixed or floating. CDO liabilities are usually LIBOR-based floating instruments. To convert a fixed rate asset into a floating rate liability, it is necessary to use either an interest rate swap or a cap.

Figure 38.4 shows how this is done. The CDO enters into a swap with an interest rate swap counterpart. The CDO pays a fixed rate coupon to the swap counterparty, and receives a LIBOR-based coupon from the swap counterparty, Figure 38.5 shows a bond-backed CDO using an interest rate cap. With an interest rate cap, the CDO makes an upfront payment, and receives a payment only if LIBOR is over a certain prespecified level. This protects the deal against the scenario in which LIBOR spikes, and the fixed rate coupons on the assets are insufficient to cover the cash flow on the liabilities.

The use of an interest rate swap or cap requires assumptions about the cash flows on the assets. If the assets run off more quickly than anticipated, the CDO can be left with the swaps in place, and no assets.

Call Provisions in CDO Transactions

The commonly used optional redemption features in CDO transactions is where the deal is callable at par by the equity holders, after a prespecified lockout. The call is generally exercised when the deal is doing very well, and

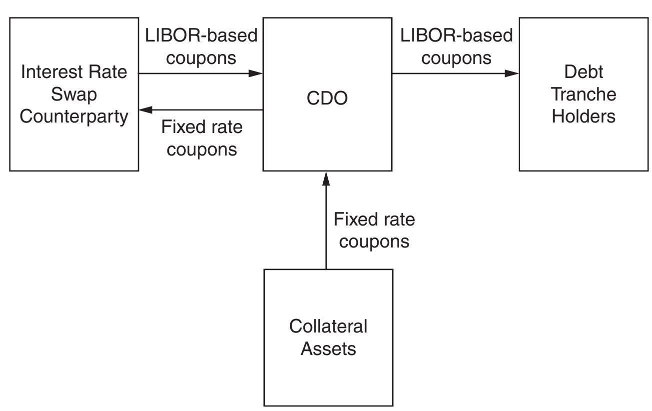


Figure 38.4 Bond-Backed CDO and Interest Rate Swap

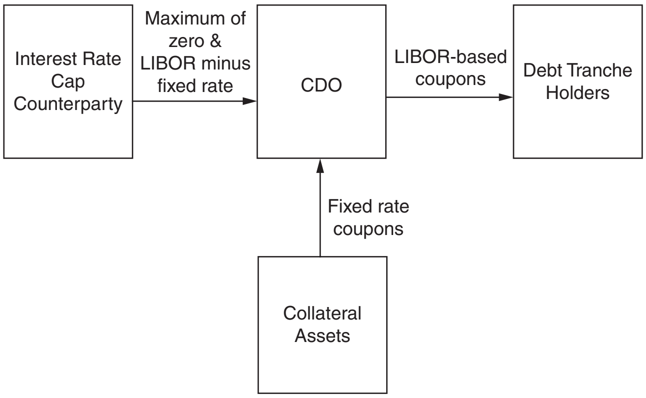


Figure 38.5 Bond-Backed CDO and Interest Rate Cap

the collateral can be liquidated at a healthy net profit. The deal is more apt to be called when the spreads on the debt tranches have narrowed. That is, the equity holders are looking at the possibility of liquidating the deal, paying off the debt holders, and putting the collateral into a new deal where the debt holders are paid a narrower spread. When evaluating CDOs that have been outstanding for a few years and are being traded in the secondary market, call provisions can be important to the valuation of the securities.

Call Protection for Bond Investors

There are many different variations of the basic CDO structure in which the deal is callable at par after a preset lock-out period. Two of the most common variations protecting bondholders are prepayment penalties and coupon step-ups.

Prepayment penalties can take two forms: Either the investor is compensated with a premium call, or there is a "make-whole" provision. The most typical premium call is an amount equal to one-half the annual coupon, which steps down over time. Essentially, the effect of the prepayment penalties is to make the call less attractive to the asset manager.

Coupon step-ups are somewhat rare in deals. If the tranche is not called on a certain date, the coupon "steps up" to a higher level. A coupon step-up is only used if the asset manager wants to signal to investors that it is unlikely that the deal will extend beyond a certain point. For example, deals backed by collateral with long legal final maturities are more apt to have a coupon step up to quell investor concerns about extension risk.

Variations of Call Provisions that Benefit Equity Holders

Not all call provisions will be exercised because the deal is going well. Sometimes if the deal is going very poorly, the equity holders may choose to liquidate because the deal is worth more "dead" than alive. This is particularly true towards the end of the deal because the expenses of running a small deal with low leverage are too high and a "clean-up call" is beneficial.

There are also customized call provisions to protect the equity holders from the whims of an asset manager. Some CDO deals have "partial calls," which allows each group of equity holders to exercise authority over their own piece of the deal. This is different from typical structures, in which the deal is only callable in whole by a majority of the equity interests. It is clear that the value of the deal on an ongoing basis will be different for the asset manager (who earns management fees) and an equity holder (who does not). In certain rare cases, a majority of equity holders may replace the asset manager. This is most common in those deals in which the asset manager does not own a piece of the equity. Both of these call provisions are meant to protect the equity holder (who is not the asset manager) at the expense of the asset manager.

SYNTHETIC ARBITRAGE CDOs

In this section, we review synthetic CDOs. More specifically, our focus is on synthetic arbitrage CDOs. A synthetic CDO does not actually own the portfolio of assets on which it bears credit risk. Instead, it gains credit exposure by selling protection via credit default swaps. In turn, the synthetic CDO buys protection from investors via the tranches it issues. These tranches are responsible for credit losses in the reference portfolio that rise above a particular attachment point; each tranche's liability ends at a particular detachment or exhaustion point. The motivation in an arbitrage synthetic CDO is investors' desire to assume tranched credit risk in return for spread.

Synthetic arbitrage CDOs come in the following forms:

- The oldest are *full-capital structure CDOs* that include a full complement of tranches from super senior to equity. These CDOs have either static reference portfolios or a manager who actively trades the underlying portfolio of credit default swaps (CDS).
- *Single-tranche CDOs* are newer, and are made possible by dealers' faith in their ability to hedge the risk, of a CDO tranche through single-name CDS. Single tranche CDOs often allow CDO investors to substitute credits and amend other terms over the course of the CDOs' life.
- *Standard tranches of credit default swap indices* are the most liquid type of CDOs. These instruments allow long-short strategies that appeal to certain types of investors.

Next we outline the features of these types of synthetic arbitrage CDOs.

Full-Capital Structure Synthetic Arbitrage CDOs

Full capital structure synthetic arbitrage CDOs come in many forms. The best way to explain the differences is to focus on two CDO types that represent the range of structural variations.

The first has a static reference portfolio of 100 investment-grade names which we will refer to as CDO #1. The second, which we refer to as CDO #2, is managed

Table 38.7 Synthetic CDO Spectrum

		CDO#1		CDO #2	
Reference pool amount:		**$1 billion**		**$1 billion**	
Number of reference entities:		**100**		**100**	
Management:		**Static**		**Managed**	
Class	**Capital Structure**	**Amount ($ million)**	**Spread (bps)**	**Amount ($ million)**	**Spread (bps)**
	Super Senior	870	6	890	6
Class A	AAA	50	50	30	48
Class B	AA	30	90	30	85
Class C	A	5	175	14	125
Class D	BBB	15	400	20	275
Class E	Equity	30		16	
	Coverage Test	None		$\frac{\text{Cash collateral}}{\text{Class A} + \text{B} + \text{C} + \text{D}} > 111\%$	
	Final maturity	5 years		5 years	
	Write-down provisions	Immediately upon default		At end of life of deal	
	Swap settlement	Cash		Physical	

with roughly the same underlying credit quality as CDO #1. Salient features of each of the two CDOs, including capital structures and spreads, are shown in Table 38.7.

Static versus Managed

Synthetic arbitrage CDOs can be done as static pools or as managed transactions. The advantage to static CDOs is that the investor can examine the proposed portfolio before closing and know that the portfolio will not change. The investors can ask that certain credits be removed from the portfolio or can decide not to invest in the CDO at all. There are also no ongoing management fees. The disadvantage to a static deal becomes apparent if an underlying credit begins to deteriorate, because no mechanism exists for the CDO to rid itself of the problem credit, which remains in the portfolio and may continue to erode.

Capital Structure

Observe from Table 38.7 that static synthetic CDO #1 has much higher equity (3% versus 1.6%) and no coverage tests. The higher equity percentage is a reflection of the absence of coverage tests. The key to understanding the smaller size of the equity tranche in CDO #2 is the structure of its interest waterfall.

First the trustee fee, the senior default swap fee, and the senior advisory fee are all paid out of the available collateral interest and CDS premium receipts. Next interest is paid to the various note holders, from Class A to Class D, in order of their seniority. Then, a coverage test is conducted. If the coverage test is passed, remaining funds are used to pay the subordinate advisory fee, and the residual cash flow goes to equity holders.

But if the coverage test is failed, cash flow is trapped in a reserve account. Cash in the CDO's reserve account is factored into the coverage test, helping the CDO to meet its required ratio. If the coverage test comes back into compliance, future excess cash flows can be released to the subordinate advisory fee and to equity holders. At the CDO's maturity, cash in the reserve account becomes part of the principal waterfall and helps to pay off tranches in order of their seniority.

Despite the different proportions of equity in the two CDOs, the credit protection enjoyed by rated tranches in each CDO is about equal. This is so because credit protection is measured not only by the amount of subordination below a tranche, but also by how high credit losses can be on the underlying portfolio before die tranche's cash flows are affected. In this case, the rated tranches from both CDOs can survive approximately the same level of default losses; the lower amount of equity in CDO #2 is compensated for by its coverage test and cash trap mechanism.

Settlement on Credit Default Swaps

Note in Table 38.7 that CDO #1 uses cash settlement on the reference pool of assets, while CDO #2 uses physical settlement. There are advantages and disadvantages to both. Cash settlement is simple and filial, thus one generally sees cash settlement in static deals. With physical settlement, the CDO has to deal with the defaulted debt that has been delivered to it. In a managed CDO, however, the manager can decide whether to sell the debt immediately or hold it in hope of realizing a .higher market value later. Physical settlement tends is more common than cash settlement in managed deals.

Equity Cash Flows and the Timing of Write-Downs

In CDO #1, equity is paid a fixed coupon, and thus has no claim on the residual cash flows of the CDO. Equity holders receive interest only on the outstanding equity balance. In CDO #2, the equity holders have a claim on all residual cash flows of the CDO.

The timing of write-downs is very different for the two CDOs. In CDO #1, there is a cash settlement whenever a credit event occurs. Thus, when a credit event occurs (1) that credit is removed from the pool; (2) the CDO pays default losses; and (3) the lowest tranche in the CDO is written down by the amount of default losses. If equity is written down to zero, further losses are written down against the next most junior tranche and so on, moving up the CDO's capital structure.

By contrast, when a credit event occurs in CDO #2, physical settlement occurs. The security can be sold, but there is no write-down until the end of the deal. At that time, the principal cash flows go through the principal waterfall, paying off first the Class A note holders and then those in Class B, C, and D. After note holders are paid, remaining funds go to the equity holders.

Because of these structural differences and investor taste, the BBB and lower classes in CDO #1 generally sell wider than they do in CDO #2. In Table 38.7, the BBB tranche is shown at LIBOR + 400 in CDO #1; it is only LIBOR + 275 in CDO #2. In CDO #1, the write-downs are immediate, and there is no way to recoup losses by better performance later in the deal's life. Moreover, if any of the classes (including the equity) incur losses, their interest is reduced accordingly.

How "Arbitrage" Are Synthetic Arbitrage CDOs?

We have called the CDOs discussed "arbitrage" CDOs. We now look at that label more closely. In some synthetic CDOs, particularly in static portfolio CDOs, the selection of underlying credits is constrained by the availability of risk at the specific bank putting together the CDO.

What do we mean by this? By this we mean that potential equity investors in a synthetic CDO go to a bank with a list of credits on which they want to sell first loss protection. In practice, the final selection of the portfolio depends upon names that the bank either is exposed to already or can become exposed to quickly.

If the bank has an imbalance in its single-name CDS book (which was caused by having sold more protection on a particular name than it has purchased), it will be interested in buying protection on that name from an "arbitrage" CDO. Sometimes the bank's desire to buy credit protection on a particular name derives from exposures built up in other activities. For example, the bank might be exposed to a certain counterparty on interest rate and currency derivatives. In that case, the bank may be interested in buying protection from a CDO. Sometimes the bank can sell protection on a particular name, thereby creating the need to buy protection from a CDO.

The issue of the availability of credit exposure gives these "arbitrage" CDOs a certain balance sheet feel. This is less true in the case of managed synthetic CDOs, where the manager can offer to sell credit protection to a number of banks. Another "arbitrage" synthetic CDO with a balance sheet favor is the CDO driven by a bank's desire to lay off the credit risk of a bond portfolio it owns. The bank thereafter becomes the funder of the bond portfolio without being the owner of its credit risk.

Single-Tranche CDOs

Single-tranche CDOs are notable for what they are not: the placement of a complete capital structure complement of tranches, from equity to super senior. Instead, a protection seller enters into one specific CDO tranche with a CDS dealer in isolation.

This arrangement creates an imbalanced position for the CDS dealer. For example, it might have bought protection on the 3% to 7% tranche of a synthetic CDO comprising 150 underlying Investment-grade names. The CDS dealer will sell protection on these names in the single-name CDS market, varying the notional amount of protection it buys from name to name, in a process called delta hedging.

While there are concerns with using delta hedging, because CDS dealers believe in its efficacy, protection sellers enjoy great flexibility in choosing the terms of single tranche CDOs. (Any losses dealers incur in delta hedging do not affect the terms or economics of the single-tranche CDO.) Protection sellers can choose the portfolio they wish to reference, as well as the attachment and detachment points of the tranche they wish to sell protection on. These factors will imply a price for that protection.

Alternatively, the protection seller can start with a premium in mind and then negotiate other terms to create a transaction furnishing that premium. Because there are only two parties to the transaction, execution can be quicker than it would be with a full-capital structure CDO encompassing many constituencies.

The single-tranche synthetic CDO can also provide flexibility over its life. As reference credits in the underlying portfolio either erode or improve in credit quality, the value of the CDO changes. If, for example, reference credits have all been severely downgraded, the value of credit protection increases because it is more likely there will be default losses. A protection seller of such a single-tranche CDO might be willing to pay a fee to terminate the CDO early rather than be exposed to default losses later.

Single-tranche CDO investors can go back to the original dealer to reverse out of a trade, or they can reverse the trade with another dealer. If investors have sold protection to dealer A, for example, they can buy protection on the exact terms from dealer B. This would leave them with offsetting trades. In many cases, dealers will allow the investor to step out of the trades completely, and the two dealers will face each other directly.

Many single-tranche synthetic CDOs have a feature where terms of the CDO are adjustable over its life. Recall the example where underlying credits have severely deteriorated. Protection sellers might be allowed to replace a soured credit with a better one for a fee. Or, instead of paying a fee, the terms of the CDO tranche might change. In exchange for getting rid of a troubled underlying credit, the attachment point might be decreased, or the detachment point might be increased, or the premium might decrease.

Standard Tranches of CDS Indices

The last type of synthetic CDO we will discuss are those whose underlyings are indices of credit default swaps (see Table 38.8). The terms of these CDO tranches are so standardized and their trading is so liquid that they are typically sold directly from the dealer's trading desk, rather than marketed via term sheets, pitch books, memorandums, and road shows. In fact, pricing on more

Table 38.8 Credit Default Swap Indices

Geographic Concentration	Main Index Name	Main Index Composition	Subindices
North America investment grade	Dow Jones CDX NA 1G	125 corporate names	5 Industries: Consumer, Energy, Financials, Industrials, and Technology/Media/Telecom High Volatility
North American high yield	Dow Jones CDX NA HY	100 corporate names	BB-rated B-rated High Beta
Europe	Dow Jones iTraxx Europe	125 corporate names	9 Industries: Autos, Consumer, Consumer Cyclicals, Consumer Non-Cyclicals, Energy, Senior Financials, Subordinate Financials, Industrials, and Technology/Media/Telecom Largest Corporates Lower Rated (aka Crossover) High Volatility
Japan	Dow Jones iTraxx CJ Japan	50 corporate names	3 industries: Capital Goods, Financials, Technology High Volatility
Asia ex-Japan	Dow Jones iTraxx Asia ex-Japan	50 corporate and sovereign names	3 Geographies: China and Taiwan, Korea, and the rest of ex-japan Asia
Australia	Dow Jones iTraxx Australia	25 corporate names	None
Emerging market	Dow Jones CDX EM	15 sovereign names	None
Emerging market diversified	Dow Jones CDX EM Diversified	40 sovereign and corporate names	3 Geographies: Asia, EEMEA, Latin America

custom synthetic CDOs often reflects prices in the standard tranche market. The reason standard tranches are so liquid is that they are based on liquid credit default swap indices.

The desire of market participants to go long or short a portfolio of underlying names at the same time led to the establishment in 2003 of rival CDS indices, Trac-X and iBoxx. These indices merged in early 2004, deepening the liquidity of the consolidated indices. New indices and subindices have since been added. The composition of the indices and subindices is provided at www.markit.com. Each name in an index is equally weighted in the indices. For the North American indices, only Bankruptcy and Failure to Pay are credit events even though Modified Restructuring is commonly a credit event in the North American market. For the European indices, Bankruptcy, Failure to Pay, and Modified-Modified Restructuring are credit events.

These broad indices are available in maturities from one to 10 years, with the greatest liquidity at 5-, 10-, and, to a lesser extent, 7-year maturities. A new index series is created every six months. At that time, the specific composition of credits in each new series is determined and a new premium level determined for each maturity. Premiums on indices are exchanged once a quarter on the 20th of March, June, September, and December. Over the life of the index, the index's premium remains fixed. To compensate for changes in the price of credit protection, an upfront payment is exchanged. This upfront payment can be regarded as the present value of the difference between the index's fixed premium and the current market premium for the index.

Indices are static and as credit events occur, protection sellers make protection payments to protection buyers, and the notional amount of the index then decreases. It is important to realize that CDS index trades are bilateral agreements. There is no exchange and only recently have there been attempts to centralize the determination of protection payments. Otherwise, protection payments are subject to individual physical settlements.

We begin with a description of how the tranches of the CDS indices are quoted and traded like liquid synthetic CDO tranches. As shown in Table 38.9, the Dow Jones CDX.NA.IG is divided up into 0% to 3%, 3% to 7%, 7% to 10%, 10% to 15%, and 15% to 30% tranches. The lower and higher percentage for each tranche represents that tranche's attachment point and detachment point, respectively. When the cumulative percentage loss of the portfolio of reference entities reaches the attachment points, investors in that tranche begin to lose their principal, and when the cumulative percentage loss of principal reaches the detachment point, those investors lose all their principal and no further loss can occur to them. For example, in Table 38.9, the Tranche 3 has an attachment point of 7% and a detachment percentage of 10%. The tranche will be used to cover the cumulative loss during the life of a CDO in excess of 7% (its attachment point) and up to a maximum of 10% (its detachment point).

For the investment-grade indices, equity tranches require an upfront payment from the protection buyer to the protection seller. After that, a fixed 500 bps per annum is exchanged. For the high-yield index, the first two tranches require upfront payments but have no running fee. The higher tranches of the indices trade solely on their

Table 38.9 Standard Tranches of CDS Indices

CDX NA IG

	Attachment/Detachment Points	Upfront Payment	Running Premium
Tranche 1	0%–3%	Yes	500 bps
Tranche 2	3%–7%	No	Yes
Tranche 3	7%–10%	No	Yes
Tranche 4	10%–15%	No	Yes
Tranche 5	15%–30%	No	Yes

iTriaxx Europe, iTraxx Asia (ex Japan), iTraxx Japan

	Attachment/Detachment Points	Upfront Payment	Running Premium
Tranche 1	0%–3%	Yes	500 bps
Tranche 2	3%–6%	No	Yes
Tranche 3	6%–9%	No	Yes
Tranche 4	9%–12%	No	Yes
Tranche 5	12%–22%	No	Yes

CDX NA HY

	Attachment/Detachment Points	Upfront Payment	Running Premium
Tranche 1	0%–10%	Yes	No
Tranche 2	10%–15%	Yes	No
Tranche 3	15%–25%	No	Yes
Tranche 4	25%–35%	No	Yes
Tranche 5	35%–100%	No	Yes

running fees. Table 38.9 gives details of tranche structure for various CDS indices.

Investors in standard tranches often engage in various forms of long/short trades. The tranche's liquidity makes them ideal for bets on relative price relationships among the tranches. Investors might sell protection on an equity or first-loss tranche and buy protection on a more senior tranche of the same index. In market parlance, they are said to be long the equity tranche and short the more senior tranche. Being long a tranche can be confusing to some investors because one has sold protection on it, but the situation is analogous to being long a bond. When one is long a bond or long a standard tranche (having sold protection), an investor abhors a default and does not want cash or synthetic credit spreads to widen.

Another popular long/short trade is to sell protection on a tranche in a longer maturity and then to buy protection on the same tranche from the same index in a shorter maturity. Hedge funds are big participants in long/short strategies via the standard tranches of credit default swap indices.

SUMMARY

CDOs incorporate ever-evolving structures that have rapidly gained acceptance in the market. In this chapter, we provide an overview of cash and synthetic CDOs, with special attention to the cash flow credit structure, credit rating agencies' methodologies, interest rate hedging, and CDO call features. No doubt, other forms of CDOs will be invented and current forms will fall into disuse. But at least for now, our CDO Structural Matrix in Table 38.2 provides a good way to categorize the different features of CDOs.

REFERENCES

Choudhry, M., and Fabozzi, F. J. (2003). Originating collateralized debt obligations for balance sheet management. *Journal of Structured and Project Finance*, Fall: 32–52.

Fabozzi, F. J., Davis, H., and Choudhry, M. (2006). *Introduction to Structured Finance*. Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., and Goodman, L. S. (eds.) (2001). *Investing in Collateralized Debt Obligations*. Hoboken, NJ: John Wiley & Sons.

Lucas, D. J., Goodman, L. S., and Fabozzi, F. J. (2006). *Collateralized Debt Obligations: Structures and Analysis*, 2nd edition. Hoboken, NJ: John Wiley & Sons.

Lucas, D. J., Goodman, L. S., Fabozzi, F. J., and Manning, R. J. (2007). *Developments in the Collateralized Debt Obligations Markets: New Products and Insights*. Hoboken, NJ: John Wiley & Sons.

CHAPTER 39

Interest Rate Futures and Forward Rate Agreements

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

STEVEN V. MANN, PhD
Professor of Finance, Moore School of Business, University of South Carolina

Abstract: Interest rate derivatives include interest rate futures, forward rate agreements, interest rate swaps, interest rate options, and interest rate caps and floors. Interest rate derivatives can be used to control the interest rate risk of a portfolio or financial institution, to speculate on the future level of interest rates or the change in the shape of the yield curve, or to hedge future borrowing costs. These instruments can be either exchange traded or traded in the over-the-counter market. Interest rate futures are exchange-traded; options can be exchange traded or traded in the over-the-counter market, so-called dealer options. The other interest rate derivatives are traded in the over-the-counter market. Exchange-traded futures on interest rates are classified by the maturity of the underlying interest rate: short-term contracts (Eurodollar futures, Fed funds futures) and long-term contracts (Treasury bond and note futures, swap futures, and municipal bond futures). A forward rate agreement is an over-the-counter derivative instrument which is essentially a forward-starting loan, but with no exchange of principal, so the cash exchanged between the counterparties depend only on the difference in interest rates.

Keywords: derivatives, notional amount, Eurodollar futures, Fed funds futures, Treasury bond futures, Treasury note futures contract, cheapest-to-deliver issue, delivery options, quality or swap option, timing option, wildcard option, swap futures contract, municipal bond futures, forward rate agreement

Derivatives are used by portfolio managers, traders, and corporate treasurers to manage and control risk. There is an array of interest rate derivatives that are used for managing and controlling interest rate risk. As with other derivatives, the instruments can be traded on an exchange or in the over-the-counter market. In this chapter we describe interest rate futures contracts and forward rate agreements. Interest rate swaps, options, caps, and floors are described elsewhere in other chapters.

We begin by discussing interest rate futures contracts. Interest rate futures contracts can be classified by the maturity of their underlying instrument. Short-term interest rate futures contracts have an underlying instrument that matures in one year or less and we discuss these first. We then discuss long-term interest rate futures, contracts where the underlying instrument exceeds one year. Finally, we describe forward rate agreements.

SHORT-TERM INTEREST RATE FUTURES CONTRACTS

The more actively traded short-term interest futures contracts in the United States are described below.

Eurodollar Futures

A Eurodollar futures contract represents a commitment to pay/receive a quarterly interest payment determined by the level of 3-month LIBOR and a notional principal of \$1 million on the settlement date. There are Eurodollar futures contracts available to trade with quarterly settlement dates (March, June, September, December) that extend out 10 years. Accordingly, it is possible for market participants to hedge or speculate on the level of 3-month LIBOR for the next decade. The contracts are settled in cash and trade on the Chicago Mercantile Exchange (CME) as well as the London International Financial Futures Exchange (LIFFE). The price of a Eurodollar futures contract is quoted as 100 minus the annualized futures 3-month LIBOR. For example, a Eurodollar futures price of 95 translates into a futures 3-month LIBOR of 5%.

The minimum price fluctuation (tick) for this contract is 0.005 or 1/2 basis point. This means that the tick value for this contract is \$12.50, which is determined as follows:

$$\text{Tick value} = \$1{,}000{,}000 \times (0.005 \times 90/360) = \$12.50$$

The Eurodollar CD futures contract is used frequently to trade the short end of the yield curve and many hedgers believe this contract to be the best hedging vehicle for a wide range of hedging situations. Moreover, the markets for *Eurodollar futures* contracts and the interest rate swaps are tightly connected. In particular, the floating-rate payments of an interest rate swap can be derived from a portfolio of Eurodollar futures contracts whose expiration dates match the swap's floating-rate reset dates.

The 90-day sterling LIBOR (London Interbank Offered Rate) interest rate futures contract trades on the main London futures exchange, LIFFE. The contract is structured similarly to the Eurodollar futures contract described above. Prices are quoted as 100 minus the interest rate and the expiration months are March, June, September, and December. The contract size is £500,000. A tick is 0.01 or one basis point and the tick value is £12.5.

The LIFFE also trades short-term interest rate futures for other major currencies including euros, yen, and Swiss franc. Short-term interest rate futures contracts in other currencies are similar to the 90-day sterling LIBOR contract and trade on exchanges such as Deutsche Terminbourse in Frankfurt and MATIF (Marché à Terme International de France) in Paris.

Fed Funds Futures Contract

Depository institutions are required to hold reserves to meet their reserve requirements. To meet these requirements, depository institutions hold reserves at their district Federal Reserve Bank. These reserves are called federal funds. Because no interest is earned on federal funds, a depository institution that maintains federal funds in excess of the amount required incurs an opportunity cost of the interest forgone on the excess reserves. Conversely, there are also depository institutions whose federal funds are short of the amount required. The federal funds market is where depository institutions buy and sell federal funds to address this imbalance. The interest rate at which federal funds are bought (borrowed) and sold (lent) is called the federal funds rate. Consequently, the federal funds rate is a benchmark short-term interest rate.

When the Federal Reserve formulates and executes monetary policy, the federal funds rate is a primary operating target. The Federal Open Market Committee (FOMC) sets a target level for the federal funds rate. Announcements of changes in monetary policy specify changes in the FOMC's target for this rate. Once the target is set, the Federal Reserve either adds or drains reserves from the banking system using open market operations so that the actual federal funds rate is, on average, equal to the target. The 30-day *federal funds futures* contract is designed for financial institutions and businesses who want to control their exposure to movements in the federal funds rate.

The federal funds futures contract began trading on the Chicago Board of Trade (CBOT) in October 1988. These contracts have a notional amount of \$5 million and the contract can be written for the current month up to 24 months in the future. Underlying this contract is the simple average overnight federal funds rate (that is, the effective rate) for the month. As such, this contract is settled in cash on the last business day of the month. Just as the other short-term interest rate futures contracts discussed above, prices are quoted on the basis of 100 minus the overnight federal funds rate for the expiration month. These contracts are market to market using the effective daily federal funds rate as reported by the Federal Reserve Bank of New York.

LONG-TERM INTEREST RATE FUTURES CONTRACTS

The most actively traded long-term (greater than one year) interest rate futures contracts are described below.

Treasury Bond Futures

The Treasury bond futures contract is traded on the CBOT. The underlying instrument for this contract is \$100,000 par value of a hypothetical 20-year coupon bond. This hypothetical bond's coupon rate is called the *notional coupon*. Currently, this notional coupon is 6%. Treasury futures contracts trade with March, June, September, and December settlement months.

The futures price is quoted in terms of par being 100. Published quotes have two parts namely the number of points (1% of par value) and the number of ticks ($^1/_{32}$ of 1% of par value). Thus, a quote for a Treasury bond futures contract of 97-16 means 97 and $^{16}/_{32}$ or 97.50. So, if a buyer and seller agree on a futures price of 97-16, this means simply that the buyer agrees to accept delivery of the hypothetical underlying Treasury bond and pay 97.50% of par value and the seller agrees to accept 97.50% of par value. Since the par value of the bond underlying the futures contract is $100,000, the futures price that the buyer and seller agree to for this hypothetical bond is $97,500 plus accrued interest.

The minimum price fluctuation for the Treasury bond futures contract is $^1/_{32}$ of 1% as noted previously which is referred to as a 32nd. The dollar value of a 32nd for $100,000 par value (the par value for the underlying Treasury bond) is $31.25. This is true because each point (1% of the par value) is worth $1,000 and each point is comprised of 32 ticks. Thus, the minimum price fluctuation is $31.25 for this contract.

We have been referring to the underlying instrument as a hypothetical Treasury bond. While some interest rate futures contracts can only be settled in cash, the seller (the short) of a Treasury bond futures contract who chooses to make delivery rather than liquidate his/her position by buying back the contract prior to the settlement date must deliver some Treasury bond. This begs the question "which Treasury bond?" The CBOT allows the seller to deliver one of several Treasury bonds that the CBOT specifies are acceptable for delivery. These contracts have multiple deliverables to avoid having a single issue squeezed and to allow for varying schedules of new issues. The term "squeeze" is used to describe a shortage of the supply of a particular security relative to the demand. A trader who is short a particular security is always concerned with the risk of being unable to obtain sufficient securities to cover their position.

The set of all bonds that meet the delivery requirements for a particular contract is called the deliverable basket. The CBOT makes its determination of the Treasury issues that are acceptable for delivery from all outstanding Treasury issues that have at least 15 years to maturity from the first day of the delivery month. For settlement purposes, the CBOT specifies that a given issue's term to maturity is calculated in complete three month increments (that is, complete quarters). For example, the actual maturity of the issue is 15 years and 5 months would be rounded down to a maturity of 15 years and 1 quarter (three months). Moreover, all bonds delivered by the seller must be of the same issue.

It is important to keep in mind that while the underlying Treasury bond for this contract is a hypothetical issue and therefore cannot itself be delivered into the futures contract, the bond futures contract is not a cash settlement contract. The only way to close out a Treasury bond futures contract is to either initiate an offsetting futures position or to deliver a Treasury issue from the deliverable basket.

Conversion Factors

The delivery process for the Treasury bond futures contract is innovative and has served as a model for government bond futures contracts traded on various exchanges throughout the world. On the settlement date, the seller of the futures contract (the short) is required to deliver the buyer (the long) $100,000 par value of a 6% 20-year Treasury bond. As noted, no such bond exists, so the seller must choose a bond from the deliverable basket to deliver to the long. Suppose the seller selects a 5% coupon, 20-year Treasury bond to settle the futures contract. Since the coupon of this bond is less than the notional coupon of 6%, this would be unacceptable to the buyer who contracted to receive a 6% coupon, 20-year bond with a par value of $100,000. Alternatively, suppose the seller is compelled to deliver a 7% coupon, 20-year bond. Since the coupon of this bond is greater than the notional coupon of 6%, the seller would find this unacceptable. In summary, how do we adjust for the fact that bonds in the deliverable basket have coupons and maturities that differ from the notional coupon of 6%?

To make delivery equitable to both parties, the CBOT uses conversion factors for adjusting the price of each Treasury issue that can be delivered to satisfy the Treasury bond futures contract. Within the deliverable basket, conversion factors are designed to make each bond approximately equally cheap to deliver if the yield curve were flat at 6%. The conversion factor is determined by the CBOT before a contract with a specific settlement date begins trading using the following formula:

$$CF = \frac{1}{1.03^{K/6}}\left[\frac{C}{2} + \frac{C}{0.06}\left(1 - \frac{1}{1.03^{2N}}\right) + \frac{1}{1.03^{2N}}\right]$$

where

CF = conversion factor
N = complete years to maturity as of the settlement month
C = annual coupon rate (in decimal form)
K = number of months that the maturity exceeds N (rounded down to complete quarters)

For example, if the maturity of a Treasury bond from the deliverable basket is 24 years and 4.5 months, K is 3 since the 4.5 months is rounded down to complete quarters, or 3 months. Further, if the maturity is 24 years and 11 months, K is 9.

The convention of rounding down to the nearest complete quarter adds a slight distortion into the calculation of the conversion factors. To see this, recall Treasury futures contracts have expiration months of March, June, September, and December. Also note that all Treasury bonds mature on February 15, May 15, August 15, or November 15. Since conversion factors are computed as of the first day of the delivery month, bonds that mature on say, August 15 are treated as if they mature on June 1 (the first delivery day of the June contract.) The Treasury's maturity is artificially shortened by $2^1/_2$ months so that there is $2^1/_2$ months of "pull to par" built into the conversion factors. As a result, for Treasury bonds with coupon rates below 6%, the conversion factors will be slightly higher than they should be. Conversely, for issues with coupon rates above 6%, the conversion factors will be slightly lower than they should be.

Table 39.1 Deliverable Basket for the December 2006 Treasury Bond Futures Contract

Coupon Rate	Maturity Date	Price	Conv. Factor	Implied Repo Rate
7.625	11/15/22	129–20	1.1640	5.09
7.25	8/15/22	125–08	1.1250	5.06
7.125	2/15/23	124–07+	1.1147	4.68
6.25	8/15/23	114–23+	1.0260	3.35
7.5	11/15/24	130–11	1.1623	2.80
7.625	2/15/25	132–02+	1.1774	2.70
6.875	8/15/25	123–15+	1.0970	1.44
6.75	8/15/26	122–21	1.0855	0.20
6.0	2/15/26	113–00+	1.0000	−0.04
6.5	11/15/26	119–22	1.0573	−0.39
6.625	2/15/27	121–15+	1.0722	−0.68
6.375	8/15/27	118–18+	1.0439	−1.53
6.125	11/15/27	115–15+	1.0146	−2.18
5.5	8/15/28	107–17+	0.9400	−3.98
6.125	8/15/29	116–10+	1.0153	−4.26
5.25	11/15/28	104–10	0.9095	−4.75
5.25	2/15/29	104–10+	0.9090	−5.01
6.25	5/15/30	118–16+	1.0310	−5.20
5.375	2/15/31	106–15	0.9210	−7.16
4.5	2/15/36	94–06+	0.7950	−14.76

The implied repo rates are obtained from Bloomberg.

Table 39.1 displays the conversion factors for each Treasury bond in the deliverable basket for the December 2006 *Treasury bond futures* contract. This information was obtained from Bloomberg. The conversion factors are reported in column 4. Note those bonds with coupon rates with coupon rates greater than 6% have conversion factors greater than one and those with coupon rates less than 6% have conversion factors less than one.

Given the conversion factor for an issue and the futures price, the adjusted price is found by multiplying the conversion factor by the futures price. The adjusted price is called the converted price.

The price that the buyer must pay the seller when a Treasury bond is delivered is called the invoice price. Intuitively, the invoice price should be the futures settlement price plus accrued interest. However, as just noted, the seller can choose any Treasury issue from the deliverable basket. To make delivery fair to both parties, the invoice price must be adjusted using the conversion factor of the actual Treasury issue delivered. The invoice price is:

$$\begin{aligned}\text{Invoice price} = {} & \text{Contract size} \times \text{Futures settlement price} \\ & \times \text{Conversion factor} + \text{Accrued interest}\end{aligned}$$

Suppose the settlement price of the 111–04 Treasury bond futures contract is 11/15/22 and the issue selected by short to deliver is the coupon bond that matures on 111–04. The futures contract settlement price of 1.1640 means 111.125% of par value or times par value. The conversion factor for this issue is 1.11125. Since the contract size is \$100,000, the invoice price the buyer pays the seller is:

$$\begin{aligned}&\$100{,}000 \times 1.11125 \times 1.1640 \times \text{Accrued interest} \\ &= \$129{,}349.50 + \text{Accrued interest}\end{aligned}$$

Cheapest-to-Deliver Issue

In selecting the issue to be delivered, the short will select from all the deliverable issues the one that will give the largest rate of return from a cash-and-carry trade. A cash-and-carry-trade is one in which a cash bond that is acceptable for delivery is purchased with borrowed funds and simultaneously the Treasury bond futures contract is sold. The bond purchased can be delivered to satisfy the short futures position. Thus, by buying the Treasury issue that is acceptable for delivery and selling the futures, an investor has effectively sold the bond at the delivery price (that is, the converted price).

A rate of return can be calculated for this trade. This rate of return is referred to as the implied repo rate and is determined by:

1. The price plus accrued interest at which the Treasury issue could be purchased.
2. The converted price plus the accrued interest that will be received upon delivery of that Treasury bond issue to satisfy the short futures position.
3. The coupon payments that will be received between today and the date the issue is delivered to satisfy the futures contract.
4. The reinvestment income that will be realized on the coupon payments between the time the interim coupon payment is received and the date that the issue is delivered to satisfy the Treasury bond futures contract.

The first three elements are known. The last element will depend on the reinvestment rate that can be earned. While the reinvestment rate is unknown, typically this is a small part of the rate of return and not much is lost by assuming that the implied repo rate can be predicted with certainty.

The general formula for the implied repo rate is as follows:

$$\text{Implied repo rate} = \frac{\text{Dollar return}}{\text{Cost of the investment}} \times \frac{360}{\text{Days}_1}$$

where Days_1 is equal to the number of days until settlement of the futures contract. Below we will explain the other components in the formula for the implied repo rate.

Let's begin with the dollar return. The dollar return for an issue is the difference between the proceeds received and the cost of the investment. The proceeds received are equal to the proceeds received at the settlement date of the futures contract and any interim coupon payment plus interest from reinvesting the interim coupon payment. The proceeds received at the settlement date include the converted price (that is, futures settlement price multiplied by the conversion factor for the issue) and the accrued interest received from delivery of the issue. That is,

$$\begin{aligned}\text{Proceeds received} = {} & \text{Converted price} \\ & + \text{Accrued interest received} \\ & + \text{Interim coupon payment} \\ & + \text{Interest from reinvesting the} \\ & \quad \text{interim coupon payment}\end{aligned}$$

As noted earlier, all of the elements are known except the interest from reinvesting the interim coupon payment. This amount is estimated by assuming that the coupon payment can be reinvested at the term repo rate. The repo rate is not only a borrowing rate for an investor who wants to borrow in the repo market but also the rate at which an investor can invest proceeds on a short-term basis. For how long is the reinvestment of the interim coupon payment? It is the number of days from when the interim coupon payment is received and the actual delivery date to satisfy the futures contract. The reinvestment income is then computed as follows:

$$\text{Interest from reinvesting the interim coupon payment} \\ = \text{Interim coupon} \times \text{Term repo rate} \times (\text{Days}_2/360)$$

where Days_2 is the number of days between when the interim coupon payment is received and the actual delivery date of the futures contract.

The reason for dividing Days_2 by 360 is that the ratio represents the number of days the interim coupon is reinvested as a percentage of the number of days in a year as measured in the money market.

The cost of the investment is the amount paid to purchase the issue. This cost is equal to the purchase price plus accrued interest paid. That is,

$$\text{Cost of the investment} = \text{Purchase price} \\ + \text{Accrued interest paid}$$

Thus, the dollar return for the numerator of the formula for the implied repo rate is equal to

$$\text{Dollar return} = \text{Proceeds received} \\ - \text{Cost of the investment}$$

Note that in practice, the cost of the investment should be adjusted because the amount that the investor ties up in the investment is reduced if there is an interim coupon payment. We ignore this adjustment here.

The dollar return is then divided by the cost of the investment.

So, now we know how to compute the numerator and the denominator in the formula for the implied repo rate. The second ratio in the formula for the implied repo rate simply involves annualizing the return using a convention in the money market for the number of days. (The money market convention is to use a 360-day year.) Since the investment resulting from the cash-and-carry trade is a synthetic money market instrument, 360 days are used.

Let's compute the implied repo rate for a hypothetical issue that may be delivered to satisfy a hypothetical Treasury bond futures contract. Assume the following for the deliverable issue and the futures contract:

Futures contract:

Futures price = 96
Days to futures delivery date (Days_1) = 82 days

Deliverable issue:

Price of issue = 107
Accrued interest paid = \$3.8904
Coupon rate = 10%
Days remaining before interim coupon paid = 40 days
Interim coupon = \$5
Number of days between when the interim coupon payment is received and the actual delivery date of the futures contract (days_2) = 42
Conversion factor = 1.1111
Accrued interest received at futures settlement date = 1.1507

Other information:

82-day term repo rate = 3.8%

Let's begin with the proceeds received. We need to compute the converted price and the interest from reinvesting the interim coupon payment. The converted price is:

$$\text{Converted price} = \text{Futures price} \times \text{Conversion factor} \\ = 96 \times 1.1111 = 106.6656$$

The interest from reinvesting the interim coupon payment depends on the term repo rate. The term repo rate is assumed to be 3.8%. Therefore,

$$\text{Interest from reinvesting the interim coupon payment} \\ = \$5 \times 0.038 \times \left(\frac{42}{360}\right) = 0.0222$$

To summarize:

Converted price	= 106.6656
Accrued interest received	= 1.1507
Interim coupon payment	= 5.0000
Interest from reinvesting the interim coupon payment	= 0.0222
Proceeds received	= 112.8385

The cost of the investment is the purchase price for the issue plus the accrued interest paid, as shown below:

$$\text{Cost of the investment} = 107 + 3.8904 = 110.8904$$

The implied repo rate is then:

$$\text{Implied repo rate} = \frac{112.8385 - 110.8904}{110.8904} \times \frac{360}{82} = 0.0771 \\ = 7.71\%$$

Once the implied repo rate is calculated for each bond in the deliverable basket, the issue selected will be the one that has the highest implied repo rate (that is, the issue that gives the maximum return in a cash-and-carry trade). The issue with the highest return is referred to as the *cheapest-to-deliver issue.* This issue plays a key role in the pricing of a Treasury futures contract.

While a particular Treasury bond may be the cheapest to deliver today, changes in interest rates, for example, may cause some other issue to be the cheapest to deliver at a future date. A sensitivity analysis can be performed to determine how a change in yield affects the cheapest to deliver bond. In particular, the sensitivity analysis identifies which bond in the deliverable basket is cheapest to deliver following various shocks to the yield curve.

Other Delivery Options

In addition to the choice of which acceptable Treasury issue to deliver—sometimes referred to as the *quality option* or *swap option*—the short has at least two more options granted under CBOT delivery guidelines. The short is permitted to decide when in the delivery month, delivery actually will take place. This is called the *timing option.* The other option is the right of the short to give notice of intent to deliver up to 8:00 P.M. Chicago time after the closing of the exchange (3:15 P.M. Chicago time) on the date when the futures settlement price has been fixed. This option is referred to as the *wildcard option.* The quality option, the timing option, and the wildcard option (in sum referred to as the *delivery options*), mean that the long position can never be sure which Treasury bond issue will be delivered or when it will be delivered. These three delivery options are summarized below:

Delivery Option	Description
Quality or swap option	Choice of which acceptable Treasury issue to deliver
Timing option	Choice of when in delivery month to deliver
Wild card option	Choice to deliver after the closing price of the futures contract is determined

Delivery Procedure

For a short who wants to deliver, the delivery procedure involves three days. The first day is the position day. On this day, the short notifies the CBOT that it intends to deliver. The short has until 8:00 P.M. central standard time to do so. The second day is the notice day. On this day, the short specifies which particular issue will be delivered. The short has until 2:00 P.M. central standard time to make this declaration. (On the last possible notice day in the delivery month, the short has until 3:00 P.M.) The CBOT then selects the long to whom delivery will be made. This is the long position that has been outstanding for the longest period of time. The long is then notified by 4:00 P.M. that delivery will be made. The third day is the delivery day. By 10:00 A.M. on this day the short must have in its account the Treasury issue that it specified on the notice day and by 1:00 P.M. must deliver that bond to the long that was assigned by the CBOT to accept delivery. The long pays the short the invoice price upon receipt of the bond.

Treasury Note Futures

There are three Treasury note futures contracts: 10-year, 5-year, and 2-year. All three contracts are modeled after the Treasury bond futures contract and are traded on the CBOT. The underlying instrument for the 10-year Treasury note contract is $100,000 par value of a hypothetical 10-year 6% Treasury note. There are several acceptable issues that may be delivered by the short. An issue is acceptable if the maturity is not less than 6.5 years and not greater than 10 years from the first day of the delivery month. The delivery options granted to the short position and the minimum price fluctuation are the same as for the Treasury bond futures contract.

For the 5-year Treasury note futures contract, the underlying instrument is $100,000 par value of a 6% notional coupon Treasury note. An issue in the deliverable basket must satisfy the following conditions: (1) an original maturity of not more than five years and three months; (2) a remaining maturity of not more than five years and three months; and (3) a remaining maturity not less than four years and two months. The minimum price fluctuation for this contract is $^1/_{64}$ of 1% of par. The dollar value of a 64th for a $100,000 par value is $15.625 ($100,000/6,400) and is therefore the minimum price fluctuation.

The underlying instrument for the 2-year Treasury note futures contract is $200,000 par value of a 6% notional coupon Treasury note. Issues acceptable for delivery must have a remaining maturity of not more than two years and not less than one year and nine months. Moreover, the original maturity of the note in the deliverable basket cannot be more than five years and three months. The minimum price fluctuation for this contract is $^1/_{128}$ of 1% of par value. The dollar value of a 128th for a $200,000 par value is $15.625 ($100,000/12,800) and is therefore the minimum price fluctuation.

Swap Futures Contracts

An interest rate swap contract is an instrument used by market participants to transform the nature of cash flows and the interest rate exposure of a portfolio or balance sheet. The contract is an agreement between two counterparties to exchange periodic interest payments. In the most common and simplest form, one party agrees to pay the other party fixed interest payments at designated dates for the life of the contract. The other party in return agrees to make interest rate payments that float with some reference rate. When quoting swaps levels in the market, the convention is for the dealer to set the floating rate equal to the reference rate (usually LIBOR) and then quote the fixed rate (called the swap rate) that will apply.

The CBOT introduced a swap futures contract in late October 2001. The underlying instrument is the notional price of the fixed-rate side of a 10-year interest rate swap that has a notional principal equal to $100,000 and that exchanges semiannual interest payments at a fixed annual rate of 6% for floating interest rate payments based on 3-month LIBOR. Interest rate swaps are discussed in Chapter 40 of Volume I. This swap futures contract is cash-settled with a settlement price determined by the ISDA benchmark 10-year swap rate on the last day of trading before the contract expires. This benchmark rate is published with a one-day lag in the Federal Reserve Board's statistical release H.15. Contracts have settlement months of March, June, September, and December just like the other CBOT interest rate futures contracts that we have discussed.

The LIFFE introduced the first swap futures contract called Swapnote®, which is referenced to the euro interest rate swap curve. Swapnotes are available in 2-, 5-, and 10-year maturities. The CME also lists a *swap futures contract* with maturities of 2, 5, and 10 years that is similar to those listed on the CBOT.

10-Year Municipal Note Index Futures Contract

A 10-year municipal note index futures contract is traded on the CBOT. The underlying for this contract is an index. The index includes between 100 and 250 high-grade tax-exempt securities. For an issue to be eligible for inclusion in the index, the issuer

- must have a triple A credit rating assigned by both S&P and Moody's
- must have a principal size of at least $50 million
- must be a component (that is, tranche) of a municipal issue with a total deal size of at least $200 million
- must have a remaining maturity of between 10 and 40 years from the first calendar day of the corresponding futures contract expiration
- must at issuance have a price of at least 90
- must pay semiannual interest at a fixed coupon rate that ranges from 3% to 9%

An issue can be callable or noncallable. However, if an issue is callable, it must have a first call date at least seven years from the first calendar day of the corresponding futures contract expiration. The issues comprising the index insured and uninsured bonds.

In constructing the index, there are three further restrictions: (1) no more than 5% of the bond in the index can be from any one issuer, (2) no more than 15% can be from any one state or U.S. territory, and (3) no more than 40% of the issues can be insured by any one issuer.

To assure that the index continues to accurately mirror the overall tax-exempt market, it is revised quarterly on the first business day of each February, May, August, and November. When the index is revised, issues that no longer meet the selection criteria explained above will are eliminated from the index.

Each day the index is priced. Because the issues comprising the index do not typically trade each day, an independent pricing service, FT Interactive Data Corporation, provides prices for the individual issues and then calculates computes the closing value of the index. At the settlement date, the parties settle in cash. Settlement is based on the final settlement value based on the value of the index as determined by FT Interactive Data Corporation. The final settlement price is calculated as follows:

$$\text{Final settlement value} = \$100{,}000[5/r + (15/r)(1 + r/200)^{-20}]$$

where r is equal to the simple average yield-to-worst of the component bonds in the index for the last day of trading, expressed in percent terms and calculated to the nearest 1/10 of a basis point (e.g., 4.85%).

FORWARD RATE AGREEMENTS

A *forward rate agreement* (FRA) is an over-the-counter derivative instrument that trades as part of the money market. In essence, an FRA is a forward-starting loan, but with no exchange of principal, so the cash exchanged between the counterparties depend only on the difference in interest rates. While the FRA market is truly global, most business is transacted in London. Trading in FRAs began in the early 1980s and the market now is large and liquid.

In effect an FRA is a forward dated loan, transacted at a fixed rate, but with no exchange of principal—only the interest applicable on the notional amount between the rate agreed to when the contract is established and the actual rate prevailing at the time of settlement changes hands. For this reason, FRAs are off-balance sheet instruments. By trading today at an interest rate that is effective at some point in the future, FRAs enable banks and corporations to hedge forward interest rate exposure.

FRA Basics

An FRA is an agreement to borrow or lend a notional cash sum for a period of time lasting up to 12 months, starting at any point over the next 12 months, at an agreed rate of interest (the FRA rate). The "buyer" of a FRA is borrowing a notional sum of money while the "seller" is lending this cash sum. Note how this differs from all other money market instruments. In the cash market, the party buying a CD, Treasury bill, or bidding for bond in the repo market, is the lender of funds. In the FRA market, to "buy" is to "borrow." Of course, we use the term "notional" because with an FRA no borrowing or lending of cash actually takes place. The notional sum is simply the amount on which the interest payment is calculated (that is, a scale factor).

Accordingly, when a FRA is traded, the buyer is borrowing (and the seller is lending) a specified notional sum at a fixed rate of interest for a specified period, the "loan" to commence at an agreed date in the future. The buyer is the notional borrower, and so if there is a rise in interest rates between the date that the FRA is traded and the date that the FRA comes into effect, she will be protected. If there is a fall in interest rates, the buyer must pay the difference between the rate at which the FRA was traded and the actual rate, as a percentage of the notional sum.

The buyer may be using the FRA to hedge an actual exposure, that is an actual borrowing of money, or simply speculating on a rise in interest rates. The counterparty to the transaction, the seller of the FRA, is the notional lender of funds, and has fixed the rate for lending funds. If there is a fall in interest rates, the seller will gain, and if there is a rise in rates, the seller will pay. Again, the seller may have an actual loan of cash to hedge or is acting as a speculator.

In FRA trading, only the payment that arises because of the difference in interest rates changes hands. There is no exchange of cash at the time of the trade. The cash payment that does arise is the difference in interest rates between that at which the FRA was traded and the actual rate prevailing when the FRA matures, as a percentage of the notional amount. FRAs are traded by both banks and corporations. The FRA market is liquid in all major currencies and rates are readily quoted on screens by both banks and brokers.

The terminology quoting FRAs refers to the borrowing time period and the time at which the FRA comes into effect (or matures). Hence, if a buyer of a FRA wished to hedge against a rise in rates to cover a 3-month loan starting in three months' time, she would transact a "3-against-6 month" FRA, or more usually denoted as a 3×6 or 3v6 FRA. This is referred to in the market as a "three-sixes" FRA, and means a 3-month loan beginning in three months' time. So correspondingly, a "ones-fours" FRA (1v4) is a 3-month loan in one month's time, and a "three-nines" FRA (3v9) is a 6-month loan in three months' time.

As an illustration, suppose a corporation anticipates it will need to borrow in 6 months time for a 6-month period. It can borrow today at 6-month LIBOR plus 50 basis points. Assume that 6-month LIBOR rates are 4.0425% but the corporation's treasurer expects rates to go up to about 4.50% over the next several weeks. If the treasurer's suspicion is correct, the corporation will be forced to borrow at higher rates unless some sort of hedge is put in place to protect the borrowing requirement. The treasurer elects to buy a 6v12 FRA to cover the 6-month period beginning six months from now. A bank quotes 4.3105% for the FRA, which the corporation buys for a £1,000,000 notional principal. Suppose that 6 months from now, 6-month LIBOR has indeed backed-up to 4.50%, so the treasurer must borrow funds at 5% (LIBOR plus the 50-basis-point spread). However, offsetting this rise in rates, the corporation will receive a settlement amount which will be the difference between the rate at which the FRA was bought (4.3105%) and today's 6-month LIBOR rate (4.50%) as a percentage of the notional principal of £1 million. This payment will compensate for some of the increased borrowing costs.

FRA Mechanics

In virtually every market, FRAs trade under a set of terms and conventions that are identical. The British Bankers Association (BBA) has compiled standard legal documentation to cover FRA trading. The following standard terms are used in the market:

- **Notional sum:** The amount for which the FRA is traded.
- **Trade date:** The date on which the FRA is transacted.
- **Settlement date:** The date on which the notional loan or deposit of funds becomes effective, that is, is said to begin. This date is used, in conjunction with the notional sum, for calculation purposes only as no actual loan or deposit takes place.
- **Fixing date:** This is the date on which the reference rate is determined, that is, the rate to which the FRA rate is compared.
- **Maturity date:** The date on which the notional loan or deposit expires.
- **Contract period:** The time between the settlement date and maturity date.
- **FRA rate:** The interest rate at which the FRA is traded.
- **Reference rate:** This is the rate used as part of the calculation of the settlement amount, usually the LIBOR rate on the fixing date for the contract period in question.

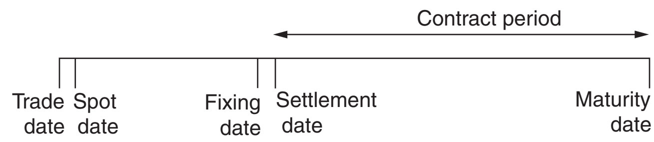


Figure 39.1 Key Dates in a FRA Trade

- **Settlement sum:** The amount calculated as the difference between the FRA rate and the reference rate as a percentage of the notional sum, paid by one party to the other on the settlement date.

These key dates are illustrated in Figure 39.1.

The spot date is usually two business days after the trade date, however it can by agreement be sooner or later than this. The settlement date will be the time period after the spot date referred to by the FRA terms: for example, a 1×4 FRA will have a settlement date one calendar month after the spot date. The fixing date is usually two business days before the settlement date. The settlement sum is paid on the settlement date, and as it refers to an amount over a period of time that is paid up front (that is, at the start of the contract period), the calculated sum is a discounted present value. This is because a normal payment of interest on a loan/deposit is paid at the end of the time period to which it relates; because an FRA makes this payment at the start of the relevant period, the settlement amount is a discounted present value sum. With most FRA trades, the reference rate is the level of LIBOR on the fixing date.

The settlement sum is calculated after the fixing date, for payment on the settlement date. We can illustrate this with a hypothetical example. Consider a case where a corporation has bought £1 million notional sum of a 1×4 FRA, and transacted at 5.75%, and that the market rate is 6.50% on the fixing date. The contract period is 90 days. In the cash market the extra interest charge that the corporate would pay is a simple interest calculation, and is:

$$\text{Extra interest charge} = \frac{6.50 - 5.75}{100} \times £1{,}000{,}000 \times (91/365) = £1{,}869.86$$

Note that in the U.S. money market, a 360-day year is assumed rather than the 365 day year used in the U.K. money market.

This extra interest that the corporation is facing would be payable with the interest payment for the loan, which (as it is a money market loan) is paid when the loan matures. Under a FRA then, the settlement sum payable should, if it was paid on the same day as the cash market interest charge, be exactly equal to this. This would make it a perfect hedge. As we noted above though, FRA settlement value is paid at the start of the contract period, that is, the beginning of the underlying loan and not the end. Therefore, the settlement sum has to be adjusted to account for this, and the amount of the adjustment is the value of the interest that would be earned if the unadjusted cash value were invested for the contract period in the money market. The settlement value is given by the following expression:

$$\text{Settlement value} = \frac{(r_{\text{ref}} - r_{\text{FRA}}) \times M \times (n/B)}{1 + [r_{\text{ref}} \times (n/B)]}$$

where

r_{ref} = the reference interest fixing rate
r_{FRA} = the FRA rate or contract rate
M = the notional value sum
n = the number of days in the contract period
B = the day-count basis (360 or 365)

The expression for the settlement value above simply calculates the extra interest payable in the cash market, resulting from the difference between the two interest rates, and then discounts the amount because it is payable at the start of the period and not, as would happen in the cash market, at the end of the period.

In our hypothetical illustration, as the fixing rate is higher than the contract rate, the buyer of the FRA receives the settlement sum from the seller. This payment compensates the buyer for the higher borrowing costs that they would have to pay in the cash market. If the fixing rate had been lower than 5.75%, the buyer would pay the difference to the seller, because the cash market rates will mean that they are subject to a lower interest rate in the cash market. What the FRA has done is hedge the interest rate exposure, so that whatever happens in the market, the buyer will pay 5.75% on its borrowing.

A market maker in FRAs is trading short-term interest rates. The settlement sum is the value of the FRA. The concept is exactly as with trading short-term interest-rate futures; a trader who buys a FRA is running a long position, so that if on the fixing date the reference rate is greater than the contract rate then the settlement sum is positive and the trader realizes a profit. What has happened is that the trader, by buying the FRA, "borrowed" money at the FRA rate, which subsequently rose. This is a gain, exactly like a short position in an interest rate futures contract, where if the price goes down (that is, interest rates go up), the trader realizes a gain. Conversely, a "short" position in an FRA that is accomplished by selling a FRA realizes a gain if on the fixing date the reference rate is less than the FRA rate.

SUMMARY

Interest rate derivatives are employed by market participants to manage and control interest rate risk. This chapter included a discussion of interest rate futures and forward rate agreements. Short-term interest rate futures are used to manage and control interest rate risk due to movements in short-term (less than one year) interest rates. Eurodollar futures contracts are among the most heavily traded contracts in the world. It is a cash settlement contract with an underlying instrument of three-month LIBOR. The federal funds futures contract allows users to control their exposure the federal funds rate.

Actively traded long-term (longer than one year) interest futures contracts include Treasury bond/note futures contract, swap futures contracts of various maturities, and the 10-year municipal note index futures contract. The underlying instrument of the Treasury bond futures contract is a notional 6% coupon, 20-year bond. Conversion factors are used to adjust the invoice price of a Treasury bond futures contract to delivery equitable to both parties. The short position has several embedded delivery options which include the following: quality, timing and wild card options. Treasury note futures contracts of 2-, 5-, and 10-years are modeled after the Treasury note futures contract.

The underlying instrument for a swap futures contract is the notional price of the fixed-rate side of a 10-year interest rate swap that has a notional principal equal to $100,000 and that exchanges semi-annual interest payments at a fixed annual rate of 6% for floating interest rate payments based on 3-month LIBOR.

The underlying for the 10-year municipal note index futures contract are 100 to 250 high-grade tax-exempt securities. The contract is a cash settlement contract.

A forward rate agreement is an over-the-counter derivative instrument which is essentially a forward-starting loan, but with no exchange of principal, so the cash exchanged between the counterparties depend only on the difference in interest rates.

The elements of an FRA are the FRA rate, reference rate, notional amount, contract period, and settlement date. The buyer of an FRA is agreeing to pay the FRA rate and the seller of the FRA is agreeing to receive the FRA rate. The amount that must be exchanged at the settlement date is the present value of the interest differential. In contrast to an interest rate futures contract, the buyer of an FRA benefits if the reference rate increases and the seller benefits if the reference rate decreases.

REFERENCES

Chance, D. M., and Brooks, R. (2007). *Introduction to Derivatives and Risk Management*, 7th edition. Mason, OH: Thomson South-Western.
Fabozzi, F. J., Mann, S. V., and Choudhry, M. (2003). *Measuring and Controlling Interest Rate and Credit Risk*, 2nd edition. Hoboken, NJ: John Wiley & Sons.
Hull, J. (1997). *Introduction to Futures and Options Markets*, 3rd edition. Englewood Cliffs, NJ: Prentice Hall.
Hull, J. (2006). *Options, Futures, and Other Derivative Securities*. Upper Saddle River, NJ: Prentice Hall.
Kolb, W., and Overdahl, J. A. (2006). *Understanding Futures Markets*. New York: Blackwell.
Pitts, M., and Fabozzi, F. J. (1990). *Interest Rate Futures and Options*, Chicago: Probus Publishing.

CHAPTER 40

Interest Rate Swaps

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

GERALD W. BUETOW, PhD, CFA
President and Founder, BFRC Services, LLC

Abstract: One of the major innovations in the financial markets has been the interest rate swap. This instrument is a derivative product that was introduced into the financial markets in the early 1980s. Interest rate swaps were first used to try to arbitrage opportunities in global capital markets. While those arbitrage opportunities quickly disappeared, the interest rate swap market continued to grow. Today, this derivative instruction provides asset managers, risk managers, corporate treasurers, and municipal treasurers with an efficient tool for controlling interest rate risk by altering the cash flow characteristics of their assets or liabilities.

Keyword: interest rate swap, plain vanilla swap, fixed-rate payer, swap fixed rate, swap rate, fixed-rate receiver, floating-rate receiver, floating-rate payer, cash flow for the swap, swap spread, swap dealer, counterparty risk, amortizing swap, accreting swap, basis swap, constant maturity swap (CMS), forward rate swap, swaption, payer's swaption, receiver's swaption

The objective of this chapter is to explain the basic features of an interest rate swap and provide an economic interpretation of this derivative instrument. Interest rate swaps include plain vanilla swaps, amortizing swaps, accreting swaps, basis swaps, constant maturity swaps, forward rate swaps, and swaptions. The valuation of interest rate swaps and the factors that affect the value of a swap are explained in Chapter 44 of Volume III for a plain vanilla swap, amortizing, and accreting swap and Chapter 45 of Volume III for a forward rate swap and a swaption.

THE PLAIN VANILLA SWAP

In an interest rate swap, two parties agree to exchange interest payments at specified future dates. The dollar amount of the interest payments exchanged is based on some predetermined dollar principal, which is called the notional principal or notional amount. The payment each party pays to the other is the agreed-upon periodic interest rate times the notional principal. The only dollars that are exchanged between the parties are the interest payments, not the notional principal.

In the most common type of swap, one party agrees to pay the other party fixed interest payments at designated dates for the life of the contract. This party is referred to as the *fixed-rate payer*. The fixed rate that the fixed-rate payer must make is called the *swap fixed rate* or *swap rate*. The other party, who agrees to make payments that float with some reference rate, is referred to as the *fixed-rate receiver*. The fixed-rate payer is also referred to as the *floating-rate receiver* and the floating-rate payer is also called the *floating-rate receiver*. The type of swap that we have just described is called a *plain vanilla swap*. The fixed-rate payer (floating-rate receiver) and floating-rate payer (fixed-rate receiver) are counterparties to the swap.

The reference rates that have been used for the floating rate in an interest rate swap are those on various money market instruments: the London interbank offered rate, Treasury bills, commercial paper, bankers acceptances, certificates of deposit, the federal funds rate, a constant maturity Treasury rate, and the prime rate. The most common is the London Interbank Offered Rate (LIBOR). LIBOR is the rate at which prime banks offer to pay on Eurodollar deposits available to other prime banks for a given maturity. Basically, it is viewed as the global cost of bank borrowing. There is not just one rate but a rate for different maturities. For example, there is a 1-month LIBOR, 3-month LIBOR, 6-month LIBOR, and so on.

To illustrate a plain vanilla interest rate swap, suppose that for the next five years party X agrees to pay party Y 6% per year (the swap fixed rate), while party Y agrees to pay party X 3-month LIBOR (the reference rate). Party X is the fixed-rate payer, while party Y is the fixed-rate receiver. Assume that the notional principal is $100 million, and that payments are exchanged every three months for the next five years. This means that every three months, party X (the fixed-rate payer) will pay party Y $1.5 million (6% times $100 million divided by 4). The amount that party Y (the fixed-rate receiver) will pay party X will be 3-month LIBOR times $100 million divided by 4. If 3-month LIBOR is 5%, party Y will pay party X $1.25 million (5% times $100 million divided by 4). Note that we divide by four because one-quarter of a year's interest is being paid. (We will be more precise about the days in the period for determining the payments in the next chapter.) This is illustrated in panel a of Figure 40.1

Swap Payments versus Cash Flows

The payments between the parties are usually netted. In our illustration, if the fixed-rate payer must pay $1.5 million and the fixed-rate receiver must pay $1.25 million, than rather than writing checks for the respective amounts, the fixed-rate party can just make a payment of $0.25 million (= $1.5 million − $1.25 million) to the fixed-rate receiver. We shall refer to this netted payment between the two parties as the *cash flow for the swap* for the period. We note that throughout the literature the terms "swap payments" and "cash flows" are used interchangeably. However, in this chapter we will use the term swap payments to mean the payment made by a counterparty before any netting and cash flow to mean the netted amount.

Swap Quote Conventions

The convention that has evolved for quoting a swap fixed rate is that a dealer sets the floating rate equal to the reference rate and then quotes the swap fixed rate that will apply. The swap fixed rate is some "spread" above the Treasury yield curve with the same term to maturity as the swap. This spread is called the *swap spread*.

Entering Into a Swap and Counterparty Risk

Interest rate swaps are over-the-counter (OTC) instruments. This means that they are not traded on an exchange. A party wishing to enter into a swap transaction can do so through either a securities firm or a commercial bank that transacts in swaps. These entities can do one of the following. First, they can arrange or broker a swap between two parties that want to enter into an interest rate swap. In this case, the investment bank or commercial bank (simply bank hereafter) is acting in a brokerage capacity. The broker is not party to the swap.

The second way in which a bank can get a party into a swap position is by taking the other side of the swap. This means that the bank is acting as a dealer rather than a broker in the transaction. Acting as a dealer (which we refer to as a *swap dealer*) the bank is a counterparty to the swap and therefore must hedge its swap position in the same way that it hedges its position in other securities that it holds. Also it means that the dealer is the counterparty to the transaction. If a party entered into a swap with a swap dealer, the party will look to the swap dealer to satisfy the obligations of the swap; similarly, that same swap dealer looks to the counterparty to fulfill its obligations as set forth in the swap agreement.

The risk that the two parties take on when they enter into a swap is that the other party will fail to fulfill its obligations as set forth in the swap agreement. That is, each party faces default risk and therefore there is bilateral *counterparty risk*.

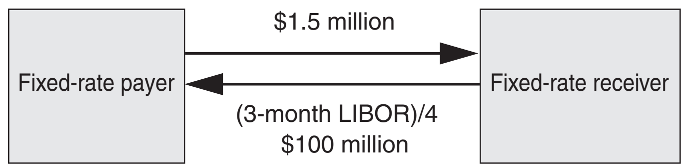
RISK/RETURN CHARACTERISTICS OF AN INTEREST RATE SWAP

The value of an interest rate swap will fluctuate with market interest rates. To see how, let's reconsider our hypothetical swap. Suppose that interest rates change immediately after parties X and Y enter into the swap. Panel a in Figure 40.1 shows the transaction. First, consider what would happen if the market required that in any 5-year swap the fixed-rate payer must pay a swap fixed rate of 7% in order to receive 3-month LIBOR. If party X (the fixed-rate payer) wants to sell its position to party A, then party A will benefit by having to pay only 6% (the original swap fixed rate agreed upon) rather than 7% (the current swap fixed rate) to receive 3-month LIBOR. Party X will want compensation for this benefit. Consequently, the value of party X's position has increased. Thus, if interest rates increase, the fixed-rate payer will realize a profit and the fixed-rate receiver will realize a loss. Panel b in Figure 40.1 summarizes the results of a rise in interest rates.

Next, consider what would happen if interest rates decline to, say, 5%. Now a 5-year swap would require the fixed-rate payer to pay 5% rather than 6% to receive 3-month LIBOR. If party X wants to sell its position to party B, the latter would demand compensation to take over the position. In other words, if interest rates decline, the fixed-rate payer will realize a loss, while the fixed-rate receiver *a. Initial position*

Swap fixed rate	=	6%
Payment frequency	=	quarterly
Reference rate	=	3-month LIBOR
Term of swap	=	5 years
Notional principal	=	$100 million
Payment by fixed-rate payer	=	$1.5 million

Every quarter



b. Interest rates increase such that swap fixed rate is 7%

Fixed-rate payer pays initial swap fixed rate of 6% to obtain 3-month LIBOR
Advantage to fixed-rate payer: pays only 6% not 7% to obtain 3-month LIBOR

Fixed-rate receiver pays 3-month LIBOR
Disadvantage to fixed-rate receiver: receives only 6% in exchange for 3-month LIBOR, not 7%

Results of a rise in interest rates:

Party	Value of swap
Fixed-rate payer	Increases
Fixed-rate receiver	Decreases

c. Interest rates decrease such that swap fixed rate is 5%

Fixed-rate payer pays initial swap fixed rate of 6% to obtain 6-month LIBOR
Disadvantage to fixed-rate payer: must pay 6% not 5% to obtain 3-month LIBOR

Fixed-rate receiver pays 3-month LIBOR
Advantage to fixed-rate receiver: receives 6% in exchange for 3-month LIBOR, not 5%

Results of a rise in interest rates:

Party	Value of swap
Fixed-rate payer	Decreases
Fixed-rate receiver	Increases

Figure 40.1 Summary of How the Value of a Swap to Each Counterparty Changes when Interest Rates Change

will realize a profit. Panel c in Figure 40.1 summarizes the results of a decline in interest rates.

INTERPRETING A SWAP POSITION

There are two classic ways that a swap position can be interpreted: (1) a package of futures (forward) contracts and (2) a package of cash flows from buying and selling cash market instruments. These interpretations will help us understand how to value swaps and how to assess the sensitivity of a swap's value to changes in interest rates. A third, more complicated interpretation, uses caps and floors. Specifically, a portfolio consisting of a long cap and a short floor struck such that the net cost is zero is equivalent to a plain vanilla swap. We do not discuss this interpretation because of the complex nature of interest rate options.

Package of Futures (Forward) Contracts

Contrast the position of the counterparties in an interest rate swap to the position of a long and short price-based interest rate futures (rate-based forward) position. The long futures (short forward position) position gains if interest rates decline and loses (gains) if interest rates rise; this is similar to the risk/return profile for a fixed-rate receiver. The risk/return profile for a fixed-rate payer is similar to that of short futures position (long forward position):

there is a gain if interest rates increase and a loss (gain) if interest rates decrease. By taking a closer look at the interest rate swap we can understand why the risk/return profile are similar.

Consider party X's position in our swap illustration. Party X has agreed to pay 6% and receive 3-month LIBOR. More specifically, assuming a $100 million notional principal, party X has agreed to buy a commodity called "3-month LIBOR" for $1.5 million. This is effectively a 3-month forward contract where party X agrees to pay $1.5 million in exchange for delivery of 3-month LIBOR. If interest rates increase to 7%, the price of that commodity (3-month LIBOR) in the market is higher, resulting in a gain for the fixed-rate payer, who is effectively long a 3-month forward contract on 3-month LIBOR. The fixed-rate receiver is effectively short a 3-month forward contract on 3-month LIBOR. There is therefore an implicit forward contract corresponding to each exchange date.

Now we can see why there is a similarity between the risk/return profile for an interest rate swap and a forward contract. If interest rates increase to, say, 7%, the price of that commodity (3-month LIBOR) increases to $1.75 million (7% times $100 million divided by 4). The long forward position (the fixed-rate payer) gains, and the short forward position (the fixed-rate receiver) loses. If interest rates decline to, say, 5%, the price of our commodity decreases to $1.25 million (5% times $100 million divided by 4) The short forward position (the fixed-rate receiver) gains, and the long forward position (the fixed-rate payer) loses.

Consequently, interest rate swaps can be viewed as a package of more basic interest rate derivatives such as price-based futures of rate-based forward contracts. The pricing of an interest rate swap will then depend on the price of a package of forward contracts with the same settlement dates in which the underlying for the forward contract is the same reference rate. This principle is used when valuing swaps.

While an interest rate swap may be nothing more than a package of forward contracts, it is not a redundant contract for several reasons. First, maturities for forward or futures contracts do not extend out as far as those of an interest rate swap; for example, an interest rate swap with a term of 10 years or longer can be obtained. Second, an interest rate swap is a more transactionally efficient instrument. By this we mean that in one transaction an entity can effectively establish a payoff equivalent to a package of futures or forward contracts. Third, the interest rate swap market has grown in liquidity since its introduction in 1981; interest rate swaps now provide more liquidity than forward contracts, particularly long-dated (that is, long-term) forward contracts.

Package of Cash Market Instruments

To understand why a swap can also be interpreted as a package of cash market instruments, consider an investor who enters into the following transaction:

- Buy $100 million par of a 5-year floating-rate bond that pays 3-month LIBOR every three months.
- Finance the purchase by borrowing $100 million for five years on terms requiring a 6% annual interest rate paid every three months.

As a result of this transaction, the investor

- Receives a floating rate every three months for the next five years.
- Pays a fixed rate every three months for the next five years. Has no initial outlay.

The cash flows for this transaction are set forth in Table 40.1. The second column of the exhibit shows the cash flows from purchasing the 5-year floating-rate bond. There is a $100 million cash outlay and then 20 cash inflows. The amount of the cash inflows is uncertain because they depend on future LIBOR. The next column shows

Table 40.1 Cash Flows for the Purchase of a 5-Year Floating-Rate Bond Financed by Borrowing on a Fixed-Rate Basis

Transaction:
- **Purchase for $100 million a 5-year floating-rate bond: floating rate = LIBOR, quarterly pay**
- **Borrow $100 million for five years: fixed rate = 6%, semiannual payments**

3-Month Period	Cash Flow (In Millions of Dollars) From Floating Rate Bond*	Borrowing Cost	Net cash flow = Same as swaps's cash flow
0	−100	+100.0	0
1	+(LIBOR1/4) × 100	−1.5	+ (LIBOR1/4) × 100 − 1.5
2	+(LIBOR2/4) × 100	−1.5	+ (LIBOR2/4) × 100 − 1.5
3	+(LIBOR3/4) × 100	−1.5	+ (LIBOR3/4) × 100 − 1.5
...	....	...	....
...	....	...	....
...	....	...	....
19	+(LIBOR19/4) × 100	−1.5	+(LIBOR19/4) × 100 − 1.5
20	+(LIBOR20/4) × 100 + 100	−100 − 1.5	+(LIBOR20/4) × 100 − 1.5

*The subscript for LIBOR indicates the 3-month LIBOR as per the terms of the floating-rate bond at time *t*.

the cash flows from borrowing $100 million on a fixed-rate basis. The last column shows the net cash flows from the transaction. As the last column indicates, there is no initial cash flow (no cash inflow or cash outlay). In all 20 of the 3-month periods, the net position results in a cash inflow of LIBOR and a cash outlay of $1.5 million. This net position is identical to the position of a fixed-rate payer.

It can be seen from the net cash flow in Table 40.1 that a fixed rate payer has a cash market position that is equivalent to a long position in a floating-rate bond and a short position in a fixed-rate-bond—the short position being the equivalent of borrowing by issuing a fixed-rate bond.

What about the position of a fixed-rate receiver? It can be easily demonstrated that the position of a fixed-rate receiver is equivalent to purchasing a fixed-rate bond and financing that purchase at a floating rate, where the floating rate is the reference rate for the swap. That is, the position of a fixed-rate receiver is equivalent to a long position in a fixed-rate bond and a short position in a floating-rate bond.

DESCRIBING THE COUNTERPARTIES TO A SWAP

The terminology used to describe the position of a party in the swap market combines cash market jargon and futures market jargon. This is not surprising given that a swap position can be interpreted as a position in a package of cash market instruments or a package of futures/forward positions. As we have said, the counterparty to an interest rate swap is either a fixed-rate payer or fixed-rate receiver.

Table 40.2 lists how the counterparties to an interest rate swap agreement are described. To understand why the fixed-rate payer is viewed as "short the bond market," and the fixed-rate receiver is viewed as "long the bond market," consider what happens when interest rates change. Being short the bond market implies that the position becomes profitable when interest rates increase; long the bond market implies that the position becomes profitable when interest rates decrease. Those who borrow on a fixed-rate basis will benefit if interest rates rise because they have locked in a lower interest rate. But those who have a short bond position will also benefit if interest rates rise. Thus, a fixed-rate payer can be said to be short the bond market since both legs of the swap become more profitable when interest rates increase. A fixed-rate receiver benefits if interest rates fall since both legs of the swap becomes more profitable when interest rates decrease. A long position in a bond also benefits if interest rates fall, so terminology describing a fixed-rate receiver as long the bond market is not surprising. From our discussion of the interpretation of a swap as a package of cash market instruments, describing a swap in terms of the sensitivities of long and short cash positions follows naturally.

Table 40.2 Describing the Parties to a Swap Agreement

Fixed-rate payer	Fixed-rate receiver
• pays fixed rate in the swap	• pays floating rate in the swap
• receives floating in the swap	• receives fixed in the swap
• is short the bond market	• is long the bond market
• has bought a swap	• has sold a swap
• is long a swap	• is short a swap
• has established the price sensitivities of a longer-term liability and a floating-rate asset	• has established the price sensitivities of a longer-term asset and a floating-rate liability

Source: Kopprasch, Macfarlane, Ross, and Showers (1991)

BEYOND THE PLAIN VANILLA SWAP

Thus far we have provided a description of a plain vanilla swap. There are other types of swap structures—simple extensions and complex structures. We describe some of these in this section.

A simple extension of the plain vanilla swap is one in which the notional principal changes based on a specified schedule. A plain vanilla swap in which the notional principal decreases over time is called an *amortizing swap*. When the notional principal increases over time, the swap is referred to as an *accreting swap*. As we will see, once we know how to value a plain vanilla swap where the notional principal is constant over the life of the swap, it is a simple matter to value one with a varying notional principal.

There are swaps where both parties pay a floating rate based on two reference rates. For example, one party can make payments based on the 3-month Treasury bill rate plus some spread and the other party make payments based on 3-month LIBOR. Swaps where both parties make floating payments like the one just described are called *basis swaps*. Swaps can have the floating leg based on a reference rate other than LIBOR. For example, the floating leg might be based on the 2-year U.S. Treasury note yield. Swaps tied to U.S. Treasuries are referred to as Constant Maturity Treasury (CMT) swaps. Other intermediate-term floating reference rates are called *constant maturity swaps* (CMSs). A CMT is an example of a CMS. CMSs can either be structured in the basis swap framework (floating for a floating) or traditional (fixed for non-LIBOR-based floating).

Two complex swap structures are (1) a swap that starts at some future date and (2) an option on a swap. A swap that starts at some future date is called a *forward start swap*. An example of a forward start swap would be one where the obligation starts now but the swap starts two years from now and matures three years later for a total of five years. The swap fixed rate for the forward start swap is determined at the inception of the obligation. Swaps that combine some or all of these characteristics are also possible.

An option on a swap gives the owner of the option the right to enter into a swap at some future date. An option on a swap is called a *swaption*. A *payer's swaption* is one in which the owner of the option has the right to enter into a swap to pay a fixed rate and receive a floating rate. A *receiver's swaption* is one in which the owner of the option has the right to enter into a swap to receive a fixed rate and pay a floating rate. The swap fixed rate is the strike rate of the swaption.

SUMMARY

An interest rate swap is a derivative instrument. In this chapter, we have explained the basic features of interest rate swaps. The parties to the contract agree to exchange interest payments at specified future dates based on a notional amount. In a plain vanilla swap, over the life of the contract, one party agrees to pay the other party fixed interest payments based on the swap rate and the other party agrees to make floating-rate payments based on a specified reference rate. A swap position can be interpreted in two ways. First, it is equivalent to a position in a package of futures (forward) contracts. Second, it is equivalent to a position in a package of cash flows from buying and selling cash market instruments.

There are different types of interest rate swaps. These include interest rate swaps, plain vanilla swaps, amortizing swaps, accreting swaps, basis swaps, constant maturity swaps, forward rate swaps, and swaptions.

REFERENCES

Brown, K. C., and Smith, D. J. (1995). *Interest Rate and Currency Swaps: A Tutorial*. Charlottesville, VA: The Research Foundation of the Institute of Chartered Financial Analysts.

Buetow, G. W., and Fabozzi, F. J. (2001). *Valuation of Interest Rate Swaps and Options*. Hoboken, NJ: John Wiley & Sons.

Chisholm, A. M. (2004). *Derivatives Demystified: A Step-by-Step Guide to Forwards, Futures, Swaps and Options*. Chichester, UK: John Wiley & Sons.

Flavell, R. (2002). *Swap and Other Derivatives*. Chichester, UK: John Wiley & Sons.

Kopprasch, R. F., Macfarlane, J., Showers, J., and Ross, D.R. (1991). The interest rate swap market: Yield mathematics, terminology, and conventions. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Securities*, 3rd edition (pp. 1189–1217), Homewood, IL: Dow Jones-Irwin.

Ludwig, M. S. (1993). *Understanding Interest Rate Swaps*. New York: McGraw Hill.

Marshall, J. F., and Kapner, K. R. (1993). *Understanding Swaps*. New York: John Wiley & Sons.

Whaley, R. E. (2006). *Derivatives: Markets, Valuation, and Risk Management*. Hoboken, NJ: John Wiley & Sons.

CHAPTER 41

Interest Rate Options and Related Products

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

STEVEN V. MANN, PhD
Professor of Finance, Moore School of Business, University of South Carolina

MOORAD CHOUDHRY, PhD
Head of Treasury, KBC Financial Products, London

Abstract: Interest rate options and related option-type products are used by market participants to control interest rate risk. There are exchange-traded and over-the-counter interest rate options. The interest rate options product most traded on exchanges is a futures option. Over-the-counter options include options on specific securities, spread options, compound options, caps, and floors. With all over-the-counter products there is counterparty risk faced by the buyer of the option or option-related product.

Keywords: option, American option, European option, Bermudan option, exotic options, nonstandard options, exchange-traded options, over-the-counter options, dealer options, futures options, counterparty risk, spread options, compound option, split-fee option, front fee, back fee, cap, floor, caplets, captions, flotions

An interest rate option is a derivative instrument that differs from an interest rate forward contract, a futures contract, and a swap in terms of its risk and return characteristics. As such, an interest rate option can be employed to control interest rate risk in ways that are either not possible or too costly to achieve using forwards, futures, or swaps. Options, like most other financial instruments, can be traded either on an organized exchange or in an over-the-counter (OTC) market. We begin the chapter with a basic description of an option and then go on to describe exchange-traded interest rate options. The most popular form of an exchange-traded option is an option on a futures contract. We then examine the various types of OTC options and derivative instruments with option-like features.

BASIC OPTION CONTRACT

An *option* is a contract in which the writer of the option grants the buyer of the option the right, but not the obligation, to purchase from or sell to the writer something at a specified price within a specified period of time (or on a specified date). The writer, also referred to as the seller, grants this right to the buyer in exchange for a certain sum of money, which is called the option price or option premium. In effect, the writer is selling a promise in exchange for the option price. Conversely, the buyer pays the option price to obtain the writer's promise. The price at which the underlying may be bought or sold is called the exercise or strike price. The date after which an option is void is called the expiration date. Our focus in this chapter is on options where the "something" underlying the option is an interest rate instrument.

When an option grants the buyer the right to purchase the designated instrument from the writer (seller), it is referred to as a call option, or call. When the option buyer has the right to sell the designated instrument to the writer, the option is called a put option, or put.

An option is also categorized according to when the option buyer may exercise the option. There are options that may be exercised at any time up to and including the expiration date. Such an option is referred to as an *American option.* There are options that may be exercised only on the expiration date. An option with this feature is called a *European option.* There are also *Bermudan options,* which are hybrids between American and European option contracts. The distinguishing feature of a Bermudan option contract is that early exercise is possible but is restricted to certain dates in the option's life.

The maximum amount that an option buyer can lose is the option price. The maximum profit that the option writer can realize is the option price. The option buyer has substantial upside return potential, while the option writer faces substantial downside risk.

There are no margin requirements for the buyer of an option once the option price has been paid in full. Because the option price is the maximum amount that the investor can lose, no matter how adverse the price movement of the underlying instrument, there is no need for margin. Because the writer of an option has agreed to accept all of the risk (and none of the reward) of the position in the underlying instrument, the writer is generally required to put up the option price received as margin. In addition, as price changes occur that adversely affect the writer's position, the writer is required to deposit additional margin (with some exceptions) as the position is marked to market.

Notice that, unlike in a futures contract, one party to an option contract is not obligated to transact. Specifically, the option buyer has the right but not the obligation to transact. The option writer does have the obligation to perform. In the case of a futures contract, both buyer and seller are obligated to perform. Of course, the buyer of a futures contract does not pay the seller to accept the obligation, while an option buyer pays the seller the option price.

Consequently, the risk and reward characteristics of the two contracts are also different. In the case of a futures contract, the buyer of the contract realizes a dollar-for-dollar gain when the price of the futures contract increases and suffers a dollar-for-dollar loss when the price of the futures contract drops. The opposite occurs for the seller of a futures contract. Options do not provide symmetric payoffs. The most that the buyer of an option can lose is the option price. While the buyer of an option retains all the potential benefits, the gain is always reduced by the amount of the option price. The maximum profit that the writer may realize is the option price; this is offset against substantial downside risk. This difference is extremely important because managers can use futures to protect against symmetric risk and options to protect against asymmetric risk.

EXCHANGE-TRADED VERSUS OTC OPTIONS

There are *exchange-traded options* and *over-the-counter options.* Exchange-traded options have two advantages. First, the exercise price and expiration date of the contract are standardized. Second, as in the case of futures contracts, the direct link between buyer and seller is severed after the order is executed because of the interchangeability of exchange-traded options. The clearinghouse associated with the exchange where the option trades performs the same function in the options market that it does in the futures market.

OTC options are used in the many situations where an institutional investor needs to have a customized option because the standardized exchange-traded option does not satisfy its investment objectives. Investment banking firms and commercial banks act as principals as well as brokers in the OTC options market.

OTC options can be customized in any manner sought by an institutional investor. There are plain vanilla options such as options on a specific Treasury issue. The more complex OTC options created are called *exotic options* or *nonstandard options.* Examples of OTC options are given later in this chapter. While an OTC option is less liquid than an exchange-traded option, this is typically not of concern since institutional investors who use OTC options as part of a hedging or asset/liability strategy intend to hold them to expiration.

In the absence of a clearinghouse, the parties to any OTC contract are exposed to *counterparty risk.* In the case of a forward contract (an OTC contract), both parties face counterparty risk since both parties are obligated to perform. Thus, there is bilateral counterparty risk. In contrast, for an OTC option, once the option buyer pays the option price, it has satisfied its obligation. It is only the seller that must perform if the option is exercised. Thus, the option buyer is exposed to unilateral counterparty risk—the risk that the option seller will fail to perform.

FUTURES OPTIONS

The underlying for an interest rate option can be a fixed income security or an interest rate futures contract. The former options are called options on physicals. In the United States, there are no actively exchange-traded options on physicals. Options on interest rate futures are called *futures options.* The actively traded interest rate options on exchanges are futures options.

A futures option gives the buyer the right to buy from or sell to the writer a designated futures contract at the strike price at any time during the life of the option. If the futures option is a call option, the buyer has the right to purchase one designated futures contract at the strike price. That is, the buyer has the right to acquire a long futures position in the designated futures contract. If the buyer exercises the call option, the writer acquires a corresponding short position in the futures contract.

A put option on a futures contract grants the buyer the right to sell a designated futures contract to the writer at the strike price. That is, the option buyer has the right to acquire a short position in the designated futures contract. If the put option is exercised, the writer acquires a corresponding long position in the designated futures contract.

Because the parties to the futures option will realize a position in a futures contract when the option is exercised, the question is: what will the futures price be? That is, at what price will the long be required to pay for the instrument underlying the futures contract, and at what price will the short be required to sell the instrument underlying the futures contract?

Upon exercise, the futures price for the futures contract will be set equal to the strike price. The position of the two parties is then immediately marked to market in terms of the then-current futures price. Thus, the futures position of the two parties will be at the prevailing futures price. At the same time, the option buyer will receive from the option seller the economic benefit from exercising. In the case of a call futures option, the option writer must pay the difference between the current futures price and the strike price to the buyer of the option. In the case of a put futures option, the option writer must pay the option buyer the difference between the strike price and the current futures price.

For example, suppose an investor buys a call option on a futures contract in which the strike price is 85. Assume also that the futures price is 95 and that the buyer exercises the call option. Upon exercise, the call buyer is given a long position in the futures contract at 85, and the call writer is assigned the corresponding short position in the futures contract at 85. The futures positions of the buyer and the writer are immediately marked to market by the exchange. Because the prevailing futures price is 95 and the strike price is 85, the long futures position (the position of the call buyer) realizes a gain of 10, while the short futures position (the position of the call writer) realizes a loss of 10. The call writer pays the exchange 10, and the call buyer receives from the exchange 10. The call buyer, who now has a long futures position at 95, can either liquidate the futures position at 95 or maintain a long futures position. If the former course of action is taken, the call buyer sells a futures contract at the prevailing futures price of 95. There is no gain or loss from liquidating the position. Overall, the call buyer realizes a gain of 10. The call buyer who elects to hold the long futures position will face the same risk and reward of holding such a position, but still realizes a gain of 10 from the exercise of the call option.

Suppose instead that the futures option is a put rather than a call, and the current futures price is 60 rather than 95. Then if the buyer of this put option exercises it, the buyer would have a short position in the futures contract at 85; the option writer would have a long position in the futures contract at 85. The exchange then marks the position to market at the then-current futures price of 60, resulting in a gain to the put buyer of 25 and a loss to the put writer of the same amount. The put buyer, who now has a short futures position at 60, can either liquidate the short futures position by buying a futures contract at the prevailing futures price of 60 or maintain the short futures position. In either case, the put buyer realizes a gain of 25 from exercising the put option.

OVER-THE-COUNTER INTEREST RATE OPTIONS

OTC interest rate options are created by commercial banks and investment banks for their clients. Dealers can customize the expiration date, the underlying, and the type of exercise. For example, the underlying could be a specific fixed income security or a spread between yields in two sectors of the fixed income market.

In addition to American- and European-type options, an OTC option can be created in which the buyer may exercise prior to the expiration date but only on designated dates. As noted, such options are referred to as Bermuda options. With an OTC option, the buyer need not pay the option price at the time of purchase. Instead, the option price can be paid at the expiration or exercise date. For such options, the option writer, as well as the option buyer, is exposed to counterparty risk.

In the OTC option market, there are plain vanilla and exotic options. Plain vanilla options are options on specific securities or on the spread between two sectors of the bond market. Exotic options have more complicated payoffs, and we do not review these in this chapter.

Options on a Specific Security

Institutional investors who want to purchase an option on a specific Treasury security or a Ginnie Mae pass-through can do so on an OTC basis. There are government and mortgage-backed securities (MBS) dealers who make a market in options on specific securities. OTC or dealer options typically are purchased by institutional investors or mortgage bankers who want to hedge the risk associated with a specific security. Typically, the maturity of the option coincides with the time period over which the buyer of the option wants to hedge, so the buyer is usually not concerned with the option's liquidity.

A popular option used by mortgage originators for hedging forward delivery is an option on a specific MBS. Typically, the underlying security is a TBA (pools to be announced) agency pass-through security (Ginnie Mae, Fannie Mae, or Freddie Mac). The settlement process in the MBS market is forward delivery. The exercise of a mortgage option means the delivery of that security in the month specified in the option. Options are of the European type.

Spread Options

Some institutional investors may have exposure not only to the level of rates but the spread between two yields. It is difficult to hedge against spread risk with current exchange-traded options. As a result, several dealer firms have developed proprietary products for such purpose. These options can be structured with a payoff in one of the following ways should the option expire in-the-money. First, there could be a cash settlement based on the amount that the option expires in-the-money. Second, there could be an exchange of ownership of the two securities underlying the option. It is difficult to structure options with a settlement based on an exchange of securities, but there are institutional investors who desire this type of structure.

Next, we discuss two types of *spread options*—an option on the yield curve and an option on the spread between MBS and Treasury securities.

Yield Curve Spread Option

The reason for the popularity of yield curve spread options is that there are many institutional investors whose performance is affected by a change in the shape of the yield curve. As an example of a yield curve spread option, consider the Goldman Sachs' product called SYCURVE. This option represents the right to buy (in the case of a call option) or sell (in the case of a put option) specific segments of the yield curve. "Buying the curve" means buying the shorter maturity and selling the longer maturity; "selling the curve" means selling the shorter maturity and buying the longer maturity. The curve is defined by the spread between two specific maturities. They could be the 2-year/10-year spread, the 2-year/30-year spread, or the 10-year/30-year spread. The strike is quoted in basis points.

The yield spread is measured by the long maturity yield minus the short maturity yield. For a call option to be in-the-money at the expiration date, the yield spread must be positive; for a put option to be in the money at the expiration date, the yield spread must be negative. For example, a 25-basis-point call option on the 2-year/10-year spread will be in-the-money at the expiration date if:

$$\text{10-year yield} - \text{2-year yield} > 25 \text{ basis points}$$

A 35-basis-point put option on the 10-year/30-year spread will be in-the-money at the expiration date if:

$$\text{30-year yield} - \text{10-year yield} < 35 \text{ basis points}$$

Yield curve options such as the SYCURVE are cash settlement contracts. In the case of the SYCURVE, if the option expires in-the-money, the option buyer receives \$0.01 per \$1 of notional amount, per in-the-money basis point at exercise. That is:

$$\text{Amount option expires in-the-money (in basis points)} \\ \times \$0.01 \times \text{notional amount}$$

For example, suppose that \$10 million notional amount of a 2-year/10-year call is purchased with a strike of 25 basis points. Suppose at the expiration date the yield spread is 33 basis points. Then the option expires 8 basis points in-the-money. The cash payment to the buyer of this option is

$$8 \times \$0.01 \times \$10{,}000{,}000 = \$800{,}000$$

From this amount, the option price must be deducted.

MBS/Treasury Spread Option

Some institutional investors seek to control the spread risk between the yield on MBSs and Treasuries. One example of an option on this spread is Goldman Sachs's MOTTO (mortgages over Treasury) option. The buyer of a MOTTO call option benefits if MBSs outperform Treasuries; the buyer of a MOTTO put option benefits if Treasuries outperform MBSs.

As noted earlier in discussing MBS options, the structuring of MOTTO options is complicated by the nuances of the MBS market. For the particular Treasury, the calculation of its yield at the expiration date is straightforward given its price at the expiration date. However, at the expiration date, while the market price of a generic agency MBS with a given coupon rate is known, its yield is not uniquely determined. The yield depends on the prepayment assumption, which determines the particular security's cash flow. This yield is called the cash-flow yield, and the prepayment assumption is commonly called the prepayment speed. Each MBS dealer has a proprietary prepayment model to project the speed. One important factor in a prepayment model is the yield level relative to the coupon rate paid on the mortgages in the underlying mortgage pool. Thus, the yield on an MBS depends on the prepayment speed, which, in turn, depends on the yield level.

One possible way to handle this problem is to specify at the outset of the option the prepayment speed that should be used to determine the yield on an MBS given the Treasury yield at the expiration date. Specifically, the higher the Treasury yield, the lower the prepayment speed. However, it is not only the yield level but also the shape of the yield curve that affects the prepayment speed. Structuring a MOTTO such that the prepayment speed for all possible combinations of yield curves and yield levels would be difficult. Consequently, a MOTTO is structured so that an in-the-money option at the expiration date can be settled by the exchange of the two underlying securities.

COMPOUND OPTIONS

A *compound option* or *split-fee option* is an option to purchase an option. We can explain the elements of a compound option by using a long call option on a long put option. This compound option gives the buyer of the option the right but not the obligation to require the writer of the compound option to sell the buyer a put option. The compound option would specify the following terms:

1. The day on which the buyer of the compound option has the choice of either requiring the writer of the option to sell the buyer a put option or allowing the option to expire. This date is called the extension date.
2. The strike price and the expiration date of the put option that the buyer acquires from the writer. The expiration date of the put option is called the notification date.

The payment that the buyer makes to acquire the compound option is called the *front fee.* If the buyer exercises the call option in order to acquire the put option, a second payment is made to the writer of the option. That payment is called the *back fee.* An option that allows the option buyer to purchase a put option is called a caput. A Cacall grants the option buyer the right to purchase a call option.

Compound options are most commonly used by mortgage originators to hedge pipeline risk. They can also be used in any situation when a manager needs additional time to gather information about the need to purchase an option.

CAPS AND FLOORS

An important option combination in debt markets is the *cap* and *floor,* which are used to control interest rate risk exposure. Caps and floors are combinations of the same types of options (calls or puts) with identical strike prices but arranged to run over a range of time periods. The main instruments used to control interest-rate risk, including short-dated interest-rate futures and forward-rate agreements (FRAs). For example, a corporation that desires to protect against a rise in future borrowing costs could buy FRAs or sell futures. These instruments allow the user to lock in the forward interest rate available today. However, such positions do not allow the hedger to gain if market rates actually move as feared/anticipated. Hedging with FRAs or futures can prevent loss but at the expense of any extra gain. To overcome this, the hedger might choose to construct the hedge using options. For interest rate hedges, primary instruments are the cap and floor. (The terms "cap" and "floor" are not to be confused with floating-rate note products that have caps and/or floors that restrict how much a floater's coupon rate can float.)

Caps and floors are agreements between two parties whereby one party, for an up-front fee, agrees to compensate the other if a designated interest rate (called the reference rate) is different from a predetermined level. The party that benefits if the reference rate differs from a predetermined level is called the buyer, and the party that must potentially make payments is called the seller. The predetermined interest rate level is called the strike rate. An interest rate cap specifies that the seller agrees to pay the buyer if the reference rate exceeds the strike rate. An interest rate floor specifies that the seller agrees to pay the buyer if the reference rate is below the strike rate.

The terms of an interest rate agreement include (1) the reference rate, (2) the strike rate that sets the cap or floor, (3) the length of the agreement, (4) the frequency of reset, and (5) the notional amount (which determines the size of the payments). If a cap or a floor is in-the-money on the reset date, the payment by the seller is typically made in arrears.

Some commercial banks and investment banks now write options on interest rate caps and floors for customers. Options on caps are called *captions.* Options on floors are called *flotions.*

Caps

A cap is essentially a strip of options. A borrower with an existing interest-rate liability can protect against a rise in interest rates by purchasing a cap. If rates rise above the cap, the borrower will be compensated by the cap payout. Conversely, if rates fall the borrower gains from lower funding costs and the only expense is the upfront premium paid to purchase the cap. The payoff for the cap buyer at a reset date if the value of the reference rate exceeds the cap rate on that date is as follows:

$$\begin{aligned}&\text{Notional amount} \times (\text{Value of the reference rate} - \text{Cap rate})\\&\quad \times (\text{Number of days in settlement period}/\\&\quad \text{Number of days in year})\end{aligned}$$

Naturally, if the reference rate is below the cap rate, the payoff is zero.

A cap is composed of a series of individual options or *caplets.* The price of a cap is obtained by pricing each of the caplets individually. Each caplet has a strike interest rate that is the rate of the cap. For example, a borrower might purchase a 3% cap (London Interbank Offered Rate [LIBOR] reference rate), which means that if rates rise above 3%, the cap will pay out the difference between the cap rate and the actual LIBOR. A one-year cap might be composed of a strip of three individual caplets, each providing protection for successive three-month periods. The first three-month period in the one-year term is usually not covered, because the interest rate for that period, as it begins immediately, will be known already. A caplet runs over two periods—the exposure period and the protection period. The exposure period runs from the date the cap is purchased to the interest reset date for the next borrowing period. At this point, the protection period begins and runs to the expiration of the caplet. The protection period is usually three months, six months, or one year, and will be set to the interest rate reset liability that the borrower wishes to hedge. Therefore, the protection period is usually identical for all the caplets in a cap.

Floors

It is possible to protect against a drop in interest rates by purchasing a floor. This is exactly opposite of a cap in that a floor pay outs when the reference rate falls below the strike rate. This would be used by an institution that wished to protect against a fall in income caused by a fall in interest rate—for example, a commercial bank with a large proportion of floating-rate assets. For the floor buyer, the payoff at a reset date is as follows if the value of the reference rate at the reset date is less than the floor rate:

Notional amount × (Floor rate − Value of the reference rate)
× (Number of days in settlement period/
Number of days in a year)

The floor's payoff is zero if the reference rate is higher than the floor rate.

Collars

The combination of a cap and a floor creates a *collar,* which is a corridor that fixes interest payment or receipt levels. A collar is sometimes advantageous for borrowers because it is a lower cost than a straight cap. A collar protects against a rise in rates and provides some gain if there is a fall down to the floor rate. The cheapest structure is a collar with a narrow spread between cap and floor rates.

Risk and Return Characteristics

In an interest rate cap and floor, the buyer pays an up-front fee, which represents the maximum amount that the buyer can lose and the maximum amount that the seller of the agreement can gain. The only party that is required to perform is the seller of the interest rate agreement. The buyer of an interest rate cap benefits if the reference rate rises above the strike rate because the seller must compensate the buyer. The buyer of an interest rate floor benefits if the reference rate falls below the strike rate because the seller must compensate the buyer.

How can we better understand interest rate caps and interest rate floors? In essence these contracts are equivalent to a package of interest rate options. As with a swap, a complex contract can be seen to be a package of basic contracts—options in the case of caps and floors.

The question is what type of package of options is a cap and a floor. It depends whether the underlying is a rate or a fixed income instrument. If the underlying is considered a fixed income instrument, its value changes inversely with interest rates. Therefore:

- For a call option on a fixed income instrument:
  1. Interest rates increase → fixed income instrument's price decreases → call option value decreases
  and
  2. Interest rates decrease → fixed income instrument's price increases → call option value increases
- For a put option on a fixed income instrument:
  1. Interest rates increase → fixed income instrument's price decreases → put option value increases
  and
  2. Interest rates decrease → fixed income instrument's price increases → put option value decreases

To summarize:

	When Interest Rates Increase	**When Interest Rates Decrease**
Value of long call	Decreases	Increases
Value of short call	Increases	Decreases
Value of long put	Increases	Decreases
Value of short put	Decreases	Increases

For a cap and floor, the situation is as follows:

	When Interest Rates Increase	**When Interest Rates Decrease**
Value of short cap	Decreases	Increases
Value of long cap	Increases	Decreases
Value of short floor	Increases	Decreases
Value of long floor	Decreases	Increases

Therefore, buying a cap (long cap) is equivalent to buying a package of puts on a fixed income instrument, and buying a floor (long floor) is equivalent to buying a package of calls on a fixed income instrument. In contrast, if the underlying is viewed as an option on an interest rate, then buying a cap (long cap) is equivalent to buying a package of calls on interest rates. Buying a floor (long floor) is equivalent to buying a package of puts on interest rates.

SUMMARY

An option is a contract in which the writer of the option grants the buyer the right, but not the obligation, to purchase from or sell to the writer something at a specified price within a specified period of time (or on a specified date). The option buyer pays the option writer (seller) a fee, called the option price. A call option allows the option buyer to purchase the underlying from the option writer at the strike price; a put option allows the option buyer to sell the underlying to the option writer at the strike price.

Interest rate options include options on fixed income securities and options on interest rate futures contracts, called futures options. There are exchange-traded options and OTC options. The actively traded exchange-traded options are futures options. OTC interest rate options are customized by dealers for their clients in terms of the expiration date, the underlying, and the type of exercise. An OTC option can be created in which the buyer may exercise prior to the expiration date but only on designated dates (so-called modified American or Atlantic or Bermuda options). An OTC option can be created whereby the buyer pays the premium at the expiration date.

There are OTC options on specific securities. There are OTC options on the spread between two yields. Spread options can be structured with a payoff that is either cash settled or requires an exchange of ownership of the two securities underlying the option. Two common spread options are options on the yield curve and options on the spread between mortgages and Treasuries.

A compound option (also called a split-fee option) is an option to purchase an option. The front fee for a compound option is the initial payment that the buyer makes. The back fee for a compound option is the fee paid by the buyer if the option is exercised.

An interest rate cap is an agreement whereby the seller agrees to pay the buyer if the reference rate exceeds the strike rate. An interest rate floor is an agreement whereby the seller agrees to pay the buyer if the reference rate is below the strike rate. The terms of a cap and floor set forth the reference rate, the strike rate, the length of the agreement, the frequency of reset, and the notional principal amount. An interest rate collar can be created by combining an interest rate cap and an interest rate floor. In an interest rate cap and floor, the buyer pays an up-front fee, which represents the maximum amount that the buyer can lose and the maximum amount that the seller of the agreement can gain.

Buying a cap is equivalent to buying a package of puts on a fixed income security, and buying a floor is equivalent to buying a package of calls on a fixed income security.

REFERENCES

Babbel, D., Bouyoucos, P., and Strickler, R. (1996). Capping the interest rate risk in insurance products. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Options* (pp. 437–463), Burr Ridge, IL: Irwin Professional Publishing.

Barr, W.A. (1996). Options on mortgage-backed securities. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Options* (pp. 127–143), Burr Ridge, IL: Irwin Professional Publishing.

Bhattacharya, A. K. (1996). Interest-rate caps, floors, and compound options. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Options* (pp. 143–163), Burr Ridge, IL: Irwin Professional Publishing.

Brauer, J. S., and Goodman, L. S. (1996). Hedging with options and option products. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Options* (pp. 345–368), Burr Ridge, IL: Irwin Professional Publishing.

Dickstein, A., and Braus, J. (2004). *Understanding and Using Mortgage Options*. Merrill Lynch Fixed Income Strategy Publication, April.

Fabozzi, F. J., Mann, S. V., and Choudhry, M. (2003). *Measuring and Controlling Interest Rate and Credit Risk*, 2nd edition, Hoboken, NJ: John Wiley & Sons.

Gartland, W. J., Letica, N. C., and Fabozzi, F. J. (1996). Overview of fixed-income contracts. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Options* (pp. 3–35), Burr Ridge, IL: Irwin Professional Publishing.

Gastineau, G. L. (2003). Exotic (nonstandard) options on fixed-income instruments. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Options* (pp. 49–74), Burr Ridge, IL: Irwin Professional Publishing.

McDermott, S. (1996). A survey of spread options for fixed-income investors. In F. J. Fabozzi (ed.), *The Handbook of Fixed Income Options* (pp. 75–126), Burr Ridge, IL: Irwin Professional Publishing.

Pitts, M., and Fabozzi, F. J. (1990). *Interest Rate Futures and Options*, Chicago: Probus Publishing.

Prendergast, J. R. (2003). The complexities of mortgage options. *Journal of Fixed Income*, March: 7–24.

Prendergast, J. R. (2006). Mortgage options. In F. J. Fabozzi (ed.), *The Handbook of Mortgage-Backed Securities*, 6th edition (pp. 1023–1040), New York: McGraw Hill.

CHAPTER 42

Introduction to Credit Derivatives

VINOD KOTHARI
Independent financial consultant and Visiting Faculty, Indian Institute of Management, Kolkata, India

Abstract: Credit derivatives are devices that provide for trading in generic credit risk of an entity, asset, or bunch of entities or bunch of assets. Credit risk is the risk inherent in credit, and credit is the very basis of our present society. Credit derivatives were first introduced in the early 1990s and are part of the market for financial derivatives. Since credit derivatives are presently not traded on any of the organized exchanges, they are a part of the over-the-counter (OTC) derivatives market. Credit derivatives include credit default swaps, total return swap, credit-linked notes, and credit spread options, as well as the fast-growing world of portfolio synthetic trades structured either as bespoke collateralized debt obligations (CDOs) or as index trades referenced to baskets of entities or asset-backed securities. Though still a relatively small part of the huge market for OTC derivatives, credit derivatives are growing faster than any other OTC derivative.

Keywords: credit derivatives, credit asset, reference entity, reference obligation, reference asset, reference portfolio, protection buyer, protection seller, risk seller, risk buyer, credit events, premium, deliverable obligations, protection payments, credit event payments, physical settlement, cash settlement, synthetic asset, unfunded asset, single name credit default swap, single name derivative, portfolio derivative, portfolio default swap, static portfolio, dynamic portfolio, structured credit trades, structured portfolio trades, index-based derivative, index trades, credit default swap indices, nth to default in a basket, fixed-rate payer, floating-rate payer, valuation method, first loss risk, credit default swap, total return swap, credit-linked notes, credit spread options, securitization, synthetic securitizations

Credit derivatives are derivative contracts that seek to transfer defined credit risks in a credit product or bunch of credit products to the counterparty to the derivative contract. The counterparty to the derivative contract either could be a market participant or could be the capital market through the process of securitization. The credit product might either be exposure inherent in a *credit asset* such as a loan, or might be generic credit risk such as bankruptcy risk of an entity. As the risks and rewards commensurate with the risks are transferred to the counterparty, the counterparty assumes the position of a virtual or synthetic holder of the credit asset.

The counterparty to a credit derivative product that acquires exposure to the risk synthetically acquires exposure to the entity whose risk is being traded by the credit derivative product. Thus, the credit derivative trade allows people to trade in the generic credit risk of the entity, without having to trade in a credit asset such as a loan or a bond. Given the fact that the synthetic market does not have several of the limitations or constraints of the market for cash bonds or loans, credit derivatives have become an alternative parallel trading instrument that is linked to the value of a firm—similar to equities and bonds.

Coupled with the device of securitization, credit derivatives have been rendered into investment products. Thus, investors may invest in credit-linked notes and gain credit exposure to an entity or a bunch of entities. Securitization linked with credit derivatives has led to the commoditization of credit risk. Apart from commoditization of credit risk by securitization, there are two other developments that seem to have contributed to the exponential growth of credit derivatives—index products and structured credit trading.

In the market for equities and bonds, investors may acquire exposure to either a single entity's stocks or bonds or to a broad-based index. The logical outcome of the increasing popularity of credit derivatives was credit derivative indices. Thus, instead of gaining or selling exposure to the credit risk of a single entity, one may buy or sell exposure to a broad-based index, or subindices, implying risk in a generalized, diversified index of names.

The idea of tranching or structured credit trading is essentially similar to that of seniority in the bond market—one may have senior bonds, pari passu bonds, or junior bonds. In the credit derivatives market, this idea has been carried to a much more intensive level with tranches representing risk of different levels. These principles have been borrowed from the structured finance market. Thus, on a bunch of 100 names, one may take either the first 3% risk, or the 3% to 7% slice of the risk, or the 7% to 10% slice, and so on.

The combination of tranching with the indices leads to trades in tranches of indices, opening doors for a wide range of strategies or views to take on credit risk. Traders may trade on the generic risk of default in the pool of names or may trade on correlation in the pool, or the way the different tranches are expected to behave with a generic upside or downside movement in the credit spreads, or the movement of the credit curve over time, and so on.

Quite often, the development of the hedge fund industry has been associated with the development of credit derivatives. Hedge funds are prominent in credit derivatives trades, particularly in case of the lower tranches of the structured credit spectrum. The hedge fund industry represents the segment of investor capital that is least regulated, risk neutral, out to seize opportunities arising out of mispricing, and so on. As the credit derivatives trades are almost completely unregulated and offer opportunities of short trades in credit that is difficult to accomplish in the bond market, the credit derivatives industry provides an excellent playing ground to hedge funds.

DERIVATIVES: THE BUILDING BLOCK OF CREDIT DERIVATIVES

The development of credit derivatives is a logical extension of the ever-growing array of derivatives trading in the market. The concept of a derivative is to create a contract that transfers some risk or some volatility. This risk or volatility may relate to the price or performance of a reference asset, event, a market price or any other economic or natural phenomenon. Such trade in risk does not mean a trade in the reference asset. The reference may remain with someone who is a complete stranger to the derivative contract. However, the derivative trade closely mimics the risks and returns of holding the underlying asset, or at least a segment thereof. Thus, derivatives bring about a completely independent trade in the risks/returns of an asset. For example, a trade in options or futures in equities may run completely independent of trades in equity shares.

Credit derivatives apply the same notion to a *credit asset*. A credit asset is the asset that a provider of credit creates, such as a loan given by a bank, or a bond held by a capital market participant. A credit derivative enables the stripping of the loan or the bond from the risk of default (or more risks, depending on the nature of the derivative), such that the loan or the bond can continue to be held by the originator or holder thereof, but the risk gets transferred to the counterparty. The counterparty buys the risk obviously for a premium, and the premium represents the reward to the counterparty.

Thus, credit derivatives essentially use the derivatives format to acquire or shift risks and rewards in credit assets, namely, loans or bonds, to other financial market participants. Like capital market derivatives, credit derivatives make it possible to hold a credit asset and either remove the risks in holding it and replace the same by a pure counterparty risk or risk is a safer asset. Reciprocally, credit derivatives make it possible to not hold a credit asset and yet synthetically create the position of risk and reward in a credit asset or portfolio of assets. (Note that the terms "synthetic transfer," "synthetic exposure," and "synthetic lending" use "synthetic" as opposed to real or natural. For example, a "synthetic transfer" would mean a transfer that is not really a transfer, but achieves the same purpose artificially or synthetically.)

SECURITIZATION: THE OTHER BUILDING BLOCK

Much of the significance that credit derivatives enjoy today is because of the marketability imparted by securitization. Credit derivatives would have mostly been a closely held esoteric market, but for the introduction of the securitization device to commoditize a credit derivative and bring it to the capital market.

Securitized credit derivatives, or synthetic securitization, is a device of embedding a credit derivative feature into a capital market security so as to transfer the credit risk into the capital markets. In the case of synthetic securitizations, the protection against the risk is ultimately provided by the capital markets.

The synthesis of credit derivatives with securitization techniques has complemented each other. Credit derivatives have acquired a new meaning when they were turned into marketable securities using securitization techniques; securitization, however, got a new impetus by opening up possibilities of keeping a whole portfolio of credit assets on the books and yet transfer the credit risks of the portfolio. Many erstwhile securitizers, particularly in Europe and Asia, prefer synthetic securitizations to cash transfers.

MEANING OF CREDIT DERIVATIVES

A credit asset is the extension of credit in some form: normally a loan, accounts receivable, installment credit, or a financial lease contract.

Every credit asset is a bundle of risks and returns: Every credit asset is acquired to make certain returns on the asset, and the probability of not making the expected return is the risk inherent in a credit asset. The credit asset may, of course, end up in a full or partial loss, which is also a case of volatility of return in that the return is negative.

There are several reasons why a credit asset may not generate the expected return to the holder: delinquency, default, losses, foreclosure, prepayment, interest rate movements, exchange rate movements, and so on.

A credit derivative contract intends to create a trade in either some risk or all the risk of volatility of return in a credit asset, without transferring the underlying asset. For example, if Bank A enters into a credit derivative with Bank B relating to a loan sitting on the balance sheet of Bank A, Bank B bears the risk, of course for a fee, inherent in the asset held by Bank A. A couple of significant points need to be noted here.

First, we made a reference to transfer of risk in a loan or portfolio of loans held by Bank A. Credit derivatives are essentially derivative deals, and for any derivative deal, it is not necessary that the reference asset must actually be held by any of the counterparties. For example, to buy a put on an equity share, it is not necessary for the put buyer to hold the equity share. Similarly, in order for Bank A to transfer the risk of a loan taken by a particular obligor, it is not necessary for Bank A to have actually given a loan to the obligor. In other words, without Bank A actually holding any credit exposure in the obligor, Bank A may sell the risk (that is, buy protection) and Bank B may buy the risk (that is, sell protection). The purpose of the protection buyer in a derivatives deal is not necessarily hedging—the protection buyer may be buying protection for trading purposes, that is, to be able to benefit from widening of spreads over time.

Second, in most cases, the transaction of credit derivatives is not referenced to particular loans—it is referenced to the generic risk of default of an entity. In other words, a credit derivative views credit risk as an independent commodity by itself and creates a trade in the credit risk of an entity.

The premium that Bank B earns for selling protection is representative of the credit risk premium being priced on the asset. Thus, the protection seller by selling protection is earning the credit spread, and is exposed to the risk of default of the reference entity. The position of the protection seller is equivalent to that of an actual lender.

Credit derivatives may thus be defined as arrangements that allow one party (the protection buyer or originator) to transfer, for a premium, the defined credit risk, or all the credit risk, computed with reference to a notional value, of a reference asset or assets, which it may or may not own, to one or more other parties (the protection sellers).

Quick Guide to Basic Jargon

The subject matter of a credit derivative transaction is a credit asset, that is to say, an asset or contract that gives rise to a relationship of a creditor and debtor. However, credit derivatives are usually not related to a specific credit asset but trade in the generic risk of default of a particular entity. The entity whose risk of default is being traded in is commonly referred to as the *reference entity.* There are cases where the credit derivative is linked not to the general default of the reference entity but the default of specific asset or portfolio of assets. This is called the *reference obligation, reference asset*, or the *reference portfolio.*

The party that wants to transfer the credit risks is called the *protection buyer* and the party that provides protection against the risks is called the *protection seller.* The two are mutually referred to as the counterparties. Protection buyer and protection seller may alternatively be referred to as the *risk seller* and the *risk buyer,* respectively.

We have mentioned above that it is not necessary for the protection buyer to actually own the reference asset: he might either be using the credit derivative deal as a proxy to transfer the risk of something else that he holds, or may be doing so for trading or arbitraging reasons. Irrespective of the motive, a derivative deal does not necessitate the holding of the reference asset by either of the counterparties, by which it is also obvious that the protection buyer need not hold the reference asset of the same value or for the same tenure for which the derivative deal is written.

Therefore, like most other derivatives, credit derivatives are written for a notional value, usually in denominations of $1 million. The premium to be paid by the protection buyer, and the protection payment to be made by the protection seller, are both computed with reference to this notional value. For the same reason, the tenure of the credit derivative does not have to coincide with the tenure of the credit asset.

Since the derivative deal focuses on the credit risk, it is necessary to define the credit risk. This is done by defining *credit events.* Credit events are the specific events on the happening of which protection payments will be made by the protection seller to the protection buyer. Parties may define their credit events; in OTC transactions taking place under the standard documentation of the International Swap and Derivatives Association (ISDA) standard documentation, credit events are chosen from out of the list of credit events specified by the ISDA. In the case of a total-rate-of-return swap, a type of a credit derivative discussed later, the entire credit risk of volatility of returns from a credit asset, without reference to the reasons therefore, is transferred to the protection seller, and therefore, the definition of credit events is relevant only for termination of the swap on its occurrence.

The *premium* is what the protection buyer pays to the protection seller over the tenure of the credit derivative. If there is no credit event during the tenure of the deal, the protection buyer pays the premium, and on efflux of time, the deal is closed. If there is a credit event, there will be protection payment due by the protection seller to the protection buyer, and the deal is closed without waiting for the tenure to be over. The *protection payments* or *credit event payments* are what the protection seller has to pay to the protection buyer should the credit event happen. The protection payment is either the outstanding par value plus accrued interest (computed with reference to the notional value) of the reference asset, or the difference between such par value plus accrued interest and the post-credit-event market value of the reference asset. In the former case, the protection buyer delivers the reference asset to the protection seller (called *physical settlement*) and the latter case, there is no transfer of the credit asset (called *cash settlement*) as the protection seller merely compensates the protection buyer for the losses suffered due to the credit event.

In either case, the protection payments are not connected with the actual losses suffered by the protection buyer.

In case the terms between the parties have fixed physical settlement as the mode, the protection buyer shall be required to deliver a defaulted obligation of the reference entity on default. Generally, the definition of such defaulted obligations is broad enough to allow the protection buyer to buy from out of several available obligations of the reference entity. Such obligations are called *deliverable obligations.* Both reference obligations and deliverable obligations are defined usually by characteristics. Hence, any obligation of the reference entity that satisfies the characteristics listed will be deliverable obligation. Quite obviously, the protection buyer will have the motivation to deliver the *cheapest-to-deliver* obligations.

For example, let us suppose a bank has an outstanding secured loan facility of $65 million, payable after seven years, given to a certain corporation, say X Corporation. The bank wants to shed a part of the risk of the said facility, say $50 million, and enters into a credit derivative deal with a counterparty, the protection seller. The bank is the protection buyer in this deal. The derivative deal is done for a notional value of $50 million for X Corporation as the reference entity and, say, with a tenure of five years. The reference obligation is "senior unsecured loans or bonds of the reference entity." The parties agree to physical settlement. In this deal, the bank will pay a premium of 80 basis points to the protection seller for the full term of the contract, that is, five years if a credit event does not occur. If a credit event occurs, the bank stops making payments up to the date of the credit event and seeks protection payment.

The type of credit derivative described in this illustration is called a *credit default* swap or simply *default swap* and is the most common form of a credit derivative.

In our example, the bank is buying protection basically for hedging purposes. However, it may be noted that there are mismatches between the actual loan held by the bank and the derivative. The amount of the loan is $65 million, where the notional value of the derivative is only $50 million. The actual loan is a secured loan facility, while the reference asset for the credit derivative is a senior unsecured loan. The term of the loan is seven years, while the term of the derivative is five years. We wish to emphasize that there may be complete disconnect between the actual credit asset, if at all held by the protection buyer, and the credit derivative. For the purpose of our discussion, it would be all the same if the protection buyer did not have any loan given to X Corporation, and was simply trying to buy protection hoping to make a profit when the premium for buying protection against X Corporation went above 80 basis points.

Since the transaction of credit derivative is referenced to "senior unsecured loans or bonds of X Corporation," the credit events (as defined by the parties) will be triggered if there is such event on any of the obligations of X Corporation that satisfy the characteristics listed for the reference obligations. Generally speaking, if there is a default on any of the loans or bonds of X Corporation, or if X Corporation files for bankruptcy, it would trigger a credit event.

The obvious purpose of the party buying protection in this case is to partially hedge against the risk of default of the exposure held by the protection buyer. The protection buyer, the bank in our example, actually holds a secured loan, but buys protection for a senior unsecured loan for two reasons. First, since the market trades in general risk of default of X Corporation, the defaults are typically defined with reference to unsecured loans as they are more likely to default than secured loans. Second, for the protection buyer the protection is stronger when it is referenced to an inferior asset than the one actually held by the bank in our example.

The protection seller is earning a premium of 80 basis points by selling protection. This party, of course, is exposed to the risk of default of X Corporation. In normal course, to create the same exposure, the protection seller would have to lend out money to X Corporation. In this case, the protection seller has acquired the exposure without any initial investment (see later in this chapter about funded derivatives). The objective of the protection seller might be simply to create and hold this exposure as a proxy for a credit asset to X Corporation. Alternatively, the protection seller might also be viewing the transaction as a trade: this party would stand to gain if the cost of buying protection against X Corporation declines to below 80 basis points. The protection seller may encash this gain either by buying protection at the reduced price, or by other means.

If the credit event does not happen over the five-year term of the contract, the derivative expires with the protection buyer having made periodic premium payments to the protection seller. If the credit event does happen, the protection buyer may choose to make a physical settlement. In that case, the protection buyer may well deliver an unsecured bond of X Corporation, as evidently, the possible recovery on the secured loan that X Corporation is holding will be better than the market price of the unsecured bonds of X Corporation. Thus, if the protection buyer purchases such bonds at a price of 30%, he would stand to make 70% of the notional value because the protection seller will obligated to pay to the protection buyer the par value of the defaulted assets that satisfy the characteristics of the deliverable obligations. The protection buyer may continue to hold the secured loan and recover it through enforcement of security interests or otherwise.

Synthetic Lending

Through a credit derivative contract, the protection buyer transfers defined credit risks of a reference asset to the protection seller. Assuming the protection buyer holds the reference asset, as is the case in the example above, what is the impact of the derivative on the protection buyer? The protection buyer still holds the reference asset, but has now transferred the defined credit risks. Instead, the protection buyer now has a risk on the protection seller. Should a defined credit event take place, the protection buyer is not concerned with receiving interest or principal on the reference obligation from the obligor; the protection buyer is rather concerned about getting the protection payment from the protection seller. So, there is a substitution of obligor risk by counterparty risk.

As far as the protection seller is concerned, the protection seller has not bought the reference asset, but is exposed to risks and rewards of the reference asset. Should the reference asset not default, the protection seller continues to get the premium that is obviously based on the credit risk of the obligor, and is therefore a reward related to the reference obligor. Should the credit event take place, the protection seller is exposed to the risk of having to make protection payments.

In other words, the protection seller has assumed risk and reward in the obligor, without actually lending to the obligor. The obligor is now the synthetic asset of the protection seller, as by the derivative contract, the protection buyer has synthetically substituted obligor exposure by counterparty exposure, and the protection seller has synthetically created a new asset, namely, exposure in the obligor.

Credit derivatives deals provide a new opportunity of synthetically creating assets—without actually creating a portfolio or lending. Instead of originating a loan, virtually the same position can be created synthetically by selling protection. (Note that this will be even more true in the case of total-rate–of-return swaps, discussed later, where the parties replicate the actual cash flows from a reference obligation.)

The credit asset so created is referred to as *synthetic asset* or *unfunded asset.*

Motivations of the Parties

The motivations of the protection buyer in our above example are easily understandable—the bank wants to transfer the risk of holding the exposure in X Corporation without transferring the asset. But a primary question arises on the motivation of the protection seller: Why would that party be willing to write protection on something never actually created by him.

Briefly speaking, credit derivatives have provided an easy way for banks to gain exposures in credit risks without having to actually create assets. Consider a bank, say Bank A, that specializes in lending to the office equipment industry. Over the years, this bank has acquired a specialized knowledge of the office equipment industry. Suppose further that there is another bank, Bank B, that, say, specializes in the textiles industry. Both these banks are specialized in their own industries, but both suffer from the risks of portfolio concentration. Bank A is concentrated in the office equipment industry and Bank B is focused on the textiles industry. Understandably, both banks should diversify their credit portfolios to be safer.

One obvious option for both of bank is as follows: Bank A should invest in an unrelated portfolio, say textiles, while Bank B should invest in a portfolio of credits in which it has no credit exposure, say, the office equipment industry. Doing so would involve inefficiency for both the banks, as Bank A does not know enough about the textiles industry as Bank B does not know anything about the office equipment industry.

Here, credit derivatives offer an easy solution: Without transferring their portfolio or reducing their portfolio concentration, both banks could buy into the risks of each other's portfolios by credit derivative deals. By doing so, both banks would have diversified their risks. Moreover, both banks have diversified their returns, as the premiums being earned by the derivative contract represent return from the portfolio held by the other bank.

The above example has depicted credit derivatives being a bilateral transaction—as a sort of a bartering of credit risks. As a matter of fact, credit derivatives can be completely marketable contracts: The credit risk inherent in a portfolio can be securitized and sold in the capital market just like any other capital market security. So anyone who buys such a security is inherently buying a fragment of the risk inherent in the portfolio, and the buyers of such securities are buying a fraction of the risks and returns of a portfolio held by the originating bank.

Credit derivatives allow parties who are completely strangers to the banking market to eat into the rewards and bear the risks of banking assets, which would be otherwise not be possible. For example, a capital market participant buying a synthetic security with an embedded derivative feature gets to create a synthetic loan asset. An insurance company would not have been allowed to enter the banking market at all, but credit derivatives enable it to sell protection, which is synthetically the same as writing a loan itself. This discussion also reveals how credit derivatives could replicate credit assets in different markets and geographies.

Credit derivatives succeed in creating a new derivative product parallel to a cash bond or obligation. This synthetic product can have structured or leveraged risk/reward positions, and therefore, can be a device for the markets to allow structured trading in a credit asset without, of course, investing in the asset at all.

ELEMENTS OF A CREDIT DERIVATIVE

Bilateral Deals and Capital Market Deals

A credit derivative may be a transaction between two counterparties, or may be a capital market transaction. Bilateral transactions between parties or dealers are normally referred to as OTC deals, since they take place between parties on an over-the-counter basis, as opposed to exchange-traded derivatives. The other possible format of a credit derivative deal is embedding the derivative into some capital market instrument, and offering such instrument to investors in the capital market.

The most basic distinction between capital market deals and counterparty or OTC deals is based on who the counterparty is. Obviously, the counterparty for any credit derivative deal is a specific party, and it is impossible to envisage a credit derivative where the "capital market" is the counterparty. However, capital market transactions intend to transfer the exposure to the capital market instruments by interposing special purpose vehicles (SPVs). In a capital market transaction, the risk is first transferred by the protection buyer to the SPV, which in turn transmits the risk into the market by issuing securities that carry an embedded derivative feature.

A credit derivative deal might either be linked with a single reference entity, called a *single-name default swap,* or a portfolio of entities, called a portfolio default swap. Since the market is essentially OTC, it is intermediated by dealers and brokers. For well-known reference entities, the market is quite liquid and bid-ask spreads are quite fine. Another very liquid part of the market is standardized index trades, discussed later.

Sometimes, credit derivative deals are embedded into capital market securities to make it an investment product. This takes the form of collateralized debt obligations (CDOs). CDOs might relate either to a pool of assets sitting on the balance sheet of a bank (called a balance sheet CDO) or a bunch of reference entities drawn from the market (called an arbitrage CDO).

OTC deals and capital market deals differ in terms of pricing as well—the pricing of OTC deals is based on prices quoted for the specific reference entity or index in the market. The risk is assessed and priced by market mechanism, which may inherently adopt one or more models for pricing credit derivatives. In a bespoke capital market transaction, the obligor portfolio is mostly diversified and the risk is assessed by the quality and extent of diversity of the pool. The pricing of the risk transfer is mostly implied by the negative carry inherent in the assets and liabilities of the SPV—that is, the rate of return that the investments of the SPV fetch, and the weighted average coupon of the liabilities.

Reference Asset or Portfolio

From the viewpoint of obligor specification, there are two types of credit derivatives: a single-obligor derivative (or *single-name derivative*) and a *portfolio derivative.* As implied by the name, a single-obligor credit derivative refers to an obligation of a specific named obligor, whereas a portfolio trade refers to specific obligations of a portfolio of obligors.

In either case, the reference is to obligations of the reference entity, such as an unsecured loan, or unsecured bond of the obligor. Parties may define the obligation either by making it specific such as a particular loan or a particular bond issue, or give a broad generic description—such as any loan or any bond and so on. Most of the OTC transactions are referenced to a generic senior unsecured loan of the reference entity, which is primarily chosen as representative of the risk of default, mostly leading to a bankruptcy, of an obligor on a plain unstructured credit.

In the case of portfolio default swaps, the portfolio may be a *static portfolio* or a *dynamic portfolio.* As implied by its name, a static portfolio is one in which the constituents of the obligor portfolio will remain fixed and known over time. In the case of a dynamic portfolio, though the total value of reference portfolio remains fixed, its actual composition may change over time as new obligors may be introduced into the pool, usually for those that have been repaid or prepaid, or those that have been removed due to failure to comply with certain conditions. It is obvious that the selection of the names forming part of the dynamic portfolio will be based on definite selection criteria, elaborately laid down in the transaction documents, so as to ensure that the reinstatement of obligors over time does not change the portfolio risk.

Structured Portfolio Trade

Where the credit derivative deal relates to a portfolio, it is possible to create tranches of the risk arising out of it. We have earlier briefly discussed the concept of tranches. Hence, it is possible for the protection buyer to come up with several tranches—say, junior, mezzanine, and senior tranche, or say 0% to 4%, 4% to 8% tranche, and so on. The protection buyer may either buy protection on all these tranches, or one or more than one of these. Such trades are called *structured credit trades,* or *structured portfolio trades.* The word "structured" puts such trades in line with other segments of structured finance, such as securitization. The word "structured" also implies that the number and sizing of the tranches are structured to suit investors' appetite for risk and urge for returns.

Basket Trades

Another common variety of a structured credit derivatives prevailing in the market is called a *basket derivative,* where the reference asset is a basket of obligations, and the credit event is *nth to default in a basket,* let us say, first to default in a basket of 10 obligors. So the deal is referenced to a basket of 10 defined obligors, each with a uniform notional value, and when any one out of the basket becomes the first to default, the protection payments will be triggered, and thereafter, the deal is terminated. Effectively, this might be a very efficient way of buying protection against a portfolio of 10 assets, while paying a much smaller premium. This is because the joint probability of more than one obligor defaulting in a basket of 10 obligors is very small, while the probability of any one of the 10 defaulting is much higher. So the losses of the protection seller are limited to only one of the 10 obligors, while at the same time providing needed protection against a larger portfolio to the protection buyer.

At times, parties might even transact a basket deal where protection is bought for the second-to-default obligor. The intent here is that the first or threshold risk will be borne by the protection buyer, but any subsequent loss after the first default will be transferred to the protection seller. Conceptually, the protection buyer has limited losses to the first default in the portfolio, seeking protection from the protection seller for the second default. The third or subsequent default in the portfolio is unprotected, but that is only a theoretical risk, as the probability of three defaults in an uncorrelated portfolio is nominal. Likewise, one may think of an nth to default basket swap.

Basket default swaps, like all portfolio trades, are structured with the parties taking a view on the inherent correlation in the basket. The higher the correlation in the basket, the risk of the first-to-default protection seller comes down, and that of the second-to-default protection seller goes up.

Index-Based Credit Derivative Trades

The idea of portfolio credit trades, structured or otherwise, was carried further with the introduction of the index trades and gained tremendous popularity. A single-name credit derivative allows the parties to trade in credit risk of a particular entity. A portfolio derivative allows parties to transact trade in the credit of a broad-based portfolio, say, a portfolio of 100 U.S. corporates. The selection of these 100 U.S. corporates may be done by the person who structures the transaction. However, to allow parties to trade on a common portfolio, index trades construct a standard pool of *n* number of names (or securities), and allows various traders to trade in such common portfolio. The common portfolio is known as the *index,* in line with indices of equities, bonds, or other similar securities. The advantage of index trades is that they allow the carrying out of structured trades in a generalized portfolio, so capital market participants may take views on the general corporate credit environment in the United States, or Europe, or so on. In view of their advantage over bespoke portfolio trades, that is, portfolios of names selected by the structurer, index trades have quickly grown to become a very large component of the credit derivatives market.

Protection Buyer

The protection buyer is the entity that seeks protection against the risk of default of the reference obligation. The protection buyer is usually a bank or financial intermediary that has exposure to credit assets, funded or unfunded. In such a case, the primary objective of a protection buyer is to hedge against the credit risks inherent in credit assets. The credit assets in case of OTC transactions are mostly corporations, or sovereigns, primarily emerging-market sovereigns. In the case of several CDOs, the assets can diversified obligor pools representing a broad cross-section of exposure in various industries. There have been several cases where risks on a portfolio of a very large number of obligors have been transferred through derivatives, for example, small- and medium-enterprise (SME) loans, auto leases, and so on.

At times, dealers could be buying protection for shorting credit assets, for the purpose of arbitraging by selling protection or otherwise gaining by way of a widening of credit spreads on the reference entity. Buying protection is the same as going short on a bond. The protection buyer gains if the credit quality of the reference entity worsens. One may also visualize that usually, among the bond market, equity market, and the credit derivatives market, there is a degree of correlation. Hence, the protection buyer shorts exposure on the entity by buying protection.

Buying of protection is also seen by the market as a convenient way of synthetically transferring the loan, while avoiding the problems associated with actual loan sales. Sale or securitization of loans involves various problems, depending on the jurisdiction concerned, relating to obligor notification, partial transfers, transfer of security interests, further lending to the same borrower, and so on. (Apart from the procedural issues related to transfer of loan portfolios, a major legal risk in a loan sale is generically referred to as the "true sale" risk, that is, the possibility that the sale of the loans will either be disregarded by a court or undone by a consolidation of the transferee with the transferor. For a detailed discussion on the true sale problems, see Kothari [2006].) Synthetic transfers, in contrast, avoid all of these problems, as the reference asset continues to stay with the originator.

In credit derivatives documentation, the protection buyer is also referred to as the *fixed-rate payer.* Perhaps this term is the remnant of the interest rate swap documentation.

Protection Seller

Earlier, we discussed briefly the motivations of the protection seller. To reiterate, the protection seller is mainly motivated by yield enhancement, or getting to earn credit spreads from synthetic exposures where direct creation of loan portfolios is either not possible or not feasible. In OTC transactions, the major protection sellers are insurance companies, banks, hedge funds, equity funds, and investment companies. In the case of CDOs, the protection sold is embedded in securities that are mostly rated, and the investors acquire these securities based on their respective investment objectives.

The protection seller may also be taking a trading view and expecting the credit quality of the reference entity to improve. Selling protection is equivalent of longing a bond—as the quality of the underlying entity improves, the protection seller stands to gain.

In credit derivatives documentation, the protection seller is also referred to as the *floating-rate payer.*

Funded and Unfunded Credit Derivatives

Typically, a credit derivative implies an undertaking by the protection seller to make protection payments on the occurrence of a credit event. Until the credit event happens, there is no financial investment by the protection seller. In this sense, a credit derivative is an unfunded contract.

However, quite often, for various reasons, parties may convert a credit derivative into a funded product. This may take various forms, such as:

- The protection seller prepays some kind of estimate of protection payments to the protection buyer, to be adjusted against the protection payments, if any, or else, returned to the protection seller
- The protection seller places a deposit or cash collateral with the protection buyer which the latter has a right to appropriate, in the case of protection payments.
- The protection buyer issues a bond or note which the protection seller buys, with a contingent repayment clause entitling the protection buyer to adjust the protection payments from the principal, interest, or both, payable on the bond or note.

The purpose of converting an unfunded derivative into a funded form may be variegated: It could either be a simple collateralization device for the protection buyer or may be the creation of a funded product that features a derivative and is therefore a restructured form of the original obligation with reference to which the derivative was initially written. When the funded derivative takes the form of a fixed income security, it is referred to as a credit-linked security or *credit-linked note,* which implies that a credit derivative has been embedded in a fixed-income security.

Credit Event

Credit events are the contingencies or the risk of being transferred in a credit derivative transaction. There are certain credit derivatives, such as total-rate-of-return swaps, where the reference to credit event is merely for closing out the transaction because the cash flows are swapped regularly; but most credit derivative deals refer to an event or events, upon the happening of which protection payments will be triggered.

The ISDA's standard documentation lists and elaborates different credit events for different types of credit derivative deals. For standard credit derivatives, there are six credit events: bankruptcy, failure to pay, obligation default, obligation acceleration, repudiation or moratorium, and restructuring. Parties are free to choose one or more credit events. If the parties use a non-ISDA document, they can define their own credit events as well. In most capital market transactions, credit events are given a structured meaning by the parties.

In OTC trades, the most common credit events are bankruptcy, failure to pay, and restructuring. Restructuring as a credit event has had a checkered history in the credit derivatives business, as a mere restructuring is not a case of default in common banking or credit parlance, and yet triggers protection payments in the case of credit derivatives. If a protection buyer holds a loan that gets restructured, say, with the borrower seeking extension of maturity by something like two years, theoretically, the protection buyer has not lost much money (except maybe on account of impairment of credit of the borrower), and may still seek compensation by delivering a cheapest-to-deliver asset of the reference entity that he may acquire from the market. To put reasonable curbs on what may be delivered pursuant to a restructuring event, ISDA documentation gives certain options to parties, essentially in the form of maturity limitations of the deliverable obligations.

It is quite possible for credit derivatives trades to not include restructuring as a credit event at all—for example, index trades do not include restructuring.

In the case of credit derivatives on asset-backed securities, the generic definitions of "bankruptcy" and "failure to pay" would obviously not be applicable; hence, there are unique credit events in the case of such contracts.

Notional Value

Earlier, we discussed the relevance of notional value in a derivative deal. Like all derivative deals, credit derivatives also refer to a notional value as the reference value for computing both the premium and the protection payments. Notional values are generally standardized into denominations of $1 million. However, capital market transactions can use their own nonstandard notional values.

There are certain derivatives in which the notional value is not fixed—it declines over time. This is where the derivative is linked with an amortizing loan or an asset-backed security where the underlying asset pool consists of amortizing assets.

Premium

The *premium* is the consideration for purchasing protection that the protection buyer pays to the protection seller over time. The premium is normally expressed in terms of basis points (bps). For example, a premium of 85 bps means on a notional value of $1 million the protection buyer will pay to the protection seller $8,500 as the annual premium. The premium is normally settled on a quarterly basis but typically accrues on a daily basis.

The premium may not be constant over time—there might be a step-up feature, meaning the premium increases after a certain date. This might be either to reflect the term structure of credit risk or simply for a perfunctory regulatory compliance as discussed next.

Tenure

The tenure is the term over which the derivative deal will run. The tenure comes to an end either by the efflux of time or upon happening of the credit event, whichever is earlier. For portfolio derivatives, the credit event on one of the obligors may not lead to termination of the derivative.

As we discussed earlier, the tenure of the credit derivative need not coincide with the maturity of the actual exposure of the protection buyer. However, for regulatory purposes, conditions for capital relief curtail the benefit of capital relief where there is a maturity mismatch between the tenure of the underlying credit asset and that of the credit derivative. So, the common practice in transactions where the protection buyer intends to seek a capital relief, but where the protection seller wants to give protection for only three years while the underlying exposure is for five years, is to quote a rate for three years, with a step-up after year 3, with an option to terminate with the protection buyer. The protection buyer will terminate the transaction due to the step-up feature, effectively getting protection only for three years, while theoretically, for regulatory purposes, the exposure is fully covered for five years.

Loss Computation

If a credit event takes place, the protection seller must make compensatory loss payments to the protection buyer, as in case of a standard insurance contract. However, the significant difference between a standard insurance contract and a credit derivative is that for the latter, it is not important that the protection buyer must actually suffer losses; nor is the amount of actual loss relevant. Losses of the protection seller are also known as the protection payment.

The loss computation and the payments required to be made by the protection seller are a part of the "settlement" of the contract. Obviously, the losses of the protection seller will depend on the settlement method—physical or cash. Where the terms of settlement are cash, the contract will provide for the manner of computing losses. Here, the loss is the difference between the par value of the reference asset (that is, the notional value plus accrued interest, as per terms of the credit), less the fair value on the valuation date. Most of the reference assets will not have any deterministic market values as such. Consequently, the method of computing the fair value is described in the contract in details. If the reference asset is something like a senior unsecured loan, the market value may be determined by taking an average of the quotes given by several independent dealers. Typically, the quotes are taken on more than one date, and, therefore, there are various valuation methods applicable, such as highest or average highest.

As significant as specifying the valuation method is the specification of the valuation date. Usually, a cooling-off period is allowed between the actual date of happening of an event of default and the valuation date. This is to allow for the knee-jerk reaction of the market values to be mitigated and more rational pricing of the defaulted credit asset to take place.

Computation of losses is not required for a type of derivative called binary swaps or fixed recovery swaps, where the protection seller is required to pay a particular amount to the protection buyer, irrespective of the actual losses or valuation.

Threshold Risk or Loss Materiality Provisions

Credit derivative contracts may sometimes provide for a threshold risk, up to which the losses will be borne by the protection buyer, and it is only when the losses exceed the threshold limit that a claim will lie against the protection seller. This is also called a materiality loss provision, under the understanding that only material losses will be transferred to the protection seller, even though the threshold limit may be quite high and not necessarily prevent immaterial losses from being claimed from the protection seller. In such cases, the more appropriate term is *first loss risk,* where the first loss risk up to the specified amount is borne by the protection buyer and it is only losses above the first loss amount that are transferred to the protection seller.

Cash and Physical Settlement

Settlement arises when the credit events take place. The terms of settlement could be either cash settlement or physical settlement. In the case of cash settlement, the losses computed as discussed above are paid by the protection seller to the protection buyer, and there is no transfer of the reference asset by the protection buyer. With physical settlement, the protection buyer physically delivers, that is, transfers an asset of the reference entity that satisfies the criteria for a deliverable obligation, and gets paid the par value of the delivered asset, limited, of course, to the notional value of the transaction. The concept of deliverable obligation in a credit derivative is critical, as the derivative is not necessarily connected with a particular loan or bond. Being a transaction linked with generic default risk, the protection buyer may deliver any of the defaulted obligations of the reference entity. However, to prevent against something like equity or other contingent securities from being delivered, transaction documents typically specify the characteristics of the deliverable obligations.

The general belief in the credit derivatives market is that losses of the protection seller are less in the case of a physical settlement than in the case of cash. This belief is quite logical, since the quotes in the case of a cash settlement are made by potential buyers of defaulted assets, who also hope to make a profit in buying the defaulted asset.

Physical settlement is more common where the counterparty is a bank or financial intermediary who can hold and take the defaulted asset through the bankruptcy process or resolve the defaulted asset. Physical settlement is, however, quite problematic where there are plenty of outstanding transactions referenced to an entity. This situation is almost certain to arise in the case of entities included in popular indices. When several protection buyers scout the market for buying defaulted assets, there might be a short squeeze in the market and an artificial inflation in the price of the defaulted security. In appreciation of these difficulties, the market has of late started moving in the direction of cash settlements or fixed recovery trades.

QUICK INTRODUCTION TO THE TYPES OF CREDIT DERIVATIVES

The following is a quick introduction to the various types of credit derivatives.

Credit Default Swap

A *credit default swap* can literally be defined as an option to swap a credit asset for cash, should it default. A credit default swap is essentially an option bought by the protection buyer and written by the protection seller. The strike price of the option is the par value of the reference asset. Unlike a capital market option, the option under a credit default swap can be exercised only when a credit event takes place.

In a credit default swap, if a credit event takes place, the protection buyer at his option may swap the reference asset or any other deliverable obligation of the reference obligor, either for cash equal to the par value of the reference asset, or get compensated to the extent of the difference between the par value and market value of the reference asset.

Credit default swaps are the most important type of credit derivative in use in the market.

Total Return Swap

A credit default swap protects the protection buyer against losses when a credit event happens. However, a credit event is a rare event. The holder of a credit asset is not merely concerned with losses in the event of default, but mark-to-market losses, since the latter is more frequent.A credit asset might continue to give mark-to-market losses for quite some time before it actually ripens into a default.

As the name implies, a total-rate-of-return swap or *total return swap* is a swap of the total return out of a credit asset swapped against a contracted prefixed return. The idea in a total-rate–of-return swap is to protect the protection buyer against mark-to-market losses as well; hence, the parties swap the total return from the reference credit asset or pool of assets. The total return out of a credit asset is reflected by the actual earnings realized from the reference asset plus the actual appreciation/depreciation in its price over time. The total returns from a credit asset may be affected by various factors, some of which may be quite extraneous to the asset in question, such as interest rate movements. Nevertheless, the protection seller in a total return swap guarantees a prefixed spread to the protection buyer, who in turn agrees to pass on the actual collections and actual variations in prices on the credit asset to the protection seller.

So periodically, the protection buyer swaps (the actual return on a notional value of the reference asset), in lieu of (a certain spread on a reference rate, say, LIBOR + 60 bps).

Credit-Linked Notes

A credit-linked note (CLN) is a securitized form of credit derivative that converts a credit derivative into a funded form. Here, the protection buyer issues notes or bonds that implicitly carry a credit derivative. The buyer of the CLN sells protection and prefunds the protection sold by way of subscribing to the CLN. Should there be a credit event payment due from the protection seller, the amounts due on the notes/bonds on account of credit events will be appropriated against the same, and the net, if any, will be paid to the CLN holder. A CLN carries a coupon which represents the interest on the funding and the credit risk premium on the protection sold, that is to say, the protection inherently sold via the CLN is compensated in the form of the coupon on the CLN. Obviously, the maximum amount of protection that the CLN holder provides is the amount of principal invested in the CLN.

Credit Spread Options

A *credit spread option* is basically a call or put option on an asset exercisable based on a certain spread. The call or put is an option with the holder, who is the protection buyer. Let us say a protection buyer agrees with the protection buyer that should the spread of a particular bond exceed a particular spread over LIBOR (referred to as the strike spread), then the protection buyer will have the option, as usual, of either a physical settlement of the reference obligation at the strike spread, or net settlement.

The option to put the asset can be said to be the option to call a predetermined spread. In other words, the protection buyer intends to protect a particular spread over a base rate and indicates a negative view on the reference obligation. On the contrary, if the protection buyer holds a positive view on the reference obligation, he may enter into an option to call the asset or put the spread.

Credit spread options are not related to events of default as, understandably, the movement in spreads can be related to various factors besides credit events.

In regulatory standards of most countries, credit spread options are not considered for regulatory capital relief. See, for example, paragraph 8.2.1 of FSA, the United Kingdom's regulatory requirements on credit derivatives state that "protection bought using a credit spread option is ignored for capital purposes." That they are not eligible for regulatory capital relief is a major reason why spread options have not become as popular as the other types of credit derivatives.

CREDIT DERIVATIVES AND TRADITIONAL FINANCIAL GUARANTEE PRODUCTS

Credit derivatives, particularly credit default swaps, have very close affinity with some traditional financial guarantee contracts such as:

- Bond insurance
- Letters of credit
- Revolving credit
- Financial guarantees

Credit Derivatives and Guarantees

The traditional guarantee contract provides for payment by the guarantor to the creditor in case of a default by the debtor. Credit derivatives, particularly credit default swaps, might have an apparent similarity with traditional guarantees. However, the similarity goes no further.

In a traditional guarantee, the intent of the guarantor is to protect the creditor from losses and put the creditor at par with what would have been received had the original debtor not defaulted. Thus, the payments by a guarantor are typically due only:

- When the principal debtor has defaulted.
- To the extent of the loss or damage suffered by the creditor.

Credit default swaps, however, are not limited to "default" as such but generally extend to cover events such as bankruptcy, compromise, and restructuring. Besides, for credit default swaps, the payments to be made by the protection seller might be either a prefixed amount or based on a valuation, which may or may not equal to the damage suffered by the protection buyer.

Another significant difference lies in the fact that a guarantee is always a trilateral contract: the guarantor, debtor, and creditor are all parties to the contract of guarantee. Credit default swaps, however, are purely a contract between the protection buyer and seller, and the obligor may not come to know about the contract at all.

The differences between traditional guarantees and credit derivatives are summarized in Table 42.1.

Table 42.1 Differences between Traditional Guarantees and Credit Derivatives

	Credit Default Swap	Financial Guarantee
Nature of the contract	A contract whereby the protection seller makes predefined payments to the protection buyer on happening of certain events. In contract law parlance, it is an independent contract, neither a contract of guarantee, nor indemnity.	A contract whereby the guarantor will pay the sums due and payable by the principal debtor on the failure of the latter to pay. In contract law parlance, it is a contract of guarantee.
Parties to the contract	The protection seller and the protection buyer. There is no contractual relationship with the obligor and the protection seller.	The guarantor (protection provider), surety (protection seeker), and the principal debtor (obligor). There is a contractual relationship between the guarantor and the obligor.
Consideration	Payment of certain fees or premium by the protection buyer to the protection seller.	Consideration needs to exist between the guarantor and the principal debtor—normally, a guarantee commission.
Assumption of rights against the obligor	Upon default, unless the protection buyer delivers the asset to the protection seller, the latter has no rights against the obligor.	As per law, if the guarantor makes payment of any sum due by the principal debtor, he becomes the creditor of the principal debtor for the sum so paid.
Nature of protection	Protection is provided against predefined credit events, not limited to defaults.	Protection is normally provided against default by the obligor.
Nature of payments upon default	Where the predefined credit events take place, the protection seller is to make the predefined credit event payments to the protection buyer.	Where the default by obligor takes place, the surety is first expected to proceed against the obligor. Having exhausted remedies, the surety can claim defaulted payments from the guarantor.
Relationship between the protection provided and the obligation	Credit default swaps are not necessarily connected with the existence and extent of the payment obligation of the obligor: While the obligation may be different, the default swap might be referenced to a different asset. The notional amount for the swap might also differ from the actual obligation.	Guarantees are necessarily connected with a specific obligation of the obligor.
Tradability	Credit default embedded in credit-linked notes are tradable.	Guarantees are bilateral contracts and are not tradable.
Pricing	Credit default swaps are priced by the market.	Guarantees are priced bilaterally.
Marking to market	CDS are marked to market.	Guarantees are not marked to market.
Documentation	Standard documentation as developed by ISDA.	No standard documentation.

CREDIT DERIVATIVES AND SECURITIZATION

Securitization is the device whereby financial assets such as receivables are converted into marketable securities and are offered to investors, usually with credit enhancements. As a generic process, securitization refers to the very process of converting something that is not a marketable security into one; the term "asset securitization" is sometimes used specifically to refer to the application of the device to converting assets into securities.

Asset securitization and credit derivatives are contradictory but have been used as mutually complementary. An asset securitization results in the transfer of assets, mostly while the risks are retained by the originator in the form of the credit enhancements. In the case of credit derivatives, there is no transfer of assets, but a mere transfer of risks. Securitization results in the creation of liquidity, while credit derivatives are unfunded as far as the protection buyer is concerned.

However, securitization and credit derivatives have joined hands to result in *synthetic securitizations,* which can be viewed as a securitization of a credit derivative, that is, conversion of a credit derivative into marketable securities. Synthetic securitization has provided wider use and far-reaching effect to credit derivatives, while at the same time providing greater flexibility for those seeking to use asset securitization.

SUMMARY

Credit derivatives have brought about a market where credit risk of entities can be traded independent of loans or bonds of the particular entities. Credit risk has thus become a commodity, and credit derivatives have effectively commoditized credit risk. The increased liquidity in the credit derivatives market for popularly traded names has created a market that parallels that for equities and bonds, and investors and traders may trade in credit default swaps with the same trading or investing intent as in case of equities and bonds. Index trades have further enabled traders to take a view on generalized portfolios of credits. Credit derivatives were essentially envisaged as hedging products but have actually become important tools of trading. The most common type of liquid credit derivative is a credit default swap, but total-rate-of-return swaps and option trades are also common. Portfolio default swaps, referenced to pools of names, are importantly linked to the correlation inherent in the names in the pool, besides the credit quality of those names.

REFERENCES

Anson, M. J. P., Fabozzi, F. J., Choudhry, M., and Chen, R-R. (2004). *Credit Derivatives: Instruments, Pricing, and Applications*, Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J., Davis, H., and Choudhry, M. (2007). *Introduction to Structured Finance*. Hoboken, NJ: John Wiley & Sons.

Kothari, V. (2002). *Credit Derivatives and Synthetic Securitisation: Guide to Commodisation of Credit Risk*. Kolkata, India: Academy of Financial Services.

Kothari, V. (2006). *Securitization: The Financial Instrument of the Future*, 3rd edition. Singapore: John Wiley & Sons.

Lucas, D. J., Goodman, L. S., and Fabozzi, F. J. (2006) A framework for evaluating trades in the credit derivatives market. *Journal of Trading*, Fall: 58–69.

Schönbucher, P. J. (2003). *Credit Derivatives Pricing Models: Models, Pricing and Implementation.*, Hoboken, NJ: John Wiley & Sons.

CHAPTER 43

Fixed Income Total Return Swaps

MARK J. P. ANSON, PhD, JD, CPA, CFA, CAIA
President and Executive Director of Nuveen Investment Services

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

MOORAD CHOUDHRY, PhD
Head of Treasury, KBC Financial Products, London

REN-RAW CHEN, PhD
Associate Professor of Finance, Rutgers University

Abstract: The total return on a bond, bond portfolio, or bond index is taken into account interest income and any capital gain or loss realized. In the fixed income market, derivative instruments that allow an investor to obtain exposure to the total return of a bond, bond portfolio, or bond index without the actual purchase of the underlying is available. This derivative instrument is a total return swap. Similarly, a total return swap can be used to short the underlying without the need to borrow it.

Keywords: total return swap, total return bond index swap, total return index swap, swap buyer, swap seller, interest rate swap, basis swap, funding leg, synthetic repo

A total return swap is a swap in which one party makes periodic floating rate payments to a counterparty in exchange for the total return realized on a reference asset (or underlying asset). In the fixed income market, reference asset could be a credit-risky bond, a reference portfolio consisting of bonds or loans, or an index representing a sector of the bond market. We first explain how a total return swap can be used when the reference asset is a credit-risky bond and a loan. While these types of total return swaps are more aptly referred to as total return credit swaps, we will simply refer to them as total return swaps. When the bond index consists of a credit risk sector of the bond market, the total return swap is referred to as a *total return bond index swap* or in this chapter as simply a *total return index swap*. We will explain how a total return index swap offers asset managers and hedge fund managers greater flexibility in managing a bond portfolio. (For the valuation of fixed income total return swaps, see Chapter 48 of Volume III).

ECONOMICS OF A TOTAL RETURN SWAP

A total return of a reference asset includes all cash flows that flow from it as well as the capital appreciation or

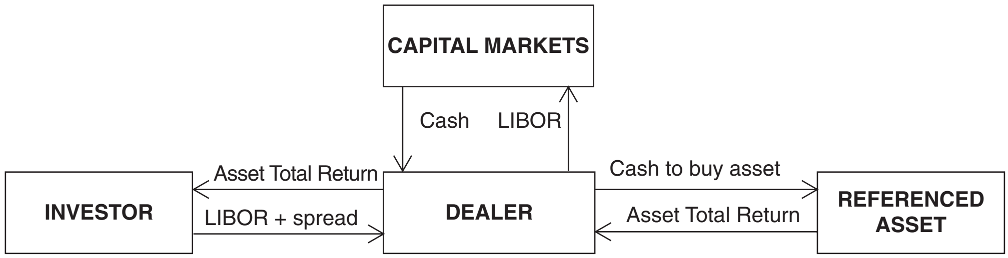


Figure 43.1 Total Return Swaps

depreciation of the reference asset. The floating rate is a reference interest rate (typically the London Interbank Offered Rate [LIBOR]) plus or minus a spread. The party that agrees to make the floating rate payments and receive the total return is referred to as the total return receiver or the *swap buyer*; the party that agrees to receive the floating rate payments and pay the total return is referred to as the total return payer or swap buyer. Total return swaps are viewed as unfunded credit derivatives, because there is no up-front payment required.

If the total return payer owns the underlying asset, it has transferred its economic exposure to the total return receiver. Effectively, then, the total return payer has a neutral position that typically will earn LIBOR plus a spread. However, the total return payer has only transferred the economic exposure to the total return receiver; it has not transferred the actual asset. The total return payer must continue to fund the underlying asset at its marginal cost of borrowing or at the opportunity cost of investing elsewhere the capital tied up by the reference assets.

The total return payer may not initially own the reference asset before the swap is transacted. Instead, after the swap is negotiated, the total return payer will purchase the reference asset to hedge its obligations to pay the total return to the total return receiver. In order to purchase the reference asset, the total return payer must borrow capital. This borrowing cost is factored into the floating rate that the total return receiver must pay to the swap seller. Figure 43.1 diagrams how a total return credit swap works.

In Figure 43.1 the dealer raises cash from the capital markets at a funding cost of straight LIBOR. The cash that flows into the dealer from the capital markets flows right out again to purchase the reference asset. The asset provides both interest income and capital gain or loss depending on its price fluctuation. This total return is passed through in its entirety to the investor according to the terms of the total return swap. The investor, in turn, pays the dealer LIBOR plus a spread to fulfill its obligations under the swap.

From the dealer's perspective, all of the cash flows in Figure 43.1 net out to the spread over LIBOR that the dealer receives from the investor. Therefore, the dealer's profit is the spread times the notional amount of the total return swap. Furthermore, the dealer is perfectly hedged. It has no risk position except for the counterparty risk of the investor. Effectively, the dealer receives a spread on a riskless position.

In fact, if the dealer already owns the reference asset on its balance sheet, the total return swap may be viewed as a form of credit protection that offers more risk reduction than a credit default swap. A credit default swap has only one purpose: To protect the investor against default risk. If the issuer of the reference asset defaults, the credit default swap provides a payment. However, if the underlying asset declines in value but no default occurs, the credit protection buyer receives no payment. In contrast, under a total return swap, the reference asset owned by the dealer is protected from declines in value. In effect, the investor acts as a "first loss" position for the dealer because any decline in value of the reference asset must be reimbursed by the investor.

The investor, however, receives the total return on a desired asset in a convenient format. There are several other benefits in using a total return swap as opposed to purchasing a reference asset itself. First, the total return receiver does not have to finance the purchase of the reference asset itself. Instead, the total return receiver pays a fee to the total return payer in return for receiving the total return on the reference asset. Second, the investor can take advantage of the dealer's "best execution" in acquiring the reference asset. Third, the total return receiver can achieve the same economic exposure to a diversified basket of assets in one swap transaction that would otherwise take several cash market transactions to achieve. In this way, a total return swap is a much more efficient means for transacting than via the cash market. Finally, an investor who wants to short a credit-risky asset such as a corporate bond will find it difficult to do so in the market. An investor can do so efficiently by using a total return swap. In this case the investor will use a total return swap in which it is a total return payer.

There is a drawback of a total return swap if an asset manager employs it to obtain credit protection. In a total return swap, the total return receiver is exposed to both credit risk and interest rate risk. For example, the credit spread can decline (resulting in a favorable price movement for the reference asset), but this gain can be offset by a rise in the level of interest rates.

Total Return Swap Compared to an Interest Rate Swap

It is worthwhile comparing market conventions for a total return swap to that of an interest rate swap. A plain vanilla or generic interest rate swap involves the exchange of a fixed-rate payment for a floating-rate payment. A *basis swap* is a special type of interest rate swap in which both parties exchange floating-rate payments based on a different reference interest rate. For example, one party's payments may be based on 3-month LIBOR, while the other parties payment is based on the 6-month Treasury rate. In a total return swap, both parties pay a floating rate.

The quotation convention for a generic interest rate swap and a total return swap differ. In a generic interest rate swap, the fixed-rate payer pays a spread to a Treasury security with the same tenor as the swap and the fixed-rate receiver pays the reference rate flat (that is, no spread or margin). The payment by the fixed-rate receiver (that is, floating rate payer) is referred to as the *funding leg*. For example, suppose an interest rate swap quote for a 5-year, 3-month LIBOR-based swap is 50. This means that the fixed-rate payer agrees to pay the 5-year Treasury rate that exists at the inception of the swap and the fixed-rate receiver agrees to pay 3-month LIBOR. In contrast, the quote convention for a total return swap is that the total return receiver receives the total return flat and pays the total return payer a interest rate based on a reference rate (typically LIBOR) plus or minus a spread. That is, the funding leg (that is, what the total return receiver pays includes a spread).

Illustration

Let's illustrate a total return swap where the reference asset is a corporate bond. Consider an asset manager who believes that the fortunes of XYZ Corporation will improve over the next year so that the company's credit spread relative to U.S. Treasury securities will decline. The company has issued a 10-year bond at par with a coupon rate of 9% and therefore the yield is 9%. Suppose at the time of issuance, the 10-year Treasury yield is 6.2%. This means that the credit spread is 280 bps and the asset manager believes it will decrease over the year to less than 280 bps.

The asset manager can express this view by entering into a total return swap that matures in one year as a total return receiver with the reference asset being the 10-year, 9% XYZ Corporation's bond issue. For simplicity, assume that the total return swap calls for an exchange of payments semiannually. Suppose the terms of the swap are that the total return receiver pays the 6-month Treasury rate plus 160 bps in order to receive the total return on the reference asset. The notional amount for the contract is $10 million.

Suppose that at the end of one year the following occurs:

- The 6-month Treasury rate is 4.8% initially.
- The 6-month Treasury rate for computing the second semiannual payment is 5.4%.
- At the end of one year the 9-year Treasury rate is 7.6%.
- At the end of one year the credit spread for the reference asset is 180 bps.

First, let's look at the payments that must be made by the asset manager. The first swap payment made by the asset manager is 3.2% (4.8% plus 160 bps divided by two) multiplied by the $10 million notional amount. The second swap payment made is 3.5% (5.4% plus 160 bps divided by two) multiplied by the $10 million notional amount. Thus,

First swap payment paid:	$10 million × 3.2% = $320,000
Second swap payment paid:	$10 million × 3.5% = $350,000
Total payments:	$670,000

The payments that will be received by the asset manager are the two coupon payments plus the change in the value of the reference asset. There will be two coupon payments. Since the coupon rate is 9% the amount received for the coupon payments is $900,000.

Finally, the change in the value of the reference asset must be determined. At the end of one year, the reference asset has a maturity of 9 years. Since the 9-year Treasury rate is assumed to be 7.6% and the credit spread is assumed to decline from 280 bps to 180 bps, the reference asset will sell to yield 9.4%. The price of a 9%, 9-year bond selling to yield 9.4% is 97.61. Since the par value is $10 million, the price is $9,761,000. The capital loss is therefore $239,000. The payment to the total return receiver is then:

Coupon payment	= $900,000
Capital loss	= $239,000
Swap payment	= $661,000

Netting the swap payment made and the swap payment received, the asset manager must make a payment of $9,000 ($661,000 − $670,000).

Notice that even though the asset manager's expectations were realized (that is, a decline in the credit spread), the asset manager had to make a net outlay. This illustration highlights one of the disadvantages of a total return swap noted earlier: The return to the investor is dependent on both credit risk (declining or increasing credit spreads) and market risk (declining or increasing market rates). Two types of market interest rate risk can affect the price of a fixed income asset. Credit-independent market risk is the risk that the general level of interest rates will change over the term of the swap. This type of risk has nothing to do with, the credit deterioration of the reference asset. Credit-dependent market interest rate risk is the risk that the discount rate applied to the value of an asset will change based on either perceived or actual default risk.

In the illustration, the reference asset was adversely affected by market interest rate risk, but positively rewarded for accepting credit dependent market interest rate risk. To remedy this problem, a total return receiver can customize the total return swap transaction. For example, the asset manager could negotiate to receive the coupon income on the reference asset plus any change in value due to changes in the credit spread. Now the asset manager has expressed a view exclusively on credit risk; credit independent market risk does not affect the swap value. In this case, in addition to the coupon income, the asset manager would receive the difference between the present value of the reference asset at a current spread of 280 bps and the present value of the reference asset at a credit spread of 180 bps.

APPLICATIONS OF A TOTAL RETURN SWAP

An asset manager typically uses a credit default swap to hedge a credit exposure. However, a total return swap is typically used to increase credit exposure. A total return swap transfers all of the economic exposure of a reference asset to the total return receiver. In exchange for accepting this exposure, the total return receiver pays a floating interest rate to the total return payer.

Total return swap applications fall into three categories:

1. Asset managers using a total return swap for leveraging purposes.
2. Asset managers using a total return swap as a more transactionally efficient means for implementing a portfolio management strategy.
3. Managers of bank portfolios using a total return swap as an efficient vehicle for transferring credit risk and as a means for reducing capital charges.

Below, we provide two applications of total return swaps and further when total return index swaps are discussed.

Creating a Synthetic Repo

There are a number of reasons why asset managers may wish to enter into total swap arrangements. As noted above, one of these is to reduce or remove credit risk. Using total return swaps as a credit derivative instrument, a party can remove exposure to an asset without having to sell it. In a vanilla total return swap the total return payer retains rights to the reference asset, although in some cases servicing and voting rights may be transferred. This assumes that the reference asset is on the payer's balance sheet.

The total return receiver gains an exposure to the reference asset without having to pay out the cash proceeds that would be required to purchase it. As the maturity of the swap rarely matches that of the reference asset, in a positive yield curve environment the swap receiver may gain from the positive funding or carry that derives from being able to roll over short-term funding of a longer-term asset. The total return payer on the other hand benefits from protection against interest rate and credit risk for a specified period of time, without having to liquidate the asset itself. At the maturity of the swap the total return payer may reinvest the asset if it continues to own it, or it may sell the asset in the open market. In this respect a total return swap is in essence a *synthetic repo*.

A total return swap agreement entered into as a credit derivative is a means by which banks can take on unfunded off-balance sheet credit exposure. Higher-rated banks that have access to London Interbank Bid Rate (LIBID) funding can benefit by funding on-balance-sheet assets that are credit protected through a credit derivative such as a total return swap, assuming the net spread of asset income over credit protection premium is positive.

A total return swap conducted as a synthetic repo is usually undertaken to effect the temporary removal of assets from the balance sheet. This may be desired for a number of reasons, for example if the institution is due to be analyzed by credit rating agencies, or if the annual external audit is due shortly. Another reason a bank may wish to temporarily remove lower-credit-quality assets from its balance sheet is if it is in danger of breaching capital limits in between the quarterly return periods. In this case, as the return period approaches, lower quality assets may be removed from the balance sheet by means of a total return swap, which is set to mature after the return period has passed.

However, this is a semantic point associated with the motivation of the total return payer. If effected for regulatory capital reasons a total return swap is akin to a synthetic repo; if effected for credit speculation reasons it becomes a credit derivative.

Use in the Bank Loan Market

Let's use an actual case to see how a total return swap can be employed in the bank loan market (This illustration is an expanded discussion of a bank loan swap presented by Keith Barnish, Steve Miller, and Michael Rushmore [1997].) Consider the details of a 3-year swap on a term bank loan. A large AA insurance company purchased a 3-year total return swap on a $10 million piece of Riverwood International's Term Loan B. Term Loan B was actually a tranche of $250 million, but the insurance company only wanted credit exposure to a portion of the term loan.

This demonstrates one of the advantages of a credit derivative in general: customization. An investor may like the credit risk of a particular bank loan tranche, but may not have sufficient appetite for the whole loan. A total return credit swap allows the investor to choose a big or small piece of credit exposure depending on the investor's appetite for the credit risk. Furthermore, the term loan had a maturity of 10 years, while the holding period horizon of the insurance company was three years. Therefore, the total return swap can accommodate the insurance company's investment horizon while the term loan does not.

The seller of the swap (that is, the total return payer) was a large institutional bank. In order for the insurance company to purchase the total return swap, the bank effectively loaned the insurance company the $10 million notional amount of the swap. The bank in fact did not disburse $10 million to the insurance company, but instead charged the insurance company interest on $10 million dollars as if the bank had loaned the full amount. In this transaction, the bank charged the insurance company LIBOR + 75 bps. Since the insurance company's normal borrowing rate was 12.5 bps over LIBOR, the bank effectively charged the insurance company a swap processing fee of 62.5 bps, equivalent to $62,500 on an annual basis. In addition to the annual fee, the insurance company was required to put up $1 million of collateral as security for the effective loan. This $1 million was invested in U.S. Treasury securities.

In return for paying this fee, the insurance company received the total return on the Riverwood International term loan. The total return included the floating interest on the term loan of LIBOR + 300 bps plus any gain or loss in market value of the loan. In sum, the bank passed through the swap to the insurance company all of the interest payments and price risk as if the insurance company had the term loan on the asset side of its balance sheet.

The benefit to the insurance company was the net interest income earned on the swap. The insurance company agreed to pay LIBOR + 75 bps to the bank in return for LIBOR + 300 bps received from the Riverwood International term loan. The annual net interest income from the swap paid to the insurance company was:

$$\$10{,}000{,}000 \times [(\text{LIBOR} + 300 \text{ bps}) - (\text{LIBOR} + 75 \text{ bps})]$$
$$= \$10{,}000{,}000 \times 2.25\% = \$225{,}000$$

Provided that Riverwood International did not default on any portion of the term loan, the insurance company also received the interest income on the Treasury securities.

Why would the bank want to enter into this transaction? Perhaps, the bank bit off more than it wanted to chew when it purchased the full tranche from Riverwood International. The total return swap with the insurance company allowed the bank to reduce its credit exposure and collect a fee. In effect, the bank got paid to reduce its credit risk.

And what about the insurance company? Was this a good deal for it? The answer is yes if we consider the alternative to the total return swap. Assume, that instead of the total return swap, the insurance company could have purchased a $10 million portion of the Riverwood International term loan at its normal financing cost of LIBOR + 12.5 bps, held the term loan on its balance sheet for three years, and then sold it at the end of its holding period. The question we need to answer is which alternative provided a greater return: the total return swap or the outright purchase of the term loan?

Tabel 43.1 details the holding period returns to the two alternatives. In the first case, the insurance company borrows $1 million at its normal financing rate to purchase the Treasury security collateral and receives three annual net payments of $225,000 from the bank as well as interest income on the Treasury securities. Additionally, in year 3, the insurance company receives back the $1 million of collateral. These cash flows are discounted at the insurance company's cost of capital of 3-year LIBOR + 12.5 bps.

In the second case, the insurance company receives the full payment of LIBOR + 300 bps on the term loan, but must finance the full $10 million for three years. It receives an annual cash flow of $950,000, and sells its investment at the end of three years for $10 million.

To keep the analysis simple, assume that the insurance company bought a 3-year U.S. Treasury note as collateral with a maturity equal to the tenor of the swap and with an annual coupon of 6.00%, that 1-year LIBOR remains constant at 5.78125%, and that there is no change in value of the Riverwood International term loan. The discount rate for present value purposes is 5.90625% (LIBOR + 12.5 bps).

Under the swap, the insurance company will receive each year a cash flow of $225,000 from the bank and $60,000 from the Treasury note. In addition, in year 3, the

Table 43.1 Investment Returns for a Total Return Bank Loan Credit Swap

Assumptions		
Asset	$10,000,000 bank term loan	
Maturity	Three years	
1-year LIBOR	5.78125% (constant)	
3-year Treasury	6.00%	
Discount rate	5.90625%	
Term loan value remains constant		
	Investment Alternatives	
	Credit Swap	**Purchase Term Loan**
Initial investment	($1,000,000)	($10,000,000)
Annual cash flows (loan value remains constant)		
Year 1	$285,000	$950,000
Year 2	285,000	950,000
Year 3	1,285,000	10,950,000
Present value of annual cash flows	$1,604,983	$10,961,833
Net present value	$604,983	$961,833
IRR	29%	9%
Initial investment	($1,000,000)	($10,000,000)
Annual cash flows (loan value declines by $1,000,000)		
Year 1	$285,000	$950,000
Year 2	285,000	950,000
Year 3	285,000	9,950,000
Present value of annual cash flows	$763,132	$10,120,431
Net present value	($236,868)	$120,431
IRR	−7%	6%

insurance company will receive back its $1 million collateral contribution. Under the outright purchase of the term loan, the insurance company will receive each year a cash flow of $950,000. At the end of three years the insurance company sells the term loan in the market for its original investment of $10 million. Table 43.1 details these assumptions as well as a comparison of the cash flows for each alternative.

As can be seen from Table 43.1, the outright purchase of the term loan results in a higher net present value than the total return swap. The net present value for the term loan is $961,833 and for the total return swap it is $604,983, a difference of $356,850. However, the total return swap requires a much smaller capital requirement than the outright purchase of the term loan. Even though the total return swap results in lower total cash flows, it provides an internal rate of return (IRR) that is three times greater than that of the term loan purchase.

This example demonstrates the use of leverage in a total return swap. The smaller capital commitment of the total return swap allows the insurance company to earn a higher rate of return on its investment than the outright purchase of the term loan. In fact, the leverage implicit in this total return swap is 10:1. Economically, the total return swap is more efficient because it allows the insurance company to access the returns of the bank loan market with a smaller required investment.

However, what if the value of the term loan had declined at the end of three years? Assume that over the 3-year holding period, the value of the Riverwood International bank loan declined in value to $9 million. With the total return swap arrangement, the $1 million loss in value would wipe out the posted collateral value. At the end of year 3, the insurance company would receive only the cash flow from the interest income, $225,000 from the swap, and $60,000 in interest from the posted collateral.

Under the purchase scenario, the insurance company would receive back $9 million of its committed capital. Additionally, in each year the insurance company would receive the $950,000 interest income from the term loan. Table 43.1 also compares the two investment choices under the assumption of a $1 million decline in loan value.

Under the total return swap, the net present value of the investment is now a negative $236,868. Conversely, a decline in loan value of $1 million still leaves the purchase scenario with a positive net present value of $120,431. Comparing the IRR on the two investments, we now see that the total return swap yields a negative IRR of −7%, while the purchase of the term loan yields a positive IRR of 6%—slightly more than the insurance company's cost of borrowed funds. Table 43.1 demonstrates that the embedded leverage in the total return swap can be a double-edged sword. It can lead to large returns on capital, but can also result in rapid losses.

TOTAL RETURN INDEX SWAPS

Thus far our focus has been on a single reference asset. Total return index swaps are swaps where the reference asset is the return on a market index. The market index can be an equity index or a bond index. Our focus will be on bond indices.

Broad-based bond market indices such as the Lehman, Salomon Smith Barney, and Merrill Lynch indexes have subindexes that represent major sectors of the bond market. For example there is the Treasury and agency sector, the credit sector (that is, investment-trade corporate bonds, at one time referred to as the corporate sector), the mortgage sector (consisting of agency residential mortgage-backed securities), the commercial mortgage-backed securities (CMBS) sector, and the asset-backed securities (ABS) sector. The non-Treasury sectors offer a spread to Treasuries and are hence referred to as "spread sectors." The spread in the mortgage sector is primarily compensation for the prepayment risk associated with investing in this sector. Spread to compensate for credit risk is offered in the credit spread sector, of course, and the CMBS and ABS sectors. There are also indexes available for other credit spread sectors of the bond market: high-yield corporate bond sector and emerging market bond sector. Thus, a total return index swap in which the underlying index is a credit spread sector allows an asset manager to gain or reduce exposure to that sector.

Below, we discuss the flexibility offered asset managers and hedge fund managers by using total return swaps in which the index is a credit spread sector of the bond market.

Indexing a Credit Spread Sector by an Active Asset Manager

Bond portfolio strategies range from indexing to aggressive active strategies. The degree of active management can be quantified in terms of how much an asset manager deviates from the primary risk factors of the target index. A bond indexing strategy for a sector involves creating a portfolio so as to replicate the issues comprising the target sector's index. This means that the indexed portfolio is a mirror image of the target sector index or, put another way, that the *ex ante* tracking error is close to zero.

Why would an asset manager pursuing an active portfolio management strategy want to engage in an indexing strategy for a credit sector of the target index? Suppose that the asset manager's target index is the Lehman Brothers U.S. Aggregate Bond Index. Suppose further that the asset manager skills are such that she believes she can add value in the mortgage, CMBS, and ABS sectors but has no comparative advantage in the credit (corporate sector). The asset manager in this case can underweight the credit sector. However, the risk is that the credit sector will perform better than the other sectors in the target index and, as a result, the asset manager will underperform the target index. An alternative is to be neutral with respect to the credit sector and make active bets within the sectors of the target index that the asset manager believes value can be added. This approach requires that the asset manager follow an indexing strategy for the credit sector of the target index. However, in pursuing this strategy of creating a portfolio to replicate the credit sector, the asset manager will encounter several logistical problems.

First, the prices for each issue in the credit sector used by the organization that publishes the sector index may not be execution prices available to the asset manager. In fact, they may be materially different from the prices offered by some dealers. In addition, the prices used by organizations reporting the value of sector indexes are based on bid prices. Dealer ask prices, however, are the ones that the manager would have to transact at when constructing or rebalancing the indexed portfolio. Thus there will be a bias between the performance of the sector index and a portfolio that attempts to replicate the sector index that is equal to the bid-ask spread.

Furthermore, there are logistical problems unique to certain sectors in the bond market. For the credit sector, which consists of investment-grade corporate bonds, there are typically more than 4,000 issues. Because of the illiquidity for many of the issues, not only may the prices used by the organization that publishes the index be unreliable, but also many of the issues may not even be available.

Third, as bonds mature, their shrinking duration will force them out of this index. This will create natural turnover and higher transaction costs. Last, bonds pay consistent coupons that must be reinvested in the index.

In the absence of a total return swap, there are two methodologies that have been used to construct a portfolio to replicate the index representing the credit sector: stratified sampling methodology and the variance minimization methodology. With the stratified sampling approach (or also called the cellular approach) to indexing, the sector index is divided into cells representing the primary risk factors. The objective is then to select from all of the issues in the index one or more issues in each cell that can be used to represent that entire cell. The total dollar amount purchased of the issues from each cell will be based on the percentage of the index's total market value that the cell represents. For example, if X% of the market value of all the issues in the credit sector index is made up of single-A-rated corporate bonds, then X% of the market value of the replicating portfolio should be composed of single-A-rated corporate bond issues. The number of cells that the asset manager uses will depend on the dollar amount of the portfolio to be indexed. In indexing a portfolio of less than $50 million, for example, using a large number of cells would require purchasing odd lots of issues. This increases the cost of buying the issues to represent a cell, and thus would increase the *ex ante* tracking error. Reducing the number of cells to overcome this problem increases *ex ante* tracking error because the major risk factors of the indexed portfolio may differ materially from those of the index. For corporate bonds, for example, there is the concern of downgrade risk of individual corporate issues that would adversely affect tracking error. Figure 43.2 shows the findings of a Lehman Brothers study that demonstrates how many issues must be purchased to minimize tracking error due to downgrade risk (see Dynkin, Hyman, and Konstantinovsky, 2002). As can be seen, if only a few issues are selected tracking error is high.

The variance minimization methodology is a more complicated approach than stratified sampling. This approach requires using historical data to estimate the variance of the tracking error for each issue in the index. The objective then is to minimize the variance of the tracking error in constructing the replicating portfolio.

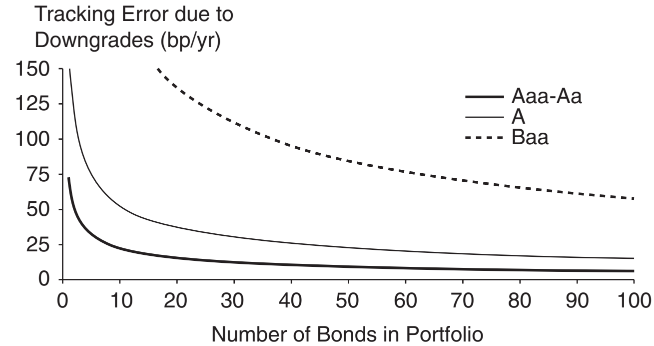


Figure 43.2 Risk due to Downgrades as a Function of Portfolio Size—by Credit Quality
Source: Exhibit 14 in Dynkin, Hyman, and Konstantinovsky (2002), p. 100. This copyrighted material is reprinted with permission from Institutional Investor, Inc., *Journal of Portfolio Management*, 225 Park Avenue South, New York, NY 10003.

The more efficient solution may be simply to use an total return index swap where the credit sector to be indexed is the underlying index for the swap.

Active Strategies

Active bond portfolio strategies involve constructing a portfolio that deviates from the target index. There are various strategies that can be employed. For example, one strategy is to construct a portfolio that is intentionally different from the duration of the target index based on the view of the asset manager regarding future interest rates. Another is to overweight a sector of the index based on the asset manager's view of the relative performance of the sectors comprising the index. For example, if the credit sector is expected to outperform the other sectors, an asset manager may wish to overweight that sector. The asset manager can monetize this view by entering into a total return swap as the total return receiver. Again, as noted earlier, this is an efficient way to replicate the performance of the index.

Hedge funds manager can use total return swaps to create leverage in the same way described earlier when we showed how a synthetic repo can be created for a credit-risky bond. Moreover, suppose instead that a hedge fund manager believes that the credit sector will have a negative return. The manager can monetize this view by selling a total return swap. The advantage of the total return swap is that the credit sector can be shorted, a task that is extremely difficult and costly to do for individual bond issues in the credit sector.

Risk Control

Total return swaps can be sued as effective risk control instruments. Interest rate swaps can be used to control the duration of the portfolio. Total return swaps can be used to control the spread duration of a portfolio and, more specifically, the credit spread duration of a portfolio, that is the sensitivity of a portfolio to changes in credit spreads. Hedging a position with respect to credit spread risk means creating a cash and total return swap position whereby the credit spread duration is zero. An asset manager would want to hedge a portfolio that has exposure to credit spread risk if the credit spread duration of the portfolio differs from that of the target index. Total return swaps can be used to bring the portfolio's credit spread risk duration in line with the credit spread risk of the target index.

SUMMARY

In a total return swap, the total return receiver (or swap buyer) agrees to make to floating-rate payments on designated dates to the total return payer (or swap seller) in exchange for the total return realized on a reference asset. In the fixed income market, the reference asset can be a credit-risky bond, a reference portfolio, or an index representing a sector of the bond market. Total return swaps can be used by fixed income managers for leveraging purposes or to more efficiently implement a portfolio strategy. In addition, total return swaps are an efficient vehicle for allowing bank portfolio managers to transfer credit risk and thereby reduce capital charges. Total return index swaps can be used for a wide range of bond portfolio strategies, ranging from indexing to aggressive active strategies.

REFERENCES

Anson, M. J. P., Fabozzi, F. J., Choudhry, M., and Chen, R-R (2004). *Credit Derivatives: Instruments, Pricing, and Applications*. Hoboken, NJ: John Wiley & Sons.

Barnish, K., Miller, S., and Rushmore, M. (1997). The new leveraged loan syndication market. *Journal of Applied Corporate Finance* (Spring): 79–88.

Dynkin, L., Hyman, J., and Konstantinovsky, V. (2002). Sufficient diversification in credit portfolios. *Journal of Portfolio Management* (Fall): 89–114.

Goodman, L. S., and Fabozzi, F. J. (2005). CMBS total return swaps. *Journal of Portfolio Management*, Special Issue on Real Estate: 162–167.

Kolb, R. and Overdahl, J. (2007). *Futures, Options and Swaps.* Oxford: Blackwell.

CHAPTER 44

Bond Market Transparency

DANIEL E. GALLEGOS
Principal, Barclays Global Investors

CHRIS BARR, CFA
Principal, Barclays Global Investors

Abstract: Unlike equities, most U.S. fixed income securities are traded over the counter, not on formal exchanges like the New York Stock Exchange. This convention, combined with the large number of debt issues outstanding, is largely responsible for the lack of price transparency that exists in the fixed income markets. Poor transparency contributes to investor differences in bond valuations as well as other inefficiencies that may lead to economic losses for market participants and ultimately inhibit business development. To address the inefficient nature of fixed income pricing, the bond market has evolved, aided by advances in technology, innovations in pricing methodologies, and modification of regulatory requirements.

Keywords: price transparency, bid/offer, price indications, multidealer-to-client electronic platforms, Trade Reporting and Compliance Engine (TRACE), Municipal Securities Rulemaking Board (MSRB), Real-time Transaction Reporting System (RTRS), matrix pricing, best execution, fiduciary, option-adjusted spread (OAS)

In this chapter, we discuss the benefits of price transparency for investors, the causes and effects of poor price transparency in the major fixed income sectors, arguments that support and oppose further transparency, and a few key elements of these markets that continue to undergo significant changes.

BENEFITS OF AN OPEN SYSTEM

Price transparency can be defined as the extent to which information is available to easily establish the correct market value of a security. It is important for a variety of reasons. For individual investors, transparency gives the market participant reassurance that they were fairly charged for a particular security. For institutional investors, it can help fulfill fiduciary duty by reducing costs and ensuring accuracy in performance reporting. For the market in general, price transparency helps increase activity, which in turn fosters participation, liquidity, and new product development.

Transparency has allowed for great precision in attributing skill in markets such as equities, where the value of stock selection or of a particular portfolio strategy is borne out readily by virtue of indisputable price identification. The bond market, however, has historically been plagued by unclear pricing, making it more difficult for market participants to conduct trades efficiently, let alone assess their worth. In fact, it has been a topic of concern since the inception of fixed income trading in the early 1900s, and has intensified over the last decade or two.

Unfortunately for investors, the nature of the bond market lends itself to price ambiguity. Fixed income is a principal market that operates on Wall Street's terms: broker/dealers buy at the bid and sell at the offer. Investors abide by these terms as they buy at the offer side and sell at the bid side. Although corporate and municipal bonds traded on exchanges in the early 1900s, the market has long since been characterized by dealer *bid/offer* facilitation and subsequent price opaqueness. For years, bond managers have complained about the inefficient nature of pricing bonds: manually calling multiple brokers for *price indications* on whole portfolios (with markets open and prices moving) while their equity counterparts effortlessly garner unambiguous valuations from the financial website of their choosing. Or worse, going to sell a bond position only to find the mark is "stale" and no longer worth what the custodial report quoted.

The Texan artist Jack White once said, "Art is only worth what people will pay for it." If a quote from history is worth a page of logic, then the bond market, like life, imitates art—the price of a bond should reflect what someone is willing to pay for it. The problem is that many bonds do not trade frequently enough to accurately know what people will pay for them. For example, a look at the National Association of Securities Dealers (NASD) trade histories for the constituents of the Lehman Brothers U.S. Corporate Index shows that only a little over half of the line items had traded during the previous day. And these are not small issues; the index rules require a minimum issue size of $250 million.

So if a trade price is not available, why not just do the bond math to get a market valuation? After all, a bond is nothing more than a future stream of cash flows—discount each one of those flows by some prevailing rate, add them up, and voila! the result is a bond price that everyone can agree on.

In theory, this should work if the assumptions regarding the prevailing discount rate are shared. In some cases they are. U.S. government securities—Treasuries, for example—trade very close to the theoretical value calculated from readily available interest rate information that, for the most part, everyone can agree on. In fact, Treasuries are so transparent that other bonds, like corporates, are priced according to their yield spreads.

This is the point at which different levels of transparency can be observed. Although the Treasury yield is transparent, this corporate "credit spread" represents the additional yield the investor requires above and beyond a "riskless" Treasury rate to compensate for the risk that the bond issuer may default on the obligation. The spread may vary for different sizes, structures, and maturities, and therefore it may be subject to interpretation. In addition, the bond may include embedded options, such as call or prepayment provisions, covenants, restrictions, and esoteric cash-flow structures, all of which change the complexion of the bond and require assumptions regarding creditworthiness, interest rate volatility, prepayment behavior, and liquidity. Each of these variables entails assumptions that give rise to different levels of price transparency.

DISPARITY OF MARKET PRICING

But many bonds do seem to have transparent pricing. U.S. Treasuries trade throughout the day at levels quoted electronically: what you see on the broker screens is what you get in the markets. Or is it?

Case in point: Institutional investors are typically offered Treasuries significantly cheaper than those offered to individual investors. The difference in offering prices is easily observed on two quoting sources: a retail brokerage account and an institutional quote system. And even that price may vary owing to the size and nature of the trade (odd-lot or block order, agent or principal). In fact, different levels of price transparency exist in each of the major fixed income sectors: government, mortgage, corporate, and municipal.

Much of the lack of transparency within each of the fixed income sectors is a function of liquidity—it is hard to accurately assess the market value of a bond if the market is not active. Figure 44.1 provides average daily trade volume for the different sectors of the Lehman Brothers U.S. Aggregate Bond Index, arguably the most popular benchmark for U.S. fixed income strategies. As the graph

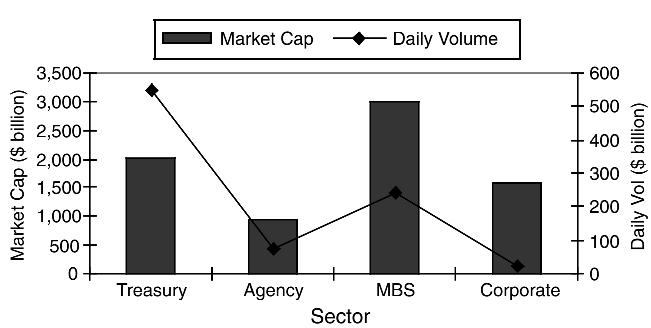


Figure 44.1 Lehman Brothers U.S. Aggregate Bond Index as of 9/30/2006: Market Capitalization versus Daily Volume
Source: The Bond Market Association, "Average Daily Trading Volume in the US Bond Markets," September 30, 2006.

suggests, there are dramatic differences in liquidity between sectors. Many names do not typically trade on a daily basis, and some do not even trade weekly. By contrast, virtually all of the stocks in the Russell 3000 Index, a broad-capitalization equity benchmark, trade at least daily, with the average stock trading more than a million shares per day.

To gain insight into this liquidity spectrum, we can compare the major fixed income sectors. But it's important to keep in mind that, within each of these sectors, there exists at least one subsector whose price transparency varies widely from the others.

U.S. Treasury and Agency Securities

U.S. Treasury and agency securities are perceived to have the greatest liquidity and price transparency in the fixed-income markets. Despite the complexities of many government bond structures, (roughly half of the agency bonds in the Lehman Aggregate are callable), this sector has always enjoyed the highest degree of price transparency, largely due to the guaranteed status (implied or actual), and it trades over $600 billion per day. In fact, the U.S. government market was the first to raise the curtain on pricing when, in 1991, a 24-hour global electronic reporting system called GovPX was developed for U.S. Treasury and other government securities. But investors were really rewarded in 2000 when a syndicate of 18 broker/dealers joined to form an electronic dealer-to-customer auction platform called TradeWeb, where executable bids and offers are clearly and unambiguously listed for all subscribers to participate. According to TradeWeb, virtually all of the Treasury market and much of the Agency market can currently trade electronically.

Corporate Bonds (Investment and Non–Investment Grade)

Unlike Treasuries, whose average issue size in the Lehman Aggregate is around $15 billion, the average size of a corporate bond in the same index is about $600 million, and it trades far less frequently. Despite this relative gap in liquidity, significant changes to the corporate bond market over the past few years have fostered increased price transparency. Like governments, corporates now have venues to trade electronically. *Multidealer-to-client electronic platforms* (such as MarketAxess) allow "best bid/offer" facilitation for the credit markets, contributing a valid source of accurate price dissemination. Electronic coverage of the corporate market, however, is not as comprehensive as the government market. Only 12% of the investment-grade corporate market trades are done by electronic execution, totalling roughly $3.6 trillion a year, according to MarketAxess.

The credit derivative market also plays a pivotal role. Credit default swaps (CDSs), designed to transfer the credit exposure of fixed income products between parties, are the most widely used credit derivatives, and represent a revolutionary way of gaining transparency in the credit markets. Since many corporate bond investors are only interested in accessing a debt issuer's credit risk (the compensation received in the form of credit spread), the $5+ trillion CDS market is fertile ground for eliminating price ambiguity from issuers with multiple maturities, coupons, and structures. After all, a $600 million 6% coupon Ford bond should be no more or less creditworthy than a $3 billion 7% Ford bond. In fact, many believe the CDS market leads the cash markets in corporate price relevance.

But the biggest improvement in corporate price transparency came as a result of the evolution of the NASD's *Trade Reporting and Compliance Engine* (*TRACE*) program. TRACE facilitates the mandatory reporting of over-the-counter secondary-market transactions for eligible fixed-income securities by requiring brokers to report trades on eligible bonds to the TRACE system within 15 minutes. With the hopes of providing better price transparency to the market, all broker/dealers who are NASD member firms have an obligation to report transactions in corporate bonds to TRACE under a Securities and Exchange Commission (SEC)-approved set of rules. TRACE became operational on July 1, 2002, and now encompasses 95% of all corporate issues according to the NASD.

The TRACE system has made great strides, but price inconsistencies still exist. Substantial disparities can be observed on TRACE for even the largest and most liquid issues, owing primarily to trade size (unlike the equity markets, bond liquidity typically increases with trade size). And as mentioned above, the method for establishing corporate bond prices (evaluating the creditworthiness of the issuer and then assigning different probabilities for receiving the cash flows which the bond is structured to pay) can be an analytically intensive process which few investors have the resources to accomplish, and with ample disagreement among those who can. Just like Treasuries, institutional investors receive more favorable pricing than individual investors, and large differences in corporate prices can be realized depending on the counterparty. Furthermore, a study by Edwards, Harris, and Piwowar (2005) suggests that liquidity (hence transparency) also decreases with quality, especially as ratings drop below BBB into high yield. A review of the NASD trade histories for the constituents of the Merrill Lynch High Yield Index reveals that only about one-third trade every day.

While the TRACE system represents a significant step toward increased price transparency in the fixed income markets, it does not apply to all sectors, especially the largest sector: mortgage securities.

Mortgage-Related Securities

Representing more than 35% of the Lehman Brothers U.S. Aggregate Bond Index's nearly $9 trillion market capitalization, mortgage-backed securities (MBSs) make up the largest sector of the fixed income market. The mortgage markets are made up of many distinct security types, each having their own level of transparency. Generalizations regarding the ease of pricing cannot be made, regardless of the security descriptive characteristics. The same sophistication that is needed in determining expected cash flows in the corporate markets applies to the MBS sector.

However, in addition to creditworthiness, the probability of principal prepayment must be understood and projected.

As a whole, the mortgage markets experience the greatest amount of deviation in terms of price transparency. Agency TBAs (generic "to be announced" pass-through securities issued by Ginnie Mae, Fannie Mae, and Freddie Mac) enjoy the same liquidity as Treasuries, with trillions of dollars traded electronically (also on TradeWeb). Many derivatives structures, like collateralized mortgage obligations (CMOs), are created from mortgage pools. Since the structures derive their value from the underlying collateral, the sensitivity of pricing to prepayment assumptions can be magnified. And finally, when the additional level of credit complexity must be accounted for, as found in whole loan or private-label subordinate notes, it is difficult for even the most sophisticated investor to price these securities.

Municipals

The municipal market has long been the poster child for bad behavior in price transparency. In fact, a 2004 study by the SEC showed individual investors typically faced bid/offer spreads of nearly 2% for average municipal bond trades—almost 100 times larger than U.S. Treasury spreads! This is due primarily to the facts that municipalities have myriad liquidity and creditworthiness factors for their more than 1.5 million issues, their markets are highly fragmented and regionalized, and they are characterized by relatively low trading volumes. The *Municipal Securities Rulemaking Board* (*MSRB*), however, has been instrumental in facilitating increased price transparency by requiring publicly disseminated trade information on par with TRACE. Since January 2005, the *Real-time Transaction Reporting System* (*RTRS*) receives municipal trade reports within 15 minutes of execution.

Other

There are a number of other structures, including those that utilize leverage, exhibiting various levels of price transparency. Structured credit products such as collateralized debt obligations (CDOs)—collateralized bond obligations (CBOs) and collateralized loan obligations (CLOs)—have become increasingly popular, and derive their value from underlying collateral and cash-flow priority. The complexity and limited liquidity of these structures restrict their transparency to all but a select group of sophisticated analysts, but this is changing due to growing interest from participating investors.

BUYER BEWARE: ALL PRICES ARE NOT CREATED EQUAL

There is no doubt that actual transaction data could provide the most certainty for deriving the fair value of a security. However, because some fixed income securities trade so infrequently, and transaction data may be too stale to be relevant, many third-party vendors exist to supply investors with bond prices based on a variety of sources and methodologies. Each of these may establish prices using different methods, the most common of which include market indications, *matrix pricing*, and *option-adjusted spread* (*OAS*) calculations.

Market Indications

Talking directly to brokers for dealer quotes is a reliable way to establish price transparency by removing ambiguity. This method is performed simply by calling brokers and trade participants who specialize in specific sectors of the market and soliciting indicative bids and/or offers for a particular lot size. This could very well be the most-time-consuming method used to derive fair value. But in many ways, it offers the best sense of price transparency for those securities that trade infrequently, as the broker will ostensibly be providing an indication of interest.

A drawback to this method is that these price levels are not necessarily firm indications as to where the bonds might trade, but rather a best guess of the securities' market value. The indications may also vary according to who is making the inquiry. A large transactional investor, for example, will likely receive more accurate indications than a small, infrequent investor. And of course it may be prohibitively time consuming, rendering it impractical for large portfolios.

Matrix Pricing

A matrix-pricing system is another method in which indicative prices can be established. Matrix pricing makes it possible to price a large number of bonds within a short period of time. It also has the added benefit of providing indicative price capability for bonds that are thinly traded. This is accomplished by using a liquid subset of bonds from which to extract indicative spread data.

A pricing matrix is constructed using recent execution data; from this data, credit spreads to specific benchmarks are created and updated as new data become available. For this to be effective, bonds are segregated by characteristics such as sector, industry, term, rating, size, and so on. Once spread levels have been stratified, like bonds can be priced relative to a rate benchmark, such as Treasuries, using the matrix. Though this method is not as accurate as real-time execution levels, price changes attributable to spreads tend to be lower than changes attributable to rates, making it an appropriate tool for creating general price indications.

What matrix pricing neglects to consider are the security-specific risks associated with individual issuers, such that the spreads used to build the matrix may not reflect the actual securities priced from them. Also, matrices may have stale values that do not accurately reflect the current market environment for various structures. And finally, the more esoteric structures, such as derivatives or levered securities, may be wildly mispriced on these systems owing to even the slightest structural differences that cannot be captured with the limited amount of granularity within a particular matrix.

Option-Adjusted Spread Calculation

Whereas a matrix can offer indications of spreads attributable to sector, industry, rating, and maturity, option-adjusted spread (OAS) calculators are often used to price bonds with embedded options. Option value is a function of many variables, some fairly intuitive, others more complex. The value of a simple European call option, for example, is predominantly a function of strike, term, and volatility. These factors are easily accessible and are routinely incorporated into binomial lattice models. Since the assumptions regarding the time and level of exercise tend to be binary (that is, corporations are assumed to call an obligation any time it makes economic sense to do so), there tends to be, for the most part, only modest variance in plain-vanilla option valuation. Prepayment options, on the other hand, like those found in MBSs, must include not only rate factors but also behavioral assumptions regarding prepayments, which may be far less uniform and frequently incorporate complex stochastic interest rate–path-generating models that represent an enormous barrier to transparency for the average investor. Even though many securities require the input of an OAS calculator, this is a technical tool, and one that is demanding of specialized skill.

THE DEBATE ON PRICE TRANSPARENCY

Although it may seem clear at first glance that transparency is a good thing for all securities markets, there are two sides of the debate for bond market participants.

Arguments in Favor of Price Transparency

The primary reason for increasing the level of price transparency in the fixed income markets is to protect the investor from buying or selling a security at a price that is materially different from its true market value. It is not uncommon to find a news story about an unscrupulous broker taking advantage of an uninformed retail investor.

And although institutional investors are better equipped than individual investors to address the market inefficiencies that riddle the bond market, poor price transparency can negatively impact any portfolio's performance in the form of transaction costs. This is of particular importance to institutional money managers bound by the fiduciary responsibility of *best execution* (more on the concept of best execution below).

Just how negative is the impact of poor transparency? In 2005, armed with TRACE data, the SEC staff in the Office of Economic Analysis studied transaction costs before and after transaction reporting and concluded that for the period observed, costs were 5 basis points lower overall with price transparency. Further, $1 billion could have been realized by investors if all bonds were transparent for the whole year, suggesting increased transparency had cut some costs almost in half! While spreads are indeed a function of more than just price transparency, there appears to be evidence to support the hypothesis that transparency reduces costs.

The benefits of price transparency, however, are greater than just reduced transaction costs. Timely dissemination of accurate price information can help all investors value individual securities, evaluate investment strategies, and objectively monitor performance. Also, higher levels of price transparency can only help foster investor confidence, encouraging higher bond market participation from the small U.S.-based individual investor to the large international institutional investor. Lifting the fog on bond prices would likely spur an increase in market efficiency and new product innovation while deterring improper trade practices.

Arguments Opposed to Price Transparency

Brokerage trading desks make money primarily by providing liquidity, either consistently (through flow trading) or opportunistically (with proprietary trading). The difference is subtle, but it offers some insight as to the motivations of some market participants who oppose better transparency.

If the desk's goal is solely to maximize profit opportunistically, there is a clear incentive for a trader to bid low and offer high, without being so far off the perceived market as to scare away investors. Therefore the less transparency there is for the "perceived market," the more the trader is advantaged by the lack of information. The business model for this type of desk is based on the premise that inefficiencies can be effectively exploited by staying ahead of the competition, continually buying low and selling high.

An opponent of better transparency might argue, "Why is it that in other industries in which those who take risks and are able to identify unrealized value are entitled to commensurate rewards?" A broker's logic supporting this argument is that publicized trade prints (that is, the TRACE program) advertise exactly where a bond trades to the next potential buyer, leaving inadequate room for even a minimal markup, and leaving the broker holding the risk without receiving upside compensation, further adding insult to injury. Some dealers argue that if they can no longer mark up to where they feel they are being fairly compensated, liquidity will subsequently dry up. The NASD is proposing "contemporaneous cost" rules to regulate dealer markups (the current regulatory guideline generally limits markups to 5% with justification).

While the arguments for both sides of the debate continue, the trend toward increasing price transparency has not halted.

A RAPIDLY EVOLVING MARKET

The march toward increasing price transparency in the fixed income markets has brought about a sea change in the way bonds are viewed and traded. Among those elements that are rapidly evolving, it is reporting, best execution practices, and various trading methods that are undergoing significant changes.

Execution Reporting

As investors are becoming more aware of the consequences of inadequate price transparency, analysis of the execution data has become much more rigorous and informative, and various methods of evaluating the execution data are being demanded from investors.

To accommodate this demand, many electronic trading platforms are developing reports to evaluate the effectiveness of bond traders and the price levels being offered. Some of the methods of evaluation include comparing executions versus composite levels, percentage use of particular brokers, and evaluations of whether the best price was achieved for each execution. This is done in the hopes of both increasing transparency as well as evaluating the source responsible for the execution. Such methods of evaluation and increased scrutiny have dictated the development of third-party evaluation systems and the establishment of best execution committees.

Best Execution

Most asset managers and fiduciary trade execution providers have adopted so-called best execution practices that govern their daily trading procedures by committing them to pursue the most favorable price in the market at the time of the trade. Best execution requires market participants to use reasonable due diligence to ensure customers receive favorable prices under prevailing market conditions.

Often used to source liquidity is the auction process. An auction process is a trading method that sources numerous counterparties' price indications simultaneously (detailed below). Best execution practices can also include the strategy of engaging a single source of liquidity as opposed to multiple providers, depending on the expected impact caused by the trade strategy. Best execution is not a performance guarantee; rather, it helps define a consistent approach to sourcing liquidity.

Trading Styles

In order to alleviate the risk of executing a security at a value that differs greatly from its true market value, many institutional traders have adopted the use of an auction-style trade process. This process offers the advantage of seeking liquidity from multiple sources with the intent of better targeting the most interested parties.

Historically this process, known as "bid wanted in comp," was prohibitively time consuming, as bid lists were communicated by fax, and solicitations were made one at a time by telephone. However, the auction style is again becoming more widely used among most of the electronic trading platforms, and for those securities that are not traded on electronic venues, investors are creating systems to allow the same process to take place. It has been proven that this method provides the greatest amount of transparency into the current market and allows sellers and buyers to have increased confidence in their execution levels.

Finally, in an effort to improve liquidity and transparency, the NYSE Group received SEC approval of exemption to trade unlisted debt securities, adding nearly 6,000 bonds to its current inventory.

SUMMARY

Price transparency can be defined as the ability to easily establish the correct market value of a particular security. This is not always an easy task because different levels of price transparency exist in each of the major fixed income sectors. And within each of these sectors, there exists at least one subsector whose price transparency varies widely from the others. Though it may be easy to establish or obtain an indicative price level on a security, this doesn't necessarily mean that the price reflects the true market value.

To help address these issues, many pricing sources are now available, each of which establishes prices using a different method. Some of the more common methods include: matrix pricing, option-adjusted spread (OAS) calculators, price talk indications, and execution data. Not all of these methods are based on current trade data, and because of this, the development of systems like the NASD's TRACE program (which provides immediate trade transparency on eligible securities) is particularly significant.

Arguments are made both supporting and opposing the increase of bond market price transparency. Bond pricing continues to become more transparent, however, aided by advancements in technology, innovations in pricing methodologies, and changes to regulatory requirements. As this continues, investors will be equipped with more information with which to base decisions and more opportunities to add value. After all, as Warren Buffett once said, "Price is what you pay. Value is what you get."

REFERENCES

Bessembinder, H., Maxwell, W., and Venkataraman, K. (2006). Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. *Journal of Financial Economics* 82: 251–288.

Bloomfield, R., and O'Hara, M. (1999). Market transparency: Who wins and who loses? *Review of Financial Studies* 12, 1: 5–13.

Chakravarty, S., and Sarkar, A. (2003). Trading costs in the U.S. corporate, municipal and Treasury bond markets. *Journal of Fixed Income* 13, 1: 39–48.

Chen, L., Lesmond, D. A., and Wei, J. (2004). Corporate yield spreads and bond liquidity. Working paper.

Chordia, T., Sarkar, A., and Subrahmanyam, A. (2005). An empirical analysis of stock and bond market liquidity. *Review of Financial Studies* 18, 1: 85–130.

Edwards, A. K., Harris, L. E., and Piwowar, M. S. (2005). Corporate bond market transparency and transaction costs. Working paper.

Green, R. C., Hollifield, B., and Schürhoff, N. (2004). Financial intermediation and the costs of trading in an opaque market. International Center for Financial Asset Management and Engineering. Research paper no. 130.

Hong, G., and Warga, A. (2000). An empirical study of bond market transactions. *Financial Analysts Journal* 56, 2: 32–46.

Hotchkiss, E. S., and Ronen, T. (2002). The informational efficiency of the corporate bond market: An intraday analysis. *Review of Financial Studies* 15, 5: 1325–1354.

Kalimipalli, M., and Warga, A. (2002). Bid/ask spreads, volume and volatility in the corporate bond market. *Journal of Fixed Income* 11, 4: 31–42.

Levitt, A. (1998). Testimony of Chairman Arthur Levitt Before the House Subcommittee on Finance and Hazardous Materials, Committee on Commerce, Concerning Transparency in the United States Debt Market and Mutual Fund Fees and Expenses, September 29. Retrieved from www.sec.gov.

Schultz, P. (2001). Corporate bond trading costs: A peek behind the curtain. *Journal of Finance* 56 (April): 677–698.

CHAPTER 45

Bond Spreads and Relative Value

MOORAD CHOUDHRY, PhD
Head of Treasury, KBC Financial Products, London

Abstract: Bond spreads are used to determine relative value in bonds that are not credit risk free. This relative value is a measure of the risk premium return implied in the bond yield. To calculate bond spread, we may use one of four different measures, which are described and illustrated in this chapter. Investors can also determine relative value in the cash market by comparing its yield spread with that observed in the synthetic market, which is represented by the credit default swap premium payable for the same reference name. The spread difference between the two markets is known as the credit default swap basis and the existence of the basis implies arbitrage opportunities between the two markets.

Keywords: asset swap, credit default swaps (CDSs), basis, bond spread, interpolated spread, relative value, swap spread, zero-volatility spread (Z-spread)

In this chapter we consider the methods by which the value of a bond can be ascertained by comparing its yield to that of another bond or benchmark interest rate. Whichever method is used to calculate it, this measure is important for investors, as it is an indication of *relative value* and, as such, the main method by which they gauge whether it is worthwhile to hold the bond and if its risk-reward profile is acceptable. The methods that can be used reflect the differences in measuring yield in cash and derivative markets, and include the *interpolated spread,* Treasury spread, and asset-swap spread. We also discuss the difference in yield between cash and synthetic credit markets, known as the credit derivative *basis,* and how this influences the measurement of bond relative value.

BOND SPREADS

Investors measure the perceived market value, or relative value, of a corporate bond by measuring its yield spread relative to a designated benchmark. This is the spread over the benchmark that gives the yield of the corporate bond. A key measure of relative value of a corporate bond is its *swap spread.* This is the basis-point spread over the interest rate swap curve, and is a measure of the credit risk of the bond. In its simplest form, the swap spread can be measured as the difference between the yield to maturity of the bond and the interest rate given by a straight-line interpolation of the swap curve. In practice, traders use the asset swap spread and the *zero-volatility spread (Z-spread)* as the main measures of relative value. The government *bond spread* is also used. In addition, now that the market in synthetic corporate credit is well established, using credit derivatives and *credit default swaps (CDSs),* investors may consider the cash-CDS spread as well, which is the basis and which we consider in greater detail later.

The spread that is selected is an indication of the relative value of the bond, and a measure of its credit risk. The greater the perceived risk, the greater the spread should be. This is best illustrated by the credit structure of interest rates, which will (generally) show AAA- and AA-rated bonds trading at the lowest spreads and BBB, BB, and lower-related bonds trading at the highest spreads. Bond spreads are the most commonly used indication of the risk-return profile of a bond.

Swap Spread and Treasury Spread

In this section we consider the Treasury spread, asset swap spread, Z-spread, and basis. A bond's swap spread is a measure of the credit risk of that bond, relative to the interest rate swaps market. Because the swaps market is traded by banks, this risk is effectively the interbank market, so the credit risk of the bond over and above bank risk is given by its spread over swap rates. This is a simple calculation to make, and is simply the yield of the bond minus the swap rate for the appropriate maturity swap.

The spread over swaps is sometimes called the I-spread. It has a simple relationship to swaps and Treasury yields, shown here in the equation for corporate bond yield:

$$Y = I + S + T \tag{45.1}$$

where

Y = yield on the corporate bond
I = I-spread or spread over swap
S = swap spread
T = yield on the Treasury security (or an interpolated yield)

In other words, the swap rate itself is given by $T + S$.

The I-spread is sometimes used to compare a cash bond with its equivalent CDS price, but for straightforward relative value analysis it is usually dropped in favor of the asset swap spread, which we look at later in this section.

Of course, the basic relative value measure is the Treasury spread or government bond spread. This is simply the spread of the bond yield over the yield of the appropriate government bond. Again, an interpolated yield may need to be used to obtain the right Treasury rate to use. The bond spread is given by:

$$BS = Y - T$$

Using an interpolated yield is not strictly accurate because yield curves are smooth in shape, and so straight-line interpolation will produce slight errors. Despite this, the method is still commonly used by market practitioners.

Asset Swap Spread

An *asset swap* is a package that combines an interest rate swap with a cash bond, the effect of the combined package being to transform the interest rate basis of the bond. Typically, a fixed-rate bond will be combined with an interest rate swap in which the bondholder pays fixed coupon and receives floating coupon. The floating coupon will be a spread over the London Interbank Offered Rate (LIBOR) (see Choudhry et al., 2001). This spread is the asset swap spread and is a function of the credit risk of the bond over and above interbank credit risk. (This is because in the interbank market, two banks transacting an interest rate swap will be paying/receiving the fixed rate and receiving/paying LIBOR flat.) Asset swaps may be transacted at par or at the bond's market price, usually par. This means that the asset swap value is made up of the difference between the bond's market price and par, as well as the difference between the bond coupon and the swap fixed rate.

The zero-coupon curve is used in the asset swap valuation. This curve is derived from the swap curve, so it is the implied zero-coupon curve. The asset swap spread is the spread that equates the difference between the present value of the bond's cash flows, calculated using the swap zero rates, and the market price of the bond. This spread is a function of the bond's market price and yield, its cash flows, and the implied zero-coupon interest rates. (Bloomberg refers to this spread as the "gross spread.")

For example, on August 10, 2005, a U.K. pound sterling (GBP)-denominated corporate bond, GKN Holdings 7% 2012, was observed to have a an asset swap spread of 121.5 basis points. This is the spread over LIBOR that will be received if the bond is purchased in an asset swap package. In essence, the asset swap spread measures a difference between the market price of the bond and the value of the bond when cash flows have been valued using zero-coupon rates. The asset swap spread can therefore be regarded as the coupon of an annuity in the swap market that equals this difference.

Zero-Volatility Spread

The conventional approach for analyzing an asset swap uses the bond's yield-to-maturity (YTM) in calculating the spread. The assumptions implicit in the YTM calculation (see Chapter 17 in Volume I) make this spread problematic for relative analysis, so market practitioners use what is termed the "zero-volatility" or "Z-spread" instead. The Z-spread uses the zero-coupon yield curve to calculate spread, so is a more realistic and effective spread to use. The zero-coupon curve used in the calculation is derived from the interest rate swap curve.

Put simply, the Z-spread is the basis-point spread that would need to be added to the implied spot yield curve such that the discounted cash flows of a bond are equal to its present value (its current market price). Each bond cash flow is discounted by the relevant spot rate for its maturity term. How does this differ from the conventional asset swap spread? It differs essentially in its use of zero-coupon rates when assigning a value to a bond. Each cash flow is discounted using its own particular zero-coupon rate. A bond's price at any time can be taken to be the market's value of the bond's cash flows. Using the Z-spread, we can quantify what the swap market thinks of this value, that is, by how much the conventional spread differs from the Z-spread. Both spreads can be viewed as the coupon of a swap market annuity of equivalent credit risk of the bond being valued.

In practice, the Z-spread, especially for shorter-dated bonds and for better credit-quality bonds, does not differ greatly from the conventional asset-swap spread. The Z-spread is usually the higher spread of the two, following the logic of spot rates, but not always. If it differs greatly, then the bond can be considered to be mispriced.

Taking the same bond mentioned earlier and as at the same date, we observed a number of different spreads for the bond. The main spread of 151.00 bps was the spread over the government yield curve. This is an interpolated spread over the appropriate benchmark sovereign bond. The asset swap spread was 121.5 bps as stated earlier, while the Z-spread was 118.8 bps. When undertaking relative value analysis, for instance if making comparisons against cash funding rates or the same company name CDS, it is this lower spread that should be used.

The Z-spread is closely related to the bond price, as shown by

$$P = \sum_{i=1}^{n} \left[\frac{C_i + M_i}{(1 + ((Z + S_i + T_i)/m))^i} \right] \tag{45.2}$$

where

n = number of interest periods until maturity
P = bond price
C = coupon
M = redemption payment (so bond cash flow is all C plus M)
Z = Z-spread
m = frequency of coupon payments

In effect, this is the standard bond price equation, with the discount rate adjusted by whatever the Z-spread is; it is an iterative calculation. The appropriate maturity swap rate is used, which is the essential difference between the I-spread and the Z-spread. This is deemed to be more accurate, because the entire swap curve is taken into account rather than just one point on it. In practice, though, as we have seen in the preceding example, there is often little difference between the two spreads.

To reiterate, then, using the correct Z-spread, the sum of the bond's discounted cash flows will be equal to the current price of the bond.

We illustrate the Z-spread calculation in Figure 45.1. This is done using a hypothetical bond, the XYZ plc 5% of June 2008, a three-year bond at the time of the calculation. Market rates for swaps, Treasury and CDS are also shown. We require the spread over the swaps curve that equates the present values of the cash flows to the current market price. The cash flows are discounted using the appropriate swap rate for each cash flow maturity. With a bond yield of 5.635%, we see that the I-spread is 43.5 basis points, while the Z-spread is 19.4 basis points. In practice, the difference between these two spreads is rarely this large.

We also show the Excel formula in Figure 45.1. This shows how the Z-spread is calculated; for ease of illustration, we have assumed that the calculation takes place for value on a coupon date, so that we have precisely an even period to maturity.

THE ASSET SWAP CDS PRICE

Credit default swaps provide an efficient means of pricing pure credit and, by definition, are a measure of the credit risk of a specific reference entity or reference asset. Asset swaps are well established in the market and are used both to transform the cash flow structure of a corporate bond and to hedge against interest rate risk of a holding in such a bond. As asset swaps are priced at a spread over LIBOR, with LIBOR representing interbank risk, the asset swap spread represents in theory the credit risk of the asset swap name. By the same token, using the no-arbitrage principal, it can be shown that the price of a CDS for a specific reference name should equate the asset swap spread for the same name. However, a number of factors, both structural and operational, combine to make CDSs trade at a different level to asset swaps. This difference in spread is known as the credit default swap *basis* and can be either positive (the credit default swap trading above the asset swap level) or negative (trading below the asset swap).

Asset Swap Pricing

At the inception of the market, credit derivatives were valued using the asset swap pricing technique. We will explain shortly why this approach is no longer used. However, let us first consider the theoretical reason why they should be priced using this approach.

A par asset swap typically combines the sale of an asset such as a fixed-rate corporate bond to a counterparty, at par and with no interest accrued, with an interest rate swap. The coupon on the bond is paid in return for LIBOR, plus a spread, if necessary. This spread is the asset swap spread and is the price of the asset swap. In effect, the asset swap allows market participants that pay LIBOR-based funding to receive the asset swap spread. This spread is a function of the credit risk of the underlying bond asset, which is why it could be viewed as equivalent to the price payable on a credit default swap written on that asset.

The generic pricing is given by

$$Y_a = Y_b - ir \tag{45.3}$$

where

Y_a = asset swap spread
Y_b = asset spread over the benchmark
ir = interest rate swap spread

The asset spread over the benchmark is simply the bond (asset) redemption yield over that of the government benchmark. The interest rate swap spread reflects the cost involved in converting fixed-coupon benchmark bonds into a floating-rate coupon during the life of the asset (or default swap) and is based on the swap rate for that maturity.

The theoretical basis for deriving a default swap price from the asset swap rate can be illustrated by looking at a basis-type trade involving a cash market reference asset (bond) and a default swap written on this bond. This is similar in concept to the risk-neutral or no-arbitrage concept used in derivatives pricing. The theoretical trade involves:

- A long position in the cash market floating-rate note (FRN) priced at par, and which pays a coupon of LIBOR + X basis points.
- A long position (bought protection) in a default swap written on the same FRN, of identical term-to-maturity and at a cost of Y basis points.

A1	B	C	D	E	F	G	H	I
2	**Issuer**	**XYZ plc**						
3	Settlement date	6/1/2005						
4	Maturity date	6/1/2008						
5	Coupon	5%		**YIELD**	**0.05635**			
6	Price	98.95		[Cell formula =YIELD(C4,C5,C6,C7,C8,C9,C10)]				
7	Par	100						
8	Semi-annual coupon	2		**PRICE**	**98.95000**			
9	act/act	1		[Cell formula =PRICE(C4,C5,C6,C6,C8,C9,C10)]				
10								
11	Bond yield	5.635%						
12	Sovereign bond yield	4.880%						
13	Swap rate	5.200%						
14								
15	3-year CDS price	28 bps						
16								
17	**Treasury spread**							
18	5.635 - 4.88	55 bps						
19								
20	**I-spread**							
21	5.635 - 5.20	43.5 bps						
22								
23	**Z-spread (Z)**	19.4 bps	0.00194					
24	The Z-spread is found using iteration							
25								
26								**Sum of PVs**
27	Cash flow date	12/1/2005	6/1/2006	12/1/2006	6/1/2007	12/1/2007	6/1/2008	
28	Cash flow maturity (years)	0.50	1.00	1.50	2.00	2.50	3.00	
29	0.5-year swap rate (S)	4.31%	4.84%	4.99%	5.09%	5.18%	5.20%	
30	Cash flow (CF)	2.50	2.50	2.50	2.50	2.50	102.50	
31	Discount factor	0.97797598	0.9514988	0.9261035	0.9009477	0.8758358	0.8524197	
32	(DF Calculation)	1/(1+(S+Z)/2)^1+(S+Z)/2)^2+(S+Z)/2)^3+(S+Z)/2)^4+(S+Z)/2)^5+(S+Z)/2)^6						
33	CF present value (PV)	2.445	2.379	2.315	2.252	2.190	87.373	**98.95**
34								
35								
36								
37	A Z-spread of 19.4 basis points gives us the current bond price so is the correct one							
38	Using this value, the sum of all the discounted cashflows is equal to the market price							
39								
40	**CDS Basis**							
41	28 - 19.4	8.6 bps						
42	The basis is positive in this example							

Figure 45.1 Calculating the Z-Spread, Hypothetical 5% 2008 Bond Issued by XYZ plc

The buyer of the bond is able to fund the position at LIBOR. In other words, the bondholder has the following net cash flow:

$$(100 - 100) + ((\text{Libor} + X) - (\text{Libor} + Y))$$

or $X - Y$ basis points.

In the event of default, the bond is delivered to the protection seller in return for payment of par, enabling the bondholder to close out the funding position. During the term of the trade, the bondholder has earned $X - Y$ basis points while assuming no credit risk. For the trade to meet the no-arbitrage condition, we must have $X = Y$. If $X \neq Y$, the investor would be able to establish the position and generate a risk-free profit.

This is a logically tenable argument as well as a reasonable assumption. The default risk of the cash bondholder is identical in theory to that of the default seller. In the next section we illustrate an asset swap pricing example, before looking at why in practice there exist differences in pricing between default swaps and cash market reference assets.

Asset Swap Pricing Example

XYZ plc is a Baa2-rated corporate. The seven-year asset swap for this entity is currently trading at 93 basis points; the underlying seven-year bond is hedged by an interest rate swap with an Aa2 rated bank. The risk-free rate for

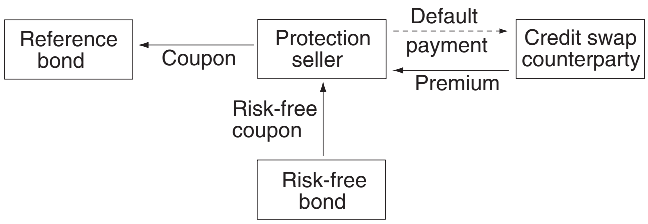


Figure 45.2 Credit Default Swap and Asset Swap Hedge

floating-rate bonds is London Interbank Bid Rate (LIBID) minus 12.5 basis points (assume the bid-offer spread is 6 basis points). This suggests that the credit spread for XYZ plc is 111.5 basis points. The credit spread is the return required by an investor for holding the credit of XYZ plc. The protection seller is conceptually long the asset, and so would short the asset to hedge its position. This is illustrated in Figure 45.2. The price charged for the default swap is the price of shorting the asset, which works out as 111.5 basis points each year.

Therefore, we can price a credit default written on XYZ plc as the present value of 111.5 basis points for seven years, discounted at the interest rate swap rate of 5.875%. This computes to a credit default swap price of 6.25%.

Reference XYZ plc
Term Seven years
Interest rate swap rate 5.875%
Asset swap LIBOR plus 93 bps

Default swap pricing:

Benchmark rate LIBID minus 12.5 bps
Margin 6 bps
Credit default swap 111.5 bps
Default swap price 6.252%

Pricing Differentials

A number of factors observed in the market serve to make the price of credit risk that has been established synthetically using default swaps to differ from its price as traded in the cash market. In fact, identifying (or predicting) such differences gives rise to arbitrage opportunities that may be exploited by basis trading in the cash and derivative markets. These factors include the following:

- Bond identity—the bondholder is aware of the exact issue that they are holding in the event of default; however, CDS sellers may receive potentially any bond from a basket of deliverable instruments that rank pari passu with the cash asset; this is the delivery option afforded the long swap holder.
- The borrowing rate for a cash bond in the repo market may differ from LIBOR if the bond is to any extent special; this does not impact the CDS price, which is fixed at inception.
- Certain bonds rated AAA (such as U.S. agency securities) sometimes trade below LIBOR in the asset swap market; however, a bank writing protection on such a bond will expect a premium (positive spread over LIBOR) for selling protection on the bond.
- Depending on the precise reference credit, the credit default swap may be more liquid than the cash bond, resulting in a lower default swap price, or less liquid than the bond, resulting in a higher price.
- CDSs may be required to pay out on credit events that are technical defaults, and not the full default that impacts a cash bondholder; protection sellers may demand a premium for this additional risk.
- The default swap buyer is exposed to counterparty risk during the term of the trade, unlike the cash bondholder.
- An issuance of new bonds by the same reference name may increase demand for credit protection, resulting in a higher CDS price.

For these and other reasons, the CDS price often differs from the cash market price for the same asset. Therefore, banks are increasingly turning to credit pricing models, based on the same models used to price interest rate derivatives, when pricing credit derivatives.

The difference between the CDS price and the asset swap spread can be observed for any number of corporate credits across all market sectors. This suggests that middle-office staff and risk managers that use the asset swap technique to independently value default swap books are at risk of obtaining values that differ from those in the market. This is an important issue for credit derivative market-making banks.

Cash-CDS Basis

To reiterate then, the CDS basis is the CDS spread minus the asset swap spread. Alternatively, it can be the CDS spread minus the Z-spread. So the basis is given by

$$B = D - CashSpread \qquad (45.4)$$

where D is the CDS price. Where $D - CashSpread > 0$ it is a positive basis, the opposite is a negative basis. There is no one accepted measure of *CashSpread*; practitioners generally use either the I-spread, asset swap spread or Z-spread. The only rule is to use the same measure consistently when conducting relative value analysis. Further observations on the efficacy of using each method are given in Choudhry (2006). Measuring the basis becomes more important when used in formulating arbitrage strategy. Changes in the basis give rise to arbitrage opportunities between the cash and synthetic markets, which can be exploited via a negative basis trade (buying the reference name cash bond and buying protection in CDS) or positive basis trade (selling the bond and selling protection). This is discussed in greater detail in Choudhry (2006).

A wide range of factors drive the basis. The existence of a non-zero basis has implications for investment strategy. For instance, when the basis is negative investors may prefer to hold the cash bond, whereas if for liquidity, supply, or other reasons the basis is positive, the investor may wish to hold the asset synthetically by selling protection using a CDS. Another approach is to arbitrage between the cash and synthetic markets, in the case of a negative basis by buying the cash bond and shorting it synthetically by buying protection in the CDS market. Investors have a range of spreads to use when performing their relative value analysis.

SUMMARY

Investors use bond spread analysis to determine the relative value of a bond compared to a benchmark bond or yield curve. The spread is measured in basis points and can be used to ascertain if the expected return from the bond is sufficient compensation for the risk profile it represents. To calculate the bond spread, we can use one of four measures. These measures are the interpolated spread over the benchmark government bond yield, the interpolated spread over the interest rate swap curve, the asset swap spread, and the Z-spread.

The development of a liquid market in credit default swaps means that there is now a yield measure for both the cash market and the synthetic market. The price of a CDS written on a specific reference name is another measure of its relative value. The difference between the cash market yield, given by any spread measure, and the synthetic price is the credit default swap basis.

REFERENCES

Andritzky, J., and Singh, M. (2007). Recovery value effect on CDS during distress. *Euromoney Structured Credit Products Handbook 2007/08* (pp. 69–77). London: Euromoney Publications.

Choudhry, M., Joannas, D., Pienaar, R., and Pereira, R. (2001). *Capital Market Instruments: Analysis and Valuation*. London: FT Prentice Hall.

Choudhry, M. (2004). *Structured Credit Products: Credit Derivatives and Synthetic Securitisation*. Singapore: John Wiley & Sons.

Choudhry, M. (2006). *The Credit Default Swap Basis.* Princeton, NJ: Bloomberg Press.

Leibowitz, M., and Homer, S. (2004). *Inside the Yield Book.* Princeton, NJ: Bloomberg Press.

Martellini, L., and Priaulet, P. (2004). *Fixed Income Securities.* Chichester: John Wiley & Sons.

Miller, T. (2007). *Introduction to Option-Adjusted Spread Analysis*, revised and expanded 3rd edition of the OAS Classic by Tom Windas. Princeton, NJ: Bloomberg Press.

CHAPTER 46

The Determinants of the Swap Spread and Understanding the LIBOR Term Premium

MOORAD CHOUDHRY, PhD
Head of Treasury, KBC Financial Products, London

Abstract: The use of interest rate swaps to manage and hedge interest rate risk exposure is widespread. The rate payable on a swap is a positive spread above the sovereign benchmark rate, reflecting interbank quality credit risk. The magnitude of this spread is dependent on various factors, including prevailing macro-level political and economic circumstances, supply and demand, the shape and level of the yield curve, and market volatility. Market practitioners analyze these factors and anticipate changes in the spread, as the extent of the spread feeds into hedge costs. Funding and hedging costs also reflect overall term premium rates. The level of the term premium in the short-term yield curve is a function of the current shape of the curve and market expectations of interest rate levels in the future.

Keywords: futures curve, interest rate swap, LIBOR curves, LIBOR spread, swap spread, term premium

An important hedging tool in bank asset-liability management (ALM) operations is the *interest rate swap*, a derivative instrument. In this chapter, we consider an important issue for interest rate analysis, the *swap spread*. Specifically, we look at the spread of the swap curve over the government bond yield curve; this subject is important because the swap spread is an indicator of value and risk premium in the market, as well as an indicator of the overall health of the economy. Understanding the determinants of the swap spread is worthwhile for ALM practitioners for this reason. In the second part of this chapter we look at a related area: the magnitude of the *term premium*. Given "normal" market conditions, what should be the extent of the term premium of the (under normal conditions, positively sloping) yield curve? The two subjects are related because they both represent a measure of risk premium in the economy at any given point in time.

DETERMINANTS OF THE SWAP SPREAD

Interest rate swaps are an important ALM and risk management tool in banking markets. The rate payable on a swap represents bank risk, if we assume that a swap is paying (receiving) the fixed swap rate on one leg and receiving (paying) London Interbank Offered Rate (LIBOR) flat on the other leg. If one of the counterparties is not a bank, then either leg is adjusted to account for the different counterparty risk; usually, the floating leg will have a spread added to LIBOR. We can see that this produces a swap curve that lies above the government bond yield curve, if we compare Figure 46.1 with Figure 46.2. Figure 46.1 is the U.S. dollar (USD) swap rates page from Tullett & Tokyo brokers as of July 3, 2006; Figure 46.2 is the U.S. Treasury yield curve for the same day. The higher

2 P1P122 Govt **TTIS**
200<Go> to view in Launchpad
10:10 **TULLETT & TOKYO** PAGE 1 / 2

USD Swaps	Bid	Ask	Time	USD Swaps	Bid	Ask	Time
IMM SWAPS				19) 15 Year	5.7380	5.7780	1:46
1) 1st	5.6900	5.7100	10:08	20) 20 Year	5.8060	5.8460	1:46
2) 2nd	5.6580	5.6790	10:08	21) 25 Year	5.8080	5.8490	3:02
3) 3rd	5.6150	5.6350	10:10	22) 30 Year	5.7740	5.8150	1:46
4) 4th	5.5800	5.6000	10:01	SEMI-ANNUAL SWAPS			
ANNUAL SWAPS				23) 2 Year	5.6310	5.6710	7/03
5) 1 Year	5.6770	5.6970	9:36	24) 3 Year	5.6120	5.6520	7/03
6) 2 Year	5.6240	5.6650	7/03	25) 4 Year	5.6210	5.6610	7:16
7) 3 Year	5.6070	5.6480	7/03	26) 5 Year	5.6410	5.6810	0:01
8) 4 Year	5.6180	5.6590	7:16	27) 6 Year	5.6600	5.7000	0:01
9) 5 Year	5.6390	5.6790	0:01	28) 7 Year	5.6770	5.7170	0:01
10) 6 Year	5.6560	5.6970	0:01	29) 8 Year	5.6910	5.7310	0:01
11) 7 Year	5.6740	5.7150	0:01	30) 9 Year	5.7080	5.7480	0:01
12) 8 Year	5.6890	5.7300	0:01	31) 10 Year	5.7270	5.7670	7:40
13) 9 Year	5.7070	5.7470	0:01	32) 11 Year	5.7370	5.7770	0:01
14) 10 Year	5.7250	5.7650	7:40	33) 12 Year	5.4570	5.4970	1:46
15) 11 Year	5.7350	5.7760	0:01	34) 13 Year	5.7620	5.8020	0:01
16) 12 Year	5.4520	5.4920	1:46	35) 14 Year	5.7720	5.8120	0:01
17) 13 Year	5.7610	5.8020	0:01	LIVE Treasury Mid-Yields &			
18) 14 Year	5.7700	5.8110	0:01	Treasury Swap Spreads -> SMKR<GO>			

Figure 46.1 Brokers USD Interest Rate Swaps Page on Bloomberg, July 3, 2006

<HELP> for explanation. P122 Govt **IYC**
Cancel: Screen not saved
YIELD CURVE – US TREASURY ACTIVES Page 2/2
DATE 7/ 4/06

	DESCRIPTION	PRICE	SRC	UPDATE	YIELD	HEDGED YIELD
3MO	1) B 0 09/28/06	B 4.8900	BGN	4:00	5.0158	5.0158
6MO	2) B 0 12/28/06	B 5.0700	BGN	4:00	5.2711	5.2711
1YR	3)					
2YR	4) T 5 1/8 06/30/08	B 99.9063	BGN	4:00	5.1748	5.1748
3YR	5) T 4 7/8 05/15/09	B 99.2969	BGN	4:00	5.1400	5.1400
4YR	6)					
5YR	7) T 5 1/8 06/30/11	B100.0625	BGN	4:00	5.1104	5.1104
6YR	8)					
7YR	9)					
8YR	10)					
9YR	11)					
10YR	12) T 5 1/8 05/15/16	B 99.7813	BGN	4:00	5.1527	5.1527
15YR	13)					
20YR	14)					
30YR	15) T 4 1/2 02/15/36	B 89.4531	BGN	4:00	5.2017	5.2017
1MO	16) B 0 07/27/06	B 4.5500	BGN	4:00	4.6261	4.6261

To change price source for securities, use <FMPS>.
To change price source for swaps, use <XDF>.
Yields are based on STANDARD settlement and are Conventional

Figure 46.2 U.S. Treasury Yield Curve on July 3, 2006

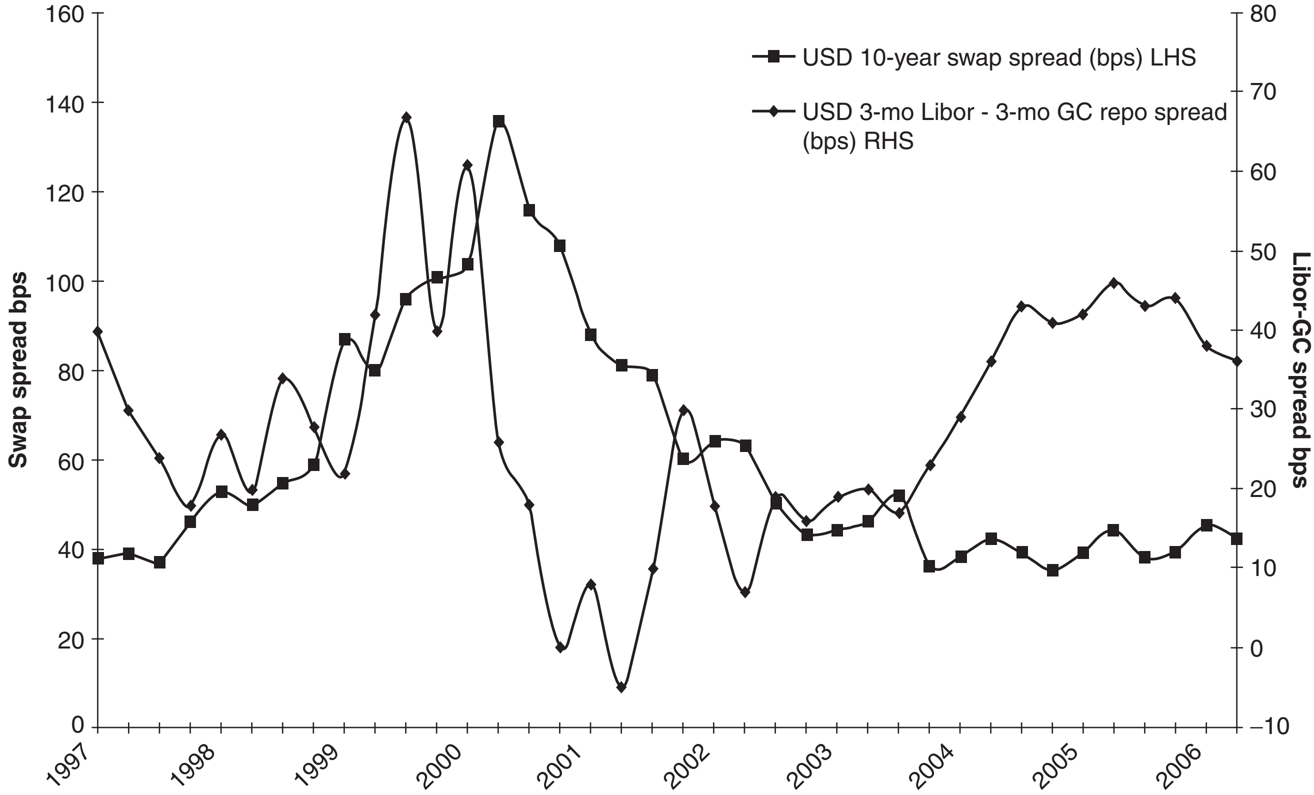


Figure 46.3 Comparison of USD Swap Spread and Three-Month LIBOR-GC Spread

rates payable on swaps represent the additional risk premium associated with bank risk compared to government risk. The spread itself is the number of basis points the swap rate lies above the equivalent-maturity government bond yield, quoted on the same interest basis.

In theory, the swap spread represents only the additional credit risk of the interbank market above the government market. However, as the spread is variable, it is apparent that other factors influence it. An ALM desk will want to be aware of these factors, because they influence swap rates. Swaps are an important (if not the most important) risk-hedging tool for banks, so it becomes necessary for practitioners to have an appreciation of what drives the swap spreads.

Determinants of the Spread

We already noted that in theory the swap spread, representing interbank counterparty risk, should reflect only the market's perception of bank risk over and above sovereign risk. Bank risk is captured in LIBOR—the rate paid by banks on unsecured deposits to other banks. (In fact, banks are more likely to pay LIMID to other banks, and the biggest banks pay London Interbank Bid Rate [LIBID]. But we can safely ignore this for the purposes of our discussion here.) So, in other words, the swap spread is meant to compensate adequately against the risk of bank default. LIBOR is the floating rate paid against the fixed in the swap transaction, and moves with the perception of bank risk.

Other factors influence the swap spread. We can illustrate this better comparing the swap spread for 10-year quarterly paying swaps with the spread between 3-month LIBOR and the 3-month general collateral (GC) repo rate. The GC rate is the risk-free borrowing rate, whereas the LIBOR rate represents bank risk again. In theory, the spread between 3-month LIBOR and the GC rate should therefore move closely with the swap spread for quarterly resetting swaps, as both represent bank risk. A look at Figure 46.3 shows us that this is not the case. Figure 46.3 compares the two spreads in the U.S. dollar market, but we do not need to calculate the correlation or the R^2 for the two sets of numbers. Even on cursory observation, we can see that the correlation is not high. Therefore, we conclude that other factors, in addition to perceived bank default risk, drive one or both spreads. These other factors influence swap rates and government bond yields, and hence the swap spread, and we consider them in this chapter.

Level and Slope of the Yield Curve

The magnitude of the swap spread is influenced by the absolute level of base interest rates. If the base rate is 10%, so that the government short-term rate is around 10%, with longer-term rates being recorded higher, the spread tends to be greater than that seen if the base rate is 5%. The shape of the yield curve has even greater influence. When the curve is positively sloping, under the expectations hypothesis explanation of the shape of the yield curve, investors will expect future rates to be higher; hence, floating rates are expected to rise. This would suggest that the swap spread will narrow. The opposite happens if the yield curve inverts.

Supply and Demand

The swap spread is influenced greatly by supply and demand for swaps. For example, greater-volume cash market instruments drive up a need for hedging instruments, which will widen swap spreads. The best example of this is corporate bond issuance; as volumes increase, the need for underwriters to hedge their bond holdings increases. Greater bond issuance also has another impact, as issuers seek to swap their fixed-rate liabilities to floating-rate. This also increases demand for swaps.

Market Volatility

Swap spreads widen during times of market volatility. This may be in times of market uncertainty (e.g., the future direction of base rates or possible inversion of the yield curve) or in times of market shock such as 9/11. In some respects, widening during periods of volatility reflects the perception of increased bank default risk. It also reflects the "flight to quality" that occurs during times of volatility or market correction: This is the increased demand for risk-free assets such as government bonds that drives their yields lower and hence swap spreads wider.

Government Borrowing

The level of government borrowing influences government bond yields, so perforce will also impact swap spreads. If borrowing is viewed as being in danger of getting out of control, or the government runs persistently large budge deficits, government bond yields will rise. All else being equal, this will lead to narrowing swap spreads.

We can see then that a number of factors influence swap spreads. An ALM or Treasury desk should be aware of these and assess them because the swap rate represents a key funding and hedging rate for a bank.

MAGNITUDE OF THE TERM PREMIUM

From an introductory reading of financial economics, we know that a positively sloping yield curve is to be expected under transparent, liquid market conditions. A combination of the expectations hypothesis, the liquidity premium and the inflation premium explains why this is so; longer-dated assets yield a higher return than shorter-dated assets. Thus, in most circumstances, we expect the one-month rate to be higher than the one-week rate, and the three-month rate to be higher than both the one-week and the one-month rates. This is confirmed at Figure 46.4, which shows the *LIBOR curves* for U.S. dollar and sterling (GBP) on May 25, 2006.

We expect that the rate on a longer term will be higher than that on a shorter term, unless we have an inverted

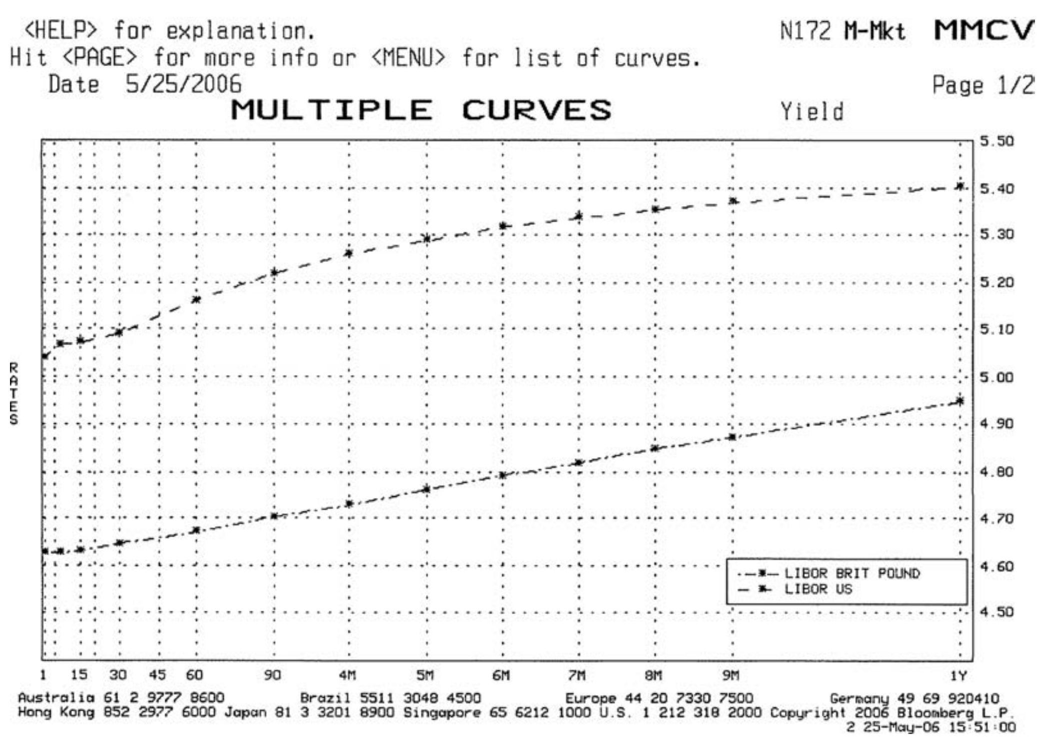


Figure 46.4 USD and GBP Money Market Yield Curves on Bloomberg, May 25, 2006
Source: © Bloomberg L.P. Reproduced with permission. All rights reserved. Visit www.bloomberg.com <http://www.bloomberg.com>

yield curve. This is because under most circumstances lenders demand a higher return for longer-dated loans as compensation for the increased inflation and credit risk exposure of longer-dated assets. But what should be the magnitude of this term premium? By how much more should a three-month deposit pay compared to a one-month deposit?

The answer to this question is not fixed and is a function of a number of factors. In a developed economy that is not subject to high inflation, the most important of these factors is probably future interest rate expectations. If we allow for this factor, we can conclude that a reasonable term premium under "normal" market conditions for the three-month rate compared to the central bank base rate is in the order of between 12 and 20 basis points. We choose the three-month rate because it is traded on a liquid futures contract (the Eurodollar and short-sterling contracts for USD and GBP, respectively) and so we can analyze the market's forward rate expectations for this tenor deposit. But the basic principles will apply to any maturity. Of course, there is no such thing as a "normal" market condition; the term premium will fluctuate daily and always reflect the interaction of a number of factors.

Illustration

On May 25, 2005, we observed the following rates for U.S. dollar:

Fed funds rate (overnight)	3.00%
Three-month LIBOR fix	3.31%

The three-month rate is 31 basis points above the overnight rate.

The same rates for pounds sterling are:

Bank of England base rate	4.75%
Three-month LIBOR fix	4.87%

The three-month rate here is at a much lower spread, only 12 basis points.

Fast-forwarding one year later to May 25, 2006, we observe the following rates:

Fed funds rate (overnight)	5.00%
Three-month LIBOR fix	5.22%
Bank of England base rate	4.50%
Three-month LIBOR fix	4.705%

The spreads here are 22 basis points for the U.S. dollar and 20.5 basis points for sterling.

We need to look at market expectations for an explanation of these term premiums. In May 2005, the market was expecting a continuation of the gradual, "measured" interest raises (the Federal Reserve's own term to describe its rate-setting policy), in clips of 25 basis points, at each meeting of the Federal Reserve. (The Federal Reserve's Open Market Committee [FOMC], which sets the USD base rate, meets every six weeks or so.) This is reflected in the positively sloping yield curve for USD money markets, as shown in Figure 46.5. This is confirmed in Figure 46.6, a graph of the Fed Funds rate for the period May 2005–May 2006, which shows that the rate was moved upward by 25 basis points at every Fed meeting up until the May 10, 2006, meeting, when the rate was raised to

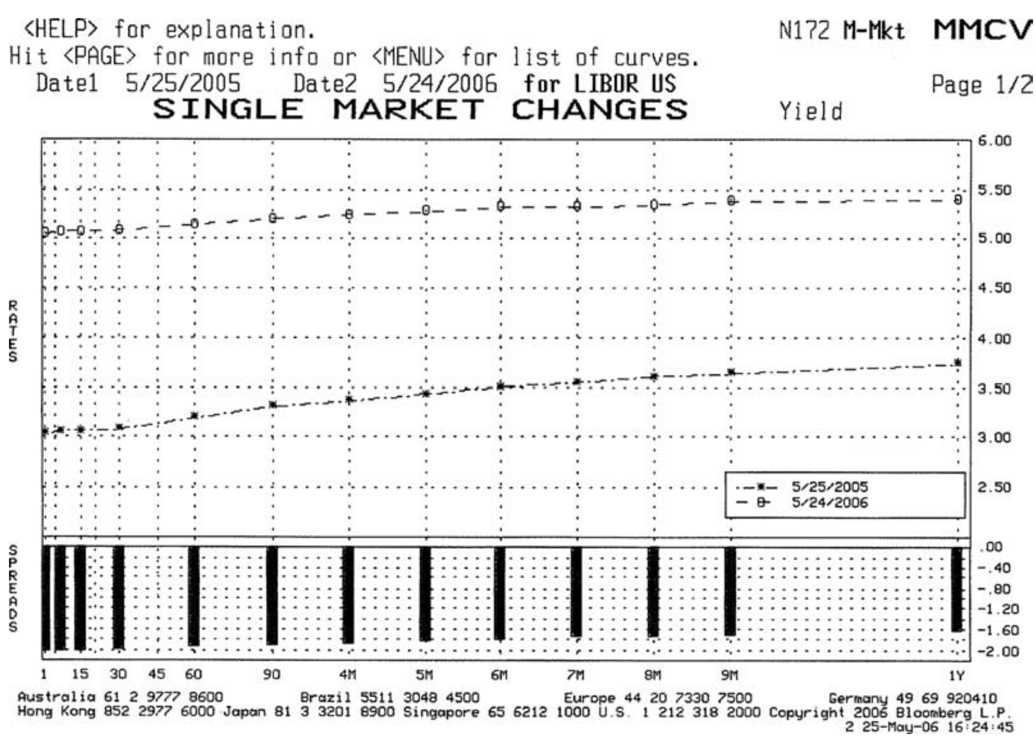


Figure 46.5 Positively Sloping USD Money Market Yield Curves: May 24, 2005 and May 25, 2006
Source: © Bloomberg L.P. Reproduced with permission. All rights reserved. Visit www.bloomberg.com <http://www.bloomberg.com>

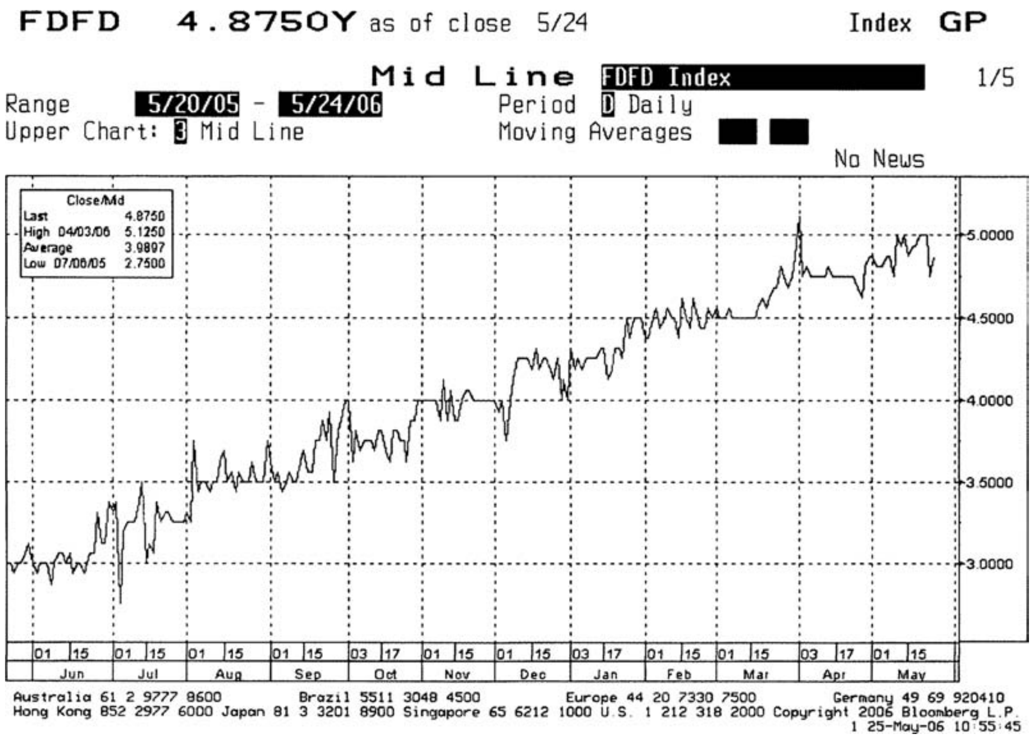


Figure 46.6 Fed Funds Rate for May 2005–May 2006
Source: © Bloomberg L.P. Reproduced with permission. All rights reserved. Visit www.bloomberg.com <http://www.bloomberg.com>

5.00%. Lenders will require a premium to reflect the expectations of higher interest rates; hence, the three-month term premium in May 2005 was 31 basis points. Figure 46.7 shows the USD three-month LIBOR rate history for the same period.

Contrast the situation with pound sterling. In May 2005, the prevailing market sentiment was that the next move in base rates would be downward. This is shown in Figure 46.8, which shows that the money market yield curve for May 25, 2005, was inverted.

Note that the curve slopes gently upward before then inverting, implying that the market expected the cut in rates to be in a period more than three months from now. However, the term premium was only 12 basis points, reflecting the negative curve. Figures 46.9 and 46.10 show the rate histories for the Bank of England base rate and GBP three-month LIBOR.

So we see that the term premium reflects the market expectations of future rates, and in an environment where the expectations are for higher rates, the premium will be higher. The opposite applies where the expectation is for lower base rates.

This begs the question, "What should the term premium be in a 'neutral' interest rate environment? That is, what should a lender demand for term funds lent out when the market does not expect rates to be stable over the next 12 months and not move up or down?

We can look at the 90-day money futures contracts for an idea of when this is the case. In May 2006, the outlook for base rates in USD and GBP was fairly stable. In the United States, the consensus was that rates would either top out at 5.00% or be raised one more time to 5.25% at the June 29, 2006, FOMC meeting. This is shown by the Eurodollar curve, which gives us the market expectations for forward three-month deposit rates. Figure 46.11 shows the Eurodollar curve on May 25, 2006. Figure 46.12 shows the LIBOR fix for the same day.

We see that the curve is essentially flat. The market expectations for 90-day money range from 5.275% in June 2006 to 5.235% in June 2007. This implies that fair value in a stable rate environment is roughly 22 basis points for U.S. dollars.

The scenario in the United Kingdom is slightly different. Figure 46.13, the short-sterling curve for May 25, 2006, shows an expectation of rising base rates in the following 12 months. We see that the expected 90-day LIBOR fix for June 2006 is 4.72%, compared to 5.070% for June 2007. In the case of sterling, there is possibly greater uncertainty compared to the United States, which was approaching the end of an obvious rising rates cycle. In the United Kingdom, only a few months previously there was commentary that the next move in rates would be down (rates had been stable since the cut to 4.50% in August 2005). This uncertainty is perhaps reflected in the term premium of 20.5 basis points—we suggest that a greater level of certainty (of the next move being a rise in rates) would have translated into a greater term premium, as we saw with USD in May 2005. Notice also how the rest of the curve is very flay after that—the June 2008 forward rate is 5.14%, a difference of only 7 basis points from the rate implied by

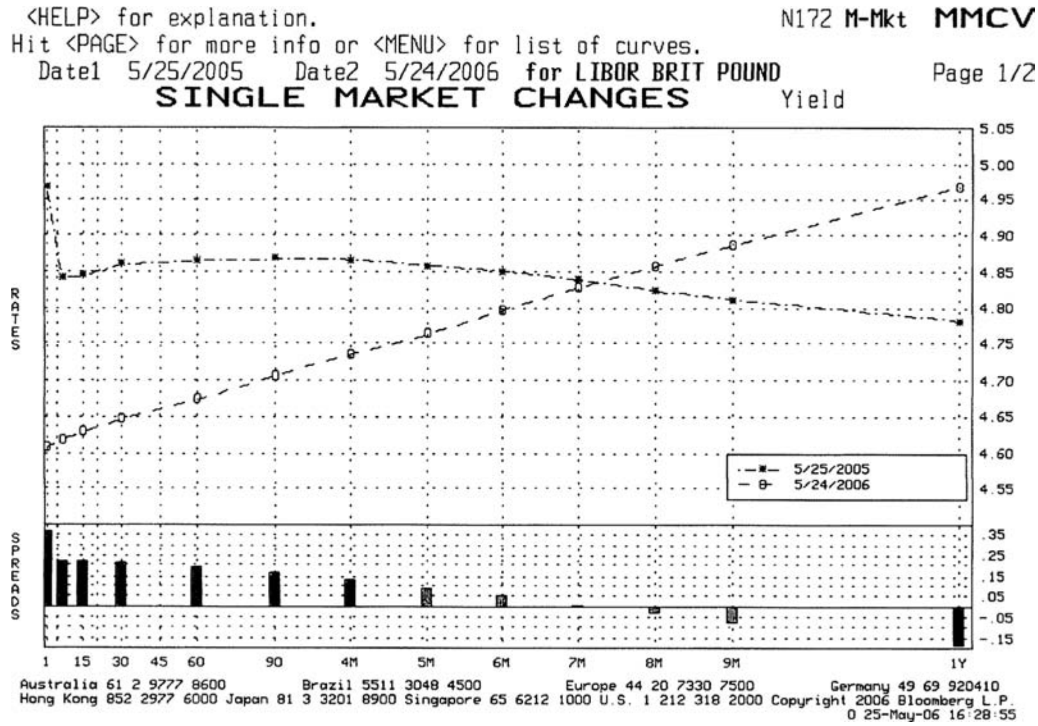


Figure 46.8 GBP Money Market Curves, 2005–2006
Source: © Bloomberg L.P. Reproduced with permission. All rights reserved. Visit www.bloomberg.com <http://www.bloomberg.com>

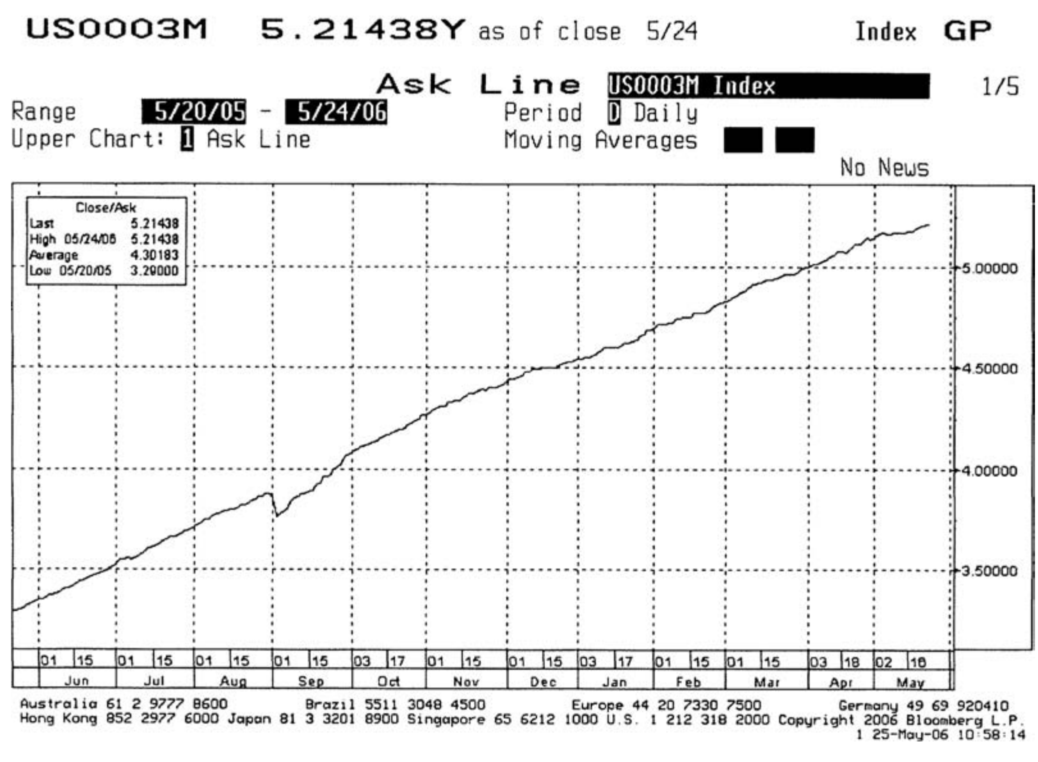


Figure 46.7 USD Three-Month LIBOR Rate for May 2005–May 2006
Source: © Bloomberg L.P. Reproduced with permission. All rights reserved. Visit www.bloomberg.com <http://www.bloomberg.com>

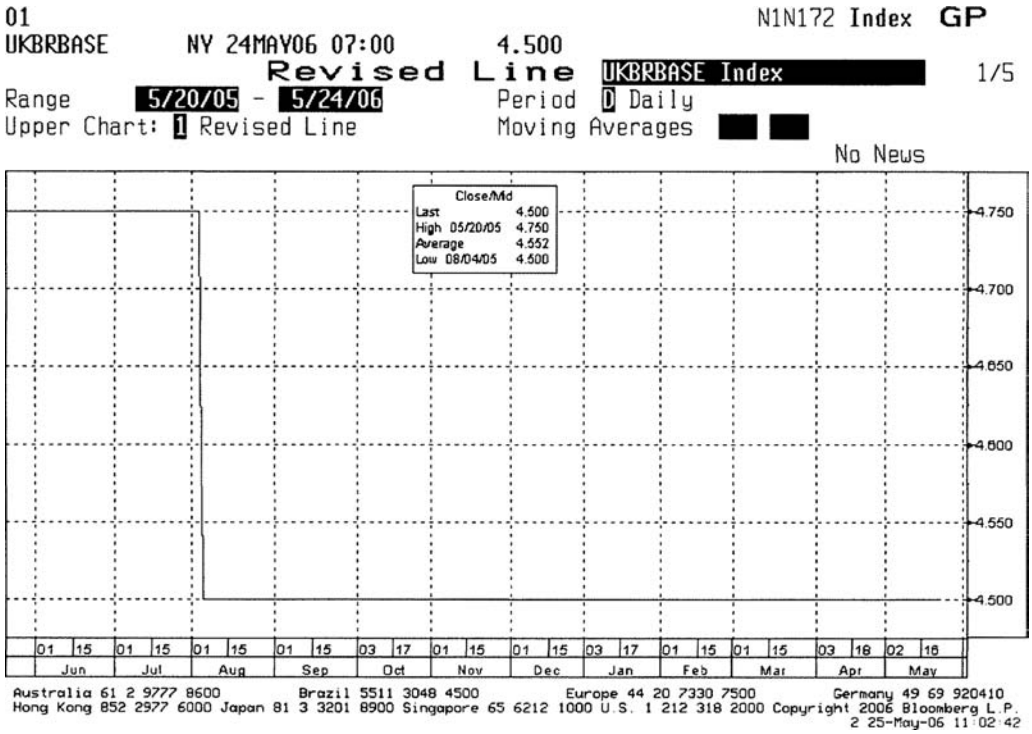
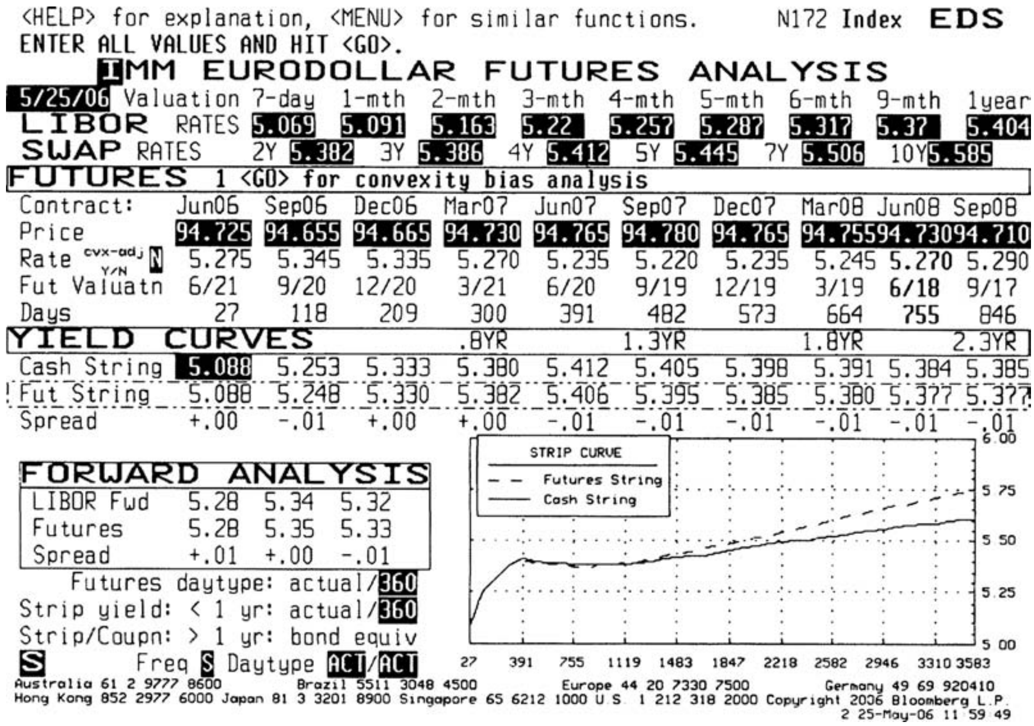


Figure 46.9 GBP Base Rate History, May 2005–May 2006
Source: © Bloomberg L.P. Reproduced with permission. All rights reserved. Visit www.bloomberg.com <http://www.bloomberg.com>

Figure 46.10 GBP Three-Month LIBOR History, May 2005–May 2006
Source: © Bloomberg L.P. Reproduced with permission. All rights reserved. Visit www.bloomberg.com <http://www.bloomberg.com>



<HELP> for explanation, <MENU> for similar functions. N172 Index **EDS**
ENTER ALL VALUES AND HIT <GO>.

IMM EURODOLLAR FUTURES ANALYSIS

5/25/06 Valuation	7-day	1-mth	2-mth	3-mth	4-mth	5-mth	6-mth	9-mth	1year
LIBOR RATES	5.069	5.091	5.163	5.22	5.257	5.287	5.317	5.37	5.404

SWAP RATES 2Y 5.382 3Y 5.386 4Y 5.412 5Y 5.445 7Y 5.506 10Y 5.585

FUTURES 1 <GO> for convexity bias analysis

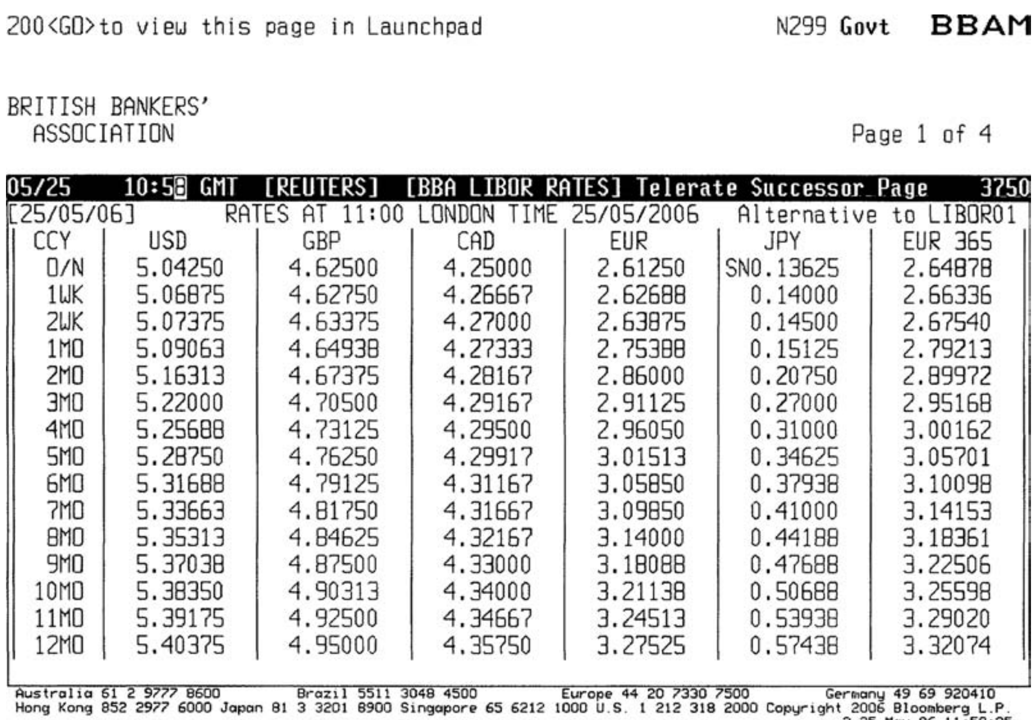
Contract:	Jun06	Sep06	Dec06	Mar07	Jun07	Sep07	Dec07	Mar08	Jun08	Sep08
Price	94.725	94.655	94.665	94.730	94.765	94.780	94.765	94.755	94.730	94.710
Rate cvx-adj Y/N N	5.275	5.345	5.335	5.270	5.235	5.220	5.235	5.245	5.270	5.290
Fut Valuatn	6/21	9/20	12/20	3/21	6/20	9/19	12/19	3/19	6/18	9/17
Days	27	118	209	300	391	482	573	664	755	846
YIELD CURVES				.8YR		1.3YR		1.8YR		2.3YR
Cash String	5.088	5.253	5.333	5.380	5.412	5.405	5.398	5.391	5.384	5.385
Fut String	5.088	5.248	5.330	5.382	5.406	5.395	5.385	5.380	5.377	5.377
Spread	+.00	-.01	+.00	+.00	-.01	-.01	-.01	-.01	-.01	-.01

FORWARD ANALYSIS

LIBOR Fwd	5.28	5.34	5.32
Futures	5.28	5.35	5.33
Spread	+.01	+.00	-.01

Futures daytype: actual/360
Strip yield: < 1 yr: actual/360
Strip/Coupn: > 1 yr: bond equiv
S Freq S Daytype ACT/ACT

Figure 46.11 Eurodollar Futures Curve on May 25, 2006



200<GO>to view this page in Launchpad N299 Govt **BBAM**

BRITISH BANKERS' ASSOCIATION

Page 1 of 4

05/25 10:58 GMT [REUTERS] [BBA LIBOR RATES] Telerate Successor Page 3750
[25/05/06] RATES AT 11:00 LONDON TIME 25/05/2006 Alternative to LIBOR01

CCY	USD	GBP	CAD	EUR	JPY	EUR 365
O/N	5.04250	4.62500	4.25000	2.61250	SNO.13625	2.64878
1WK	5.06875	4.62750	4.26667	2.62688	0.14000	2.66336
2WK	5.07375	4.63375	4.27000	2.63875	0.14500	2.67540
1MO	5.09063	4.64938	4.27333	2.75388	0.15125	2.79213
2MO	5.16313	4.67375	4.28167	2.86000	0.20750	2.89972
3MO	5.22000	4.70500	4.29167	2.91125	0.27000	2.95168
4MO	5.25688	4.73125	4.29500	2.96050	0.31000	3.00162
5MO	5.28750	4.76250	4.29917	3.01513	0.34625	3.05701
6MO	5.31688	4.79125	4.31167	3.05850	0.37938	3.10098
7MO	5.33663	4.81750	4.31667	3.09850	0.41000	3.14153
8MO	5.35313	4.84625	4.32167	3.14000	0.44188	3.18361
9MO	5.37038	4.87500	4.33000	3.18088	0.47688	3.22506
10MO	5.38350	4.90313	4.34000	3.21138	0.50688	3.25598
11MO	5.39175	4.92500	4.34667	3.24513	0.53938	3.29020
12MO	5.40375	4.95000	4.35750	3.27525	0.57438	3.32074

Figure 46.12 BBAM LIBOR Fixing on May 25, 2006

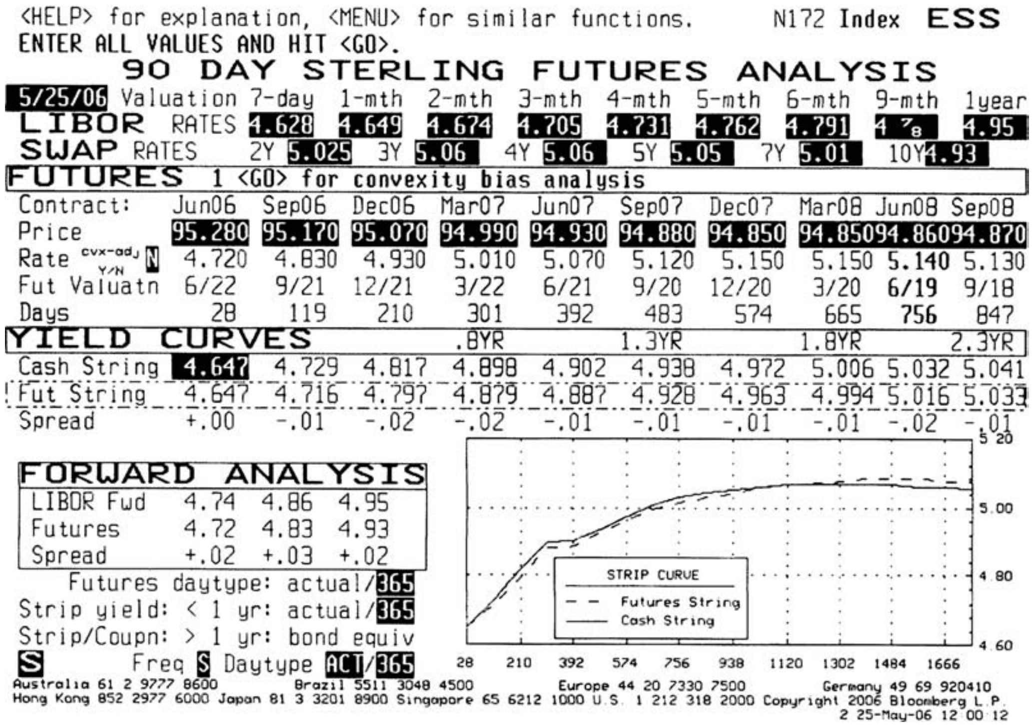


Figure 46.13 Short Sterling Futures Curve, May 25, 2006
Source: © Bloomberg L.P. Reproduced with permission. All rights reserved. Visit www.bloomberg.com <http://www.bloomberg.com>

the June 2007 contract. This is not really meaningful since rate changes are usually effected in 25-basis-point clips. In fact, it reflects the lack of market consensus on the timing and magnitude of the next base rate move.

In a stable interest rate environment, then, we would suggest that the 90-day term premium would be between 15 and 20 basis points. This can be considered fair value. Considering the forward rates implied in Figures 46.11 and 46.13, if we had a firm view in either direction, we would trade the contracts to reflect this. If we expect the base rate to be different at the time of the futures contract expiry, in our analysis we should logically build in a term premium to reflect this expected base rate, together with any further rate move expectations that we ourselves have.

OBSERVATION OF MACRO-LEVEL ECONOMIC AND POLITICAL FACTORS ON SWAP SPREADS

Banks are an important part of the global economic system, if not the most important part. It goes without saying, therefore, that efficient management of a bank's assets and liabilities feeds directly into overall economic development and national well-being. The treasury or ALM desk of a bank must perforce have a keen understanding of macro-level economic factors and the overall geopolitical situation because this drives swap spreads and the term premium. It is worth considering the impact of these factors, in general terms, on spreads and the overall level of interest rates because the ALM desk will need to take them into account as part of its strategy. Also, geopolitical events often arrive unannounced. An ability to work effectively under the circumstances prevailing in such occurrences is crucial to efficient ALM.

Events that impact the financial markets at a macro level are often termed "market shocks" or "external geopolitical events." Such events invariably result in higher market volatility. The immediate impact of this is a market sell-off and a "flight to quality," which is when investors move out of higher-risk assets such as equities and emerging-market sovereign bonds and into risk-free assets such as U.S. Treasuries and U.K. gilts. This is an almost knee-jerk reaction as investors instantly become more risk averse.

Swap spreads, which we define as the spread between the fixed-rate paid on an interest-rate swap over the yield of the government bond of similar maturity, reflect the market perception about the general health of the economy and its future prospects, as well as the overall macro-level geopolitical situation. Because the swap curve is an indicator of interbank credit quality, the swap spread can be taken to be the market perception of the health and prospects of the interbank market specifically and the bank sector generally.

Speaking generally, swap spreads widen during periods of increased market volatility. By implication, a flight-to-quality should be reflected in a widening of the spread. This is expected because investors' new risk aversion manifests itself in lower government bond yields, arising from higher demand for government bonds. However, on

Table 46.1 Correlation between the USD 10-Year Swap Spread and the CBOE VIX Index and the 10-Year U.S. Treasury Yield and the CBOE VIX Index

Event	Correlation between VIX and 10-Year Swap Spread	Correlation between VIX and 10-Year U.S. Treasury Yield
Asian currency crisis (1997–1998)	0.71	−0.52
Long-Term Capital Management and Russian debt default (June–September 1998)	0.90	−0.78
9/11 to Afghan war (September 2001–March 2002)	−0.17	−0.67
Iraq War (March–May 2003)	0.54	−0.08
Ford and GM credit rating downgrade (March–May 2005)	0.38	−0.53

Source: HBOS. Reproduced with permission.

occasion, this analysis might be overly simplistic because other micro-level factors will still be in play and can be expected to influence market rates. How can we consider the interaction between government yields, swap rates, and possible influences on the swap spread?

The research team at HBOS (2006) produced a report that suggests a novel way for us to analyze this, and we summarize the findings of that study. We require an indicator of market volatility; one measure of this for the U.S. dollar market is the VIX index. The VIX index is produced by the Chicago Board Options Exchange (CBOE) and is a proxy measure of market volatility. It uses a weighted average of implied volatilities to calculate an estimate of future volatility. An increase in the level of the index indicates increased market volatility.

We illustrate the relationship between geopolitical events and the magnitude of the swap spread by looking at the correlation between the U.S. dollar 10-year swap spread and the VIX index. Table 46.1 shows—as expected—a positive correlation between the VIX index

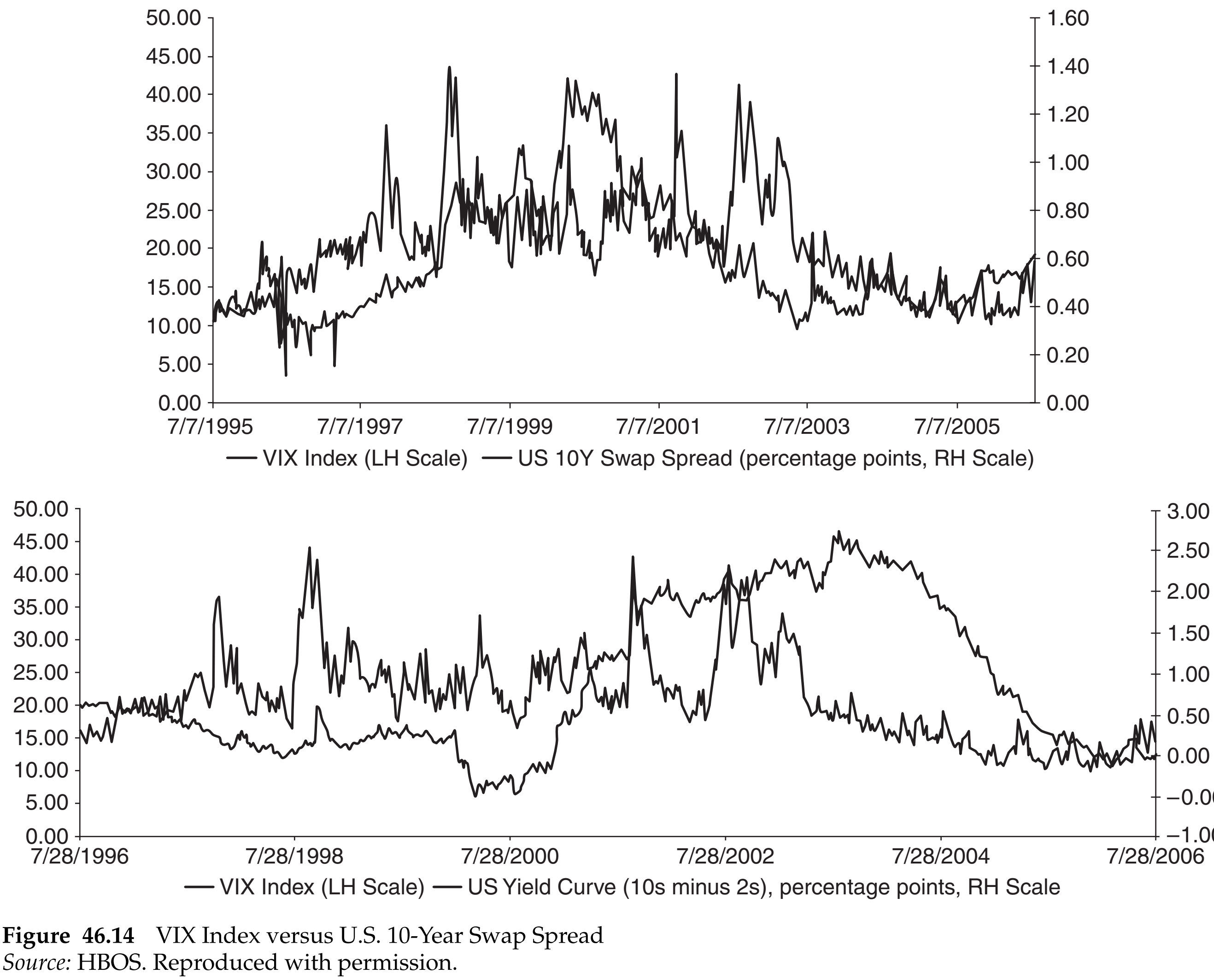


Figure 46.14 VIX Index versus U.S. 10-Year Swap Spread
Source: HBOS. Reproduced with permission.

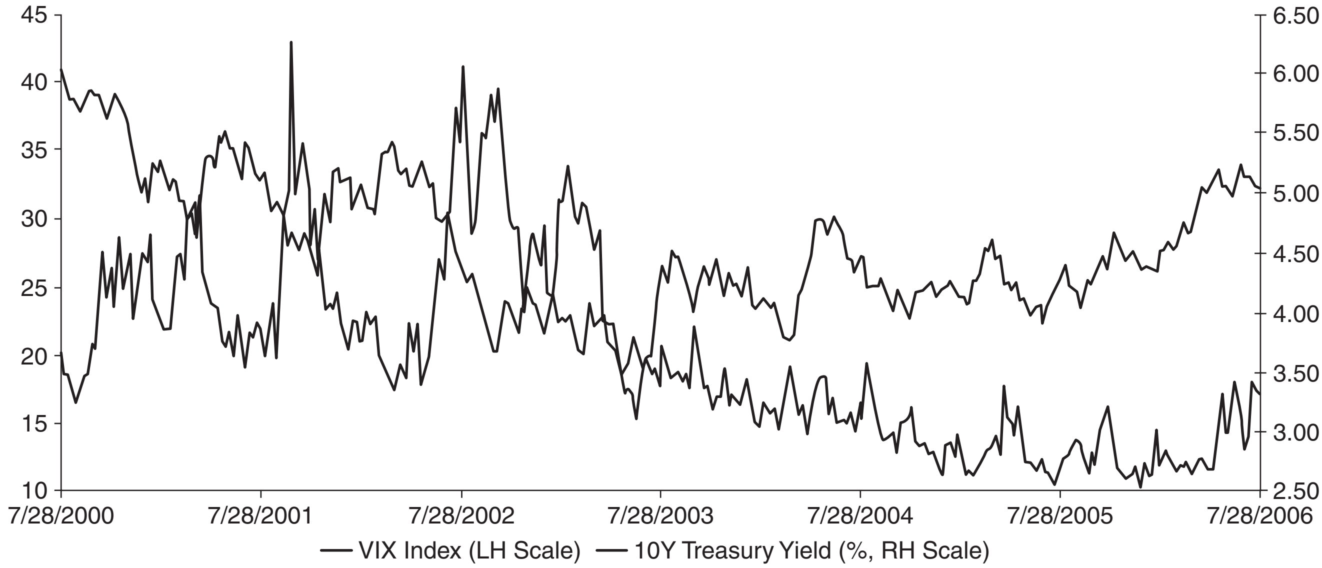


Figure 46.15 VIX Index versus U.S. 10-Year Treasury
Source: HBOS. Reproduced with permission.

and the swap spread during a period of both economic events and macro-level geopolitical events. For instance the period covers the 9/11 events as well as the Ford and GM credit rating downgrades of 2005. There is a notable exception for the period September 2001 to March 2002, when there is a negative correlation. This is our first indication that the relationship is not as simplistic as we might think. Although the geopolitical situation was negative, with the events of 9/11 leading to the U.S. war in Afghanistan, suggesting that swap spreads should widen, this was also a period of successive cuts in the U.S. base interest rate (the "Fed rate"). During this time, the swap rate fell by more than 100 basis points as the Fed rate was cut by 175 basis points. So, here, we observe that the impact of specific financial market factors was greater than macro-level geopolitical issues. Generally, though, we observe the strong positive correlation between the swap spread the volatility index.

Figure 46.14 is a chart of the spread to the level of the VIX index.

By the same analysis, we can expect a negative correlation between the U.S. Treasury yield and the VIX index level. This is generally borne out in Table 46.15. However, as with the case of the swap spread correlation, we see an occasion when other factors impact the correlation value. The low negative value for the period in 2003 leading up to and after the second Iraq war shows other factors influencing the Treasury yield. Miller and Chester (2006) of the Halifax Bank of Scotland (HBOS) report that the flight-to-quality had taken place before the war actually began and was fully priced-in to Treasury yields.

Figure 46.15 illustrates the lower government bond yields that are observed at times of higher market volatility.

The purpose of the foregoing has been to illustrate how the swap spread interacts with macro-level geopolitical factors. However, even during periods of high market tension, characterized by high levels of market volatility, the swap spread will respond also to more micro-level financial factors. Practitioners need to be aware of the nature of this interaction and allow for this in their ALM strategy, planning, and risk hedging.

SUMMARY

Interest rate swaps are important tools for use in bank asset-liability management, risk management, and interest rate risk hedging. The swap curve lies above the sovereign benchmark yield curve, this yield difference representing the risk premium of the interbank market compared to the sovereign market. The factors that drive the magnitude this premium are varied and include macro-level political and economic issues, level and slope of the yield curve, supply and demand, market volatility, and the level of government borrowing.

The shape of the short-term yield curve reflects a term premium. The level of this term premium varies with the current shape of the curve and market expectations of future interest rates. We can use the 90-day *futures curve* to infer market rate expectations, and use this to estimate what the term premium will be for future term borrowings.

REFERENCES

Burghardt, G. (2003). *The Eurodollar Futures and Options Handbook*. New York: McGraw-Hill.

Fabozzi, F. J., Mann, S. V., and Choudhry, M. (2002). *The Global Money Markets*. Hoboken, NJ: John Wiley & Sons.

Ludwig, N. (1993). *Understanding Interest Rate Swaps*. New York: McGraw Hill.

Miller, M., and Chester, A. (2006). Geo-politics returns to the limelight. *Economics Perspectives,* August 8, HBOS Treasury Services.

PART 4

Real Estate

CHAPTER 47

Real Estate Investment

SUSAN HUDSON-WILSON, CFA
Member, Boards of Hawkeye Partners, LLC, Property & Portfolio Research, Inc. and University of Vermont Endowment

Abstract: Real estate is a multidimensional investment asset class. Trading markets, geography, property types, leverage, and investment structures all define the dimensions. Each of the four quadrants of public debt, private debt, public equity, and private equity are characterized by different attributes and contribute to an investor's portfolio in differing and sometimes complimentary ways. Real estate has two special characteristics. One is that assets tend to hold some of their value even through egregious market cycles and, second, each property evidences both debtlike and equity-like investment behaviors. Despite certain difficulties with valuation and a lack of transparency international investment is an expanding field and derivatives are being explored. The relationship across the quadrants of the real estate investment universe have proven to provide ample room for creating diversified portfolios.

Keywords: apartment, appraisers, cash flow, closed-end funds, collateral, commercial mortgage-backed securities (CMBSs), conduit lenders, core, credit tenant, debt-equity hybrid, derivatives, diversification, endowments and foundations, funds of funds, family trusts, geography, hotels, indexes, industrial, inflation hedging, institutional real estate investment universe, leverage, National Counsel of Real Estate Investment Fiduciaries (NCREIF), net asset value (NAV), net lease, non-recourse, office, open-end commingled funds, opportunistic, option value, pension funds, persistence, private equity, publicly traded real estate equity market, quadrants, real estate investment trusts (REITs), real estate operating companies (REOCs), real estate risk, replacement cost, Resolution Trust Corporation (RTC), retail, risk management

Everyone knows about real estate! People own houses, stay in *hotels,* shop at malls and convenience stores, rent apartments, and store old household items in mini-storage facilities. People work in office buildings and in manufacturing plants. We travel through airports, drive over bridges, eat food from farms, and pack using boxes made of timber-generated cardboard. The wine we drink was grown on land that might be considered real estate. Real estate is a truly ubiquitous asset. If there is land involved, then the land and whatever is, or might be, on it might be considered to be part of the real estate asset class.

What is a real estate investor? Anyone can be an investor in this asset class. Anyone owning a home is an investor, even if that is not their primary motivation for owning the home. Like it or not, the price paid is not necessarily going to be the price received at the time a home is sold. Individuals dabble in real estate investment all the time, buying and operating two-unit apartment buildings, living on one floor and renting out the two other floors of a triple-decker in Philadelphia. Buying *real estate investment trusts (REITs)* in the public stock equity market, participating in a private offering that invests in mini-storage or land or vineyard development, or executing a tenant-in-common (TIC) exchange to mitigate gains taxes upon a sale of an asset—all are examples of the ways individuals can participate in the real estate investment market.

That said, for purposes of this chapter, the definition of real estate is pared down (not by much) to the ways in which institutional investors view and use real estate. The market and the types of investors are further defined below.

And, that said, it is surprising how much intuition is used by institutional investors as they develop and execute investment strategies. The "mathematics" of the real estate investment world is still under development, the tools of the trade are still unrefined, although significant strides have been made over the decade of the 1990s and the early 2000s. These tools will be described throughout the chapter.

Institutional real estate investors include public and corporate *pension funds,* academic and other *endowments and foundations,* and wealthy families and *family trusts,* both domestic and foreign. Institutional investors use real estate to achieve a variety of objectives including high absolute return, *risk management* for the overall portfolio, high income generation, *inflation hedging,* low-volatility performance, and to ensure that the portfolio accurately reflects the overall investment universe. Institutional investors are best suited to using real estate in their portfolios as there are few ways an individual can invest in what we might call institutional quality real estate. Individuals can buy REITs in the public market—this is a market that has achieved reasonable scale and presence only within the decade of the 1990s, and today there are more than a few REITs that are large enough to be included in the various stock *indexes* such as the S&P 500. REITs will be further developed later in this chapter in the section on public real estate equity.

In contrast, larger investors, such as institutions, can invest in individual properties directly or through one or more investment vehicles, REITs, private mortgage debt, and commercial *mortgage-backed securities (CMBSs).* There are domestic and international options available as well.

While many hold timber, farmland, infrastructure (such as toll roads, and airports), and some forms of residential investing to be a part of the institutional investment universe, they are excluded from the contents of this chapter. It is the premise of this author that just because something occupies land, it does not follow that therefore it is a part of the *institutional real estate investment universe.* Asset classes are defined as having constituent parts that share some behavioral characteristics that are generated by a common response to common drivers.

REAL ESTATE MARKET

The institutional real estate investment universe encompasses the four areas of *private equity,* public equity (REITs and other publicly traded stock instruments), private mortgage debt, and public debt or CMBS. Each will be developed below. The total dollar size of this universe is not known with perfect accuracy but is certainly in the area of \$4 trillion. In addition, there is another \$1 to \$2 trillion of non-investor-owned real estate. The largest segment of the investment universe is still private debt, although the public debt market is growing rapidly and is taking share from the private debt market as more loans are bundled, securitized, and sold by tranche. The equity markets combined consist of around 35% of the overall market capitalization but the public market share, after growing for 10 years, has now begun to lose ground as public to private transactions abound. It has never been entirely clear that the public market is a good and comfortable place in which to hold and value real estate assets.

Within the class real estate and within each of the so-called *quadrants* of the real estate investment universe is a wide range of types, locations, and management style and approach. To name a few of these many dimensions:

- Real estate is highly liquid (smaller holdings of REITs, for example) and highly illiquid (resort properties that are under development).
- Real estate can be found in small unit sizes (such as an individual stand-alone Starbucks store) and in very large packages (such as the office, retail, and residential complex called the Prudential Center in Boston.)
- Investors can control one building or a portfolio of buildings.

The list of property types included under the rubric "real estate" includes: *office, retail* (stand alone, community, strip, life style, mall), *apartment* (large and small complexes in suburban to dense urban locations, stick built to steel built), *industrial* buildings and industrial parks (old-fashioned 24-foot clear ceiling height to modern robot-driven logistics facilities operating 24 hours a day), and small town inns, big city conference hotels, and huge Hawaiian resorts with golf courses, spas, restaurants, and retail.

The *geography* of such assets ranges from U.S. urban, suburban, rural, resort, north, south, east, west, state, city, submarket, first tier, second tier, third tier city, international—Europe, both emerging and "old", and Asia plus Japan and Australia, the Middle East, and Latin America.

There are a myriad of investment structures within the real estate asset: direct investment, where an investor buys the building; indirect investment, where the investors' adviser buys and manages the building; commingled investing, where a pool of investors come together to buy a portfolio of buildings or an individual building; *open-end commingled funds*, where the investor can ask for his capital to be redeemed, and *closedend funds,* where the investment capital is essentially locked in for a specified period of time.

Real estate investors may either or both invest in debt instruments by lending money to others or can buy investments in mortgages issued by others to others. Alternatively, investors can borrow to *leverage* their real estate equity or debt investments.

Investors can manage their own leasing and building maintenance or can hire an adviser to do that, or they can contract with a local property manager and/or leasing agent to run the day-to-day operations of a building. In the case of a triple-net-leased building, the owner cedes all day-to-day and capital management issues to the tenant leasing the building.

KEY CHARACTERISTICS OF REAL ESTATE

Private equity real estate, the purest form of the asset and the base from which the other quadrants are generated, has two very important key characteristics. One is that real estate is a *debt-equity hybrid* and second is that it is exceedingly rare for a piece of real estate to permanently lose all of its value. Each is discussed below.

Debt-Equity Hybrid

The performance of each real estate quadrant is produced by a mix of equity-like and debt-like behaviors. Consider the classic case of a private real estate equity asset leased to a *credit tenant* with a very long-term, triple-net lease. The payments on that lease are analogous to the fixed payments usually associated with a bond, not with equity. The value of this lease to the investor fluctuates with the same factors that influence the value of a bond or a mortgage, such as interest rate movements, inflation, and the credit of the tenant. An opposing case is presented by an equity position in an empty, speculative, multitenant property. The value of that building is a function of the market forces of supply and demand for space, at that particular time and through time. As the building is more fully leased, it changes from a "pure" equity to a debt-equity hybrid, and perhaps—if it were to become fully leased to long-term tenants—becomes very debtlike. In related fashion, as the *net lease* on the building in the first example ages, the residual value of the property at lease-end becomes increasingly important, and finally the dominant, component of the building's value. Equity forces, such as space market conditions, urban economic health, tenant demand, interest rates, and the unique nature of the property, such as its location, history, visibility, and neighbors, increase their influence on the asset's value (see Booth, Cashdan, and Graff, 1989). So the "pure" equity play can become more fixed income–like and the "pure" debt play will ultimately become a pure equity play. There is a continuum of debt-equity behaviors in each asset and through the life cycle of each asset. This is a unique and important characteristic of equity real estate. This characteristic means that before one says they are invested in real estate, they need to be articulate about exactly where on the debt-equity spectrum their position lies. Later in the chapter we will see why this matters to the performance of a portfolio.

Holding Value

Even at the depths of the market in the late 1980s and early 1990s, real estate values were not, and were not even close to, zero. Values for office buildings in Texas may well have fallen by half. In Denver, brand new buildings were not leased at all and stood vacant for years waiting out the cycle. There was no income whatsoever accruing to the bankers who held the debt and were left holding the asset following foreclosure. However, the market cycle ultimately reverses itself (in most but not all cases) and the fact that there was, in the case of Denver, a building standing in a reasonable location, in the midst of an economy that had regained some vitality instantly accorded value to that asset. The reason is that the *replacement cost* of that new, vacant building is a very large number calculated in terms of both hard and soft construction costs, land value, and time to occupancy. So replacement cost, especially in times of construction cost inflation, preserves the value of many a building even through the toughest of market cycles. This rule does not hold in all cases—for example, a building in an old steel mill town when the steel mill has closed, a hotel in a resort area that has been abandoned by the guests for more easily accessed locations, or an apartment complex in a dying economy driven by a single employer that has lost the global competition for market share—all of these assets will lose value and may never regain what was lost. But, interestingly, these are the exceptions, not the rules. Think of major cities (which is where most of the real estate value is located) like Los Angeles, New York, and Miami. These cities' fortunes certainly ebb and flow, but they never go down and stay down. In almost any market condition, there is a rental rate that will clear the market and render buildings in these markets worth something to someone.

Compare those examples to the example of stock equity in the sectors of technology through the early 2000s and perhaps of biotechnology going forward. These investment options are only as strong as either their people or their ideas. There are large numbers of companies formerly based in Silicon Valley that do not exist at all today. However, the real estate in Silicon Valley had and kept at least minimal value all through the cycle and in fact is now gaining rapidly in value as a new tech cycle is upon us. The *cash flow* may well falter, a building may suffer from insufficient capital investment to maintain functionality, or a building may even become functionally obsolete, but still between the land and the physical structure, there is virtually always value to be reclaimed.

REAL ESTATE INVESTMENT

As is the case for other asset classes, investors are properly concerned with both the return of the investment and the riskiness of the investment. *Real estate risk* is generally measured as the volatility of the returns through time. Interestingly, the real use of the risk concept and math is still pretty primitive among leading real estate investors and managers. Plenty of "lip service" is paid to the notion that risk exists and should be considered in the investment decision-making process, but less real consideration is given than one would think or hope. The reason for this is that real estate investors are incredibly optimistic and are comfortable with active, complex problem solving. They do not see risk so much as they see solutions and opportunity. This is especially true as one gets closer and closer to the actual property itself. How would you measure the riskiness of a particular building? Such a measurement would be analogous to measuring the riskiness of a specific product line in the arsenal of Procter & Gamble offerings, but in real estate such measurement activity is deemed to be too mysterious, too fraught with nonquantifiable drivers, and too time consuming besides. Investment managers will talk about the market risk, the risk of uncertain capital investment requirements, and the risk of tenant credit. But they will, in the next breath, tell you why these issues are all under control. They will simulate 20 scenarios of superior and inferior performance outcomes and will select as the highest probability the one that gets the deal done.

This tendency in real estate recedes as one moves away from the specifics of the property and up to the level of the portfolio. The understanding of risk grows further when the behavior of debt investors is examined. Lenders are the most cautious—perhaps because they do not have hands-on control, at least not until post-foreclosure, when the value of the asset would have fallen below the loan balance.

Investors in real estate care a great deal about the time horizon of the investment, the faster the investment is launched and the capital is returned to the investor, in real estate as with all investing, the higher the internal rate of return will be. Investors also care because much real estate is fairly illiquid, although as the investor pool broadens and deepens the illiquidity improves, sometimes quite dramatically. Investors care about how the investment fits into their life cycle; this is true particularly for family offices and individuals and is less the case for pension funds and endowments, which are perpetuities.

Investment Characteristics of Each Quadrant

The Publicly Traded Real Estate Equity Market

The *publicly traded real estate equity market* is among the better defined of the four quadrants and encompasses around $300 billion of capitalized value, traded in deep public equity markets such as the New York Stock Exchange and the American Stock Exchange, and in global stock exchanges in Hong Kong, Australia, Tokyo, London, and Paris.

Most public real estate companies in the United States and in many other countries are structured as real estate investment trusts (REITs) or as *real estate operating companies (REOCs).* The REOC is a normally structured corporate entity that simply specializes in buying and operating real estate assets. The REOC is not entitled to the federally tax-exempt status of the REIT, but nor is it subject to the rules on distributions and asset sales to which REITs are held. A REIT, interestingly, and despite its name, is not even a trust. Rather, it is a tax election. Essentially, a REIT acts as a perpetual ownership vehicle of one or more buildings. In exchange for the exemption from all federal taxation at the REIT level (holders of REIT shares must pay federal taxes on dividends, just as do holders of non-REIT securities, and at the same rate), REITs are required to pay out at least 90% of all accrual-based accounting earnings (Block, 2002). Some REITs pay out more than required in order to defend their dividend levels and yields. In the past, there were considerable restrictions on a REIT's ability to sell buildings, but these rules have been relaxed, and REITs are allowed to run their portfolios without undue concern for impairing their tax-exempt status through sales activities.

REITs come in all property types and geographic locations with good coverage across the U.S. real estate investment universe. That being said, the distribution of the REIT universe across property types and locations differs somewhat from the distribution of the true real estate market universe. REITs have historically tended to own relatively more retail and apartment and relatively less office and industrial. The overweight in retail is caused by the fact that most mall operating companies are structured as REITs and many apartment companies of size are REITs as well. Thus, REITs as a group cannot be regarded as anything close to a replica or index of the overall real estate investable universe. REITs are what they are, allocated due to historical accident and not in accordance with any plan. In addition, since REITs only comprise less than 10% of the overall real estate investable universe, it would be an odd accident if at any time the REIT universe did perfectly replicate the larger real estate investment universe. So when investors take a position in a REIT index such as the Morgan Stanley RMS Index or the National Association of Real Estate Investment Trusts (NAREIT) Index, they should be clear that they have not acquired the larger real estate market. They have bought a smaller, idiosyncratic subselection of the larger market. Whether that "skewed" holding is a better or a worse portfolio of real estate exposures is an empirical question and the answer to the question can, and will, change through time.

REITs use moderate leverage at the entity level, introducing additional volatility to the cash flows (which are otherwise quite stable). Volatility is a consideration in REIT investing, although the level of leverage is generally very low as the analyst community has put consistent pressure on REIT managers to keep the leverage low. While leverage contributes to the creation of some minimal volatility in the cash flows, a more important source of volatility is introduced by the simple fact that REITs trade in the public equities markets along with all the other stocks. As a result, values may gyrate for reasons unrelated to cash flow changes. A small sector like the REIT sector, comprising around 3% of the overall capitalization of the U.S. stock equity market (measured against the Wilshire 5000), can be easily whipsawed by relatively large flows into and out of the real estate sector. Some of these flows are driven by investors' expectations for real estate, both positive and negative, but many of the flows are driven by investors' expectations for other sectors' relative performance, and the REITs get swept up in the rush.

Valuation in the public equity market is seemingly straightforward. The market tells you what the share price is. Now, of course, the market can be wrong, and as the real estate sector is pretty small, the market is frequently wrong, but there is opportunity in these misassesments. Investors think about whether a real estate security is over- or undervalued is by comparing the share price with the per-share *net asset value (NAV)*. NAV, however, is a difficult number to determine, as it is simply the estimated or appraised value of each building in the portfolio, less debt. Appraised values are better indications of value today than they were in the past, but they are never absolute. This is discussed further below. Also, some believe that REITs should be credited with franchise or enterprise value, above and beyond the simple sum of the values of the buildings in the REIT portfolio. Franchise value is even more difficult to assess, but it is commonly assessed for "regular" companies trading in the same exchanges. Analysts try to determine NAVs and then offer their views on whether REITs are trading at, above, or below their NAVs. When the analysts believe the REITs are above NAV, the REITs claim that the analysts' estimates are incorrect, and when the analysts think REITs are trading below NAV, the REITs are content to remain silent.

Privately Traded Real Estate Equity

For purposes of this chapter, privately traded real estate equity includes only the traditional categories of office, retail (recall, however, that much of the retail market is held in the public market), apartment, and warehouse. Our best estimates ("best" does not imply "excellent") are that the unleveraged value of the private real estate equity quadrant is near to $1 trillion, comprising approximately 16% of the value of the real estate capital market wheel. If one were to add in hotels, health care facilities such as assisted living, and mini-storage, the size of the sector would certainly expand, although not by a huge amount.

Since there is no investable index of private equity (a goal rendered even more difficult by virtue of the lack of a precise estimate of aggregate, let alone disaggregate, value), there is no potential investment with which to compare the measured allocations. In addition, private real estate equity is highly idiosyncratic and is generally accessed only in sizeable "units," and therefore very difficult to access by smaller institutional investors and individuals.

The traditional categories of private equity can be further divided into subcategories such as suburban office, strip center retail, mall retail, lifestyle center retail, pad retail, community center retail, townhouse apartments, high-rise apartments, major warehouse, secondary warehouse, manufacturing warehouse, and office showroom, as examples. The geographic distinctions are nearly as rich with primary, secondary, and tertiary urban areas; central business district; and suburban, exurban, and resort locations. Each property type and location and each combination thereof have different risks, returns, and drivers, as manifested in correlation matrices used by institutional investors to evaluate the risk characteristics of portfolios. For example, the *National Counsel of Real Estate Investment Fiduciaries (NCREIF)* Index shows a high correlation between office and industrial performance since 1978 (as measured by total return). Yet within those sectors, there is a much lower correlation between the Central Business District (CBD) office and industrial R&D space, and a similarly low correlation between CBD and suburban office space. Clearly, there are opportunities to increase returns and manage risk by wisely choosing investment markets, submarkets, property type, and subclass.

Layer onto this mélange the various stages of, and ways to participate in, private real estate equity. For example—from land banking, land development, presales, infrastructure development, merchant building, construction, leasing, operations, rehabilitation, repositioning, and sales—and there are lots of ways to play and lots of risk profiles to assume within this quadrant. Many more than is the case for public equity, where this variety of activity is far less in evidence, largely due to the pressure to maintain dividend levels.

There are three main categories of strategy within private equity—*core,* value added/enhanced core, and *opportunistic*—each quite different from the others and even presenting wide differences of activity within itself. The correlation between core and enhanced core strategies is very high, but the correlation between opportunistic strategies and core and enhanced core are considerably lower. As leverage increases, the differences between the strategies grow. This is interesting given that all of the strategies use the same basic ingredient—private real estate equity. Within each of the three primary categories of strategy are several ways to make (and, of course, sometimes lose) money. Some key strategies are:

- Land banking.
- Development.
- Getting land permitted and perhaps putting in the water, sewer, electrical, and the like infrastructure, then selling.
- Merchant building (building for an investor, for a fee, not as an equity participant).
- Presale (preselling, at an agreed upon price, then building and closing on the sale).
- Buying and operating a building for a sustained period of time.
- Buying and redeveloping a building, including retenanting, then holding or selling.
- Capital market "surfing"—anticipating where capital will seek investments—and buying ahead of the major flows, then selling quickly into them, having done little or nothing to enhance the operations of the building.
- Buying undermanaged buildings and bringing them to a market level of performance.

Valuation is a definite problem in the private equity quadrant (Fisher, 1998). Since the market is by definition private, transaction and carrying values do not have to be disclosed to anyone except the investors themselves and sometimes, depending on the terms of the adviser/investor agreement, not even to them. That being said, investors work hard to discover transaction values and a firm was founded to research and sell reasonably accurate transaction values. But this enterprise covers only properties that have been transacted; most properties are sold relatively infrequently, so the problem of interim valuation remains. Even when one is an "insider" to a private portfolio, at best the valuation (at least to the buyer and the seller) is "known" when the asset is acquired and sold. Assets that are held in portfolios are typically marked to market annually or even less frequently. *Appraisers* conduct research to assign an interim value to an asset, and while over time these valuation exercises have certainly improved, there is no guarantee that they are correct. Unfortunately, in the private market, unlike the case in the public market, there are infrequent chances to revisit what may be an incorrect valuation. Thus, until an asset is sold, the investor really does not know what its value is. For the closed-end portfolios of value-add and opportunistic strategies, interim valuation is much less of an issue as your capital is locked up until the fund is liquidated. In addition, assets in such portfolios are even more difficult to value given their complexity and risk profiles.

Privately Traded Real Estate Debt

Privately traded real estate debt includes commercial and multifamily mortgages. This sector comprises the largest portion of the real estate capital wheel at nearly $2 trillion, although it is rapidly giving ground to the public real estate debt quadrant, as will be discussed in the next section. This sector is exactly as it would appear to be—composed of loans backed by real estate *collateral.* These loans are almost always *nonrecourse* to the borrower (a unique feature of real estate lending in which the lender can seek redress for a default only on the mortgage to the property and not to any other assets that may be held by the borrower), and so lenders are highly motivated to be sure that they understand the performance attributes of the collateral. Loans can be fixed rate or floating rate, or interest only, and can include various other features like cash-flow participation, shared appreciation, and so on. The only constraints are the borrowers' and lenders' imaginations. Again, these are privately negotiated transactions, so the rules of engagement are subject to the competitive environment and the needs of each party. As the public, securitized market has developed, some increase in standardization of loan terms and documents has begun to develop in order to facilitate the securitization of a lender's portfolio should they choose to use that market.

Traditionally, private debt has been the purview of the insurance companies and the banks, but now the field has opened considerably, and anyone who has enough capital to get started can enter the field. Again, a significant driver of this shift is the advent of the public debt quadrant. Whereas once private-debt issuers had to hold those investments on their own balance sheets, now there is an active secondary market for individual mortgages and pools of mortgages. Even with this greater democratization of the commercial and multifamily mortgage world, the spreads over Treasuries for private mortgages historically have been wider than is the case for comparable risk corporate bonds and private non–real estate issues. This is a signal that the market may not quite understand how to price the risk of a private mortgage. Mortgages are underwritten on all property types in all geographic locations. Most individual mortgages are larger than most individual investors could invest in.

Publicly Traded Real Estate Debt

Publicly traded real estate debt was "invented" during, and as a government solution to, the severe distress of the real estate and economic cycle of the late 1980s. A CMBS is a security, backed by the cash flows from one or a pool of mortgages (see Esaki, de Beur, and Pearl, 2003). The security is "tranched" so that each holder of a piece of the security has a known piece of the hierarchy of rights to the cash flows and risks associated with the underlying collateral. These securities are modeled on the structure of securitized corporate debt. The market's development was spurred by the very large volume of defaulted real estate mortgages held on the books of most lending institutions when the real estate markets fell to earth in the late 1980s. With the help of the *Resolution Trust Corporation (RTC),* a quasi-governmental agency, lenders could sell their mortgages to the RTC, which then packaged them and resold them to securitizers who were then able to turn the pools into tranched securities and sell the rated securities into the institutional market. This large-scale government intervention saved the lending industry from certain demise, although not before lots of individual banks and savings and loans went out of business.

Those who bought the securities and the underlying collateral were rewarded when the real estate market cycle turned and the buildings were again able to generate positive cash flow over and above the interest rates of heavily renegotiated mortgages. Some have called this episode a major wealth transfer sponsored by the government, but for sure the financial system needed some kind of a fix. The advent of the CMBS market has truly democratized the real estate debt investment sector to the benefit of borrowers everywhere.

A new breed of *conduit lenders*—lenders who originate loans and, working with rating agencies, form securities and sell them, sometimes retaining a piece of the security on their books and sometimes just capturing an origination and securitization fee along the way—has emerged. As competition from new and traditional issuers of mortgages has increased, interest rates and spreads over corporates and Treasuries have come down. This shift in the cost of borrowing, compounded by the historical low levels of all interest rates, has made borrowing even more attractive to real estate investors.

This new market is steadily eating away at the former dominance of the private-debt market and now is within reach of $1 trillion, comprising nearly 20% of the capital wheel. In 1995, the CMBS market size was just $88.4 billion, only 5% of the real estate investment universe.

The market is growing because the idea makes sense. Real estate mortgages are being "deoligopolized," creating greater efficiency, better pricing (from the borrower's point of view) and transparency in the process. However, securitization creates instruments that are far more fractional and so sized to fit the investment parameters of more capital sources, including individual investors. Also, securitization allows for the pooling of risk across more loans and for segmenting the return-risk hierarchy, enabling each investor to participate in just the portion of the capital hierarchy with which they feel most comfortable. Whole mortgages require the investor to buy into the entire hierarchy. So a disaster in the lending business generated a brand new way to participate in real estate investing and created enormous new business opportunities out of what had been a very "clubby" part of the capital markets.

The array of strategies in the public debt market is, as is true for the whole loan market, somewhat narrow. However, there are some extremes nonetheless. Traditional fixed income investors have discovered that they can buy and hold the AAA tranche and receive results that slightly outperform the comparable corporate security with no additional risk (so far). A small number of more entrepreneurial real estate investors have chosen to hold the B and unrated pieces of the securities waiting for the spread compression that must come as the pace of delinquencies and defaults continues at record lows. In fact, considerable spread compression has already occurred, validating their faith in the market's learning curve.

The correlation between the AAA tranche and the B tranche is low, indicating that investors have interesting portfolio strategies to execute within the quadrant. Many have figured this out, and as the rest of the fixed income world catches on, you can expect to see spreads equalize across like corporate credits. Again, barring a disaster in the performance of the underlying collateral, this spread compression should continue until the relative wideness between the corporate and real estate sectors, across like credit quality, is competed away.

WHY REAL ESTATE?

Why do investors work to overcome the illiquidity, the crude valuations, and the "lumpiness" of real estate to hold it in their portfolios? There are five motivations (see Hudson-Wilson, Gordon, Fabozzi, and Anson, 2005):

1. Reduce the overall risk of an investment portfolio.
2. Realize a high absolute return.
3. Hedge inflation.
4. Reflect the larger investment universe.
5. Deliver high and sustainable cash flows to the portfolio.

Reduce Aggregate Portfolio Risk

It is well established that adding real estate to a portfolio of stock equity, fixed income, and other key assets both domestic and international causes the return of the portfolio to be, as usual, the simple weighted return of the components of the portfolio and the risk to be less than the simple weighted average of the components' risk (See Georgiev, Gupta, and Kunkel, 2003; and Markowitz, 1952). However this benefit is not evenly distributed across the parts of the real estate capital markets. And the benefit is not evenly distributed across each of the component parts of each of the quadrants. Thus, while real estate is an obvious risk mitigator and there is increasing awareness of this on the part of institutional investors, care must be taken when applying the lesson to the actual portfolio.

Realize a High Absolute Rate of Return

When real estate is viewed in the aggregate (using an index of all four quadrants, capitalization weighted through time), it is clear that the promise of using real estate as a tool to achieve absolute outperformance is not achieved: Real estate does not outperform stock equity and fixed income over very long time periods. However, and interestingly, aggregate real estate does outperform both stock equity and fixed income on a risk adjusted basis. So an investor needs to be clear about what his objective really is—absolute return without consideration of risk or absolute return subject to consideration of risk?

When real estate is decomposed into its four major parts and then further decomposed into opportunistic private equity and some value added strategies the message can change (see Hahn, Geltner, and Gerardo-Lietz, 2005). Real estate, especially opportunistic strategies, can be used to generate high absolute returns to the investor. However, manager selection becomes crucial to the successful execution of opportunistic strategies. There is wide variation in performance across managers in this arena. Interestingly, there is considerable *persistence* in performance for both high-performing opportunistic managers and low-performing managers. Unfortunately, as more vintages of funds are offered and executed by the high performers, their results tend to deteriorate. This suggests that while a manager may have a good idea or set of ideas at the start, either the market shifts against the good ideas or others in the industry catch on to and so compete away the good ideas. It also may be that the star employees working for an outperforming manager get hired by others or leave to start firms on their own, taking these good track records with them. At any rate, it is clear that it is hard to use real estate as an outperformer and really hard to achieve this over repeat funds and long stretches of time.

Hedge Inflation

Conventional wisdom maintains that real estate per se is an inflation hedge. Analysts have probed the question through time and, interestingly enough, the answer changes. (See Hartzell, Heckman, and Miles, 1987; Miles and Mahoney, 1997; and Huang and Hudson-Wilson, 2007.)

The bulk of the research has been conducted using various sources of private equity and of public securities returns. A broad and consistent conclusion is that the inflation-hedging capacity of publicly traded real estate securities is extremely limited. Another important conclusion is that the condition of the space market for real estate matters; in other words, if the vacancy rates are high and the capital markets are disinterested in holding real estate, the inflation transmission mechanisms are disabled. This is not surprising and should not deter one's interest in the issue or in the potential utility of real estate as a hedge.

If real estate in total and in terms of the four property types has or lacks inflation-hedging capability, then one's preference for portfolio structure would change, depending on one's view of future inflation. Indeed, if any of the property types acts as a successful hedge, then one would likely maintain a position in that asset despite one's views on future inflation, as it is too late to put a hedge on once an inflationary period is evident.

There are multiple places for inflation to contribute to the total return of a real estate asset—through the growth of rental and other revenue, through the growth of expenses and the issue of who bears those expenses, and through the capitalization rate, where income is transformed into value.

Most investors believe that the key to the success of the hedge is whether net income grows at the rate of inflation. If so, and if all else stays equal, the cap rate will reflect the impact of inflation, and so the value will reflect inflation. As inflation rises, the cap rate will fall and value will rise. However, what if the nominal growth of income comes short of fully reflecting inflation? Is this the only way for inflation to be reflected in performance?

No, if investors believe that real estate is an inflation hedge, then as inflation is either in evidence or is anticipated, more investors will seek access to real estate, and the risk premium will fall, reflecting the flows of capital. Thus, even if incomes do not exactly keep pace, the hedge may be effective. This is especially the case for assets in which construction costs are high and subject to inflation and where construction lead times are long. Excess supply is seen as a smaller risk to such assets, and so the flow of capital will inevitably and sensibly cause cap rates to decline.

Given these pieces of logic, it must be the case that there are important differences in inflation-hedging ability across the four major property types, and indeed there are. Office, which has long construction lead times and very expensive and heavily inflating construction costs, is a superhedge, with the hedge transmitted exclusively via the capital return. Similarly, apartment is the second best hedge among the real estate candidates. This is true of infill and urban multifamily assets, where costs and lead times are long, but it is also true because of the short lease durations allowing owners to adjust to inflation and inflationary expectations more quickly. Warehouse follows, although it is not nearly as capable as is office, and finally the retail sector evidences no hedging capability at all, through income or through value. This is especially interesting as not only has the key source of the hedge switched from income to value, but in past days, when retail leases contained percentage rent provisions, retail was the primary means by which to use real estate to hedge inflation. How times change.

Warehouse and, more acutely, retail are less interesting to investors seeking hedges, as they are characterized by short construction cycles, low relative construction cost (most suburban retail and typical warehouse construction is virtually the same), and longer leases—all negatives in terms of the transmission of inflation to performance.

An overall conclusion appropriate to any investor seeking to hedge inflation is to put a hedge in place before inflation is in evidence. Once the inflation is in evidence, investors are going to respond, adjusting the value of inflation-hedging assets and so reducing the future hedging capability of any specific asset.

Reflect the Larger Investment Universe

Given the ubiquitous nature of the real estate asset, in all of its various forms cumulating to trillions of dollars in the US alone, it is clear that a decision to *not* include real estate in a diversified mixed asset portfolio is a decision to tilt the portfolio *away* from a conceptual aggregate market index. Any large "off-market" bet would need to be rigorously defended. It is increasingly common to find real estate in well managed institutional portfolios as the attributes of the asset are better understood and as there are available more sensible and cost effective means of accessing the asset. The institutional world has come a long way since Meyer Melnikoff opened the Prudential Property Investment Separate Account in 1974. Today an investor can access real estate through REITs, separate accounts, direct investment that is self-managed, open-end commingled funds, closed-end commingled funds, private REITs, partnerships, and *funds of funds.* There are well over 500 vehicles (not counting the individual REITs) available on a variety of terms to institutional investors. Again, the field is not quite so open for individual investors, but that too may change with time.

Deliver High Cash Flows

Unlevered private equity real estate delivers net cash flows to investors that are typically three times greater than dividends from the stock market and typically 200 to 400 basis points greater than 10-year Treasury yields. Any investor with a need for cash flow to meet current liabilities will for sure have an interest in participating in the private equity real estate asset. Of course, the various quadrants are more and less robust from an income-generation perspective, but these differences are highly visible. REIT dividends exceed those for the overall stock equity market and public and private mortgage debt interest rates exceed the payout to comparably rated corporate debt.

SPECIAL TOPICS IN REAL ESTATE

Issue of International Investing

Real estate markets are truly ubiquitous, in the United States and around the world. Some institutional markets are more and less evolved, but there is investment real estate everywhere on the planet. The question is: Should an investor residing in any one location have an interest in gaining access to real estate investments, be they public or private or debt or equity, in other locations? Increasingly, the answer to this question is "yes." Despite the complexities associated with currencies, the widely differing degrees of market transparency and cultural differences in how business is transacted, real estate investors are following the lead created by their stock-equity and fixed income brethren and are actively pursuing global strategies.

In some cases, the need to move the investment program beyond one's borders is blatant—for example, the government of Singapore has long had an active global presence in real estate. Think about how much wealth the Government Investment Corporation of Singapore controls and how small the domestic real estate capital market is and so the reasoning is clear. The same can be said for the Dutch, the Swedes, the Aussies, and the Japanese. Each of these is a wealthy country with small (relative to their domestic wealth) local real estate investment universes. In addition, the smaller countries are more homogeneous in terms of investment drivers, and so there are fewer opportunities for portfolio *diversification* within the domestic market. Going global is an obvious choice.

How about the choice for more diverse economies like that of the United States? Research shows that for a U.S. real estate investor there are quite significant diversification benefits to be gained by adding Europe, Asia, and Canada to a base portfolio of U.S. office buildings located in major metropolitan areas (see Hastings and Nordby, 2007). This research also concluded that the existence of different currencies enhanced the diversification benefits even if greater performance volatility also accompanied the differences in currency. Currency variations can be at least partially hedged at some cost to the portfolio's return.

Investors seeking high absolute return might also heed the sirens' call to "go global." Emerging Asia, emerging Europe, and newly institutionalizing markets can produce outsized returns while in transition (of course, they can also move south pretty fast and by quite large magnitudes). Even in "Old Europe," where real estate had generally been held by owner/users, there are large opportunities created by the shift from owners/users to investors. In several countries (Japan, France, and Germany for example), new REIT markets are being created through national legislation, again reorganizing the manner in which property is held and making it available to investors both locally and internationally.

In the United States, it has long been conventional to finance commercial and multifamily properties and the private mortgage market is deep and well developed. Then the banking crisis of the 1980s led to the creation of the public CMBS market. In Europe, the CMBS market is developing in the absence of a financial crisis, simply because it truly is a good capital markets idea. This is an area of great investment promise as markets gain familiarity with a new investment structure and vehicle. When something is new, unfamiliarity breeds initially wide spreads.

The demands of diversification and the pursuit of higher returns have driven foreign investors to the United States and around the globe, and now U.S. investors are getting into the game. Through it all, one must remain vigilant with respect to the separation of reality, risk, and hype. A good example of a story well told to a susceptible few is that of emerging India, where the promises are huge but the reality is stark. It is often difficult to determine if the grass is really greener on the other side of the fence, or even if that green stuff is grass at all.

Issue of Leverage

While owners of residences routinely use leverage as a means to avail themselves of a home, it is less routinely used among institutional investors in real estate. This is because the use of leverage raises some important questions for institutional investors who are managing large complex portfolios and investing in multiple asset types (see Anson and Hudson-Wilson, 2003).

If lenders lend, then, of course, borrowers borrow. Collateralized leverage is the simple act of borrowing money based on the value and the security of an asset, in this case a property or a portfolio of properties. In the special case of real estate borrowings, the lender's recourse extends only to the property and not to the other assets that might be owned by the borrower. Because of this nonrecourse feature, the use of leverage is analogous to short selling the asset. The borrower takes the proceeds and reserves the option to "put" the asset to the lender at any time with no further penalties. As was mentioned earlier, while it is unusual for a property's value to go to zero and stay there, it is less uncommon for a borrower to be unable to support the mortgage payment through a rough market cycle. Moreover, sometimes a lender lends too much, based on an artificially inflated value or a "sweetheart" deal. Thus, the *option value* of the put is quite real.

An investor would choose to use leverage for one or more of the following reasons:

- To increase the total return on an asset.
- To hedge the downside risk of an asset's value.
- To enable a fixed amount of capital to be spread over a larger number of individual investments.
- To increase the yield and cash flow generated from a fixed pool of assets.
- To reduce one's exposure to an asset or a pool of assets as a way of reducing one's exposure to the asset class.
- To enhance the ability of real estate to act as a diversifier by way of the other assets in the mixed-asset portfolio (leverage exaggerates the low correlations between real estate and other standard mixed portfolio holdings).

The final reason warrants further exposition. As was discussed earlier in this chapter, real estate is a debt-equity hybrid. Thus, applying leverage to a building encumbered by one or more leases is a way of essentially "shorting out" the debtlike aspects of the asset's behavior. This then exaggerates the equity-like aspects of the asset's behavior. This has the additional effect of enhancing the diversification benefit of real estate in the context of a mixed-asset portfolio, essentially "pure" real estate equity behavior is what is left over.

Of course, leverage increases risk as measured by the volatility of the total return of the leveraged asset. There is no escaping this negative. However, one could apply leverage to assets that are, in other ways, less risky than average, thus while the application of leverage does raise volatility, it does not necessarily raise risk to an unacceptable level.

For institutional investors, which invest in multiple asset classes, there are benefits associated with using leverage on the real estate portfolio. First, the diversification benefits of the real estate portfolio itself can be enhanced with the use of leverage. Portfolio managers can gain significant allocation flexibility by using disparate amounts of leverage on different holdings to add to or subdue their risk profiles and to allow for making more investments. Second, the real estate asset produces enhanced diversification by way of stocks, bonds, and bills when it is leveraged.

However, the mixed-asset investor must also wrestle with a key question: Does it ever make logical sense for the same investor to be a borrower and a lender (by making fixed income investments) at the same time? (Certainly not, according to Benjamin Franklin). And, to compound matters, typically the rate at which the borrowing occurs will be greater than the like credit rate at which the lending occurs. So one would borrow more expensively than one would lend. (Remember the earlier discussion about credit spreads.) Don't these two approaches essentially neutralize one another while also incurring transaction costs? Of course, in the stock-equity world, this situation is routine, as it is the case that most securities are issued by companies that use leverage. The stock investor, however, does not have a choice, while the real estate investor (except the REIT investor) sometimes does have a choice.

One approach would be for the fixed income side of an institutional investor to lend to the real estate side. Unfortunately, right up until the moment where there were no problems, there would be no problems, then trouble could begin. Two sides of the same investor entity would be highly conflicted, and the situation would be awkward at best. This conflict would be exacerbated in the typical case where the portfolio manager for each asset class defends his own return and is not accountable for the overall fund level return.

Leverage should be thought of as way for the borrower to raise capital, not to use capital. If a fund is not already fully invested in highly productive investment activities, it is hard to see why leverage would make sense. Leverage can simply add to the dilemma of getting the money put to work.

Leverage, then, can be a useful portfolio management tool as long as the philosophical issues are put aside, the real estate allocation is fully invested, and the real estate and the fixed income portfolio managers are measured on their independent performance.

Emerging Issue of Derivatives

In the early 2000s, the market for commercial property *derivatives* has been emergent (see Esaki and Kotowski, 2007). Whether it will take and thrive is still a serious question, however.

In theory, there is ample room for a real estate derivatives product. Insurance companies might wish to hedge their exposure to the property and casualty lines, tenants might wish to hedge their exposure to lease renewal dates, and investors could less expensively take long or short positions in a large, lumpy, complex real estate market just by executing a swap. It all sounds good.

Several exchanges are now prepared to offer such derivatives, using several different underlying indexes of real estate market performance. The indexes being explored are:

NCREIF Property Index
Real Capital Analytics Index
REXX Real Estate Property Index
S&P Global Real Analytics
MIT Transactions Based Index (TBI)
IPD Index (Europe)

All of the indexes suffer from a common flaw. The indexes are in no way representative of the underlying real estate investment universe. The NCREIF Index (NPI) and the MIT Index (TBI) are based only on properties that are operated by the members of NCREIF, principally tax-exempt investment advisers. The NCREIF and IPD indexes are largely comprised of appraisal information on property values. The REXX Index is based on rental rates and values are inferred from those. The GRA Index is based on a small sample of transactions derived from public sources, and the RCA Index is based on a larger pool of transactions and is a better source of what is still a small subset of all properties in the U.S. real estate investment universe, those that were transacted within a particular quarter. In sum, none of these indexes are robust enough to warrant using as the base for a real trade involving real money, and none of them is even close to capturing the behavior of the true real estate investment universe. The basis risk associated with a swap constructed around any of these indexes would be massive. Thus, the risk to the counterparty would be significant, and so the spreads at which such transactions might be done are prohibitive.

Essentially, this is a good idea whose time has not yet come because the degree of transparency in the real estate markets is not yet sufficient to support it. Added to the transparency weakness is the reality that real estate returns are significantly serially correlated and exhibit low volatility. This means that the best predictor of next quarter's returns are last quarter's returns. This persistence makes the idea that there could be two interested parties with opposing views on the future of real estate market performance unlikely. In order to attempt to kick-start this industry investment, banks have been taking the other side of a handful of small index–based transactions and holding them on their own balance sheets. Good derivative markets require full transparency, low-basis risk with the bet one is trying to hedge, index volatility, and lots of reasonable uncertainty around the future directions of the market. The real estate market does not meet these criteria.

RELATIONSHIPS ACROSS THE QUADRANTS: IS REAL ESTATE REAL ESTATE?

It is clear from all that has been written in this chapter that real estate is a multidimensional asset class. In fact, this dimensionality begs the question of whether there is a real estate asset class at all.

At first blush, it makes sense to broaden the definition of real estate investment beyond a traditional private equity concept, because the key factors in real estate investment performance in the private debt and equity quadrants are reflected, to a greater or lesser degree, in investment performance in the two public debt and equity quadrants. Any real estate investment responds to a common set of influences, as well as to influences specific to each quadrant. The question is: How different is too different?

The performance of each of the assets within each quadrant is certainly tied, at some level, to the performance of the basic unit of analysis—the building. In private equity, this relationship is direct. In public equity, the only difference between public and private equity is the trading environment and the divisibility of the investment. Some very good analyses of these public/private equity relationships has been done. (See Giliberto and Mengdenl, 1996; Pagliari, Scherer, and Monopoli, 2003; and Gyourko, 2004.) All have found that when the public market pricing volatility is removed from the public market returns and the private market pricing is lagged forward to the same timing as the public market pricing, and both data sets are reweighted to reflect one another's weights in various property types, there is no meaningful difference between the performance of the public and private real estate markets.

Underlying a private mortgage is the collateral that is, again, the basic building block of the asset class—the actual building. If the building goes south, the value of the mortgage is impaired as well. Public debt or CMBS is just a market sensitive way to bundle up mortgages and structure them so that each investor can get what they need from the slice they invest in. Again, if the underlying collateral has problems, so will the issue of CMBS.

So we can see that the quadrants are related to one another at an intuitive level, but how does the data support or refute such a claim?

Empirically, the quadrants are quite different from one another. For all their theoretical sameness, the math supports a different and somewhat perplexing view. While we saw previously that the difference between public and private equity can be easily explained by the differences in each market's respective valuation systems, the impact of that difference is huge. The two sectors are actually negatively correlated. If you are an investor, you have to live with the reality of mark-to-market valuation, and so have two distinct assets in your portfolio when you have public and private equity. They are not substitutes in a portfolio context.

In addition, even across public and private debt, where you would expect pretty high correlations, we have only a 0.56. When you break down the correlations within CMBS and compare them with the whole loan, you begin to see where the breakdown occurs. The correlation between the whole loan and the AAA, AA, and A are all in the very high 0.9s. The correlation of the whole loan with the B tranche is a low 0.56. The volatility of the BB and the B tranches dominates the overall quadrant-to-quadrant comparison.

Across the public and private and debt and equity worlds, there is much room for portfolio management, with low correlations between private equity and public debt (0.42) and private debt and public equity (0.03).

So we have room for active portfolio management caused by low correlations, and we have wonderful intuitive stories rationalizing the synchronicity of quadrants within the unified asset we call real estate. But which is it? Is real estate an asset class? Or is each quadrant an asset class? Or does each quadrant belong to a different asset class, for example, public equity to the stock asset, and public debt to the fixed income asset? Is there no real estate asset class at all?

Much as a stock issued by a corporation is considered a different asset class than the bonds of that same corporation—though both investments have underlying ties to the performance of exactly the same corporation—so the empirical data shows that real estate equity investments can perform quite differently than real estate debt. While the underlying drivers are one and the same, investors benefit by understanding and capitalizing on the different behaviors of each real estate quadrant. And the question remains—Is real estate real estate?

SUMMARY

Real estate is clearly a multidimensional asset class. It is defined as consisting of four quadrants: public equity, private equity, public debt, and private debt. While each of the quadrants shares a common tie to the essential unit of real estate—the building—it is also true that the investment behavior of the quadrants is quite varied, creating within real estate diversification opportunities and even begging the question of whether real estate is an asset class or not. Real estate has two unique characteristics: Each building is a debt-equity hybrid, and it is very rare for a building's value to drop to zero and stay there. There is almost always recoverable value in real estate unlike in stock equity. Investors use real estate to reduce overall portfolio risk, add absolute return, obtain high cash flows, replicate the overall investment universe, and hedge inflation. There is growing interest in cross-border investing, despite the difficulties and sometimes absence of transparency. Leverage is increasingly a part of the real estate investment strategy, and there is a nascent derivatives market emerging.

REFERENCES

Anson, M., and Hudson-Wilson, S. (2003). Should one use leverage in a private equity real estate portfolio? *Journal of Portfolio Management*, Special Real Estate Issue, September: 54–61.

Block, R. (2002). *Investing in REITS: Real Estate Investment Trusts*, 2nd edition. Princeton: Bloomberg Press.

Booth, D., Cashdan, D. Jr., and Graff, R. (1989). Real estate: A hybrid of debt and equity. *Real Estate Review* 19, Spring: 54–62.

Esaki, H., de Beur, M., and Pearl, M. (2003). Introduction. *Transforming Real Estate Finance: A CMBS Primer*, 3rd edition. New York: Morgan Stanley.

Esaki, H., and Kotowski J. (2007). Commercial property return indexes: Choosing from the menu. *Morgan Stanley Fixed Income Research*, May 3.

Fisher, J. (1998). Appraised value versus sales price. *PPR Real Estate Portfolio Strategist* 2, 4.

Georgiev, G., Gupta, B., and Kunkel, T. (2003). Benefits of real estate investment. *Journal of Portfolio Management*, Special Real Estate Issue, September: 28–34.

Giliberto, M., and Mengden, A. (1996). REITs and real estate: Two markets reexamined. *Real Estate Finance* 13, 1: 56–60.

Gordon, J. (1994). The real estate capital markets matrix: A paradigm approach. *Real Estate Finance* 11, 3 (Fall): 7–15.

Greer, R. (2006). *The Handbook of Inflation Hedging Investments*. New York: McGraw Hill.

Gyourko, J. (2004). Real estate returns in public and private markets. *Wharton Real Estate Review* 8, 1: 34–47.

Hahn, T., Geltner, D., and Gerardo-Lietz, N. (2005). Real estate opportunity funds. *Journal of Portfolio Management*, Special Real Estate Issue, September: 143–153.

Hartzell, D., Hekman, J., and Miles, M. (1987). Real estate returns and inflation. *Journal of the American Real Estate and Urban Economics Association* 15, 1: 617–637.

Hastings, A., and Nordby, H. (2007). Benefits of global diversification on a real estate portfolio. *Journal of Portfolio Management*, Special Real Estate Issue, September: 63–73.

Huang, H., Hudson-Wilson, S., and White, S. (2007). Private commercial equity real estate returns and inflation. *Journal of Portfolio Management*, Special Real Estate Issue, September: 63–73.

Hudson-Wilson, S., Gordon, J., Fabozzi, F.J., and Anson, M. (2005). Why real estate? *Journal of Portfolio Management*, Special Real Estate Issue, September: 12–22.

Hudson-Wilson, S., and Guenther, D. (1995). The four quadrants: Diversification benefits for investors in real estate—A second look. *Real Estate Finance* 12, 2: 82–99.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance* 7, 1: 77–91.

Miles, M., and Mahoney, J. (1997). Is commercial real estate and inflation hedge? *Real Estate Finance* 13, 4: 31–45.

Pagliari, J., Scherer, K., and Monopoli, R. (2003). Public versus private real estate equities. *Journal of Portfolio Management* 29, 5: 101–111.

CHAPTER 48

Investing in Commercial Real Estate for Individual Investors

G. TIMOTHY HAIGHT, DBA
President, Menlo College and Chair of the Board, Board of Commonwealth Business Bank (Los Angeles)

DANIEL D. SINGER, PhD
Professor of Finance, Towson University

Abstract: Commercial real estate is seen as providing an attractive alternative to investing strictly in stocks and bonds. Even though commercial real estate values can be volatile their addition to an investor's portfolio of assets will reduce risk without necessarily compromising return. In fact, the advantages of personal control, financial leverage, tax shields, and an inflation hedge promise to increase the returns available to investors in commercial real estate. To be successful, commercial real estate investors must take consideration of the location, type of commercial real estate, and the stage of the local commercial real estate cycle. The greatest risk to investors of well-chosen commercial real estate is a lack of liquidity, but this problem can be minimized through the ability to successfully manage long-term compounding cash flows. The returns and risks to the commercial real estate investor are found to be sensitive to the organizational form in which the property is held. While the advantages and disadvantages of various organizational forms are considered, particular attention is paid to the limited liability company as being the vehicle of choice for most commercial real estate investors.

Keywords: location, diversification, financial leverage, operating leverage, limited liability partnership, limited liability company

Investors have a wide array of opportunities when seeking to build wealth and achieve financial independence. Each path will expose the investors to substantial rewards while also exposing them to unanticipated risks. Ultimately, one must choose the investment strategy that best suits his/her long-term goals, time horizon, temperament, and risk aversion. For many individuals, commercial real estate investing has proven to be a very rewarding choice. Over time, commercial real estate has performed remarkably well and is expected to continue to perform well over the next few decades. However, as in all investments, success in commercial real estate investing requires one to have a complete understanding of the market. The real estate market is characterized by cycles of overbuilding and price declines. Thus, timing becomes a crucial ingredient in the development of a successful wealth creation formula.

In this chapter we introduce some basic principles in real estate investing, explain the advantages and disadvantages of real estate as an investment, and the various types of business forms that can be used when investing in real estate. In the next chapter, the types of commercial real estate are explained.

IMPORTANCE OF LOCATION

Regardless of the type of commercial real estate selected, the leading factor affecting value and therefore investment performance is *location*. Real estate properties are differentiated from most other financial or real assets by their uniqueness. No two hotels are exactly alike, no two pieces of undeveloped land are alike, no two office buildings are alike, no two shopping centers are alike, and so on. Commercial real estate is not a commodity. As such one property may not be interchangeable with another. Each property is different because it is in a different physical location. This makes location one of the most important attributes of any piece of commercial real estate.

The first thing to understand about location is that location is not an absolute. There is no such thing as a generically "good" location (or a generically "bad" location.) The desirability of a particular site is relevant only in terms of its intended purpose. A property that is good for a shopping center is not necessarily good for an apartment, an office building or a factory. Assessing the value of a property always requires the strategic perspective that begins with a determination of the intended purpose for the property.

Only in that context are the actual physical attributes of that site relevant. Physical attributes of a site would include the current use of the property, its location with respect to traffic patterns, relevant zoning laws, the contour of the land, the attributes and uses of adjacent or neighboring parcels of land, the effective marketing area or impact zone of the property and trends in adjacent, neighborhood, local, or regional land use.

Another factor to consider in the valuation of commercial real estate is the impact of subjective perception. Certainly, a piece of property has an objective reality. However, that objective reality may not be as important as the subjective lens through which that property is viewed. An objective reality might describe 50 acres of rugged land surrounding a dismal swamp located 20 miles from the nearest urban area. A subjective perspective might be to consider land as a nature preserve, featuring select executive home sites surrounding ecologically important wetlands, which provide protection for a living environmental laboratory. The objective reality might be a rundown hotel adjacent to a metropolitan central business district whose desirability is threatened by crime in the neighborhood. The subjective perspective might be that the (refurbished) hotel could become a badly needed retirement community for area residents that is distinguished by its access to urban amenities and its significant architectural and historic significance. An investment in such a property could be thought of as a beacon of successful urban renewal that could revitalize the neighborhood. It is all in the perspective. A lot of highly successful commercial real estate development occurs because someone is able to think "outside the box."

IMPORTANCE OF DIVERSIFICATION

A key to successful investing, in general, is *diversification*. Specifically, diversification has that wonderful property of lowering risk without necessarily lowering gain (and often raising gain). We recommend that most investors should be diversified into real estate. The inclusion of commercial real estate into an equities portfolio may enhance the overall performance and lead to risk reduction. We further argue for diversification within the commercial real estate sector for the same reasons. To be sure, commercial real estate entails a range of investment choices from apartments, hotels and motels, office buildings, manufacturing facilities, and many more alternatives.

SPECIFIC ADVANTAGES TO INVESTING IN COMMERCIAL REAL ESTATE

Financial Leverage

"Give me where to stand, and I will move the earth," said Archimedes, referring to the notion that with a long enough lever he could move the earth itself. The power of leverage is that great. This is as true in finance as it is in physics. *Financial leverage* is simply the extent to which debt is used to finance real estate. For example, let us assume that an individual purchases an apartment building for $1 million. Further assume that the owner may put down as little as 25% of the purchase price and borrow the rest ($250,000 equity and a $750,000 mortgage). Now, let us assume that the apartment complex rises in value to $1,100,000. This results in a gain of 10%. By employing leverage, the owner experiences a gain of 40%. This is due to his $250,000 equity investment growing to $350,000. Leverage makes the investor's money work harder.

Leverage is not unique to real estate. Stockbrokers typically offer "margin" financing on stocks bought through their brokerage. However, more leverage is generally available for real estate investment because, while the commercial real estate market certainly has its ups and downs, it has nothing like the volatility of the stock market. Lenders feel more secure about their ability to recover their obligations when the value of those obligations is secured by a mortgage to real property whose value stays relatively constant.

Successful real estate investors optimize (not maximize) their leverage. The general rule is "Borrow to buy, sell for cash." More leverage can make a good investment a great investment. Wise real estate investors generally look for those properties that provide the most financing. To optimize leverage, many investors have a specific strategy that they use in identifying investment opportunities. This involves acquisition strategies that minimize the cash necessary to get into a project and divestiture strategies that look to all cash exits. Such strategies would include minimizing the down payment, borrowing the down payment, extending the life of the loan, and borrowing interest only with a balloon payment for the principal.

The reason investors want to optimize leverage, rather than maximize it, is that increased leverage brings about increased risk. In this case the additional risk comes from the fixed obligations to pay interest (and perhaps principal). Real estate investing always involves juxtaposing an uncertain cash flow coming in against a certain cash flow that must be paid out. Where this cash flow coming in is used to fund the cash flow going out (as is usually the case), this raises the possibility that the funds that were supposed to come in do not. This then puts the highly leveraged investor in a hard place. Money can fail to come in because the lessee is unable to pay, an argument with the lessee goes to court (the legal process is unbelievably slow and typically works to the disadvantage of the creditor), or the lessee, for some other reason, does not want to pay. Compelling such a person to pay is typically a long and arduous process, and while this process goes on, no money is coming in. Thus, how much leverage to use is ultimately a decision the investor makes based upon his or her preferred trade-off between risk and return.

Operating Leverage

Operating leverage is a characteristic commonly found in real estate properties due to its high proportion of fixed cost to total costs. This characteristic can be described in terms of the relationship between sales volume and profitability of a piece of property. Commercial real estate generally has a large degree of operating leverage due to its fixed costs. When fixed costs are high relative to variable costs, small increases in sales will generate large increases in profits. The other side of the coin is that large fixed costs require a substantial volume of sales to break even.

The presence of such operating leverage means that when the revenues are large, the project is wildly successful, but if the revenue is not there, disaster looms. The point about operating leverage is that very small differences in sales can make for very large differences in profits. This makes predicting the failure or success of a real estate project more difficult.

Operating leverage translates into business risk. Even where the real estate investor intends to take a very passive role in a development as a lessor, he or she is still effectively a partner with the lessee. Where the lessee is successful, the course of the lease will run successfully and both parties will be happy. Where the lessee is unsuccessful, the course of the lease will be troubled and both parties will be unhappy.

Inflation Resistance

Real estate values tend to rise with inflation. In fact, much real estate often rises faster than inflation because it is in relative limited supply compared to other consumer goods and services. Because real estate supply tends to be inelastic (insensitive to prices), as demand increases prices will rise faster in this sector.

Of course, a word of caution is necessary. Not all real estate rises in lockstep with inflation. There are variations in the price of real estate between regions, within regions, within states, within cities, and even within neighborhoods. Much depends on location and the demand for property at that location. Great care must be exercised in the selection of specific commercial real estate opportunities (location, location, location.)

Tax Advantages

Real estate ownership is encouraged by the tax system. Two important advantages come into play here. The first is interest costs. The second has to do with the concept of depreciation. Both of these factors combine to make real estate investing very attractive.

Interest costs can be fully tax deductible for any commercial real estate investment. This means the cost of funds is reduced by your marginal tax rate. As an owner, if you finance real estate at 8% and you are in the 40% tax bracket, your real cost of financing will be $8\% \times (1 - 0.4) = 4.8\%$.

The second important tax advantage to owning real estate is the ability to depreciate any property (the buildings, not the land) being rented. Depreciation is a legitimate (noncash) deduction used to offset revenue that would otherwise be subject to taxes. This means you can show a loss on your real estate investment, and, depending on how the deal is structured, use that loss to reduce your personal income, and thus lower your taxes. Anything to do with taxes tends to be a bit tricky and depreciation is no exception. Real estate rental is considered a passive activity and losses from a passive activity can only be used to offset passive income (not wages and salaries). However, if an individual actively participates in managing the rental property (as evidenced by selecting tenants, collecting rents, visiting the property, and doing maintenance—all of which are tax deductible in themselves), then the individual may deduct up to \$25,000 from earned income, provided he or she does not have adjusted gross income in excess of \$100,000 when the amount of loss that can be deducted is phased down to where adjusted gross income reaches $150,000 and no loss at all may be applied to earned income. There are a number of other constraints here having to do with marital status and the like. There is also something called an alternative minimum tax (AMT) to consider. An investor needs to consult with a tax professional to see how he or she may be impacted by the tax code. If an investor can write off $25,000 of paper losses due to depreciation and is in the 40% tax bracket, then he or she will receive a tax saving-a bottom line-of $10,000 in real dollars.

Investing in Real Estate Is Like Owning Your Own Business

Many individuals want to gain more "control" over their lives. The regimen of working for someone else, taking orders, and being subject to an array of arbitrary rules may feel stultifying. It is not uncommon for such individuals to want to "start their own business" to gain more control over their lives. For many people, this may not be a practical alternative. However, there may be another path to financial independence. Commercial real estate is an activity you control entirely. You find the opportunities, arrange the financing, bring all the elements together, and create something where there was nothing before. An individual can enter this business starting small and staying small, with the real estate investing being a profitable hobby. As an alternative, an investor can start small and over time, with a few good moves, grow his or her business into a high-paying full-time job.

Debt in an Inflationary World Is Good

Commercial real estate investors are debtors. They borrow money now to pay it back later. In an inflationary environment this confers a tremendous advantage to the buyer. In theory, interest rates adjust for the level of inflation by adding an inflation premium to the real rate of interest. In the real world, this adjustment process appears slow and uncertain. There have been a number of times within the past two decades where the rate of inflation exceeded the nominal rate of interest. Monetary history suggests a pattern in the world of modern finance where debtors have benefited from borrowing more valuable dollars and paying back with less valuable dollars.

The value of a dollar (or any unit of currency) is ultimately determined by what it will buy. What it will buy is determined by the price level of goods and services that, in turn, is determined by the demand for and supply of those goods and services. While government statistics show little inflation in the first few years of this decade, these indices do not necessarily reflect the buying pattern of real estate investors. It may be argued that broad-based indices (such as the Consumer Price Index), which rely on fixed market baskets of goods and services really understate the true level of inflation relevant to business decision makers.

There are a number of possible causes of inflation. One of the most common causes of inflation can result from the money supply increasing as a result of increasing government debt. Government debt increases because politicians basically find that, when they vote for benefits for people, they get congratulated for doing a good job by those people affected. When they vote for more taxes, they generally get voted out of office. Therefore, politicians tend to spend more without generating the needed tax revenues. The only way that can be done is to create more debt. What is the future for inflation in the United States? The effects of inflation are so powerful and pervasive that economists see inflation as a primary factor in redistributing wealth in our society. If inflation is inevitable, the real question is which side of this transfer will you be on?

Compounding Cash Flows

A hallmark of commercial real estate investment is that such investments yield compounding cash flows. Taking advantage of this requires a fairly long-term horizon, but that gets back to the tortoise and hare metaphor. An individual can go to Las Vegas, put down $10,000 on black at a casino roulette table, and double his or her money—or lose it all! The odds are against winning and there is a high degree of risk, but at least the issue is decided quickly. Or an individual can put $10,000 down on a well-located duplex apartment that will earn 21% annually over the next 15 years with very little risk. It takes a long time, but the $10,000 turns into $174,494! This is the miracle of compound interest. In finance, the tortoise not only finishes the race, the tortoise wins the race, too! Rabbits show a burst of speed that looks good for a short time, but they rarely finish the race and almost never win the race. Compounding cash flows are the surest way to wealth creation.

SPECIFIC DISADVANTAGES RELATING TO REAL ESTATE

Lack of Liquidity

Liquidity in finance refers to the ability of an asset to be exchanged for cash without loss of value. Publicly traded stocks have good liquidity. (That is the purpose of having "stock markets.") Commercial real estate investments typically do not. If you have invested in a small office building and the time has come to liquidate that investment, it cannot be done overnight, or, at least, it cannot be done overnight without great loss of value.

Of course, much will depend on prevailing supply and demand conditions. It is possible that an investor will decide to liquidate in a period of high demand and short supply. In that case, a sale may be arranged in a few weeks. If the decision is made to liquidate, when market conditions are adverse, then arranging a sale may take months or years.

Understanding of Financial Statements

Investing, in whatever form, deals with uncertainty. This is true for stocks, bonds, and, most importantly, commercial real estate. Uncertainty about the future translates into risk. Financial statements are tools of tremendous power that can be used to reduce that uncertainty. Financial statements may not lead the investor to a definitive answer. However, they will lead to a better understanding of these risks.

Income statements are often not what they seem. "The devil is in the details" is never truer than when it comes to determining what financial statements mean. Income statements attempt to show how a business performs over a specified period of time. Most commonly, income statements are presented on a yearly, quarterly, or monthly basis. If the purpose of the income statement is to provide insight into a property's performance, an immediate problem arises over what "performance" means. Performance is often discussed in generalities like "profit," "earnings," or "the bottom line." There is nothing wrong with using such terms per se. The problem is they tend to mean different things to different people. Using such terms without defining them leads to misunderstandings and misunderstandings lead to mistakes.

There are two basic kinds of income statements. They bear some similarity but are, in fact, quite different. They are most powerful when used in combination. One type of income statement depicts past performance. Thus, a property's revenues, expenses, profits, and losses are reported for a specific time period. As investors, we are generally more interested in the future than the past. Thus, the true worth of these historic financial reports is that they may give us some hints into the future. The statements can help us in constructing the second type of income statement, which is called a pro forma income statement. The astute investor must have the ability to read and understand the income statement.

Difficulties in Determining Property Value

This issue is closely related to liquidity. If real estate is inherently illiquid, that means it takes time to realize the property's value. But what is its value anyway? This is certainly an area that it is easy to disagree on.

When investors are selling a commercial property, they are really selling a stream of income. Valuing this stream of income requires two factors to be considered. First, one must quantify the stream of income itself, and secondly, one must determine the risk associated with that stream of income.

The stream is simply the net of cash inflows and outflows associated with a given real estate investment. Typically, the inflows can be from rental income while the outflows are associated with normal operating and maintenance expenses along with financial costs (interest and principal) and taxes. The income statement described earlier is a useful tool in constructing cash flow statements.

The second element (after determining income) in determining value is determining the risk associated with that value. This risk has to do with the fact that the income anticipated might not occur, or its value may in some sense be diminished. The use of a discount factor is commonly used to adjust the cash flows to take this into account. Thus, discounting that income to its present value explicitly quantifies the risk associated with income.

If a property is generating an income stream of \$10,000 per year, and that condition is expected to persist for the foreseeable future and a discount factor of 20% is considered appropriate to the risk level of that income, then the value of that property may be determined by the following equation (where n is any number of periods of time):

$$\text{Value} = \$10{,}000/(1+0.2)^1 + \$10{,}000/(1+0.2)^2 + \ldots \$10{,}000/(1+0.2)^n \tag{48.1}$$

$$\text{Value} = \$10{,}000/0.2 = (5)\$10{,}000 = \$50{,}000 \tag{48.2}$$

Equation (48.1) says that the value of the property is the present value of its income (however measured) discounted at a rate of 20%. (Stated in real estate lingo, "its value is equal to five times earnings.") This is the general rule for determining the value of every kind of commercial real estate property. That is, ultimately, its future earnings and its corresponding risks determine real estate's value. In this case the income level is determined to be \$10,000 and those earnings in the future are discounted at an annual rate of 20%. The exponential in this series allows for the compounding effect to take place.

Where disagreements over value take place (and divergent opinions are common in this area), those disagreements center either on the quantity of earnings or the quality (associated risks) of those earnings, that is, whether this property is really generating \$10,000 in income or whether there is another way to look at it. Where the buyers and sellers forecast of future earning differ, each will arrive at different valuations. Furthermore, perhaps the seller is basing his analysis on cash flow, while the buyer thinks the net income figure would be more appropriate.

The future is always hard to predict. One way to deal with the risk of the unknown is to increase the discount rate to reflect that risk. A seller might be offering the property for the \$50,000, as shown in equation (48.2), because he or she has confidence in the future ability of the property to generate that \$10,000 year after year. Potential buyers may not share that confidence. For example, potential buyers may know less about the property and thus, may have less confidence in the property's ability to generate income in the future. Thus, prospective buyers might want to discount that \$10,000 at a higher rate, say 40%, to compensate for that uncertainty. Therefore, these buyers will offer to buy the property at 2.5 (1/0.4) times earnings. When market conditions deteriorate, buyers become increasingly fearful of what the future might bring. They respond by seeing the real estate as deserving of higher discount rates. That is why prices fall on the downside of the market.

Another variation of equation (48.1) commonly encountered is where future income is likely to grow. (In equation (48.1), future income was projected to be constant.) This situation is expressed in equation (48.3).

$$\begin{aligned}\text{Value} = {} & \$10{,}000(1+g)/(1+0.2)^1 \\ & +\$10{,}000(1+g)^2/(1+0.2)^2+\ldots \\ & +\$10{,}000(1+g)^n/(1+0.2)^n\end{aligned} \tag{48.3}$$

where $g = 10\%$

$$\begin{aligned}\text{Value} &= \$10{,}000/(0.2-0.1) = (10)\$10{,}000 \\ &= \$100{,}000\end{aligned} \tag{48.4}$$

Again, equation (48.4) is just a simpler way of expressing equation (48.3), which says that the property is now worth \$100,000 (10 times earnings) because this income stream is expected to grow at 10% annually. Here again the assumptions underlying the valuation may cause differing views as to the property's value. If it is easy to disagree on the income measure to be used and what the appropriate discount rate is determined to be, then it is really easy to disagree on what the future rate of growth will be.

Equation (48.4) is the most commonly used framework to determine value. That is, the value of a commercial real estate property depends on how much income it will generate, the appropriate rate at which that income should be discounted, and how much that future is likely to grow in the future.

Overextended Borrowing

Leverage is a good thing, but too much leverage can be a bad thing. Leverage increases the potential return on a project, while at the same time increasing the risk associated with that project. This is why it is better to optimize leverage than maximize it. Too much borrowing jeopardizes the success of a real estate investment as surely as too little leverage. It is a matter of balance to be decided by the investor's taste and preference for the trade-off between risk and return.

Management Expertise Required

Where ownership of the property is direct, the commercial real estate investor is going to need to be involved with searching for the project, evaluating the project, financing the project, and (if acquired) managing the project. Even where the commercial real estate investment involves a sale–lease-back arrangement and there is no property to search for, and the evaluation is cut and dry, the project will still not manage itself. There are always ongoing issues to be dealt with between the lessor (the owner) and lessee (the tenant). Commercial real estate investment is not a passive activity. It requires active, focused, intense participation or things are likely to go terribly wrong. Commercial real estate investment is not for the detached.

BUSINESS ORGANIZATIONAL FORM

The choice of which business form to adopt, as a vehicle to invest in commercial real estate, is critical to the success of the real estate investor. Different business forms have different tax implications, different implications for investor liability, different implications for control, and different implications for cost. There is no one "best" business form. The business form adopted should be the one that best meets the individual investor's needs. The investor's needs have to do with the investor's goals, personal situation, and the particular type of investment being considered. In this context the investor must weigh the trade-offs between tax advantages, liability, control, and cost.

Perhaps the most important issue impacting the business form chosen is the potential liability for the investor by the business organization or agents of the business organization. The legal principle is *qui facet per alium, facet per se*. That is, "who acts through another, acts himself." All business forms are governed by the concept of agency. Agency is a legal relationship in which one person (real or artificial, that is, a corporation, limited liability company, and so on) represents another and is authorized to act on his or her behalf. Agency law is quite broad and covers the whole body of rules that society recognizes and enforces in regards to situations where one person acts for another. Without agency law, business could not act. Each individual would only be able to represent himself.

The form that the business takes affects the liability of the owners. To what extent is the business and the investor the same? To the extent they are the same, the investor will be responsible for torts of any agents of the business. (A tort is damage, injury, or a wrongful act done willfully, negligently, and not involving breach of contract. If the issue is a breach of contract, then the issue is dealt with in a civil suit.)

Agents of a business include employees and those to whom the business has given a power of attorney. Independent contractors are not considered agents. Principals in a business are generally responsible for the acts of their employees. (Agency can be created by contract, by conduct that implies agency or by "estoppel" (apparent authority). This means an employee of the business may bind the principal contractually, whether the employee has the actual authority to do so, as long as the employee has the apparent authority to do so. Further, even though the owner has committed no act of negligence, the principle can be held negligent if that employee is acting within the scope of their employment.

Agents are expected by law to exhibit a high degree of fidelity to the principal. This would include obeying instructions, acting with skill, protecting confidential information, and the duty to avoid a conflict of interest. The principal, in turn, has a duty to compensate the agent and inform him or her of any risks associated with the agency. When an agent acting within the scope of their employment commits a tort, both the employee, and the principal can be held liable for the tort. This is known as the doctrine of joint and several liability.

The legal framework for business organizations is created at the state level. Although the forms are similar across state lines, the keyword here is similar. A form that would provide a desired advantage in one state may well not do so in another. There is no barrier to creating a business form in a state offering the most advantages. The laws of the state in which it is formed, not the state in which it operates, govern the liability and internal affairs of a business entity. While forming the business in your home state may offer simplicity and cost savings, states such as Delaware and Nevada may, in most cases, offer superior liability and other offsetting advantages.

The following discussion deals with the attributes of these business forms in general. The specific needs of an investor should always be discussed with an experienced attorney to determine the relevance of the laws in that state to the investor's need.

Sole Proprietorship

A sole proprietorship is the easiest, most convenient, and least expensive form of business organization. Unfortunately, for the real estate investor interested in increasing and preserving wealth, it is not very good. This form of business has only one owner. There are no formal requirements to create or operate this form. The owner has unlimited, personal liability for all of the business's debts. The owner personally hires all employees, and thus the owner has unlimited, personal liability for the acts of employees. A sole proprietorship is not a separate taxpaying entity. Income is reported on the owner's personal tax return and does not require the filing of a separate tax return. For these reasons, this form of business should usually be avoided.

General Partnership

A general partnership (or simply partnership) must have two or more owners. No formal requirements are necessary to create and/or operate this form. Some states provide for the filing of "Articles of Partnership," so that the arrangement is a matter of public record. All owners have unlimited, personal liability for all of the businesses debts. All owners personally hire all employees, and thus all of the owners have unlimited, personal liability for the acts of employees. In addition, each owner has unlimited, personal liability for the acts of all of the other owners. Partnerships are a separate taxpaying entity: Income is prorated to the owners' personal tax returns and the business files only an information return with the Internal Revenue Service (IRS). Partnerships are a relatively simple business form to create and operate. Exposure to liability is so great in this form that it should not be used. Thus, this form is in a tie for the worst form of business with sole proprietorship.

Limited Partnership

A limited partnership must have two or more owners and are formally created under state law. One or more owners of the limited partners must be a general partner who has unlimited, personal liability in all of the same ways as a partner in a in a general partnership. At least one owner must be a limited partner (frequently all of the other owners will be limited partners) who has limited liability. Owners who are limited partners are prohibited from participating in the management of the business. Limited partnerships are frequently used to build tax shelters and for estate planning purposes. Income to the partnership is passed through to the owners' personal tax returns and the business files only an information return with the IRS.

Limited Liability Partnership

The *limited liability partnership* requires two or more owners as a limited partnership. This business form is formally created under state law, as is a limited partnership, but all of the owners have limited liability for the business's debts. In many states, however, this "limited liability" is less than that afforded to the owners of a limited liability company or a corporation. In some states, notably California and New York, the limited liability partnership may be used only in "professional" practices. Income is passed through to the owners' personal tax returns, and the business files only an information return with the IRS.

Limited Liability Limited Partnership

In some states, the limited partnership can register as a limited liability limited partnership that has the effect of giving the general partner limited liability. Therefore, all of the owners of the this form of business organization have limited liability for the debts of the business. This form of business organization is usually more costly to start and maintain than a limited partnership because it is subject to more formal statutory rules regarding officers and record keeping. Income is passed through to the personal tax returns of the owners and the business files only an information return with the IRS.

Registered Limited Liability Partnership

Some states provide for the creation of "registered" limited liability partnerships. This occurs where the limited liability partnership is really a general partnership that has "registered" in the limited liability partnership form to achieve some version of limited liability for all of the owners of the business.

Limited Liability Company

A *limited liability company* may have one or more owners. This business form is created by state statute where all of the owners have limited liability for the debts of the business. A limited liability company is usually less costly than a corporation to create and maintain because it has more relaxed, less burdensome rules governing operation compared to a corporation. A limited liability company is not a separate taxpaying entity: Income is reported on the personal tax returns of its owners and does not require the filing of a separate tax return when there is only one owner.

In many states, the business interests of the owners of a limited liability company are protected from the claims of the personal creditors of the owners. This advantage is not enjoyed by the limited liability partnership. This advantage may be significant for preserving wealth under adverse conditions. Therefore, the limited liability company combines into one form the best elements from a corporate entity (limited liability for all of the owners) and the general partnership (absence of formalities, low costs, tax benefits). For most commercial real estate investors, this is probably the business form of choice.

It should be noted that while owners have "limited liability" in a limited liability corporation, that limitation only means that the creditors of the corporation cannot go after the personal assets of the owner. To the extent that the owner has assets that remain in the limited liability corporation, those assets are not immune from the claims of creditors. As this form of business organization is relatively new to the business arena, it is important to form the limited liability corporation in a state (such as Delaware or Nevada) that follows the Revised Uniform Limited Partnership Act (RULPA) view in its LLC statutes.

C Corporation

Corporations are formally chartered at the state level and provide for a separation of ownership from management. They are more costly to establish and maintain that other business, but they provide unparalleled protection for the owners from claims against the business itself. Owners elect directors who have the formal responsibility for selecting and monitoring corporate management. C corporations are taxed as entities themselves and repatriate profits to their owners through dividends, which are then subject to the personal income tax (so-called double taxation).

It should be noted that while owners have "limited liability" in a C corporation, that limitation means only that the creditors of the corporation cannot go after the personal assets of the owner. To the extent that the owner has assets that remain in the corporation, those assets are not immune from the claims of creditors.

Subchapter S Corporations

Subchapter S corporations differ from C corporations only in that profits are not subject to a separate corporate tax and such profits are prorated to the various owners directly where they will be subject to the personal income tax. The term used to describe this is that Subchapter S corporations are treated as "conduits" for tax purposes.

Closed Corporations

Closed corporations have all the characteristics of a C corporation (double taxation, limited liability, etc.), but are less expensive to charter and maintain. Laws on this type of corporation vary considerable from state to state, however:

- Closed corporations generally are held by a single shareholder or closely knit group of shareholders.
- The corporation may be formed initially as a closed corporation or may amend its articles of incorporation to include this statement.
- A closed corporation's profits are taxed twice: once at the corporate level and again when profits are distributed as dividends to their shareholders. If a closed corporation meets specific IRS requirements, a corporation can file for Subchapter S corporation status and generally avoid paying tax at the corporate level.
- The shareholders of a closed corporation are personally liable for the debts and liabilities of the closed corporation only to the extent of their capital contribution.
- There are no public investors, and its shareholders are active in the conduct of the business.
- Bylaws are not required if provisions, normally included in bylaws, are included in the shareholders' agreement.

Professional Corporations

Professional corporations are designed to meet the needs of groups of professionals (physicians, dentists, lawyers, etc.) who wish to practice together and wish to organize their business association in a corporate framework. (This will have advantages in transferring and valuing ownership, in the corporation's existence separate from the owners and an indefinite life, but will also include taxation at both the corporate level and the personal level.) Control will be vested in a board of directors that is elected by the shareholders. Costs are formal registration and filing, and reporting requirements vary from state to state.

Professional Associations

Professional associations are designed to meet the needs of groups of professionals (physicians, dentists, lawyers, etc.) who wish to practice together and wish to organize their business association in a corporate framework. (This will have advantages in transferring and valuing ownership, in the corporation's existence separate from the owners, and an indefinite life, but will also include taxation at both the corporate level and the personal level.) Control will be vested in a board of directors that is elected by the shareholders. Costs are formal registration and filing, and reporting requirements vary from state to state.

Service Corporations

Service corporations are corporations designed to meet the needs of groups of professionals such as physicians, dentists, and lawyers who wish to practice together and wish to organize their business association in a corporate framework. (This will have advantages in transferring and valuing ownership, in the corporation's existence separate from the owners, and an indefinite life, but will also include taxation at both the corporate level and the personal level.) Control will be vested in a board of directors that is elected by the shareholders. Costs are formal registration and filing, and reporting requirements will vary from state to state.

Multiple Forms

It is possible to reap further advantages in terms of minimizing taxes and minimizing liability by layering different business forms for holding an investment and operating the investment. Using different forms for holding and operating an investment involves using a two-entity structure. In this type of arrangement, an operating entity will carry out the actual business functions, and a holding entity will own the major capital assets of the company, often including the operating entity itself. In this way, you can provide a nearly impermeable shield for your business assets against the claims of business and personal creditors. It is possible for business owners who desire a simplified structure to personally act as the holding entity, although in that case the liability shield will not be as strong.

The use of multiple business forms can be effective in protecting assets by minimizing the amount of vulnerable capital invested within the operating entity. Strategies that would accomplish this result would include:

- The owner's personally owning and leasing assets to the operating entity.
- A strategic combination of equity and debt funding (debt funding for the operating entity, equity funding for the holding entity).
- Encumbering the operating entity's assets with liens that run in favor of the holding entity or owner.
- Systematic withdrawals of funds as they are generated.

To avoid the problem of the limited liability being challenged by charging orders, withdrawals of funds should be done on a regular basis, following due procedures. Such withdrawals could include the use of dividends, earned salary and wages to the owners, payments to the owners on leases of property or equipment held by the owner, and factoring account receivables.

Another advantage of a two-entity structure would be to allow for proper planning for federal estate taxes. This important issue is often overlooked in the hurley-burley context of business formation. Many commercial real estate investments thrive and grow to produce tremendous wealth for the owner. Yet, in the absence of effective estate planning, much of this wealth may be paid to the federal government in the form of estate taxes, rather than to the owner's family, when the business owner dies.

Franchiser or Franchisee

A franchise is a contractual arrangement between the owner (franchisor) of some property or type of business that permits another (franchisee) to use that property or type of business. A franchise involves a relationship between the owner and the user. There are a number of situations in which a real estate investor would find it desirable to be either a franchisor or a franchisee. A real estate investor could function as a franchisor to use franchisees to provide capital for a business undertaking, provide entrepreneurship for a business, and to absorb the risk of loss. A real estate investor might wish to be a franchisee when experience or expertise in a particular business is needed, or when the franchisor's brand or goodwill is a valuable asset.

There are three basic types of franchises: (1) distributorships, (2) business systems, and (3) process systems. Distributorships involve licensing dealers to sell products such as Texaco gasoline stations. Business systems involve the use of a standard method of operation in conjunction with a brand name like Jiffy Lube. Process systems involve the ingredients and procedures used in making something like Pepsi Cola.

The relationship between the franchisor and franchisee is governed by a franchise agreement. As there is normally one large franchisor and many small franchisees, the franchisor usually has a franchise agreement prepared that is offered to prospective franchisees on a take-it-or-leave-it basis. Since franchisors have so much power relative to franchisees, courts will generally interpret any ambiguities in the franchisee agreement in favor of the franchisee. The courts will find the franchisor has an obligation of good faith in such an agreement and will generally not enforce any provision that is inherently unfair.

Entering into a franchisee agreement requires a good deal of disclosure about the nature of the franchise. Both state statutes and Federal Trade Commission rules generally requires such disclosure. Information is the commercial real estate investor's friend in considering the desirability of a franchise. It is important to evaluate all data associated with the franchise and to speak to existing franchisees about their experiences with the franchisor. Information collected should be subject to a thorough analysis of historical and pro forma income statements.

SUMMARY

Commercial real estate offers investors significant returns while providing them with means to control risks. Location continues to be the single most important determinant of real estate value. This certainly holds true for commercial real estate. Investors can enhance their diversification and thus reduce risks by including commercial real estate as part of their overall investment portfolio.

There are many advantages to investing in real estate. These include the ability to benefit from financial and operating leverage. Tax advantages also make commercial real estate and attractive investment. Historically, commercial real estate has been an effective hedge against inflation. Clearly, there are disadvantages as well. Real estate is a very illiquid investment. Thus, the holding periods can be quite long. Many investors may become overextended. Additionally, real estate investing requires a degree of management expertise.

The choice of a business form offers greatest opportunities for the investor to control risk, maximize cash flow, and minimize taxes. While there are a variety of choices available, most small (that is, less than $1 million equity) commercial real estate investors will want to choose a structure involving a limited liability company or a Subchapter S corporation. Particularly in combination, a Subchapter S corporation as the holding company and a limited liability company as the operating company offer the most effective protection for both personal assets outside the business and the investment in the business itself. Larger investors may well prefer the more formal setting of a limited liability company holding an operating, conventional C corporation.

REFERENCES

Bowman, J. L. (2005). *How to Succeed in Commercial Real Estate*. Ft. Worth, TX: Mesa House Publishing.

Fisher, S. D. (2006). *The Real Estate Investor's Handbook: The Complete Guide for the Individual Investor*. Ocala, FL: Atlantic Publishing Group.

Freedman, R. (2006). *Broker to Broker: Management Lessons from America's Most Successful Real Estate Companies*. Hoboken, NJ: John Wiley & Sons.

Gallinelli, F. (2004). *What Every Real Estate Investor Needs to Know about Cash Flow . . . and 36 Other Key Financial Measures*. New York: McGraw-Hill.

Haden, J. (2006). *The Complete Dictionary of Real Estate Terms Explained Simply: What Smart Investors Need to Know*. Ocala, FL: Atlantic Publishing Group.

Haight, G. T., and Singer, D. D. (2005). *The Real Estate Handbook*. Hoboken, NJ: John Wiley & Sons

Hudson-Wilson, S., Fabozzi, F. J., and Gordon, J. N. (2003). Why real estate? An expanding role for institutional investors. *Journal of Portfolio Management,* Special Real Estate Issue: 12–27.

Hudson-Wilson, S., Gordon, J. N., Fabozzi, F. J., Anson, M., and Gilberto, S. M. (2005). Why real estate? And . . . how? Where? And when? *Journal of Portfolio Management*, Special Real Estate Issue: 12–22.

Kolbe, P., and Greer, G. (1977). *Investment Analysis for Real Estate Decisions*, 6th edition. Chicago: Dearborn Financial Publishing.

McMahan, J. (2006.) *The Handbook of Commercial Real Estate Investing*. New York: McGraw-Hill.

Masters, N. (2006) *How to Make Money in Commercial Real Estate: For The Small Investor*. Hoboken, NJ: John Wiley & Sons.

Peiser, R. B., and Frej, A. B. (2003). *Professional Real Estate Development*, 2nd edition, Washington, DC: Urban Land Institute.

Schmitz, A., and Brett, D. (2001). *Real Estate Market Analysis: A Case Study Approach*. Washington, DC: Urban Land Institute.

CHAPTER 49

Types of Commercial Real Estate

G. TIMOTHY HAIGHT, DBA
President, Menlo College and Chair of the Board, Board of Commonwealth Business Bank (Los Angeles)

DANIEL D. SINGER, PhD
Professor of Finance, Towson University

Abstract: A wide array of opportunities is available to the potential commercial real estate investor. Such opportunities may be developed by the investor or acquired as existing facilities with a history of costs and revenues. Investment in this area is favored by the presence of tax shields from depreciation, interest rate offsets, and financial leverage. The value of existing commercial properties can be determined by capitalizing the anticipated cash flows associated with that property. The dynamic nature of the commercial real estate market generates a flow of opportunities as apartments, condos, shopping centers, restaurant sites, recreation facilities, motels, self-storage facilities, warehouses, office buildings, and manufacturing sites are continually brought to the market.

Keywords: apartments, condominiums, shopping centers, restaurant, recreation facilities, motels, self-storage facilities, parking facilities, office buildings

Commercial real estate runs the gamut from apartments, condos, shopping centers, restaurant sites, recreation centers, motels, and self-storage facilities to warehouses, office buildings, and manufacturing sites. Such investments tend to be characterized by predictable cash flows and reap the benefits of a tax shield from depreciation as well as interest cost offsets. The result is attractive risk/return opportunities for prospective investors.

As with all real estate investment, location is critical to success. However, the desirability of a particular location for a commercial real estate property will depend on the ability of that property to generate income. This will depend in turn on the type of commercial property under consideration relevant to population density and socio-economic characteristics, the transportation infrastructure, zoning laws and regulations, and the proximity of customers, suppliers, and competitors.

In this chapter we describe the different types of commercial real estate. Table 49.1 provides a summary of the advantages and disadvantages of each type.

APARTMENT COMPLEXES

Generally, *apartment* complexes can be classified as one of three types: garden, mid-rise, and high-rise apartments.

Table 49.1 Advantage and Disadvantage Comparisons by Types of Commercial Real Estate

Type of Property	Financial Leverage	Operational Leverage	Inflation Resistance	Tax Advantage	Own Business	Compounding Properties	Lack of Liquidity	Valuation Difficulty	Management Expertise
Single-family	High	Medium	High	High	High	High	Medium	Medium	Medium
Apartments	High	High	High	High	High	High	Medium	Medium	High
Condos	High	High	Medium	High	High	Medium	Medium	Medium	Medium
Time shares	Medium	Medium	Medium	Low	Medium	Low	High	High	Low
Undeveloped land	Medium	High	Medium	Very low	Medium	High	High	High	Low
Self-storage	High	High	High	Medium	High	Medium	High	Medium	Medium
Restaurants	High	High	Medium	High	High	Medium	High	High	High
Shopping centers	Medium	Medium	Medium	Medium	Medium	Medium	High	Medium	High
Athletic facilities	Medium	Medium	Medium	High	High	Medium	High	Medium	High
Office buildings	High	High	High	High	Medium	High	High	Medium	Medium
Industrial	High	Medium	Medium	Medium	Medium	High	High	Low	Low
Parking lots	High	High	High	Medium	High	High	High	High	Medium
Hotels & motels	High	Medium	High	High	High	Medium	Medium	Medium	High

Garden apartments have from one to three levels and normally contain large balconies or patio areas. Mid-rise apartment complexes range from four levels to six levels. Anything higher than six levels would be classified as high-rise. Typically, zoning laws or the value of the underlying land determines the particular type of apartment to be constructed in a given location. An apartment complex developed downtown adjacent to a central business district is typically going to be a high-rise to amortize the high price of the underlying land. Apartment complexes in suburbia are usually garden apartments, because they are perceived to better fit into their surroundings aesthetically and to make fewer demands on municipal services. This is likely to be forced by zoning laws even where there is an economic incentive to build mid-rise or high-rise complexes. As a general rule, the cost of constructing apartment units rises more than proportionately as height rises.

Condominiums

Physically, *condominiums* (condos) are individually owned units in an apartment-style building, but condominiums are really about lifestyle choices. This is true in two senses. First, the condo represents an alternative lifestyle choice to a traditional single-family residence. "Residential condos" cost less to build, usually cost less to maintain, often offer an automatic sense of community, and there is no lawn to mow! Second, there is a "destination condo" syndrome that conceptually is related to timeshare properties, but takes the ownership form of a condominium. A condo in Hawaii, Barbados, Orlando, Palm Springs, wherever—like timeshares, this is the stuff from which dreams are made. Own a little bit of paradise—who can resist?

Condominiums may be described as a hybrid form of real estate ownership. Condo owners hold a deeded title to the specific area occupied by their unit, not to the land beneath. As a result, condos are characteristically built to maximize the value of the land underneath them. Large, multistoried buildings composed of many individual units are commonplace. The owners of the different condo units jointly own the common areas such as the ground, pool, walkways, recreational areas, elevators, and the like. If the condo owner finances the property, the unit will have a separate mortgage, and the owner will pay a property tax on the unit and a pro rata share of the common area costs. A board of directors elected by the condo owners will govern the complex, making a set of rules controlling the usage of the individual units and common areas and will assess each unit a maintenance fee.

Condominiums are sometimes confused with townhouses and cooperatives. These two forms of ownership share the ownership of "commons" as condo owners, but differ in other respects. Townhouses are usually a series of single-story or multistory units that are linked to each other horizontally by common walls. Townhouse owners hold title to their units and the land beneath them, so townhouse units cannot be stacked on top of each other. Individual townhouse owners own any common property in common. Townhouse owners pay property taxes on their individual units. A property owners' association usually manages the townhouse complex and collects fees from all owners in order to maintain common areas. Cooperatives (co-ops) are formed by cooperative arrangement. A corporation holds title to all associated real estate. Buyers purchase stock in the co-op corporation and are considered shareholders, not owners of real property. Each shareholder holds a lease to his unit that runs for the life of the corporation. The corporation pays taxes. Any mortgages are normally held and paid by the corporation. All costs to operate the property are shared by shareholders. An administrative board must usually approve new cooperative shareholders.

Many successful real estate investors started by investing in condominiums. The initial out-of-pocket costs can be relatively low, and the income stream together with the tax benefits may cover substantially all of the cash outflows. Thus, the most significant contributor to return will be the property's appreciation over time. The astute investor will capitalize on this appreciation by increasing his/her leverage and acquiring additional condominiums.

TIMESHARES

The concept of a timeshare may be represented by the ownership of a particular piece of real estate for a particular time interval. (There are many different variations of this general concept.) The idea is you do not need to buy a condo in Orlando when you plan to use it only a week every spring. Instead of buying the whole condo, you buy that condo for a specific week of the year. This makes "ownership" much more affordable and available to a wider group of people. Quality timeshares (in terms of season and location) at the beginning of 2003 average \$4,000 to \$8,000 for a week's ownership at even the most desirable resorts.

With ownership you have the right to use that particular property during that particular week. While a given property might cost \$240,000 to own entirely, one may be able to buy that particular week for \$8,000. (Depending on variations in supply and demand, different weeks will have different prices.) From this perspective, the cost savings are obvious.

Timeshare owners are attracted to this form of vacation primarily by the high standards of quality accommodations and services available at the resorts that they own and exchange. It is felt that these properties are better maintained and staffed than are properties that are merely rented. In addition, the location of these properties is felt to be superior to those available through the rental market. The leading industry association, the American Resort Developers Association (ARDA), asserts that over a third of initial timeshare buyers eventually purchase additional timeshare units.

While timeshares are literally located everywhere, many timeshares are in properties with stunning, highly desired locations. The heart of London, Paris, Tokyo, and New York City all have timeshares. If the beach is your thing, from the pounding surf at Waikiki to the French Riviera, timeshares are there. Practically anywhere one could imagine as a desirable vacation destination will have timeshare properties located there. Timeshare resort amenities rival those of other top-rated resort properties and may include swimming pools, tennis, Jacuzzis, golf, bicycles, and exercise facilities. Also featured may be boating, ski lifts, restaurants, and equestrian facilities. Most timeshare resorts offer a full schedule of onsite or nearby sporting, recreational, and social activities for adults and children. The resorts are often staffed with well-trained hospitality professionals. Many timeshare resorts offer concierge services for assistance with visiting area attractions.

The attractions of timeshare ownership have an economic dimension. Timeshares offer individuals the opportunity to purchase fully furnished vacation accommodations for only a percentage of the cost of full ownership. For a one-time purchase price and payment of a yearly maintenance fee, purchasers own their property in perpetuity. The fact that owners share both the use and costs of upkeep of their unit and the common grounds of the resort property ensure that the property will be well maintained over time.

Unlike a hotel room or rental cottage, which requires payment for each use with rates that usually increase each year, ownership at a timeshare property enables vacationers to enjoy a resort, year after year. Timeshare owners may look forward to a lifetime of ownership with minimal exposure to inflation. Costs may be expected to rise by only the increase in maintenance costs.

UNDEVELOPED LAND

Undeveloped land can be classified as either raw land or developing land. Each of these has distinguishing characteristics. Raw land tends to be located in rural areas, far from existing patterns of development. Developing land is located in areas that are transitioning from a rural environment to a suburban or urban environment. The two types of undeveloped land have differing investment characteristics, but share the disadvantages of the lack of a depreciation tax shield, little opportunity for leverage, and a negative cash flow.

Considerable difficulties are encountered in trying to determine the value of undeveloped land for reasons discussed below. However, an excellent source of data on current development potential is available from *Emerging Trends in Real Estate*, an annual study by the Real Estate Research Corporation. This publication examines development potential in different areas from the perspective of existing price trends, existing business locations, demographic projections, and the attitude of local government toward development.

Undeveloped land frequently exerts an emotional appeal on an investor. "Falling in love with the land" can and does happen. However, purchasing undeveloped land because the beauty of the forest casts a spell on one is a far different behavior than purchasing the land as an investment. The strong emotional appeal of a property can cloud the business sense of an investor. Buying undeveloped land is usually easy; selling undeveloped land is usually hard. More than one investor who has been taken with the attractiveness of undeveloped property has bought high and then, in the absence of buyers, sold low.

Raw Land

The acquisition of raw land for its potential appreciation is also often overlaid with an emotional attraction. The ownership of raw land may yield investors psychic income having to do with "owning" a piece of America, or dreams of an idyllic retirement far removed from the stresses of the urban environment. This appeal is dangerous to an investor. The advertisement may have great emotional appeal to the investor who can almost hear the birds sing and the rustle of squirrels high in the oak trees. However, beauty is in the eye of the beholder. While the supply of raw land in general is quite large, each specific parcel is uniquely defined by its location and particular attributes. Effectively, the supply of that piece of land is perfectly inelastic. Demand for that piece of land may be nonexistent, low, or high.

Developing Land

Unlike raw land, developing land has little amenity value and a much more predictable future. Developing land is in a transitional area, between raw land and developed land. A classic investment in developing land might be 25 acres of farmland, but zoned for residential development, and located just outside the suburban fringe or a three-acre parcel located just off the exit and entry ramps for an interstate highway that is planned to begin building in three years. The future value of such developing land is far more predictable than that of raw land, but still not known with certainty.

Appreciation in the value of developing land results from shifting land use patterns. The trends toward more roads, increasing suburban spread, and economic growth in general are well established in our society. It is certainly conceivable that within a span of 50 years a property could have gone from agricultural cultivation along a dirt road, to residential homes along a paved two-lane road, to a strip shopping center along a four-lane highway. The difficulty of successfully investing in developing lands is that the exact pattern and timing of those shifts is very difficult to predict.

The level of economic activity or the pace of development in a geographic region is subject to the vagaries of unpredictable events. The oil crisis of the 1970s sparked a movement of population out of the Northeast and Midwest to the Sunbelt. Land values crashed in Boston in the early 1980s, and suburban development came to a dead halt. The fall in fuel oil prices in the late 1980s caused a disastrous fall in land values in Colorado and Texas. Boeing shifted its headquarters from Seattle to Chicago. The automotive industry decided to shift its production facilities from Detroit and other traditional northern areas to rural Midwest and southern communities. The Corps of Engineers decided to reverse its policy of draining the Everglades. The impact of the Cuban population on land use in Florida has been profound. Las Vegas redefined itself as a mecca for retirees. All of these events were essentially unpredictable, and all had substantial impact on the pattern of shifting land use in their region.

Whatever the overall trends in a particular region, considerable variation from that trend will be observed in specific localities. Such variation may reflect specific land use regulations, zoning laws, road patterns, the availability of public utilities (water lines, sewer districts, etc.), the availability of public services (parks, hospitals, etc.), and prior land use patterns. Predicting how those patterns will affect a specific parcel of land is difficult.

Even within a narrowly defined locality, considerable variation in land-use patterns may be observed: vacant lots on otherwise fully developed highways, apple orchards in the midst of residential neighborhoods, and apartment buildings abutting industrial sites—vacant land scattered all about with no apparent rhyme or reason, all of which work to make investing in developing land a risky enterprise.

An added danger to investing in developing land is the potential liability associated with land contaminated by hazardous waste. Under the 1980 Superfund Law, a property owner of contaminated lands (or even an ex-property owner) may be liable for clean-up costs, even if they had nothing to do with the contamination. Buying a 20-acre parcel for a shopping center, when it is currently largely undeveloped but does have a few old shacks on it, can be risky! It could turn out that one of those shacks 30 years ago was used as a foundry to make brake shoe moldings and the land is thoroughly contaminated by lead and asbestos. Developing land should always be checked against the Comprehensive Environmental Response Compensation and Liability Information (CERCLIS) list that is maintained by the federal Environmental Protection Agency (www.epa.gov/superfund/sites/query/basinstr.htm.). Each state has its own environmental protection agency that should also be contacted in this matter as they will be most familiar with the environmental issues in a specific locality.

In this contemporary world, property owners may also be beset by all manner of stakeholder claims. A 50-acre site with an old farmhouse on it is purchased to create an office park, but the farmhouse cannot be torn down because it has local historical significance. Or, as the ground is prepared for construction, a human skeleton is unearthed. It is determined that the site contains a hitherto unknown Indian graveyard. Representatives of a local Indian tribe file suit to halt construction. The possibilities are endless, overlaying all development activities with a rich layer of uncertainty.

The underlying economics of investing in developing land are not as favorable as for the other types of real estate discussed in this book. The reason lies in the fact that the vacant land cannot be depreciated. Thus, the tax shield that is available to those purchasing real estate containing depreciable facilities is not available to the purchasers of developing land. In addition, developing land will generally not generate revenue during its holding period. However, the land may well be subject to taxes and necessary expenses (e.g., taxes, preparing a land-use plan,) creating a negative cash flow during the holding period. Leverage is somewhat easier to obtain for developing land because creditors can be more certain of the land's value with actual development a close proximity.

SELF-STORAGE FACILITIES

Another type of commercial real estate investment requiring a modest cash investment is *self-storage facilities*. Self-storage facilities represent a burgeoning opportunity in real estate investment. The demand for such facilities is on the rise from both individuals and small businesses. In an increasingly affluent and materialistic society, individuals increasingly have more "stuff" than they can reasonably accommodate in their present residence. In an increasingly mobile society, individuals in transition need places to store their possessions temporarily. Small, and even some large, businesses have found self-storage facilities cost-effective ways to store their records, inventory, extra equipment, and seasonal goods.

There are currently about 30,000 self-storage facilities (SSFs) located throughout the United States. The rapid growth of the industry has resulted in progressive and accelerating change as to the function of such facilities. Self-storage facilities offer the financial leverage, tax advantages, and cash flow characteristics that characterize other commercial real estate properties. The industry began in the 1960s with properties that offered little in the way of location, convenience, and amenities such as air conditioning, heating, and 24-hour secure access. Basically, these were the equivalent of C-type office buildings. Today's self-storage facilities have evolved into centrally located facilities with elaborate amenities, the equivalent of A-type office facilities. In many locations, the demand for facilities of this type appears to be strong and will support a rate structure that makes developing such properties profitable.

SSFs are rented for the exclusive purpose of storing personal property in such a manner that the renter has access to, and control of, the property placed in storage. Self-storage is, therefore, not warehousing and presumes no possession of customers' goods. The fundamental aspect of self-storage is that it is a "self-service" operation and consumers/tenants retain the "care, custody, and control" of their personal property. The definition of "personal property" is vast, and regulations in this country typically focus on prohibited properties that may be stored as opposed to permitted properties. Prohibited properties typically include hazardous materials, perishable foods, and ammunition.

Early SSFs largely consisted of low, flat, one-story buildings typically located in lower-density suburban areas. Such facilities were generally not heated or air conditioned and were of starkly functional design. These units were originally thought of as "mini-warehouses" and often restricted by zoning ordinances to commercial and industrial zones as a result.

Modern SSFs are being built as much larger multistoried buildings with heating, air conditioning, and often offering a wide array of complimentary services. These SSFs are often designed to fit into the existing architectural setting. Increasingly, such units are being located in residential areas because they do not generate much traffic, are not noisy, do not create pollution, and place a very light burden on municipal services. These facilities are designed to meet the need for easily accessible, small-scale storage space. A typical SSF covers two to three acres and consists of five to six buildings, each containing approximately 10,000 square feet of storage space. The size of the storage units can range from 25 square feet to 600 square feet. Internal driveways provide access and parking at individual storage units.

Successful SSFs require a location well served by major highways. Well-traveled routes between commercial and residential districts, adjacent to interstate highways, and major thoroughfares abutting commercial and residential areas all show potential for this type of facility. Successful, modern SSFs are typified by a wide assortment of amenities and services that increase their value for their customers. The two absolutely critical elements for a successful SSF are that it provides good security and convenience.

Good security is provided by the sturdy construction of individual storage areas, strong doors, and door casings, locks typically supplied by the renter of the unit (to control access), door alarms, a perimeter fence, controlled access through the perimeter fence, a resident manager, bright lights, and 24/7 video and electronic surveillance of the grounds, an accurate record of individuals entering and leaving the premises.

Convenience involves 24/7 access for renters and their associates; an arrangement of facilities that facilitates access, egress, loading, and unloading; automatic interior lighting where appropriate; available loaner dollies and hand trucks; available packing, moving, mailing, and storage supplies (including tape, bubble wrap, boxes, furniture covers, etc.); outside storage for autos, trucks, recreational vehicles, boats, and equipment to complement the available inside storage; an adequate selection of different-size storage units; and clean rest-room facilities.

Another approach to SSFs gaining popularity involves the conversion of industrial or commercial property to SSFs. Depending on the area, such facilities may be conveniently located for this purpose. The property may be inexpensive because it is no longer being used for its original purpose. Thinking "outside the box" can prove rewarding in this situation. Where the facility already has heating and air conditioning, partitioning the interior can be relatively inexpensive and lead to excellent rental revenues.

RESTAURANTS

An often overlooked commercial real estate opportunity involves purchasing property servicing the *restaurant* industry. The restaurant industry is huge, accounting for almost 4% of our gross disposable product (GDP). There are approximately 890,000 restaurants operating in the United States with over 12 million employees doing over $800 billion in sales. Future growth is estimated to be between 4% and 7% annually. Americans have not lost their taste for eating out.

Not only is the industry huge, it is dynamic. Old restaurants go out of business and new restaurants spring up. How successful restaurants are depends on a wide variety of factors including broad social issues (the events of 9/11), broad economic issues (the cost of energy), industry-wide issues (mad cow disease), population trends (from the Rust Belt to the South, revitalizing inner cities), as well as the tastes and preferences of individual consumers (high-protein diets, low-carbohydrate diets, low-fat diets, organic foods, ethnic foods). All these factors contribute to a rapidly changing mix for success as a restaurant.

This is great news for commercial real estate investors. The combination of the size of the industry, along with its dynamic character, creates a mosaic of opportunities for commercial real estate investors. Investment in this industry can either be indirect (through holding a lease) or direct (through owning the property and managing the restaurant itself).

There are two basic divisions in the food service industry: (1) fast foods (characterized by both the speed of service and the limited amount of service available) and (2) full-service restaurants. These two sectors in the industry are of roughly equal size. Full-service restaurants may be further broken down into categories based on how expensive they are. These categories are (1) mid-scale (average check below \$20), (2) casual dining (average check \$20 to \$40), and (3) upscale or fine dining (average check above \$40). Each segment of the restaurant industry has its individual attributes that make for success and failure. However, a common theme among restaurants of all types is the importance of its location.

SHOPPING CENTERS

When one thinks of commercial real estate, *shopping centers* are often the first type to come to mind. Indeed, shopping centers cover a broad range of investment opportunities. Shopping centers may be invested in directly or indirectly through a partnership, general partnership, master limited partnership, limited liability corporation, Subchapter S corporation, or real estate investment trust (REIT). Rarely are large shopping centers held in the form of a regular corporation because of the problem of double taxation. Most shopping centers are owned indirectly because of the need to raise substantial capital and the advantage of spreading the risk.

Often overlooked for their inherent profitability, shopping centers frequently provide a great opportunity for the small- or medium-sized real estate investor. Shopping centers have desirable real estate investment characteristics. The largest portion of the value of the shopping center is depreciable, generating a substantial tax shield. Shopping centers generate their return through a continuous, relatively predictable cash flow, rather than requiring a one-time windfall in the distant future. Shopping centers usually can be acquired in a manner that provides for substantial leverage. Furthermore, the wide array of available types of shopping centers allows the investor to pick and choose the exact combination of risk and return that he or she prefers.

Shopping centers are defined in terms of the market served. This market can be a few square blocks for a community shopping center or cover the whole metropolitan area for a large regional shopping mall. Location, access, and traffic patterns define a mall's success relative to the market served. Securing aerial photographs that allow the shopping center to be identified within the context of its physical infrastructure can begin the analysis of a shopping center's potential. Such photographs are readily available from local or state government agencies. Traffic counts on key nearby thoroughfares can also be obtained from such agencies. Economic and demographic data by zip codes and census tracts are readily available, making it possible to pinpoint the characteristics of the market to be served.

Types of Shopping Centers

Strip Shopping Centers

A strip shopping center is 3 to 10 small, independent stores clustered on a heavily traveled road. The businesses are generally dependent on the volume of traffic and their exposure to this traffic. Convenience stores; fast-food restaurants; liquor stores; video stores; specialty produce, meats, or seafood; gas stations; dry cleaners; and so on typically dominate this environment.

Neighborhood Shopping Centers

A neighborhood shopping center is 3 to 10 small, independent stores serving a localized market. Such a shopping center is not necessarily located on a high-volume thoroughfare, although ease of access to the market served is critical. Convenience stores, fast-food restaurants, small restaurants, liquor stores, video stores, drugstores, hardware stores, toy stores, and the like typify this environment. Neighborhood shopping centers frequently may be in close proximity to apartment complexes, office buildings, hospitals, large employers, or in a densely populated urban area.

Community Shopping Centers

A community shopping center is 5 to 25 small stores with a major anchor that is a proven traffic builder. Such a shopping center is typically located at the junction of heavily traveled roads. Stores have sufficient breadth and diversity to generate strong positive externalities. The community shopping center has a mix of tenants that give it a strong destination appeal.

Inner-City Shopping Centers

An inner-city shopping center is 3 to 10 small, independent businesses located on the first floor of a large building. This shopping center serves both a neighborhood and transitional clientele. The shopping center may either be leased from the building's owners or purchased as a condominium from the building's owners.

Regional or Super-Regional Shopping Centers

A regional or super-regional shopping center is a significant concentration of retailers combined to draw shoppers from great distances. Such a shopping center contains 25 to 250 independent stores and may contain several anchors, and it is frequently located beside one or more interstate routes. The 1990s saw an overbuilt market for this type of shopping center. The resultant hypercompetitive environment resulted in the newer, even larger shopping centers doing severe damage to older shopping centers. The competition between these large centers reached such an extent as to draw some business away from community and neighborhood shopping centers in the same metropolitan area, resulting in a more difficult situation for those retailers.

A properly located shopping center with the right tenant mix affords investors excellent risk and return combinations—as long as the retail market served is not oversaturated with competitors. These opportunities are particularly attractive if the investor has the expertise to manage the shopping center itself. However, this management expertise is not absolutely necessary. In any given metropolitan area, the investor will be able to find a number of firms that specialize in managing shopping centers. It is sufficient that investors in shopping centers are entrepreneurs who act as catalysts to bring economic resources together to create something where there was nothing before.

RECREATIONAL FACILITIES

As the Baby Boomers become more and more health conscious, a larger portion of their discretionary income is being directed to recreation activities. Not surprisingly, a wide variety of facilities ranging from single-purpose indoor basketball, soccer, and ice-skating rinks, through bowling alleys, water parks, racquetball courts, weight-training rooms, and indoor and outdoor pools, to multiple-purpose family entertainment centers are being developed to accommodate this market. This provides for an excellent opportunity for individuals interested in a unique commercial real estate opportunity.

The specificity of athletic facility demand has implications for the optimal investment in a facility. The athletic facility must be properly sized. The conventional approach is to project a given-size athletic facility and build it for the least cost or to estimate demand, revenue, and profit and allow the capitalized amount of profit to determine the investment in the facility. Such approaches reflect simplistic thinking. Because the demand for the facility is not generic, but specific to groups with special socioeconomic attributes and lifestyles, there is an interaction between the market niche to be served and the investment to be undertaken.

The relationship between the investment made in the facility and the market to be served can be defined very precisely. This is normally done in a formal feasibility study that has three components: (1) a market feasibility study that identifies the viable market niches in a given market area; (2) an economic feasibility study that analyzes all cost associated with the construction of an athletic facility relevant to a given market niche target; and (3) a financial feasibility study that identifies all revenues and costs in a series of pro forma financial statements that allow an investor to judge the risk and return parameters of the facility.

OFFICE BUILDINGS

Office buildings provide excellent opportunities for discerning investors to reap excellent returns for the risk involved. Office buildings can be awesome structures with 50 stories and 1 million square feet of rentable space or modest, simple one-story buildings with 4,000 square feet of space. Whatever the size of the office building, the basic elements in the process of successfully investing in this market are the same. Find a location attractive to potential tenants, design a building that conforms to those tenants' needs, secure leases, find construction financing, find permanent financing, construct the building, and then manage the building in accordance with tenants' expectations. This is the simple recipe for success. Of course, the devil is often in the details. It is one thing to have the recipe for success, and another to successfully execute it.

The market for office buildings is highly cyclical but rather predictable. The low point of the cycle would normally constitute an excellent time to either develop an office building or buy one that is already built but in financial trouble. Even outside this office building cycle, a need may exist for an office building that has not been met in a particular location. Such a situation may offer the investor excellent prospects as well. As with other real estate investments, the life cycle of an office building from an investor's perspective goes through an acquisition stage, a holding stage, and a disposal stage.

A variety of other issues must also be considered in the development of an office building:

- **Design**. A community may often have strong feelings about appearance, construction material, and site layout. The requisite approvals from a variety of municipal agencies may be sensitive to the issue of community acceptance of the property design.
- **Use intensity**. This is often expressed by the ratio of floor space to the area of the site. The floor-area ratio is often addressed in zoning standards.
- **Access and circulation**. The intended office building must effectively be integrated into existing traffic flows as well as having a safe and efficient traffic flow of its own.
- **Traffic generation**. The impact of the proposed office building on existing traffic patterns is often a point of contention with the local community. Traffic features such as convenience, ease of access, and quality of roads are often the focus of community concerns about the maintenance of those features.
- **Parking**. Office buildings must have ample parking to satisfy the desires of their prospective tenants, the concerns of local residents, and the requirements of zoning ordinances.
- **Sewer and water availability**. Local sewer and water facilities may be at or close to capacity, requiring costly investments in infrastructure. This issue may be formally addressed by the municipal government through the imposition of "impact" fees, or the developer may have to fund such facilities as are necessary to make the project feasible.
- **Environmental considerations**. Environmental considerations increasingly impinge on land development processes. Environmental considerations would include any relevant natural features of the site (wetlands, floodplains, endangered species, etc.) as well as the more general impact variables covered by clean air, clean water, and environmental hazard regulation.

- **Disability accommodations**. The Americans with Disabilities Act (ADA) requires that facilities such as office buildings be accessible and usable by persons with disabilities. This means public entrances (such as sidewalks), parking, and meeting places must not have barriers to disabled individuals.

Types of Office Buildings

Trophy Buildings

A single tenant, willing to pay more for a unique shape and floor plans, unusual building design, and an outstanding location, is typically attracted to "trophy buildings." Such buildings are characterized by the best quality of materials and workmanship and enjoy top-quality maintenance and management. Examples of such buildings include the PPG headquarters building in Pittsburgh and the Bank of America headquarters in San Francisco. Such buildings are always considered Class A.

Character Buildings

Character buildings are typically smaller office buildings (less than 10 floors) that are created by an investor or developer to display a sense of personal accomplishment. Such buildings are typically named after that investor or developer and are intended to stand as a monument to his or her accomplishments. Such buildings are normally well constructed, well located, and well maintained, but are not constructed on as lavish a scale as trophy buildings. Character buildings may be either Class A or Class B.

Class A Buildings

Class A buildings generally constitute the best buildings available in a given market. Such buildings are well located, attractive, and well maintained, and considered highly desirable by prospective tenants. They feature excellent elevators, mechanical systems and air-control systems. Class A office buildings are frequently occupied by high-quality, prestigious tenants.

Class B Buildings

Class B buildings are constructed along utilitarian lines using standard construction techniques to create as much rentable space as possible for a given cost. The design and layout of such buildings would be considered adequate, but is primarily functional in nature. These buildings also feature adequate elevators, mechanical systems, and air-control systems. Maintenance services and building management are average.

Class C Buildings

Class C office buildings are typically unrefurbished older buildings or older buildings with limited refurbishment. Their location may be inferior. Building maintenance may be substandard. Mechanical, heating, and air-conditioning systems may have problems. The building tenants are noticeably inferior to those found in higher-class buildings and in less desirable areas.

PARKING LOTS

Among the commercial real estate plays that are often overlooked is investing in parking lots. However, this type of commercial real estate opportunity deserves a closer look. Parking lots are cash-generating machines. Parking lots can be low-cost, low-technology facilities that are simple to operate. Parking lots can also be high-cost, high-technology facilities requiring a sizable staff and sophisticated management controls. The hallmark characteristic of parking lots is their high ratio of fixed to variable expenses. This means that the key to successfully developing (or purchasing) a parking lot is revenue estimation.

The potential revenue of a given parking lot is driven by location and constrained by competition. Location is always of great importance in determining the value of property, never more so than in the case of parking lots. In any suburban or urban area, there is always plenty of free parking—but it is just not in the right place. The demand for parking lots is very location specific. You are either next to the convention center or three miles away. One of these locations will not substitute for the other.

Parking lots generate heavy externalities (benefits that do not accrue to the parking lot itself). Urban hotels, stadiums, shopping mall, hospitals, and large office complexes cannot exist without *parking facilities*. As a result, parking lots are frequently integrated into the development of such population-intense facilities. Often times, the integrated parking facility will be an important revenue generator in its own right. Municipal governments may see the need to provide parking facilities as part of an infrastructure to support its commercial, industrial, and residential population. We are a nation on wheels. To live in our culture is to be on the go. The large bulk of the population finds the automobile a necessity, and as they go from here to there, they need to park that automobile.

Parking Market Segments

Short-Term/Transient Parkers

Short-term/transient parkers may need parking for a fraction of an hour or several hours. On an individual basis, their demand for parking may be only occasional or sporadic. As a group, their demand may by quite predictable and regular. Generally, these parkers pay by the hour and the revenue schedule is set up to "front load" the cost to the patrons. That is, the first fraction of an hour, or the first few hours, is priced much higher than successive hours. Since short-term and transient parkers by definition have only a limited need for the consumption of a parking space, this pricing structure will have a favorable impact on the revenues of the parking facility. Where the demand for this type of parking is predictable, this is the most profitable type of service the parking facility can provide.

Demand from this source tends to have the smallest physical drawing area. The patron demand for parking is a direct and immediate function of the patron's destination. Therefore, this demand will be limited to arising from destinations in the immediate vicinity of the parking facility.

Where demand permits, the goal of the revenue structure is normally to capture 25% of the maximum daily rate (MDR) in the first hour and 50% of the MDR in the first three hours. It may thus be possible to get three or four patrons in a given spot during a 24-hour period, potentially doubling the MDR revenues for that spot.

Early-Bird Parkers

The demand from early-bird parkers arises from the desire of the parking facility to fully utilize its available capacity. Early-bird parkers are price-sensitive shoppers who are able to exercise discretion in determining where they park. If there are competitors nearby in the effective market area, then these parkers must be attracted by offering price discounts. Early-bird parkers have more elastic demand curves than short-term/transient parkers. This market is served by specifying a specific time slot for their parking (e.g., in by 8:00 A.M. out by 5:00 P.M.) and offering these parkers a 40% to 60% discount off the MDR.

Special-Event Parkers

Examples of special-event parking might include a sporting event, a parade, a circus, a concert, and the like. Such time-and-destination-specific parking tends to be highly inelastic. If the parking facility and the site of the event are in close physical proximity, this often means a flat fee can be charged equivalent to the MDR for the event because parking is event dependent rather than time dependent. Where the parking facility is located on the outer edge of the effective market area, price concessions may be warranted to attract parkers.

Saturday and Sunday

In most urban settings, parking demand is weakest on the weekend. Offices and business are closed, and the need for short-term and transient parking declines. If the effective marketing area of the parking facility includes attractive retail destinations, price concessions may stimulate the use of the facility. If the local retail destinations are large, they might even be interested in subsidizing parking fees to encourage their business. Downtown retail shopping districts are often at a severe disadvantage to suburban shopping malls on the weekends. Many parking facilities use low flat rates to entice weekend parkers.

Holidays

The demand for holiday parking shares many of the same attributes as the demand for parking on weekends. Offices and business are closed, and the need for short-term and transient parking declines. An exception to this may occur over the Christmas holiday season, where the demand for retail shopping is so great that parking demand will approach or exceed normal weekday demand. Under these circumstances, the normal rate schedule would apply.

Monthly-Contract Users

The demand for this market segment arises from potential patrons with employment in the effective marketing area. Depending on the degree of competition, such demand tends to be inelastic. Because the cost is large and planned to the parker, price shopping will occur if at all possible. This situation results in a wide dispersion of rates whose ultimate determination will depend on the particular circumstances of a specific facility. Monthly-contract parking rates vary between 10 and 20 times the maximum daily rate. At 21.7 times the MDR, the entire work year of 260 parking days is covered. Providing amenities such as desirable locations within the parking facility, special entrances, special exits, and expedited ticketing often mitigates the expense to monthly contract users.

HOTELS AND MOTELS

Hotels and *motels* provide wonderful opportunities for passive investors looking for excellent risk-return opportunities, or for active investors short on cash but long on a desire to work hard to build sweat equity in a business. Despite some unevenness in demand, opportunities for successful hotels and motels are likely to increase in the future. Travel and guest lodging are luxury goods in an ever more affluent society. Our increasingly mobile lifestyle assures a constant increase in the demand for lodging services.

Overall, the "hospitality" business (which offers lodging and food) has tended to the strong secular growth characterizing an affluent society one might expect. Americans will travel on just about any excuse, including trips to resorts, the ocean, or national and local parks; visiting relatives; class reunions; weddings and funerals, or just to sight-see. However, the industry is prone to shocks affecting the confidence of travelers or the cost of travel. The events of 9/11 had a significant impact on travel throughout the nation, as well as New York City itself. In the spring of 2003, sniper killings in the Washington, D.C., and northern Virginia area resulted in a dramatic fall in lodging occupancy in that region. The problems with the cost and availability of gasoline in the 1970s also had a significant impact on the demand for away-from-home lodging. While the supply of gasoline does not appear to be the problem that it was in the 1970s, the price will likely have a significant effect on the prospects of this investment class.

The industry is characterized by a large number of segments providing more or less good substitutes for each other in a given lodging market. Hotels may be huge skyscrapers, costing hundreds of millions of dollars, capable of hosting thousands of guests located in the center of the city next to a convention center or sports arena. At the other end of the spectrum, lies the 10-unit, "ma and pa" guest cottage on the side of a stream in a rural area adjacent to a state forest. It is important to note that there are at least a dozen market segments between these two extremes, which cater to a wide variety of tastes, preferences, and expectations.

INDUSTRIAL SITES

Industrial sites are seen to offer a potentially good risk and return combination for investors. When the industrial firm is unable or unwilling to finance its own properties, their alternative is to have an investor supply that capital in the form of a build-to-suit leasing arrangement. There is little doubt that this arrangement favors the industrial firm, especially where it is subject to a high tax rate. However, it is often advantageous to the commercial real estate investor as well.

Risks exist for the real estate investor in this situation, but it can be controlled for with foresight. The first risk to be encountered by the investor is that of nonrenewal of the lease. Because the taxing circumstances will normally dictate the use of an operating lease, the lease cannot be fully amortized over the property's life. The investor can address this issue by designing the building so that it has use in a wide number of applications and that it is located in an area that would be attractive to other industrial firms.

A second risk could be default by the lessee. This risk can be addressed by assessing the financial solvency of the lessee prior to committing funds. In the event of default, the risk to the lessor is minimized by the fact that the property is retained and the industrial firm's creditors cannot attach the property.

SUMMARY

All in all, commercial real estate has performed very well historically. Further, commercial real estate opportunities include the obvious plays such as hotels, shopping centers, and office buildings. However, the less obvious categories such as parking lots, self-storage facilities, and recreational facilities may also fit well into one's investment portfolio. Other types of commercial real estate not discussed in this chapter are timberlands (see Wilson, 2000) and agricultural real estate (see Wilson, 2000).

REFERENCES

Beyard, M. D., and O'Mara, W. P. (1999). *Shopping Center Development Handbook*. Washington, DC: Urban Land Institute.

Coyle, S. M. (2000). The U.S. office market. In S. Hudson-Wilson (ed.), *Modern Real Estate Portfolio Management* (pp. 75–84). Hoboken, NJ: John Wiley & Sons.

Fenker, R. M. (1996). *The Site Book: A Field Guide to Commercial Real Estate Evaluation*. Ft. Worth, TX: Mesa House Publishing.

Geltner, D. M. (2007). *Commercial Real Estate Analysis and Investments*, 2nd edition. New York: Southwestern Publishing Co.

Haight, G. T., and Singer, D. D. (2005). *The Real Estate Handbook*. Hoboken, NJ: John Wiley & Sons.

Parking Consultants Council. (2000). *The Dimensions of Parking*. Washington, DC: Urban Land Institute.

Real Estate Brokerage Managers Council. (1996). *Real Estate Office Management*. Chicago: Dearborn Financial Publishing.

Rutes, W. A., Penner, R. H., and Adams, L. A. (2001). *Hotel Design, Planning, and Development*. New York: W. W. Norton & Co.

Senn, M. (ed). (2005). *Commercial Real Estate Transactions Handbook*, 3rd edition. New York: Aspen Law and Business.

Schmitz, A., and Brett, D. (2001). *Real Estate Market Analysis: A Case Study Approach*. Washington, DC: Urban Land Institute.

White, J. R., and Gray, K. D. (eds). (1996). *Shopping Centers and Other Retail Properties: Investment, Development, Financing, and Management*. New York: John Wiley & Sons.

Wilson, J. D. (2000). Agricultural real estate. In S. Hudson-Wilson (ed). *Modern Real Estate Portfolio Management* (pp. 165–174). Hoboken, NJ: John Wiley & Sons.

Wilson, J. D. (2000). Timberland investments. In S. Hudson-Wilson (ed). *Modern Real Estate Portfolio Management* (pp. 149–164). Hoboken, NJ: John Wiley & Sons.

Zankel, M/L. (2000). *Negotiating Commercial Real Estate Leases*. Ft. Worth, TX: Mesa House Publishing.

CHAPTER 50

Commercial Real Estate Loans and Securities

REBECCA J. MANNING
Vice President, Harbor Asset Management

DOUGLAS J. LUCAS
Executive Director and Head of CDO Research, UBS

LAURIE S. GOODMAN, PhD
Co-head of Global Fixed Income Research Manager of U.S. Securitized Products Research, UBS

FRANK J. FABOZZI PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: Commercial real estate is a cyclical industry, subject to local, regional, and national economic conditions. It is also a capital intensive business, requiring funding for initial development as well ongoing maintenance and improvements. Consequently, commercial real estate finance has traditionally been dominated by banks, life insurance companies and private investors with the long-term investment horizons and the access to capital required by the industry. However, commercial real estate finance has evolved into a public market with more liquidity and transparency, which has attracted a broader range of investors. In addition, commercial real estate–related investments have become more complex, bringing new opportunities and new risks to investors.

Keywords: commercial real estate loans, whole loans, A-notes, B-notes, mezzanine loans, commercial mortgage-backed securities (CMBS), REIT securities, loan origination, master servicer, special servicer, call protection mechanisms, prepayment risk, extension risk, interest shortfalls

A *commercial real estate loan* is secured by a commercial real estate property, such as an office building or by an interest in the entity that owns the property. The principal and interest on the loan are generally paid from cash flows generated by the property. Real estate borrowers, or sponsors, will take out loans to purchase properties, refinance existing debt, or add on to an existing loan.

Over the years, commercial real estate finance has evolved from simple first-lien mortgage loans on commercial real estate properties to a variety of different types of loans and real estate–related securities. The most common real estate loans and securities in the market today include:

- Commercial real estate loans
  - Whole loans and A-notes
  - B-notes
  - Mezzanine loans
  - Preferred equity
- Commercial mortgage-backed securities
- REIT securities

While all of these investments are on some level supported by real estate properties, their risks are considerably different, depending not only on the type of loan or security, but also on the underlying property type, geographic location, and tenant concentration, to name a few differences.

In this chapter, we explain different types of commercial real estate (CRE) loans and securities, analyzing the structures, investment considerations, and the risks of CRE loans and CMBS and REIT securities. (For a discussion of the historical performance of commercial real estate loans and CMBS, see Chapter 9 in Lucas, Goodman, Fabozzi, and Manning [2007].)

LOAN ORIGINATION

To obtain a loan on a commercial property, a sponsor typically turns to a commercial loan originator. Originators include commercial banks, insurance companies, real estate investment trusts (REITs), commercial mortgage-backed security (CMBS) conduits, and CRE collateralized debt obligations (CDOs). Originators may keep the loans for their own portfolios, or sell the loans in the secondary market. Others, particularly CMBS conduits, may also serve as warehouses, collecting a pool of loans, often referred to as "conduit loans," to later be securitized as CMBS.

Loan originators underwrite loans and determine the appropriate loan structure and terms based on the results of their due diligence. The performance of the loan is often related to the quality of the underwriting done by the originator. In fact, rating agencies will look at the performance history of the loans underwritten by the originator when assigning a rating to a new loan.

Many of the loans originated today are pooled to create CMBS. CMBS issuers perform their own due diligence on each loan in the pool. In addition, they look at the pool on an aggregate basis, assessing portfolio risks such as concentrations of property type, geography, and loans. The pool of loans is then tranched into individual securities and sold to third-party investors. We discuss CMBS later in this chapter.

Underwriting a Loan

An originator's due diligence includes verifying a property's value, cash flow, and credit quality. Originators typically require and review:

- Current property appraisals.
- Current leases and rent rolls.
- Tax filings and bank statements.
- Tenant credit quality.
- Site inspections.
- Environmental and engineering reports from reputable firms.
- Title insurance and other legal property documents.
- Lockbox provisions requiring that all revenues generated by the property be collected by a trustee, who first pays all operating expenses, debt service, and any other expenses. Excess cash flow is then distributed to the sponsor.
- Escrow accounts holding cash reserves to meet unexpected cash shortfalls. The amount typically equals one month's debt service, real estate taxes, property insurance, and sometimes re-leasing costs.
- Reserve accounts holding cash reserves for property maintenance and pending repairs.

A lender, particularly the most senior lender in the property's capital structure, often requires cash management provisions, such as lockboxes and escrow/reserve accounts on highly leveraged properties. Higher leverage increases the stress on a property's cash flows, ultimately increasing the risk and severity of losses. Cash management provisions are important controls to ensure that the sponsor and property managers operate and maintain the property efficiently.

Most CRE loans are nonrecourse. That is, in the event of default, the lender's claim is to the property only; the sponsor is not personally responsible to cover any losses. However, originators usually require nonrecourse carveouts, holding sponsors personally liable for fraud, misrepresentation, misappropriation, and environmental issues. Most loans also require environmental indemnifications protecting lenders from third-party claims related to property environmental conditions.

Typically, a good originator has expertise not only in real estate, but also in the particular type of property (office, industrial, etc.), the local real estate market, and the type of financing desired. Good originators also perform thorough due diligence on every property underwritten. The performance of loans previously underwritten by an originator can provide insight into the quality of that originator's underwriting practices.

The continued strength of the real estate market and the resulting demand for CRE loans have made the loan origination business much more competitive. As a result, some originators have relaxed their underwriting standards, for example, taking on loans with higher leverage, making more aggressive property performance assumptions, or waiving reserve requirements. Given this trend, the quality and motivation of the originator are increasingly important. For instance, originators looking for market share may be more willing to relax their underwriting. While relaxed standards alone may not spell disaster, the terms of the loan, that is, the interest rate and covenants, should be appropriate for the higher levels of risk.

The Master and Special Servicers

At origination, a *master servicer* and a *special servicer* are appointed. In the event of a short-term cash shortfall, the master servicer advances principal and interest payments to the lender and pays real estate taxes and insurance premiums up to the amount the servicer is likely to recover. The master servicer also monitors documents required by the loan, such as annual property performance reports. For performing these ongoing services, the master servicer earns a fee based on a percentage of the outstanding principal balance of the loan.

The special servicer is appointed to resolve issues relating to a delinquent or defaulted loan. Usually, the master servicer hands over a loan to the special servicer when the loan is more than 60 days delinquent. The special servicer's role is to maximize the amount recovered from a defaulted loan and minimize loan losses. The special servicer is compensated by a fee on the principal balance of the assets it is monitoring (often twice the master servicer's fee), an additional workout fee for loans in default, as well as a percentage of the loan's principal and interest recovered through workout. Typically, the special servicer also has an equity interest in the property, increasing the motivation for successfully working out and remedying the defaulted loan.

PROPERTY-LEVEL LOANS

The most basic commercial real estate loan is a first-lien mortgage loan. The first-lien mortgage loan (also known as the "mortgage" or "*whole loan*") is the senior-most loan secured by the property. At origination, a mortgage's principal balance is typically 65% to 80% of a property's appraised value, commonly referred to as the LTV, or loan-to-value. The mortgage can be split into a senior and a subordinate piece, the *A-note* and the *B-note*.

The remaining 20% to 35% is the sponsor's equity interest in the property. However, a sponsor typically targets a 0% to 15% equity interest, depending on the sponsor's motivation. To increase the leverage on a property, a sponsor can take out a mezzanine loan. A *mezzanine loan* is a senior participation in the equity in the property. The loan is not secured by the property itself, but by an interest in the entity that owns the property (the sponsor). A mezzanine loan essentially reduces the sponsor's equity interest in the property. The loan can raise total leverage on the property to 85% to 100% LTV. In other words, the sum of the mortgage(s) plus the mezzanine loan can equal 85% to 100% of the property's appraised value. As a result, the sponsor's equity interest in the property can be reduced to 0% to 15% of the property's value.

A sponsor may also take out a second-lien mortgage on a property to reduce the equity contribution. Similar to a mezzanine loan, a second-lien mortgage is junior to the first-lien mortgage. However, unlike a mezzanine loan, a second-lien mortgage is secured by the property directly, rather than by an interest in the property's equity. As such, a second-lien mortgage increases the senior debt's risk of default and loss more so than a mezzanine loan. Therefore, first-lien mortgage lenders rarely allow sponsors to take out second-lien mortgages.

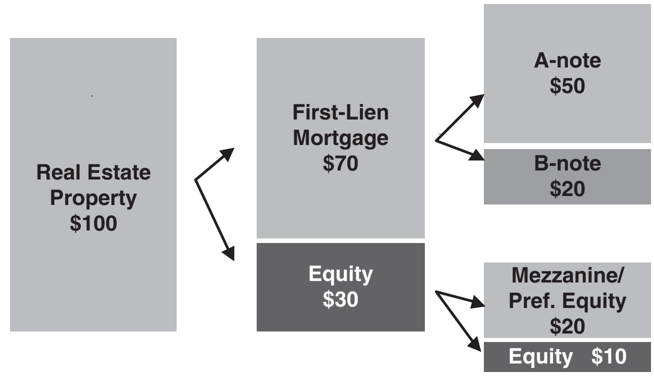


Figure 50.1 Typical Property Capital Structure

Some mortgages also prevent a sponsor from taking out a mezzanine loan. In such cases, the sponsor can instead issue preferred equity. Preferred equity in a property is essentially the same as an equity interest, but has a senior claim on the excess cash flow available after servicing the loans on the property.

Figure 50.1 illustrates a single property's capital structure. Each type of property loan, from A-note through mezzanine, has distinct terms, structure, and risks that we will discuss in the next few sections.

First-Lien Commercial Mortgage Loans

First-lien commercial mortgage loans range from $300,000 to $1 billion. They are typically 10-year balloon loans with 30-year amortization schedules, although there is an increasing number of interest-only loans. Most commercial mortgage loans are fixed rate. Generally, interest rates are 75 to 150 basis points above the 10-year Treasury, but they vary depending on leverage and other property-specific factors.

Prepayment Risk and Extension Risk

Unlike residential mortgages, commercial mortgages have low prepayment risk, thanks to the numerous call protection mechanisms built into the loan terms. Call protection mechanisms include the following:

- **Lockout.** Prepayments are prohibited during a 2- to 5-year lockout period.
- **Yield maintenance.** Equivalent to a "make-whole" premium in corporate bonds. To prepay a loan, the sponsor must make the lender whole. The yield maintenance cost is equivalent to the present value of all the future cash flows due on the loan, discounted at the then-prevailing yield of a comparable maturity U.S. Treasury.
- **Defeasance.** To defease a loan, the sponsor must pledge to the lender U.S. Treasury securities that generate cash flows equal to the cash flows due to the lender under the terms of the loan. From the lender's viewpoint, cash flows are the same, but the underlying collateral will be upgraded to U.S. Treasury securities.

- **Prepayment penalty points.** In prepaying a loan, a sponsor pays the lender a fee equal to a set percentage of 1% to 5% of the outstanding loan balance. Penalty points usually decrease as the remaining life of the loan decreases, e.g., 5% in Year 6, 4% in Year 7, 3% in Year 8, and so on.

For a more detailed discussion of these mechanisms, see Cheng, Cooper, and Huang 1999.

The most common call protection mechanism is a combination of lockout for the first five years, followed by defeasance, which remains in effect until approximately six months before maturity. The sponsor then has a six-month window to refinance the loan without penalty.

Call protection mechanisms lessen the economic incentive to refinance a commercial mortgage. Essentially, a sponsor is likely to prepay a commercial mortgage only if the property is being sold and the gain on the property exceeds the cost of prepaying the mortgage.

Although there is little prepayment risk in commercial mortgages, there is refinancing risk. The 10-year balloon structure of mortgage loans makes extending the loan past the 10-year maturity unlikely, which is a plus for many investors. If the mortgage is not paid off or refinanced at maturity, the loan is in default.

The ability to refinance at maturity, however, depends on several factors that are often out of the sponsor's and the lender's control, such as prevailing interest rates, the strictness of current underwriting requirements, credit conditions, and property occupancy at the time of refinancing. Some loan originators allow short-term extensions, but historically there are significant disincentives for extending. However, in strong markets, originators may relax these disincentives, thereby introducing greater extension risk into commercial mortgages.

Some commercial mortgages are partially or fully interest-only loans, which contribute to extension risk. An interest-only loan faces more extension risk because its principal has not amortized, therefore leaving a larger outstanding loan balance to refinance.

The A/B Structure

A mortgage is often split into a senior and junior participation, the A-note and the B-note. In the A/B structure, as it is called, the A-note has a senior claim on cash flows generated by a property, while the B-note has a subordinate claim on cash flows. (The first lien mortgage can also have an A/B/C structure. This is similar to the A/B structure, but the C-note becomes the most junior note in the structure. Since the concept is very similar, and less common, than the A/B structure, we focus our attention in this chapter on the A/B structure.)

Payment of principal and interest can either be pro rata or sequential. The desired rating on the A-note determines the size of each piece. Typically, the A-note is sized for a BBB or BBB– rating, while the B-note is below investment grade or unrated. The B-note provides credit enhancement and essentially reduces the leverage of the A-note. For example, while the LTV of the entire mortgage may be 80%, the LTV of the A-note would only be 65%, assuming an 80/20 split between the A-note and B-note.

The A-note is usually placed in a trust for securitization. The B-note, on the other hand, is held by a third party, often an experienced real estate investor. In the worst-case scenario, the B-note holder could essentially become the equity owner of the property. Therefore, the B-note holder is usually experienced in underwriting, monitoring, and, if need be, remedying property performance. The B-note holder is compensated for that increased risk position, though returns depend on the underlying property's characteristics, most notably its leverage. Spreads on B-notes can range from as low as 75 basis points to more than 1,000 basis points above Treasuries.

If a sponsor defaults on a mortgage loan, the A-note holder has the right to foreclose and take possession of the property. In this case, the B-note holder loses all collateral securing the note, and is essentially left with an equity interest in the property. This process can take 6 to 18 months, during which time the value of the property may deteriorate, thus increasing the B-note holder's risk of losses. To avoid this scenario, the B-note holder is granted specific rights, which are outlined in the participation agreement and the pooling and servicing agreement.

For instance, in the event of a mortgage default, the B-note holder has the right to cure the default. The B-note holder would thus pay the principal and interest due to the A-note holder, plus any accrued interest, legal fees, and advances. B-note holders are likely to exercise this right if there is sufficient property value above the principal of the A-note and the B-note, or if the property is a transitional property.

A transitional property is a poorly performing property, where performance and value can be increased by an improvement in the overall real estate market, improving the property through capital expenditures or new leases, or by replacing the existing property manager. In such cases, the B-note holder typically has experience in turning around properties. The B-note holder has 3 to 6 months to exercise her right to cure the mortgage.

The B-note holder also has the right, if the mortgage is in default, to buy out the A-note holder. In this case, the B-note holder becomes the senior mortgage lender and gains full control over the entire debt structure of the property, and can foreclose on the property at any time. To exercise this right, the B-note holder must pay the A-note holder the value of the A-note plus accrued interest, legal fees, and advances. A B-note holder with workout experience is likely to choose this option if the economics make sense.

If the B-note holder chooses *not* to exercise either the right to cure the defaulted mortgage or the right to buy out the A-note holder, she still has the right to approve the special servicer and the terms of the workout plan for the defaulted loan. This allows the B-note holder some control over the workout process, which directly impacts the level of potential losses the B-note holder may realize.

In addition to rights in the event of default, the B-note holder has predefault rights. For instance, the B-note holder typically has the right to approve:

- Annual property budgets.
- Key tenant leases.

- Property management and leasing agents.
- Transfer of the property by the equity holders.
- Escrow/reserve disbursements.

Approval rights over the budget, leases, property management, leasing agents, and property transfer give the B-note holder some control over the property's performance. Rights over escrow/reserve disbursements provide the B-note holder with protection over improper cash flow distributions to equity holders. B-note holders can also institute an excess cash flow trap to redirect cash flows from equity holders if the property's cash flows trip a specified trigger. The excess cash trigger is typically set at a minimum debt service coverage ratio, a measure we discuss later in this chapter.

Mezzanine Loans

Mezzanine loans are the junior-most loans in a property's capital structure. They enable a sponsor to increase a property's leverage, raising total loan-to-value ratio (LTV) to 85% to 100%. These loans typically have a minimum size of $3 million.

Mezzanine loan terms depend on the sponsor's motivation. If current rates are high, the sponsor may opt for a first lien mortgage with a low LTV, then supplement it with a short-term mezzanine loan to reduce her equity contribution. This arrangement allows the sponsor to refinance the mezzanine loan (or the entire mortgage, if the economics work) at a later date when rates are lower, or when the property is performing better. Alternatively, a sponsor may choose a mezzanine loan coterminus with the mortgage to take maximum advantage of an arbitrage opportunity. Therefore, maturities on mezzanine loans range from 18 months to 10 years; some are amortizing, some interest-only, depending on the property and the sponsor's preferences.

As mentioned earlier, a mezzanine loan is not secured by the property itself, but by an interest in the entity that owns the property. It is the first loan to absorb any losses or cash flow shortfalls. Therefore, mezzanine lenders demand a higher interest rate than A-note holders, sometimes significantly higher (upwards of 1,000 basis points) depending on the property and its leverage. Historically, mezzanine loans are held by experienced third-party real estate investors.

Similar to B-note holders, a mezzanine lender has specific rights to protect her investment and minimize losses. These rights are outlined in the intercreditor agreement between the mortgage lenders and the mezzanine lender. For example, if a sponsor defaults on the mezzanine loan, while the first lien mortgage is still current, the mezzanine lender has the option of foreclosing on the sponsor and taking control of the property (subject to the terms of the property's existing mortgage).

Foreclosing on the sponsor is generally quick, taking 60 to 90 days rather than the 6 to 18 months it would take an A- or B-note holder to foreclose on the property. Therefore, the mezzanine lender can gain control of the property more quickly than a B-note holder if property performance goes south. The quicker a lender can take control of a property in default, the sooner she can take actions to remedy or turn around the property, thus minimizing potential losses.

In the event the sponsor defaults on both the mortgage and the mezzanine loan, either the A-note or B-note mortgage lender can foreclose on the property. In this case, the sponsor no longer owns the property; the foreclosing lender does. The sponsor therefore has no collateral, and since the mezzanine lender is secured by an interest in the sponsor, the value of the mezzanine loan goes to zero.

To protect the mezzanine lender in the event of default on the mortgage, the mezzanine lender has rights similar to those of a B-note holder. First, the mezzanine lender has the right to cure the mortgage. This is identical to the right of the B-note holder, but the mezzanine lender has an unlimited amount of time to exercise this right, whereas the B-note holder has to exercise within three to six months. The mezzanine lender also has the option to buy out the mortgage from the mortgage lenders and take control of the property's entire capital structure, with the right to foreclose at any time. Which right the mezzanine lender exercises depends on her real estate expertise.

A mezzanine lender also has predefault rights, similar, though junior to the B-note holder's predefault rights as outlined above. The mezzanine lender also has the right to approve any refinancing of the mortgage. Furthermore, the mortgage lender is not allowed to make any changes to the mortgage loan documents that would be detrimental to the mezzanine lender, such as raising the mortgage rate.

Other CRE Loans

The strong performance of real estate assets over the past decade has increased demand for alternative types of CRE loans. The most common include construction loans, condo- and co-op conversion loans, and land loans. Each carries distinct risks that require additional consideration over the more traditional mortgage and mezzanine loans.

Construction loans, for example, are secured by properties that are under construction. These properties are therefore non–cash flowing, or are generating very little cash. The sponsor usually sets up a reserve account at loan origination which pays the loan's interest. Loan repayment is contingent on construction completion, at which point permanent financing (or temporary bridge financing if the property is not yet leased at completion) is put in place.

These loans tend to be floating rate with maturities of 12 to 36 months and are funded in stages as construction costs are incurred. The loan amount is determined by the construction budget plus a 10% to 20% contingency. Loan performance depends on the sponsor's credit quality and her expertise in managing and monitoring the construction process.

Rating agencies generally consider alternative loans to be riskier than first lien mortgages, so more careful underwriting and monitoring is required. Inclusion of these types of loans in CRE and CMBS CDOs often results in higher subordination requirements.

COMMERCIAL MORTGAGE-BACKED SECURITIES

Commercial mortgage-backed securities (CMBS) are backed by a static pool of commercial mortgage loans, the vast majority of which are A-notes. The pool of loans is tranched into a number of rated tranches, and principal and interest payments received from the underlying loans are used to pay principal and interest to the tranches sequentially by seniority. Any losses experienced by the underlying loan pool are absorbed, in order, by the most junior tranches.

Figure 50.2 illustrates the typical structure of a CMBS transaction. The AAA-rated tranche makes up a large portion of the debt structure, generally around 90%, and can be time-tranched into 5- and 10-year securities. Interest-only securities are often included in the structure as well.

CMBS deals appoint a master servicer to monitor the cash flows coming from the underlying loans and going out to the tranches. In the event there is insufficient cash to make all scheduled payments, the master servicer will advance principal and interest. Advancing will continue as long as these amounts are deemed recoverable. A special servicer is also appointed to handle any loans that are more than 60 days delinquent.

Property-type diversification is one of the principal benefits of securitization, since the performance of each property type is impacted by different sets of risks. Properties securing commercial real estate loans include office buildings, industrial buildings or warehouses, apartment buildings, hotels, and retail properties such as strip malls. The concentrations of different property types vary slightly over time, depending on collateral performance. In general, office, retail, and multifamily properties tend to dominate most CMBS, historically accounting for two thirds of the collateral in CMBS deals.

CMBS investors as a whole are a diverse group, but particular types of investors are drawn to different tranches of CMBS deals. Real money investors, financial institutions, insurance companies, etc., tend to buy investment-grade tranches. Traditional real estate investors, hedge funds, and CRE CDOs enter the mix further down the capital structure. Buyers of tranches rated BB+ and below, collectively referred to as the "B-piece," are typically real estate investors, with expertise to underwrite the loan portfolio and accurately assess the risks of the B-piece's first loss position. Top B-piece buyers include LNR Partners, American Capital Strategies, ARCap, J. E. Roberts, and CWCapital.

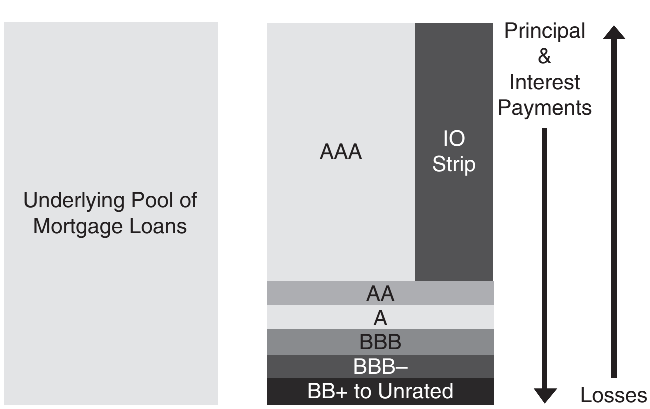


Figure 50.2 Typical CMBS Structure

Prepayment and Extension Risk

The *prepayment* stability offered by CMBS attracts investors looking for real estate exposure without the negative convexity found in residential mortgage-backed securities. Unless interest rates drop dramatically or property values soar, refinancing of the underlying loans is uncommon, due to the call protection mechanisms on the loans.

However, in a robust commercial real estate market, prepayments in the form of defeasance increase, as sponsors look to cash in on property price appreciation. Defeasance, however, is actually a plus for CMBS, as cash flows to the tranches remain the same, with their source becoming Treasury securities, thus raising the credit quality of the tranche's underlying collateral.

The sequential pay structure of CMBS itself provides additional prepayment protection to the junior tranches of the deal. The AA-rated tranche cannot be paid down before the AAA tranche is completely paid. And the A tranche cannot be paid down before the AA tranche, and so on. Therefore, the A tranche is guaranteed to remain outstanding in full at least until the AA tranche is completely paid down. However, principal losses from loan defaults impact the bottom of the CMBS structure upward, so the principal balance of the more junior tranches may be reduced due to principal writedowns.

Unfortunately, CMBS have *extension risk*, due to the 10-year balloon maturities of the underlying loans. If a loan cannot be refinanced and it defaults, the loan enters a workout period that can extend principal recovery for months, even years. The most subordinate tranches of a CMBS transaction bear the most extension risk, as they are the last to receive principal payment.

Interest Shortfalls

Interest shortfalls can be a concern to CMBS tranches, especially to the below investment grade and unrated tranches. Interest shortfalls occur when the CMBS transaction has insufficient funds to pay interest due tranches. Shortfalls are more common in CMBS transactions than other asset-backed transactions, yet they are usually less serious occurrences.

In other asset-backed securities, basis mismatch is the chief factor in interest shortfalls. Basis mismatch frequently occurs in rising interest rate environments when the coupons payable to the tranches reset at higher rates before the coupons on the underlying collateral reset, thus causing an interest shortfall. This type of interest shortfall is not a concern in CMBS, because CMBS have little, if any, inherent basis risk, since both the underlying loans and the coupons payable on the tranches are typically fixed rate.

Most interest shortfalls in CMBS are due to loan delinquencies and defaults. Except when the default is severe enough to cause a loss on the loan, interest will be recovered in part or in full when the sponsor cures the loan or when the property is liquidated. The CMBS master servicer advances funds to cover interest shortfalls that are likely to be recovered. Due to high collateral concentrations and long workout periods for defaulted loans, CMBS tranches can experience interest shortfalls for long periods of time, yet retain high likelihood of recovery.

Most CMBS interest shortfalls are recoverable, but some are not. For instance, the master servicer will not cover shortfalls resulting from delinquent loans where the underlying property has been reappraised to a negative LTV. In this case, the CMBS tranches, particularly the lowest rated, will absorb the shortfall as a loss. The master servicer does not cover shortfalls due to nonreimbursable costs such as litigation expenses or workout fees, either. Unrecoverable interest shortfalls are concerns for CMBS investors. Recent high-profile interest shortfalls in a few CMBS transactions have caused tranche downgrades and principal losses, with a few tranches losing 100% of their principal, thus increasing investor concerns over interest shortfalls.

CMBS deals with interest shortfall issues include: Asset Securitization Corp. 1996-D2 due to delinquent properties being re-appraised at lower values; Morgan Stanley Capital I 1998-CF1, JP Morgan Chase Commercial Mortgage Securities Corp. 2003-FL1, LB Commercial Mortgage Trust Series 1998-C4, and Bear Stearns Commercial Mortgage Securities Series 2001-TOP2, all due to unrecoverable fees.

Types of CMBS Deals

CMBS provide investors with exposure to a diversified pool of commercial real estate loans. Variations in loan size, sponsor, property type, geographic location, leverage levels, and the like, all contribute to the diversification benefits of CMBS. These characteristics underlie the performance of the collateral and as a result impact the CMBS tranche ratings and subordination levels.

CMBS deals are often categorized into four groups, depending on the type of loans underlying the deal, as shown in Table 50.1. Conduit loans are mortgage loans originated by conduit lenders for the sole purpose of securitizing them. These loans tend to be small in size, generally less than $10 million, but they can be larger. Almost all conduit CMBS deals are fusion deals which consist of a diverse pool of small loans as well as a small number of larger loans. Conduit CMBS deals have historically dominated the CMBS market.

Large loans are property loans that are greater than $35 million in size. Credit tenant leases (CTLs) are often the result of sale-leaseback transactions. The tenant sells corporate-owned real estate and enters into a long-term lease on the property or properties. The leases are structured such that the default risk of the lease is tied to the credit rating of the tenant. The tenant's default risk is fundamentally different than the default risk of a traditional commercial real estate loan, and can often gain a higher credit rating. Single-asset CMBS deals contain just one, large loan on a single property, such as, a large high-rise office building. Single-borrower CMBS deals contain a pool of loans on properties with the same borrower or sponsor. Often pool loans are cross-collateralized, so that if one loan defaults, the lender has recourse to any or all of the pool's properties.

Table 50.1 CMBS Deal Categories

Deal Categories	Description
Conduit loans	
Traditional	Pool of loans where no loan is greater than 10% of the total principal of the deal.
Fusion	Pool of loans where a few of the loans are greater than 10% of the total principal of the deal.
Large loan	Pool of loans where several loans are greater than 10% of the total principal of the deal.
Credit tenant leases	Pool of loans secured by property leases.
Single asset/ Single borrower	One loan secured by a single property or a pool of loans from the same borrower.

REIT SECURITIES

Real estate investment trusts (REITs) are entities that buy, develop, manage, and sell real estate assets. A special feature of REITs is that they qualify as pass-through entities which are exempt from corporate level taxes. To qualify as a REIT, the entity must pay out dividends equal to 90% of its taxable income and more than 75% of its total assets must be in real estate. A REIT generates income through the operation and management of real estate assets. Sales of asset held less than four years cannot exceed 30% of the REIT's net income. Therefore, a REIT is clearly a "buy and hold" entity, not an asset flipper or a trader.

REITs fall into three broad categories: equity REITs, mortgage REITs, and hybrid REITs. Equity REITs own and operate a portfolio of real estate properties, usually focusing on a particular type of property such as office buildings. Mortgage REITs invest in, and in some cases originate, mortgage loans and mortgage-backed securities. Hybrid REITs combine the investment strategies of equity and mortgage REITs by investing in both properties and mortgages. Equity REITs dominate the REIT market, accounting for about 95% of total REIT market capitalization.

The REIT capital structure consists of secured bank loans, unsecured debt, preferred stock, and equity. Some REITs have also issued trust preferred securities (TruPS). (For more information on the mechanics of TruPS, see Chapter 8 in Lucas, Goodman, Fabozzi, and Manning 2007.) Unsecured REIT debt and TruPS have significant covenants to protect investors. A typical covenant package includes the following:

- Total debt cannot exceed 60% of total assets.
- Unencumbered assets must be at least 150% of unsecured debt.
- Secured debt cannot exceed 40% of total assets.
- Interest coverage must be greater than 1.5×.

Given these covenants, BBB-rated unsecured REIT debt is comparable to single-A-rated CMBS debt given the similar leverage and interest coverage levels. However, a REIT's asset portfolio, and therefore its financial ratios, can change over time, unlike the static pool of assets securing CMBS debt. Also, the REIT's debt is unsecured, while CMBS debt is secured by a pool of first mortgages. Therefore, unsecured REIT debt will likely be rated below CMBS debt that has similar leverage and interest coverage levels.

REIT securities are purchased by a variety of investors, from insurance companies, mutual funds, and CDOs to individual retail investors. REIT securities provide investors with exposure to a diversified pool of real estate-related assets with little to no negative convexity, as opposed to investments in residential mortgage-backed securities (RMBS). In addition, REITs resemble corporates more than CMBS or RMBS, which opens them up to a large investor base.

EVALUATING CREL AND CMBS

Analysis of commercial real estate investments, whether investments in B-notes or CMBS tranches, begins with an analysis of the underlying property, followed by an analysis of the loan terms. For CMBS investments, additional analysis is needed at the bond or equity level. We discuss these three types of analysis next.

Property-Level Analysis

The first step in analyzing real estate investments is a property-level analysis. Understanding the property, from the credit quality of the third-floor tenant to the conditions of the local and general economies, is fundamental in assessing the financial condition of the property. Property-level analysis includes many of the elements of underwriting a loan on a property, such as property appraisals, tenant and lease review, comparable property analysis, and the like. These components are used to estimate the stabilized cash flow of the property.

A property's value can be derived from its stabilized cash flow by applying a capitalization rate appropriate for the property. A capitalization rate, or cap rate, is essentially an idealized unlevered risk-adjusted return. Embedded in the cap rate are assumptions about the relative quality of the property, the cash-flow volatility common to that type of property, comparable property yields, yields on other types of investments, and so on. The value of the property is its stabilized cash flow divided by the appropriate cap rate. The lower the cap rate, the higher the resulting property value.

Loan-Level Analysis

Analyzing a CRE loan centers on two key metrics: debt service coverage ratio (DSCR) and loan-to-value (LTV). DSCR is the property's cash flow, less tenant improvements, leasing commissions, and necessary capital expenditures; divided by the debt service on the property's loans. DSCR is considered by the rating agencies to be the best indicator of default probability. The higher the DSCR, or the more cash flow a property has to cover debt service, the greater the property's ability to withstand adverse conditions before defaulting on any loan. S&P reports that the average DSCR on existing commercial mortgage loans is around 1.5×, although many loans originated today are sized at 1.2× (see Thompson, Kay, and Ramkhelawan, 2006). For the second quarter of 2006, 21% of the loans in Moody's rated conduit CMBS were sized at 1.2× or less DSCR, up from just 6.3% a year earlier (see Philipp, Obias, Dent, and Rubock, 2006).

LTV is calculated as the principal loan balance divided by the estimated value of the property. Rating agencies consider the LTV to be the best indicator of loss severity in the event of default. The lower the LTV, or the lower the amount of the loan as a percentage of the property's value, the lower the odds that the loan will suffer losses in the event of default. LTVs can range from 65% to 90%+. S&P reports that the average LTV on securitized first-lien commercial mortgage loans is 69%, although many loans are originated at 80% LTV (see Thompson, Kay, and Ramkhelawan, 2006).

A loan's default probability and loss severity are used to determine expected losses on the loan. When rating agencies calculate expected loss, they often apply qualitative adjustments to the metrics. For example, Fitch may adjust a property's default probability upward if the property's cash flows tend to be volatile. The default probability may also be adjusted upward if the loan is floating rate, as floating-rate loans introduce more variability into the debt service costs. A property's loss severity may be adjusted either upwards or downwards based on the type of loan, be it a whole loan or a mezzanine loan, the strength of the loan covenants, loan amortization, additional debt, and so on. Another consideration is the thickness of the debt. A small piece of debt at the bottom of the capital structure is more likely to be wiped out, even if overall losses are small. Finally, Fitch will adjust the resulting expected losses for reserves, the quality and underwriting practices of the loan originator, potential environmental issues, and so on.

For B-notes, mezzanine loans, and whole loans that have not been securitized, prepayment risk becomes an issue. The financial condition of the property at origination plays an important role in the likelihood of loan prepayment. Most B-notes, mezzanine loans, and whole loans that end up in CRE CDOs are secured by interests in transitional properties or highly leveraged properties. The cash flows on these types of properties tend to be more volatile, and therefore the financing costs tend to be higher. Upon stabilization, the loans can be refinanced on more favorable terms. In such cases, the loan terms tend to provide more prepayment flexibility than do the loan terms on stabilized properties.

Ideally, investors in CMBS and CRE CDOs, especially noninvestment grade and equity investors, will perform both property-level and loan-level analysis on every loan underlying the CMBS or the CDO. However, this type of in-depth analysis is not always possible, and in the case of a CDO, this is partially why investors pay management fees. Nonetheless, prudent investors will do substantial homework on the underlying properties and loans. Analysis on properties, or sponsors, that make up a large percentage of the pool backing the CMBS or CDO will give investment-grade tranche investors some confidence in the performance of the overall pool, while analysis of the entire pool is best for non-investment-grade and equity investors.

CMBS Bond-Level Analysis

For CMBS investments, additional bond-level analysis is required. This includes looking at the pool of loans in a CMBS trust as a whole and assessing the collateral concentrations. Property-type, geographic, and loan-type concentrations are all important pool characteristics that impact the likelihood and correlation of defaults as well as losses. The ratings of the underlying collateral and the pool's weighted average rating are important pool characteristics, as well. In addition, the rating agencies calculate a CMBS pool's Herfindahl score, which is a measure of the effective number of assets in the pool, and accounts for concentrations due to loan size. The Herfindahl score, per Moody's, is calculated as follows for a pool consisting of N assets:

$$\text{Herfindahl score} = \frac{1}{\sum_{i=1}^{N}\left(\frac{\text{Principal balance of asset } i \text{ in the post}}{\text{Aggregate principal balance of the pool}}\right)^2}$$

Rating agencies determine required credit enhancement using these pool metrics, as well as the pool's overall DSCR and LTV.

In addition to analyzing the underlying pool, bond-level analysis requires cash flow modeling. Cash-flow modeling incorporates the specific structure of the CMBS, including the protective effects of overcollateralization and cash flow diversion mechanisms. With an accurate model of the CMBS tranche, cash flows can be tested for their response to various levels of default, recovery, prepayment, and other factors.

SUMMARY

As the commercial real estate market has evolved, so have CRE investments. In this chapter, we have reviewed the different types of CRE investments from first-lien mortgages to CMBS. We looked at different structures and showed how CRE loans and securities provide investors with levered exposure to price appreciating assets with the ability to customize that real estate exposure via credit enhancement and diversification. Then we discussed several factors to consider when investing in CRE, including property-level analysis, loan-level analysis, and bond-level analysis.

REFERENCES

Cheng, D., Cooper, A. R., and Huang, J. (1999). *Understanding prepayments in CMBS deals*. In F. J. Fabozzi and D. P. Jacob (eds.), *The Handbook of Commercial Mortgage-Backed Securities* (pp. 147–158). Hoboken, NJ: John Wiley & Sons.

Esaki, H., and Goldman, M. (2005). Commercial mortgage defaults: 30 years of history. *CMBS World* 6, 4: 1 21–29.

Fabozzi, F. J. (ed.) (2001). *Investing in Commercial Mortgage-Backed Securities*. Hoboken, NJ: John Wiley & Sons.

Lucas, D. J., Goodman, L. S., Fabozzi, F. J., and Manning, R. J. (2007). *Developments in the Collateralized Debt Obligations Markets: New Products and Insights*. Hoboken, NJ: John Wiley & Sons.

Philipp, T., Obias, P., Dent, P., and Rubock, D. (2006). US CMBS and CRE CDO 2Q 2006 review: Credit metrics and spreads send conflicting signals. *Moody's Investors Service*, July 31.

Snyderman, M. (1991). Commercial mortgages: Default occurrence and estimated yield impact. *Journal of Portfolio Management* 18, 1: 82–87.

Thompson, E., Kay, L., and Ramkhelawan, G. (2006). Defaults and losses of U.S. commercial mortgage loans: Year-end 2005 update reveals improved credit performance. *Standard & Poor's*, June 8.

CHAPTER 51

Commercial Real Estate Derivatives

JEFFREY D. FISHER, PhD
Dunn Professor of Real Estate, Indiana University

DAVID GELTNER, PhD
George Macomber Professor of Real Estate Finance, MIT

Abstract: Despite the tremendous growth in the use of derivatives for commodities, stocks, interest rates, currency and other applications, the availability of derivatives for commercial real estate has been limited. When you consider that real estate assets comprise over one-third of the value of all of the underlying physical capital in the United States and the world, the potential for real estate derivatives is impressive. It is therefore not surprising that in recent years real estate derivatives have begun to develop, as market participants have realized the role that derivatives can play, investment banks have been willing to offer derivatives, and new indices have been developed that are designed to meet the needs of the evolving real estate derivatives market.

Keywords: derivatives, forward, swaps, indices, hedge, alpha, short, structured notes, counterparty

This chapter discusses the type of derivatives now being offered for commercial real estate including total return *swaps, forward* contracts, and *structured notes.* Such products address several of the classical problems that have been raised regarding real estate investment, including: high transactions and management costs, lack of liquidity, inability to sell *short*, and difficulty making well-diversified property investments whose returns are measured in a manner comparable to those of stocks and bonds. This chapter also discusses the fundamentals of real estate return indices used to support derivatives including the both appraisal-based and transactions-based indices. A discussion of pricing of commercial real estate derivatives is provided in Chapter 52 of Volume III.

USES AND USERS OF COMMERCIAL REAL ESTATE DERIVATIVES

There are a myriad of potential uses of real estate *derivatives* by different market participants. Examples are exposure to the real estate asset class, *hedge* existing exposure, harvesting *alpha*, portfolio balancing, real value investing, and efficient leverage.

Exposure to the Real Estate Asset Class

Derivatives provide a way for investors to get exposure to the commercial real estate asset class relatively quickly, with relatively low transaction or management costs and relatively high diversification. This can be particularly useful for investors who lack the expertise to either purchase and manage individual properties directly, or find and manage specialized investment managers or real estate property funds. For example, a foreign investor who wants immediate, well-diversified exposure to the U.S. real estate market may want to take a long position on a real estate derivative such as a forward contract or a swap that is based on a national real estate index. Purchasing the derivative results in the equivalent of exposure to a well-diversified portfolio of properties and hence very little if any unsystematic risk without incurring the costs of purchasing and managing properties. Similarly, a small pension fund may want exposure to a well-diversified portfolio of real estate but lacks the scale to purchase enough individual properties to be well diversified by property type and location, and lacks the expertise to choose among property funds with their various investment management and transaction fees.

Investment managers who find they are over exposed to the real estate asset class, perhaps because real estate has performed well compared to their stock and bond portfolio, or because they have a relatively bearish outlook for real estate, may want to take a short position in a derivative to reduce their exposure to real estate without the need to sell properties, or until transactions can be completed on the sale of properties (which can take time to market and close). Shorting the derivative can also lock in profits made in the real estate market so the investment manager doesn't risk a drop in value before the properties can be sold. Lenders and originators of commercial mortgage-backed securities, exposed to either "warehouse" or portfolio risk, can hedge using a short positions in forwards or swaps, or by purchasing a put option, based on real estate indices. Credit default swaps can also be designed that result in a payoff to the party purchasing the swap that is triggered by the index's declining below a certain level.

Harvesting Alpha

Real estate investment managers who have the expertise to acquire, manage, and sell properties so as to persistently outperform the real estate market can monetize such positive alpha without selling properties, and produce profitable returns even when the real estate market turns down, by using the short position in the derivative to effectively "cover" their real estate market exposure, a "risk management" tool that acts effectively like real estate market value "insurance." This allows the investment manager to focus on their area of specialized expertise and comparative advantage, dealing and managing in the real estate market, regardless of the current ebbs or flows in the capital markets.

Portfolio Balancing

Real estate portfolio managers may also feel that their allocation to different property types or geographic locations has gotten out of balance. For example, they may feel that they are overexposed to office properties and under exposed to retail properties. They may enter into a swap with a *counterparty* where they pay the office returns on an index of office properties and receive the return on an index of retail properties. Similarly, an investor could swap returns on an index of properties in the east with an index of properties in the west.

Relative Value Investing

Hedge funds and other more opportunistic investors may feel that they can identify which property sectors or geographic locations will outperform others. Thus, they may enter into different long and short positions on derivatives to try to capture the perceived mispricing. They would not necessarily have any desire to own and manage the physical real estate.

Efficient Leverage

As forward and futures contracts do not in themselves require up-front cash investment, such derivatives can be used in effect to take levered positions in real estate if the investor does not fully cover the derivative position with bond investment. Depending on circumstances, this may present a lower-cost method of levering the investment, compared to traditional real estate debt.

EXAMPLE: FORWARD CONTRACT

A foreign investor wants to quickly get exposure to the U.S. real estate market to diversify into the United States but does not have the time and expertise to identify individual properties and be sure he is also diversified within the United States. He enters into a long position on a two-year forward contract based on a national real estate index. The index is currently at 100. He has seen forecasts for the index ranging from 105 to 115 in two years. He agrees on a forward price of 105 that he will pay at the end of the two years in order to receive a payment based on the actual change in the index. The contract pays \$500,000 times the index value. No cash payment is made today, although a margin or bond may be required. The magnitude of the required margin or bond posting is relatively small and may earn interest. The required posting would normally be related to the likely magnitude of change in the value of the index over the relevant derivative contract period, rather than to the magnitude of the overall notional amount of the trade, and thus allows the investor to obtain very high effective leverage unless the notional amount of the trade is otherwise covered by up-front cash investment (e.g., in bonds).

Suppose that at the end of the two years the index is 115 (upper end of forecast). The investor will receive $\$500{,}000 \times (115 - 105) = \$5$ million. However, if at the end of the two years the index is 95 (bad forecast!) the investor will pay $\$500{,}000 \times (95 - 105) = -\$5$ million.

There will also be a counterparty to the above transaction who has the short position—the other side of the position the foreign investor took. The short position receives the opposite cash flows in the previous example, receiving \$5 million when the index is 95 and paying \$5 million when it is 115. The short might be, for example, a commercial mortgage-backed security (CMBS) issuer who wants to hedge its warehouse risk, a hedge fund that believed the low end of the forecast was more likely, or an investment manager seeking to "harvest alpha" (explained next). The CMBS issuer trying to hedge "warehouse risk" (loan pools or securities held temporarily awaiting sale) would probably prefer to use a periodically cash-settled swap rather than a two-year forward (because "warehoused" loans are not held very long, though the CMBS issuer may typically always have some warehoused loans on hand). Swaps will be described shortly.

HARVESTING ALPHA EXAMPLE

A specialized real estate asset management fund believes it can purchase, manage, and sell properties so as to consistently outperform the real estate index that underlies the derivative (and with same risk), based on the manager's specialized expertise. They want to harvest this positive "alpha" from these excess returns whether the market is up or down. Since the investment manager cannot control the market, but can (presumably) control its alpha (based on its specialized expertise), the idea is for the manager to profit from the activity they can control and are particularly good at, while laying off risk exposure to factors they cannot control. This is a classical type of "risk management" for an investment management firm. To hedge exposure on $50 million worth of properties the manager owns, for example, the manager would sell (short) $50 million notional value of the forward contract on the index that we described in the previous example. For that portion of the fund's property holdings, the fund is "market neutral": they have laid off their "beta" market risk exposure by their offsetting positions in the forward short and their covering property holdings. This leaves them with only their alpha, the difference between their property performance and the market (index) performance, and with any "basis risk," systematic or nonsystematic differences between the *ex post* performance of their property holdings and the market (index) not due to the manager's actions.

Suppose at the end of the two years the fund's portfolio increased in value by 20% (including income reinvested in the fund). Suppose the index rose to 115 over the two years (that is, the fund beat the index by 500 basis points).

Appreciation on portfolio	$10,000,000
Loss on short futures	5,000,000
Net gain	$5,000,000

Suppose at the end of the two years the hedge fund's portfolio decreased in value by 2%, while the index decreased to 95 (that is, the fund beat the market by 300 basis points).

Loss on portfolio	$1,000,000
Gain on short futures	$5,000,000
Net gain	$4,000,000

The fund thus gains in this example between $4 and $5 million whether the market increases or decreases, based purely on the positive alpha obtained on the fund's properties. In this extreme example of fully hedging the $50 million amount (and with no basis risk), the fund has been turned into an "alpha machine" that makes (or loses) money purely on its differential performance relative to the index, a differential that purely reflects the fund manager's particular expertise and skill at the property and deal level relative to the index. This "disarticulates" performance based on real estate expertise from performance based on the movements and forces and flows of the broader financial capital market that may move the real estate asset market one way or another at any given time.

CAPITAL RETURN SWAP EXAMPLE

An open-ended fund has funds to invest but has not identified properties they wanted to purchase. They believe that the return on an index that tracks changes in property values will be stronger over the next two years than most market participants believe. They decide to take a long position in a real estate index capital return as a swap where they receive the index capital return and pay a fixed leg each quarter. Recall that the capital return is the change in property value. Suppose they can purchase the capital return and pay a fixed leg of 50 basis points. The notional amount of the swap is $100 million. Suppose the actual capital return over the next eight quarters is as shown in Table 51.1. In the first quarter the capital return is 2% so the fund receives 2% of $100 million or $2 million. They pay 0.5% of $100 million or $500,000 on the fixed leg. Thus, they net $1.5 million. Note that in the last four quarters they end up paying money because the capital return did not cover the fixed leg. They end up netting zero over the eight quarters, no doubt not as well as they had hoped in this case, but this reflects the real estate market risk that is represented in the index. Perhaps the market performed worse than this investor had hoped, or perhaps they agreed to a fixed leg that was too high.

REAL ESTATE INDICES

Creating derivatives for commercial real estate requires the availability of indexes that are the basis for calculating the payoffs to the parties in the derivative transaction. The oldest index for commercial real estate investment performance in the United States is the NCREIF Property Index (NPI) published by the National Council of Real Estate Investment Fiduciaries. The NPI is an appraisal-based index that has returns available on a quarterly basis since 1978 and as of the end of 2006 included almost $250 billion in real estate. More recently other indexes have been created to meet the needs of having a viable derivative market in the United States, including *indices* based on real estate transactions developed initially at the Massachusetts Institute of Technology (MIT).

In order to have good derivative contracts, we need good indices underlying the contracts. A property derivative

Table 51.1 Actual Capital Gain Returns for Illustration

Quarter	Capital Return	Long Receives ($ million)	Long Pays ($ million)	Net to Long ($ million)
1	2.00%	$2.0	$0.5	$1.5
2	1.50%	$1.5	$0.5	$1.0
3	1.00%	$1.0	$0.5	$0.5
4	0.50%	$0.5	$0.5	$0.0
5	−1.00%	−$1.0	$0.5	−$1.5
6	0.00%	$0.0	$0.5	−$0.5
7	0.50%	$0.5	$0.5	$0.0
8	−0.50%	−$0.5	$0.5	−$1.0

contract is no better than the index on which it is based. It is probably impossible to have a perfect index to use for commercial real estate derivatives. Unlike stock indices that can be used for futures contracts, it is not possible to invest in all or even a few of the properties used for a real estate index because the properties are held by many different investors in different types of investment vehicles that are privately held. Furthermore, properties do not transact on a frequent basis like stocks to be able to simply measure the change in value of each property in the index based on daily, monthly, quarterly, or even annual transaction prices. There are two main ways of dealing with the fact that the same property does not transact frequently. The first is to have an index based on appraisals of the property on a quarterly basis. This is the basis for the NCREIF property index mentioned above. The second way of creating an index is to base it on the transactions that do occur for properties and have the model control for the varying time between sales of properties.

No single index is likely to be best for all trading purposes. The informational complementarities of different types of commercial property indexes, combined with the diversity and heterogeneity in the U.S. commercial property market and real estate industry, suggests that there can be value from having more than one type of index available. Use of derivatives in "arbitrage" trading across indices can be a source of profit, price discovery, and liquidity.

Real estate indices, especially appraisal-based indices, tend to be more predictable than stock market indices. Derivative prices can reflect forecasts for the underlying index. Commodity futures contracts have always reflected consensus forecasts of where the corresponding commodity spot markets are headed. Because there is momentum in a real estate index, the equilibrium (or "fair") pricing of its derivatives in the derivatives market must reflect the index predictability implied by such momentum. This differs from typical stock market index derivatives in which the underlying indices have relatively little momentum and the stock shares on which the indices are based are directly traded in liquid cash (or "spot") markets, allowing execution of arbitrage between the futures and spot markets.

INCOME AND CAPITAL RETURNS

Periodic total returns for commercial real estate that reflect overall investment performance come from both the current cash flow generated by properties (income return) and changes in the capital value of the properties between the beginning and end of each index reporting period (capital return). Compared to capital returns and to most financial series, income returns are very nearly constant over typical trading periods. This is because in long-lived assets such as real property the current income per period is at least an order of magnitude smaller than the capitalized asset value. In the NPI, the quarterly volatility of the capital return between 1978 and 2006 was 1.7% versus only 0.3% for the income return. (This compares to quarterly volatility over the same historical period of 0.8% for Treasury bills, 6.8% for real estate investment trust [REIT] stocks, and 7.7% for the Standard & Poor's (S&P) 500 large-cap stock index.) Figure 51.1 shows the NPI return components (income, capital, and total) from 1978 through 2006 (quarterly unleveraged returns), revealing how the income return is essentially constant compared to the capital or total return components.

If the underlying index reports the total return (as in appraisal-based indices such as the NCREIF or Investment Property Databank [IPD] indices), then derivatives can be structured based on either the total return or just the capital return. However, even if the underlying index reports only the capital return, derivatives can effectively be used to create the total return synthetically, because virtually all of the index total return volatility is in the capital return component alone. In the NPI, over 116 quarters during 1978–2006, the capital return and total return were correlated +99%, with essentially equal volatilities of 1.7% each. For example, a structured note in which the investor funds up front the fixed leg of a capital return swap will effectively provide the investor with the index total return, as we will see later when we discuss swap pricing.

APPRAISAL-BASED INDICES

The first regularly produced commercial property price indexes were appraisal based, and designed for benchmarking institutional real estate investment manager performance. These include the NPI in the United States and the IPD Index in Great Britain, among others worldwide. In a traditional appraisal-based index all of the properties in the index population are reappraised frequently, and the index periodic returns are based on the average (usually value weighted) of those appraisal-based returns each period. This is similar to the way many institutional real estate investment funds "mark to market" their asset

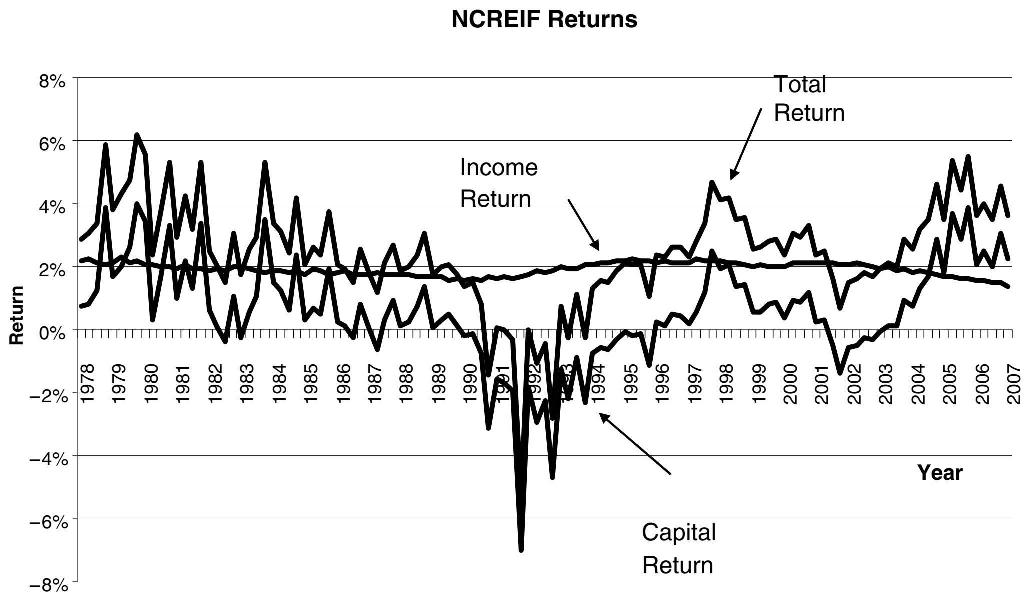


Figure 51.1 NCREIF Property Index Returns

values and correspondingly report quarterly returns to their investors. Of course, the NPI reflects property-level returns (unlevered, and before any fund-level or management expenses and fees to which investors are subject).

While such traditional appraisal-based indices can be excellent tools for benchmarking investment manager performance, and this in itself gives them a particular use in derivatives of interest to such managers, they do have some inherent problems from the perspective of a broader derivative support role. The appraisal process tends to be somewhat subjective and backward looking (perhaps more so in the United than in Britain). This tends to impart a lag bias to the property values and the index returns. Furthermore, in the case of the NCREIF Index in the United States, not all properties are reappraised every period that the index is reported, and this adds an additional "stale appraisal" effect into the index. In the NCREIF Index, at least during some periods of its history, greater frequency of reappraisals in the fourth calendar quarter has imparted an artificial seasonality to the index (the index can tend to "spike" in the fourth quarter). It must also be recognized that, at least as of the early 2000s, the NCREIF Index represents a relatively narrow segment of the population of U.S. properties. In 2006 the NCREIF population of properties consisted of less than 10% of the commercial properties in the United States, a much smaller percentage than the IPD Index represents in Britain. For example, in 2006 the NPI included less than $30 billion of property sales, whereas the Real Capital Analytics Inc. (RCA) database recorded over $330 billion of commercial property sales tracking only sales of greater than $2.5 million. As of the end of 2006, the NPI was tracking some $250 billion worth of property, whereas J. P. Morgan Asset Management's "Real Estate Universe" report estimated the total value of U.S. commercial real estate at that time to be some $6.7 trillion, or over 25 times the NCREIF population value (although this included corporate real estate and small "mom-and-pop" properties as well as the larger properties covered by the RCA database). For smaller market segments, there may be only a few NCREIF properties available in the index, and their specific identities will be known to at least some potential participants in the derivatives marketplace.

The above problems are of less concern for purposes of benchmarking institutional real estate portfolios that are marked to market using appraised values, but they can be more problematic for broader derivative support purposes. If the lag in the index causes it to still rise when the real estate market turns down (or vice versa), this can be confusing to parties trying to use the index to hedge or speculate on such market movements. Derivative pricing when the index is lagged needs to reflect the lag, and that may make price discovery more difficult, potentially hampering liquidity in the derivative market, although in principle the lag can be relatively easily reflected in the derivative price (especially if indices that are not lagged are also available as information sources). Even if the index lag is taken into account in the derivative price, if the derivative contract expires before the lagged price movement is fully reflected in the index, then the hedge will not be complete, presenting a type of "basis risk" for the user of the derivative. Thus, for a variety of reasons, futures traders may prefer indexes that lead the appraisal-based indices in time, and in which the true volatility is not dampened, as such volatility can be a source of potential profit that might motivate some derivative traders.

TRANSACTIONS-BASED INDICES

An alternative to appraisal-based indices is to have an index based on transactions (sales) of properties. In principle, such indices can be based on the entire population of commercial properties, because all properties potentially

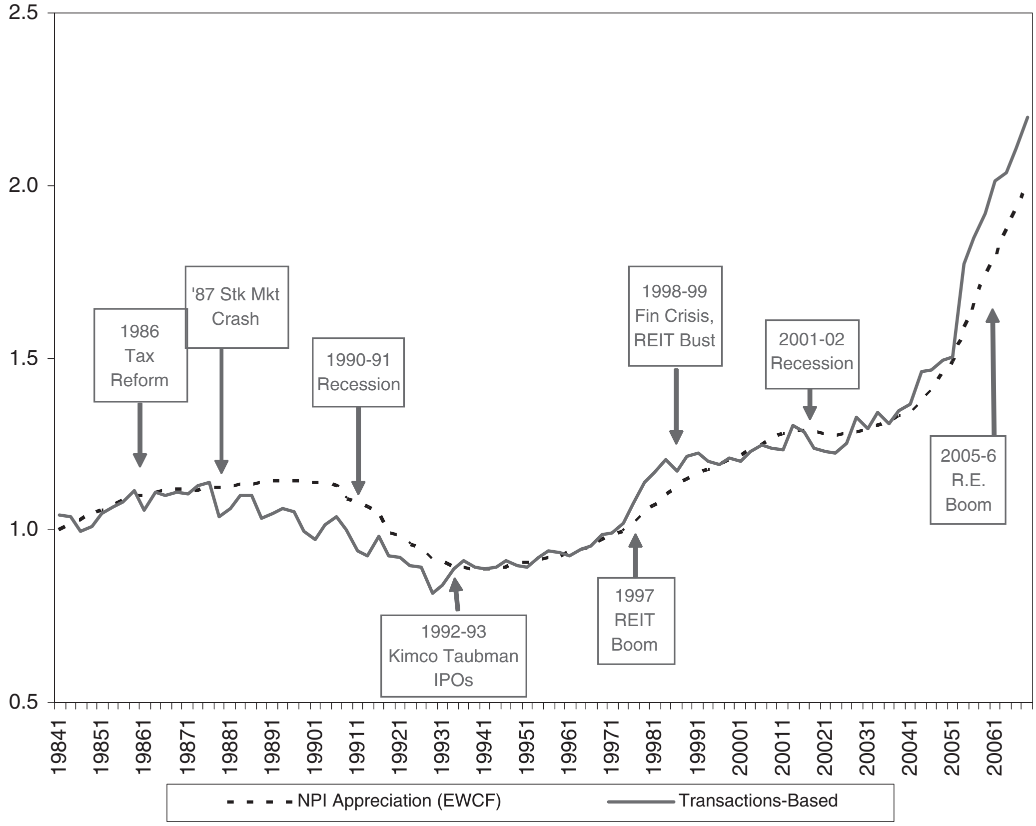


Figure 51.2 Appraisal-Based vs. Transaction-Based Capital Value Index Based on NCREIF: 1984–2006
Source: Fisher, Geltner, and Pollakowski (2007).

transact (providing a random price sample of the population each period), whereas only certain specialized portfolios of properties are regularly marked to market in the U.S. using appraisals. Transactions-based indices can be good bases for derivatives provided the indexes are carefully constructed based on sufficient quantity and quality of transactions observations data and state-of-the-art statistical procedures to control for "apples versus oranges" differences in properties trading in different periods and to minimize "noise" or random deviations from the property population prices.

There are two major procedures to calculate transactions-based indices in a statistically rigorous manner: (1) the repeat-sales regression procedure and (2) the hedonic value model. Both procedures address the fundamental problem in the construction of a transactions based real estate price index, the fact that the properties that transact in one period are generally not the same as the properties that transacted in the previous period, making a direct comparison of prices "apples versus oranges." The two procedures address this issue in different ways.

The hedonic procedure models property prices as a function of various characteristics of the properties that affect their value, such as age, size, location, building quality, and so on. By regressing property transaction prices onto these "hedonic characteristics" of the properties that sell, and controlling for or keeping track of the time of the sale, one constructs a "constant-quality" price-change index or an index that tracks property market price changes controlling for differences in the properties that transact at the different points in time. The MIT Center for Real Estate began publishing the first regularly produced hedonic index of commercial property in 2006, in cooperation with NCREIF, based on the prices of the properties sold from the NCREIF Index. This transactions-based index uses the recent appraised values of the sold properties as a "composite" indicator of the hedonic characteristics of the properties, controlling in this way for cross-sectional differences in the sold properties. Because this transactions based index was based on the same underlying population of properties as the NPI, it can present a good "apples-to-apples" comparison of the difference between a transactions-based and an appraisal-based index. This comparison over the historical period from 1984 to 2006 is shown in Figure 51.2. The comparison gives an indication of the typical differences between a transactions based and an appraisal based index. Note that the transactions based version of the NCREIF Index is a bit more volatile, and tends to slightly lead the NPI in time (in terms of the timing of major turning points in the index history).

Repeat-sales indices use a different approach to address the "apples-versus-oranges" problem. As the name suggests, repeat-sales indices rely on individual properties selling more than once, so that the change in price between sales provides an indication of how same-property values have changed over time. The index is thus based on the type of price changes that investors in properties actually experience, and the same type of price changes that stock market indices are based on. Stock market indices are also based on comparing the transaction prices of stock shares in one period with the transaction prices of similar shares in the previous period. As stock shares are homogeneous (a share of IBM that traded this month is the same as a share of IBM that traded last month), the result is comparable to a "same property" price change index such as the repeat-sales transactions-based indices. It should also be noted that stock share prices reflect the value added by the corporation not paying out all of its cash in dividends, but reinvesting some in the corporation. This is analogous to the effect of capital improvement expenditures in real estate. Thus, repeat-sales indices aimed at tracking property prices do not generally try to remove the effect of capital improvement expenditures (although normally data filters are applied to eliminate property sale pairs that would reflect major development, redevelopment, or rehabilitation of the properties). This is in some contrast to appraisal-based indices that may subtract capital expenditures from the appreciation return reported by the index.

The statistical process used to calculate repeat-sales indices takes into consideration the time between the same-property sales and appropriately allocates the price change to each period that the index is reported, based on information from other repeat sales occurring over various time frames. Repeat sales is the approach used in widely quoted housing price indices such as the S&P/Case-Shiller housing index on which the Chicago Mercantile Exchange (CME) launched futures trading in 2006. A simple numerical example of how the calculation process works is presented in the appendix to this chapter.

The first regularly published repeat-sales transactions based index for commercial property was developed by the MIT Center for Real Estate based on data from the firm Real Capital Analytics Inc. (RCA) and launched in 2006. This index was based on a much broader property population than the appraisal-based NCREIF Index, as the RCA database attempted to track all commercial property sales in the United States of over $2.5 million, whereas the NPI tracked only the NCREIF members' properties.

SUMMARY

This chapter has reviewed the nature and mechanics of the major real estate equity index derivative products and their use and usefulness. It has also presented the fundamentals of real estate return indices, including the important differences between the two major types of indices: appraisal-based and transactions-based indices.

APPENDIX: NUMERICAL EXAMPLE OF HOW THE REPEAT-SALES INDEX WORKS

In this appendix we present a simple numerical example of the mechanics of how the repeat-sales regression procedure works to construct an index of periodic capital returns based on same-property price changes. In so doing, we will also highlight some key features of the repeat-sales model that are not intuitively obvious, such as how the model can detect a downturn in the market even when all of the individual property investments are producing a positive return over their holding periods, and how no single period's return estimate is based only on the second-sales occurring in that period alone.

To understand how the repeat-sales regression (RSR) index construction process works, you must step back briefly and recall some basic statistics. You may recall that regression analysis is a statistical technique for estimating the relationship between variables of interest. In a regression model, a particular variable of interest, referred to as the dependent variable, is related to one or more other variables referred to as explanatory variables. The regression model is presented as an equation, with the dependent variable on the left-hand side of the equal sign and a sum of terms on the right-hand side, consisting of the explanatory variables each multiplied by a parameter that is estimated by the regression and that relates each explanatory variable to the dependent variable. For example, if the dependent variable is labeled "Y" and there is a single explanatory variable labeled "X," then a simple regression model of Y as a function of X would be expressed as:

$$Y = aX$$

The model says that the value of the variable Y equals the value of the variable X times the parameter "a," and we would use the regression analysis of relevant empirical data to estimate what is the value of "a.". This process is referred to as "estimation" of the regression, or "calibrating" the model.

How can this technique enable the development of a real estate price index? Let's take a very simple numerical example. Suppose that the true returns in the market are respectively: 0%, +10%, and −5%, in three consecutive periods (say, 2011, 2012, and 2013). Thus, a true price index starting out at 1.00 at the end of 2010 would remain at 1.00 at the end of 2011, jump to 1.10 in 2012, and then fall back to 1.045 in 2013 [as $(1.045 - 1.10)/1.10$ is −5%]. Now suppose we have three property repeat-sales observations involving altogether at least one sale in each of the three years, with each being consistent with the true returns but in which no one observation can directly reveal any one period's return because the properties are held across more than one period. Property 1 is bought at the beginning of 2011 for $100,000 and sold after three years at the end of 2013 for $104,500. Property 2 is also bought at the beginning of 2011, but for $200,000 and sold at the end of 2012 for $220,000 (held for two years). Property 3 is bought at the beginning of 2012 for $300,000 and sold at the end of 2013 for $313,500 (also held two years). This is summarized in the Table 51.2 and Figure 51.3, where the figure indicates both the true market price index (the solid line) and the capital returns achieved by each of the three investors in these three properties (dashed lines).

Now let's apply the RSR model to this problem. Let the dependent variable, "Y," be the natural log of the ratio of the second sale price divided by the first sale price, for each repeat-sale pair. Thus, the first repeat-sales observation, based on Property 1, has a Y value of the log of 1.045.

Table 51.2 Prices Observed at Ends of Years

	2006	2007	2008	2009
True price index	1.00	1.00	1.10	1.045
True capital return		0%	10%	−5%
Property 1	$100,000	No Data	No Data	$104,500
Property 2	$200,000	No Data	$220,000	No Data
Property 3	No data	$300,000	No data	$313,500

Similarly, the second repeat-sales observation, based on Property 2, has a Y value of the log of 1.10, and so on.

On the right-hand side of our RSR model, instead of just one variable, "X," let there be three variables, corresponding to the three consecutive periods of time for which we want to construct the index periodic returns. Let us label these "X2011," "X2012," and "X2013." These right-hand-side variables are what are called "dummy variables," which means they take on a value of either zero or one. The "X2011" variable stands for the year 2011. It takes the value of one if 2011 is after the year of the first sale and before or including the year of the second sale in the repeat-sales observation (in other words, if the dummy variable's year is during the property investor's holding period between when he bought and sold the property of the observation in question); otherwise, this dummy variable has a value of zero. Similarly, "X2012" takes the value of one if 2012 is after the year of the first sale and before or including the year of the second sale. Thus, the price observation data described previously gives the RSR estimation data in Table 51.3. For example, for the repeat-sales observation corresponding to Property 3, $313,500/$300,000 is 1.045, and the natural log of this value happens to be about 4.4%, which is therefore the Y value for that observation in the RSR estimation database.

Our regression equation can now be expressed as:

$$Y = a_{2011}(X2011) + a_{2012}(X2012) + a_{2013}(X2013)$$

Now recall from statistics that the estimation of a regression model—that is, the "calibration" of the value of the parameters in the preceding equation—is mathematically the solution of a system of simultaneous equations. Each

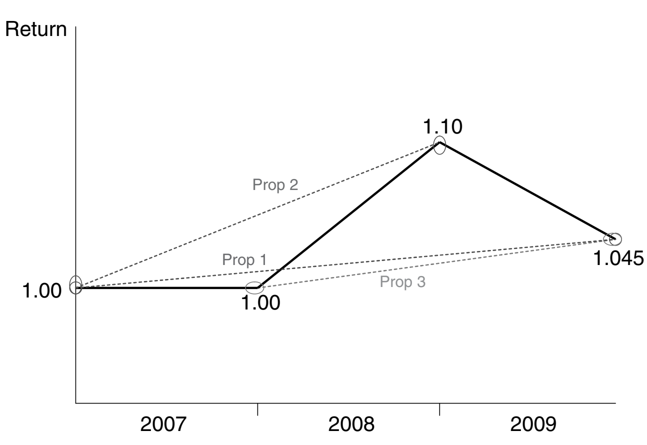


Figure 51.3 Example of Repeat-Sales Regression Model

Table 51.3 RSR Estimation Data

	Y Value = $LN(P_s/P_f)$	X2011 Value	X2012 Value	X2013 Value
Observation 1	LN(1.045)	1	1	1
Observation 2	LN(1.10)	1	1	0
Observation 3	LN(1.045)	0	1	1

equation corresponds to one "observation," one data point in the database used to estimate the regression model. Thus, in our present example, we have three equations, one corresponding to each row (each repeat-sales observation) in Table 51.3. The three equations are:

$$LN(1.045) = a_{2011}(1) + a_{2012}(1) + a_{2013}(1) \quad (51.1)$$
$$LN(1.100) = a_{2011}(1) + a_{2012}(1) + a_{2013}(0) \quad (51.2)$$
$$LN(1.045) = a_{2011}(0) + a_{2012}(1) + a_{2013}(1) \quad (51.3)$$

which equates to:

$$LN(1.045) = a_{2011} + a_{2012} + a_{2013}$$
$$LN(1.100) = a_{2011} + a_{2012}$$
$$LN(1.045) = a_{2012} + a_{2013}$$

We thus have three linear equations with three unknowns (a_{2011}, a_{2012}, and a_{2013}, representing the true log price ratios in each of the three periods). Such a system can always be solved, and in this case the solution can be found as follows:

Use equation (51.2) to derive: $a_{2012} = LN(1.1) - a_{2011}$. Then plug this into equation (51.3) to obtain: $a_{2013} = LN(1.045) - LN(1.1) + a_{2011}$. Now plug both of these into equation (51.1) to obtain:

$$LN(1.045) = a_{2011} + [LN(1.1) - a_{2011}] + [LN(1.045) - LN(1.1) + a_{2011}], \rightarrow$$
$$a_{2011} = LN(1.045) - LN(1.045), \rightarrow$$
$$a_{2011} = 0.$$

Now plug this result back into equations (51.2) and (51.3) to obtain: $a_{2012} = LN(1.1)$, and $a_{2013} = LN(1.045) - LN(1.1)$. The result that $a_{2011} = 0$ simply means that the estimated price index level did not change during 2011. From the definition of logarithms we have $0 = LN(1)$, and algebraically we can express this as LN(1/1). Similarly, we can express $a_{2012} = LN(1.1)$ as LN(1.1/1). Thus, the implied log price ratios of the price index ending values divided by its beginning values each year are:

For 2011 = a_{2011} = LN(1/1)
For 2012 = a_{2012} = LN(1.1/1)
For 2013 = a_{2013} = LN(1.045/1.1)

Exponentiating these values, we arrive at the implied straight level price index as of the end of each year as follows:

2006 = 1.000
2007 = 1.000
2008 = 1.100
2009 = 1.045

with the resulting implied price-change percentages (capital returns):

2007 = 0%
2008 = +10%
2009 = −5%

Thus, we see that the repeat-sales model has derived the true capital return in each period, even though no single repeat-sale price change observation corresponded to any one year. The model correctly derived the negative return in 2013, even though none of the repeat-sale observations used in the estimation showed a negative price change in itself. In other words, all of the three investors made a positive resale gain over their holding periods. Note also that the estimation of the returns in each of the three periods was affected by all three of the repeat-sale observations. For example, the estimate of the negative 5% return in 2013 was determined in part by the +10% return obtained on Property 2, even though that property's second sale occurred prior to the beginning of 2013.

While this is a simple numerical example, the type of result shown here is general. In principle, the repeat-sales model only requires one sales observation per period (either a first or second sale) in order to be able to estimate the true return each period, even though no single repeat-sale pair corresponds to any one period. And the model uses all observations to estimate every period's return. Thus, it is not correct to think that the estimated return in the current period is determined solely or in isolation by the second-sale observations that occur only in the current period.

Of course, in the real world, individual transaction prices will be dispersed randomly around the average (normalized) sale price at any given time, which makes index estimation a statistical process. The existence of more than one observation (hence more than one equation) in each period of time enables such estimation to be optimized in various ways, as is done in actual RSR indexes.

REFERENCES

Bailey, M., Muth, R., and Nourse, H. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association* 58: 922–942.

Court, A. (1936). Hedonic price indices with automotive examples. In *The Dynamics of Automobile Demand*. Detroit: General Motors Corporation.

Fisher, J. D. (2005). Introducing the NPI based derivative: New strategies for commercial real estate investment and risk management. *Journal of Portfolio Management*, Special Issue on Real Estate: 1–9.

Fisher, J. D., Geltner, D., and Webb, R. B. (1994). Value indices of commercial real estate: A comparison of index construction methods. *Journal of Real Estate Finance & Economics* 9, 2: 137–164.

Fisher, J, Geltner, D., and Pollakowski, H. (2007). A quarterly transactions-based index of institutional real estate investment performance and movements in supply and demand. *Journal of Real Estate Finance and Economics* 34, 1: 2007.

Geltner, D., and Pollakowski, H. (2006). A set of indexes for trading commercial real estate based on the Real Capital Analytics Database. Report by the MIT Center for Real Estate.

Geltner, D., and Fisher, J. D. (2007). Pricing and index considerations in commercial real estate derivatives.

Goodman, L. S., and Fabozzi, F. J. (2005). CMBS total return swaps. *Journal of Portfolio Management*, Special Issue on Real Estate: 162–167.

Griliches, Z., and Adelman, I. (1961). On an index of quality change. *Journal of the American Statistical Association* 56: 535–548.

Rosen, S. (1974). Hedonic prices and implicit markets. *Journal of Political Economy* 82, 1: 33–55.

PART 5

Alternative Investments

CHAPTER 52

Alternative Asset Classes

MARK J. P. ANSON, PhD, JD, CPA, CFA, CAIA
President and Executive Director of Nuveen Investment Services

Abstract: Alternative assets are not always easy to describe. Before an asset class can be considered "alternative," the primary asset classes must be defined, and then what differs is considered "an alternative asset." The bottom line is that alternative assets fall outside the scope of normal investment portfolios. They provide risk and return characteristics that are distinctly different from traditional portfolios of stocks and bonds. This provides not only excellent portfolio diversification, but also the ability to earn returns that may exceed that of stocks and bonds. However, access to alternative assets is less straightforward than it is to purchase stocks and bonds; therefore, alternative assets are often overlooked because they require more work to invest in than traditional asset classes.

Keywords: asset class, hedge fund, private equity, capital assets, store of wealth, economic inputs, strategic asset allocation, tactical asset allocation, asset location, trading strategy, alternative beta

Part of the difficulty of working with alternative asset classes is defining them. Are they a separate asset class or a subset of an existing asset class? Do they hedge the investment opportunity set or expand it? Are they listed on an exchange or do they trade in the over-the-counter market?

In most cases, alternative assets are a subset of an existing asset class. This may run contrary to the popular view that alternative assets are separate asset classes. However, we take the view that what many consider separate "classes" are really just different investment strategies within an existing asset class.

In most cases, they expand the investment opportunity set, rather than hedge it. Finally, alternative assets are generally purchased in the private markets, outside of any exchange. While hedge funds, private equity, and credit derivatives meet these criteria, we will see that commodity futures prove to be the exception to these general rules.

Alternative assets, then, are just alternative investments within an existing asset class. Specifically, most alternative assets derive their value from either the debt or equity market. For instance, most hedge fund strategies involve the purchase and sale of either equity or debt securities. Additionally, hedge fund managers may invest in derivative instruments whose value is derived from the equity or debt market.

One can classify five types of alternative assets: hedge funds, commodity and managed futures, private equity, credit derivatives, and corporate governance. This is the classification used in Anson (2006). Hedge funds and private equity are the best known of the alternative asset world. Typically, these investments are accomplished through the purchase of limited partner units in a private limited partnership. Commodity futures can be either passive investing tied to a commodity futures index or active investing through a commodity pool or advisory account. Private equity is the investment strategy of investing in companies before they issue their securities publicly, or taking a public company private. Credit derivatives can be purchased through limited partnership units, as a tranche of a special-purpose vehicle, or directly through the purchase of credit default swaps or credit options. Corporate governance is also a form of shareholder activism designed to improve the internal controls of a public company.

Yet, before we can discuss alternative assets we need to provide a definition for the term "asset class." We start by defining the major asset classes and then work our way to defining what is an "alternative asset."

SUPER ASSET CLASSES

There are three super asset classes: capital assets, assets that are used as inputs to creating economic value, and assets that are a store of value (see Greer, 1997).

Capital Assets

Capital assets are defined by their claim on the future cash flows of an enterprise. They provide a source of ongoing value. As a result, capital assets may be valued based on the net present value of their expected returns.

Under the classic theory of Modigliani and Miller (1958), a corporation cannot change its value (in the absence of tax benefits) by changing the method of its financing. Modigliani and Miller demonstrated that the value of the firm is dependent on its cash flows. How those cash flows are divided up between shareholders and bondholders is irrelevant to firm value.

Capital assets, then, are distinguished not by their possession of physical assets, but rather, by their claim on the cash flows of an underlying enterprise. Hedge funds, private equity funds, credit derivatives, and corporate governance funds all fall within the super asset class of capital assets because the value of their funds are all determined by the present value of expected future cash flows from the securities in which they invest.

As a result, we can conclude that it is not the types of securities in which they invest that distinguishes hedge funds, private equity funds, credit derivatives, or corporate governance funds from traditional asset classes. Rather, it is the alternative investment strategies that they pursue that distinguishes them from traditional stock and bond investments.

Assets that Can be Used as Economic Inputs

Certain assets can be consumed as part of the production cycle. Consumable or transformable assets can be converted into another asset. Generally, this class of asset consists of the physical commodities: grains, metals, energy products, and livestock. These assets are used as economic inputs into the production cycle to produce other assets, such as automobiles, skyscrapers, new homes, and appliances.

These assets generally cannot be valued using a net present value analysis. For example, a pound of copper, by itself, does not yield an economic stream of revenues. Nor does it have much value for capital appreciation. However, the copper can be transformed into copper piping that is used in an office building or as part of the circuitry of an electronic appliance.

While consumable assets cannot produce a stream of cash flows, we demonstrate in our section on commodities that this asset class has excellent diversification properties for an investment portfolio. In fact, the lack of dependency on future cash flows to generate value is one of the reasons why commodities have important diversification potential vis-à-vis capital assets.

Assets that Are a Store of Value

Art is considered the classic asset that stores value. It is not a capital asset because there are no cash flows associated with owning a painting or a sculpture. Consequently, art cannot be valued in a discounted cash flow analysis. It is also not an asset that is used as an economic input because it is a finished product.

Art requires ownership and possession. Its value can be realized only through its sale and transfer of possession. In the meantime, the owner retains the artwork with the expectation that it will yield a price at least equal to that which the owner paid for it.

There is no rational way to gauge whether the price of art will increase or decrease because its value is derived purely from the subjective (and private) visual enjoyment that the right of ownership conveys. Therefore, to an owner, art is a store of value. It neither conveys economic benefits nor is used as an economic input, but retains the value paid for it.

Gold and precious metals are another example of a store-of-value asset. In the emerging parts of the world, gold and silver are a significant means of maintaining wealth. In these countries, residents do not have access to the same range of financial products that are available to residents of more developed nations. Consequently, they accumulate their wealth through a tangible asset as opposed to a capital asset.

However, the lines between the three super classes of assets can become blurred. For example, gold can be leased to jewelry and other metal manufacturers. Jewelry makers lease gold during periods of seasonal demand, expecting to purchase the gold on the open market and return it to the lessor before the lease term ends. The gold lease provides a stream of cash flows that can be valued using net present value analysis.

Precious metals can also be used as a transformable/consumable asset because they have the highest level of thermal and electrical conductivity among the metals. Silver, for example, is used in the circuitry for most telephones and light switches. Gold is used in the circuitry for televisions, cars, airplanes, computers, and rocketships.

Real Estate

We provide a brief digression to consider where real estate belongs in our classification scheme. Real estate is a distinct asset class, but is it an alternative one? For purposes of this book, we do not consider real estate to be an alternative asset class. The reasons are several.

First, real estate was an asset class long before stocks and bonds became the investment of choice. In fact, in times past, land was the single most important asset class. Kings, queens, lords, and nobles measured their wealth by the amount of property that they owned. "Land barons" were aptly named. Ownership of land was reserved only for the most wealthy of society.

However, over the past 200 years, our economic society changed from one based on the ownership of property to the ownership of legal entities. This transformation occurred as society moved from the agricultural age to the industrial age. Production of goods and services became the new source of wealth and power.

Stocks and bonds were born to support the financing needs of new enterprises that manufactured material goods and services. In fact, stocks and bonds became the "alternatives" to real estate instead of vice versa. With the advent of stock-and-bond exchanges, and the general acceptance of owning equity or debt stakes in companies, it is sometimes forgotten that real estate was the original and primary asset class of society.

In fact, it was only 25 years ago in the United States that real estate was the major asset class of most individual investors. This exposure was the result of owning a primary residence. It was not until the long bull market started in 1983 that investors began to diversify their wealth into the "alternative" assets of stocks and bonds.

Second, given the long-term presence of real estate as an asset class, several treatises have been written concerning its valuation. Finally, we do not consider real estate to be an alternative asset class as much as we consider it to be an additional asset class. Real estate is not an alternative to stocks and bonds—it is a fundamental asset class that should be included within every diversified portfolio. The alternative assets that we consider in this book are meant to diversify the stock-and-bond holdings within a portfolio context.

ASSET ALLOCATION

Asset allocation is generally defined as the allocation of an investor's portfolio across a number of asset classes (see Sharpe, 1992). Asset allocation, by its very nature shifts the emphasis from the security level to the portfolio level. It is an investment profile that provides a framework for constructing a portfolio based on measures of risk and return. In this sense, asset allocation can trace its roots to modern portfolio theory and the work of Harry Markowitz (1959).

Asset Classes and Asset Allocation

Initially, asset allocation involved four asset classes: equity, fixed income, cash, and real estate. Within each class, the assets could be further divided into subclasses. For example, stocks can be divided into large capitalized stocks, small-capitalized stocks, and foreign stocks. Similarly, fixed income can be broken down into U.S. Treasury notes and bonds, investment-grade bonds, high-yield bonds, and sovereign bonds.

The expansion of newly defined "alternative assets" may cause investors to become confused about their diversification properties and how they fit into an overall diversified portfolio. Investors need to understand the background of asset allocation as a concept for improving return while reducing risk.

For example, in the 1980s the biggest private equity game was taking public companies private. Does the fact that a corporation that once had publicly traded stock but now has privately traded stock mean that it has jumped into a new asset class? Furthermore, public offerings are the primary exit strategy for private equity; public ownership begins where private equity ends (see Horvitz, 2000). Therefore, it might be argued that private equity is just an extension of the equity markets where the dividing boundary is based on liquidity.

Similarly, credit derivatives expand the fixed income asset class, rather than hedge it. Hedge funds also invest in the stock-and-bond markets but pursue trading strategies very different from a traditional buy-and-hold strategy. Commodities fall into a different class of assets than equity, fixed income, or cash, and will be treated separately in this book.

Finally, corporate governance is a strategy for investing in public companies. It seems the least likely to be an alternative investment strategy. However, it can be demonstrated that a corporate governance program bears many of the same characteristics as other alternative investment strategies (see Anson, 2006).

Strategic versus Tactical Allocations

Alternative assets should be used in a tactical rather than strategic allocation. Strategic allocation of resources is applied to fundamental asset classes such as equity, fixed income, cash, and real estate. These are the basic asset classes that must be held within a diversified portfolio.

Strategic asset allocation is concerned with the long-term asset mix. The strategic mix of assets is designed to accomplish a long-term goal such as funding pension benefits or matching long-term liabilities. Risk aversion is considered when deciding the strategic asset allocation, but current market conditions are not. In general, policy targets are set for strategic asset classes, with allowable ranges around those targets. Allowable ranges are established to allow flexibility in the management of the investment portfolio.

Tactical asset allocation is short term in nature. This strategy is used to take advantage of current market conditions that may be more favorable to one asset class over another. The goal of funding long-term liabilities has been satisfied by the target ranges established by the strategic asset allocation. The goal of tactical asset allocation is to maximize return.

Tactical allocation of resources depends on the ability to diversify within an asset class. This is where alternative assets have the greatest ability to add value. Their purpose is not to hedge the fundamental asset classes, but rather to expand them. Consequently, alternative assets should be considered as part of a broader asset class.

An example is credit derivatives. These are investments that expand the frontier of credit risk investing. The fixed-income world can be classified simply as a choice between U.S. Treasury securities that are considered to be default free, and spread products that contain an element of default risk. Spread products include any fixed income investment that does not have a credit rating on par with the U.S. government. Consequently, spread products trade at a credit spread relative to U.S. Treasury securities that reflects their risk of default.

Credit derivatives are a way to diversify and expand the universe for investing in spread products. Traditionally, fixed income managers attempted to establish their ideal credit risk-and-return profile by buying and selling traditional bonds. However, the bond market can be inefficient and it may be difficult to pinpoint the exact credit profile to match the risk profile of the investor. Credit derivatives can help to plug the gaps in a fixed income portfolio, and expand the fixed income universe by accessing credit exposure in more efficient formats.

Efficient versus Inefficient Asset Classes

Another way to distinguish alternative asset classes is based on the efficiency of the market place. The U.S. public stock-and-bond markets are generally considered to be the most efficient marketplaces in the world. Often, these markets are referred to as "semi-strong efficient." This means that all publicly available information regarding a publicly traded corporation, both past information and present, is fully digested in that company's traded securities.

Yet inefficiencies exist in all markets, both public and private. If there were no informational inefficiencies in the public equity market, there would be no case for active management. Nonetheless, whatever inefficiencies do exist, they are small and fleeting. The reason is that information is easy to acquire and disseminate in the publicly traded securities markets. Top-quartile active managers in the public equity market earn excess returns (over their benchmarks) of approximately 1% a year.

In contrast, with respect to alternative assets, information is very difficult to acquire. Most alternative assets (with the exception of commodities) are privately traded. This includes private equity, hedge funds, and credit derivatives. The difference between top-quartile and bottom-quartile performance in private equity can be as much as 25%.

Consider venture capital, one subset of the private equity market. Investments in start-up companies require intense research into the product niche the company intends to fulfill, the background of the management of the company, projections about future cash flows, exit strategies, potential competition, beta testing schedules, and so forth. This information is not readily available to the investing public. It is time consuming and expensive to accumulate. Furthermore, most investors do not have the time or the talent to acquire and filter through the rough data regarding a private company. One reason why alternative asset managers charge large management and incentive fees is to recoup the cost of information collection.

This leads to another distinguishing factor between alternative investments and the traditional asset classes: the investment intermediary. Continuing with our venture capital example, most investments in venture capital are made through limited partnerships, limited liability companies, or special-purpose vehicles. It is estimated that 80% of all private equity investments in the United States are funneled through a financial intermediary.

Investments in alternative assets are less liquid than their public market counterparts. Investments are closely held and liquidity is minimal. Furthermore, without a publicly traded security, the value of private securities cannot be determined by market trading. The value of the private securities must be estimated by book value or appraisal, or determined by a cash flow model.

Constrained versus Unconstrained Investing

During the great bull market from 1981 to 2000 the asset management industry only had to invest in the stock market to enjoy consistent, high, double-digit returns. During this heyday, investment management shops and institutional investors divided their assets between the traditional asset classes of stocks and bonds. As the markets turned sour at the beginning of the new millennium, asset management firms and institutional investors found themselves "boxed in" by these traditional asset class distinctions. They found that their investment teams were organized along traditional asset class lines, and their investment portfolios were constrained by efficient benchmarks that reflected this "asset box" approach.

Consequently, traditional asset management shops have been slow to reorganize their investment structures. This has allowed hedge funds and other alternative investment vehicles to flourish because they are not bounded by traditional asset class lines—they can invest outside the benchmark. These alternative assets are free to exploit the investment opportunities that fall in between the traditional benchmark boxes. The lack of constraints allows alternative asset managers a degree of freedom that is not allowed the traditional asset class shops. Furthermore, traditional asset management shops remain caught up in an organizational structure that is bounded by traditional asset class lines. This provides another constraint because it inhibits the flow of information and investment ideas across the organization.

Asset Location versus Trading Strategy

One of the first and best papers on hedge funds by Fung and Hsieh (1997) shows a distinct difference in how mutual funds and hedge funds operate. They show that the economic exposure associated with mutual funds is defined primarily by *where* the mutual fund invests. In other words, mutual funds gain their primary economic and risk exposures by the location of the asset classes in which they invest. Thus, we get large-cap active equity funds, small-cap growth funds, Treasury bond funds, and the like.

Conversely, Fung and Hsieh show that hedge funds' economic exposures are defined more by *how* they trade. That is, a hedge fund's risk and return exposure is defined more by a trading strategy within an asset class than it is defined by the location of the asset class. As a result, hedge fund managers tend to have much greater turnover in their portfolios than mutual funds.

Alternative Beta and the Efficient Frontier

Strategic asset allocation (SAA) revolves around the most efficient combination of stocks, bonds, and other asset classes to achieve the best return and risk trade-off. This is the concept behind charting the efficient frontier—the most efficient trade-off between risk and return given a mix of asset classes. In this sense, SAA is all about capturing the systematic risk premiums that exist for investing in different asset classes. However, if additional asset classes can be added to the mix, the efficient frontier can be "pushed out" to provide a greater range of risk and return opportunities for an investor.

This is another way to consider alternative assets—as an alternative source of beta that is different from the traditional mixture of stocks and bonds. Access to alternative assets can provide new systematic risk premiums that are distinctly different than that obtained from stocks and bonds. Commodities are a good example—they provide a different risk exposure than the stocks or bonds. Consequently, the risk premium associated with commodities is less than perfectly correlated with the traditional financial markets. This is a form of "alternative beta." Investing in alternative assets does not have to focus exclusively on the quest for excess returns; it can also look at the diversification properties of alternative assets when blended with a traditional portfolio of stocks and bonds. Alternative beta can be a form of added value through diversification properties instead of a desire for excess return.

Asset Class Risk Premiums versus Trading Strategy Risk Premiums

Related to the idea of trading strategy versus investment location is the notion of risk premiums. You cannot earn a return without incurring risk. Traditional investment managers earn risk premiums for investing in the large-cap value equity market, small-cap growth equity market, high-yield bond market—in other words, based on the location of the asset markets in which they invest.

Conversely, alternative asset managers also earn returns for taking risk, but the risk is defined more by a trading strategy than it is an economic exposure associated with the systematic risk contained within broad financial classes. For example, hedge fund strategies such as convertible arbitrage, statistical arbitrage, and equity market neutral can earn a "complexity" risk premium (see Jaeger, 2002).

These strategies buy and sell similar securities expecting the securities to converge in value overtime. The complexity of implementing these strategies results in inefficient pricing in the market. Additionally, many investors are constrained by the *long-only* constraint—their inability to short securities. This perpetuates inefficient pricing in the marketplace which enables hedge funds to earn a return.

SUMMARY

This chapter was meant as an introduction to the different kinds of asset classes that exist for investment portfolios. To be considered an "alternative" asset class, an investment must demonstrate one of the following: a different trading strategy, a risk premium based on active trading rather than systematic market risk, an exploitation of cracks in the financial markets, a tactical application to add excess return, or a systematic risk premium that is different from that derived from stocks and bonds. Any of these characteristics can distinguish an alternative asset from a traditional asset class. The trick is to use both to extract the greatest performance for the portfolio.

REFERENCES

Anson, M. J. P. (2006). *Handbook of Alternative Assets*, 2nd edition. Hoboken, N. J.: John Wiley & Sons.

Fung, W. and Hsieh, D. A. (1997). Empirical characteristics of dynamic trading strategies: The case of hedge funds. *Review of Financial Studies* 10, Summer: 275–302.

Greer, R. (1997). What is an asset class anyway? *Journal of Portfolio Management* 23, 2:83–91.

Horvitz, J. (2000). Asset classes and asset allocation: Problems of classification. *Journal of Private Portfolio Management* 2, 4:27–32.

Jaeger, L. (2002). Managing risk in alternative investment strategies, *Financial Times*, London: Prentice Hall.

Markowitz, H. M. (1959). *Portfolio Selection*. Cowles Foundation, New Haven, Conn.: Yale University Press.

Modigliani, F., and Miller, M. (1958). The cost of capital, corporation finance, and the theory of investment. *American Economic Review* XLVIII, June: 433–443.

Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *Journal of Portfolio Management* 18, 2: 7–19.

CHAPTER 53

Hedge Funds

MARK J. P. ANSON, PhD, JD, CPA, CFA, CAIA
President and Executive Director of Nuveen Investment Services

Abstract: The global hedge fund market is estimated at $1.4 trillion. This capital has flowed into hedge funds because of their unrestricted ability to invest across asset classes, to go both long and short securities, and to invest without the burden or constraint of a traditional benchmark. However, investing in hedge funds is not easy. They may pursue esoteric strategies and may take on considerable leverage to boost their returns. Nonetheless, their popularity persists because of their ability to generate returns from pricing discrepancies in the financial markets. Still, successful investing in hedge funds requires considerable due diligence—a process that requires both time and patience.

Keywords: hedge fund, market directional, corporate restructuring, convergence trading, opportunistic, equity long/short, market timers, short sellers, distressed securities, merger arbitrage, event driven, fixed income arbitrage, convertible bond arbitrage, market neutral, statistical arbitrage, relative-value arbitrage, global macro, fund of funds, due diligence, absolute return

The phrase "hedge fund" is a term of art. It is not defined in the Securities Act of 1933 or the Securities Exchange Act of 1934. Additionally, "hedge fund" is not defined by the Investment Company Act of 1940, the Investment Advisers Act of 1940, the Commodity Exchange Act, or, finally, the Bank Holding Company Act. Even though the Securities and Exchange Commission (SEC) has attempted (unsuccessfully) to regulate hedge funds, it has yet to define the term *hedge fund* within its security regulations. So what is this investment vehicle that every investor seems to know but for which there is scant regulatory guidance?

As a starting point, we turn to the *American Heritage Dictionary*, 3rd edition, which defines a hedge fund as:

> An investment company that uses high-risk techniques, such as borrowing money and selling short, in an effort to make extraordinary capital gains.

This is a good start; however, many hedge fund strategies use tightly controlled, low-risk strategies to produce consistent but conservative rates of return and do not "swing for the fences" to earn extraordinary gains.

We define *hedge fund* as a privately organized investment vehicle that manages a concentrated portfolio of public and private securities and derivative instruments on those securities, that can invest both long and short and can apply leverage.

In this chapter we will discuss the various types of hedge funds according to the investment strategies that they pursue and considerations in investing in hedge funds.

HEDGE FUNDS VERSUS MUTUAL FUNDS

Within this definition there are six key elements of hedge funds that distinguish them from their more traditional counterpart, the mutual fund.

First, hedge funds are private investment vehicles that pool the resources of sophisticated investors. One of the ways that hedge funds avoid the regulatory scrutiny of the SEC or the Commodity Futures Trading Commission (CFTC) is that they are available only for high-net-worth investors. Under SEC rules, hedge funds cannot have more than 100 accredited investors in the fund. An accredited investor is defined as an individual that has a minimum net worth in excess of $1 million, or income in each of the past two years of $200,000 ($300,000 for a married couple) with an expectation of earning at least that amount in the current year. Additionally, hedge funds may accept no more than 500 "qualified purchasers" in the fund. These are individuals or institutions that have a net worth in excess of $5 million.

There is a penalty, however, for the privacy of hedge funds. They cannot raise funds from investors via a public offering. Additionally, hedge funds may not advertise broadly or engage in a general solicitation for new funds. Instead, their marketing and fund-raising efforts must be targeted to a narrow niche of very wealthy individuals and institutions. As a result, the predominant investors in hedge funds are family offices, foundations, endowments, and, to a lesser extent, pension funds.

Second, hedge funds tend to have portfolios that are much more concentrated than their mutual fund brethren. Most hedge funds do not have broad securities benchmarks. One reason is that most hedge fund managers claim that their style of investing is "skill based" and cannot be measured by a market return. Consequently, hedge fund managers are not forced to maintain security holdings relative to a benchmark; they do not need to worry about "benchmark" risk. This allows them to concentrate their portfolio on only those securities that they believe will add value to the portfolio.

Another reason for the concentrated portfolio is that hedge fund managers tend to have narrow investment strategies. These strategies tend to focus on only one sector of the economy or one segment of the market. They can tailor their portfolio to extract the most value from their smaller investment sector or segment. Furthermore, the concentrated portfolios of hedge fund managers generally are not dependent on the direction of the financial markets, in contrast to long-only managers.

Third, hedge funds tend to use derivative strategies much more predominantly than mutual funds. Indeed, in some strategies, such as convertible arbitrage, the ability to sell or buy options is a key component of executing the arbitrage. The use of derivative strategies may result in nonlinear cash flows that may require more sophisticated risk management techniques to control these risks.

Fourth, hedge funds may go both long and short securities. The ability to short public securities and derivative instruments is one of the key distinctions between hedge funds and traditional money managers. Hedge fund managers incorporate their ability to short securities explicitly into their investment strategies. For example, equity long/short hedge funds tend to buy and sell securities within the same industry to maximize their return but also to control their risk. This is very different from traditional money managers that are tied to a long-only securities benchmark.

Fifth, many hedge fund strategies invest in nonpublic securities, that is, securities that have been issued to investors without the support of a prospectus and a public offering. Many bonds, both convertible and high yield, are issued as what are known as "144A securities." These are securities issued to institutional investors in a private transaction instead of a public offering. These securities may be offered with a private placement memorandum (ppm), but not a public prospectus. In addition, these securities are offered without the benefit of an SEC review as would be conducted for a public offering. Bottom line: with 144A securities it is buyer beware. The SEC allows this because, presumably, large institutional investors are more sophisticated than that average, small investor.

Finally, hedge funds use leverage, sometimes, large amounts. In fact, a lesson in leverage is described in this chapter with respect to Long-Term Capital Management. Mutual funds, for example, are limited in the amount of leverage they can employ; they may borrow up to 33% of their net asset base. Hedge funds do not have this restriction. Consequently, it is not unusual to see some hedge fund strategies that employ leverage up to 10 time their net asset base.

We can see that hedge funds are different than traditional long-only investment managers.

CATEGORIES OF HEDGE FUNDS

It seems like everyone has their own classification scheme for hedge funds (see, for example, L'habitant [2004] and Nicholas [2000]). This merely reflects the fact that hedge funds are a bit difficult to "box in"—a topic we will address further when we examine a number of the hedge fund index providers. For purposes of this book, we try to break down hedge funds into broad categories, as depicted in Figure 53.1.

We classify hedge funds into four broad buckets: market directional, corporate restructuring, convergence trading, and opportunistic. Market directional hedge funds are those that retain some amount of systematic risk exposure. For example, equity long/short (or, as it is sometime called, equity hedge) are hedge funds that typically contain some amount of net long market exposure. For

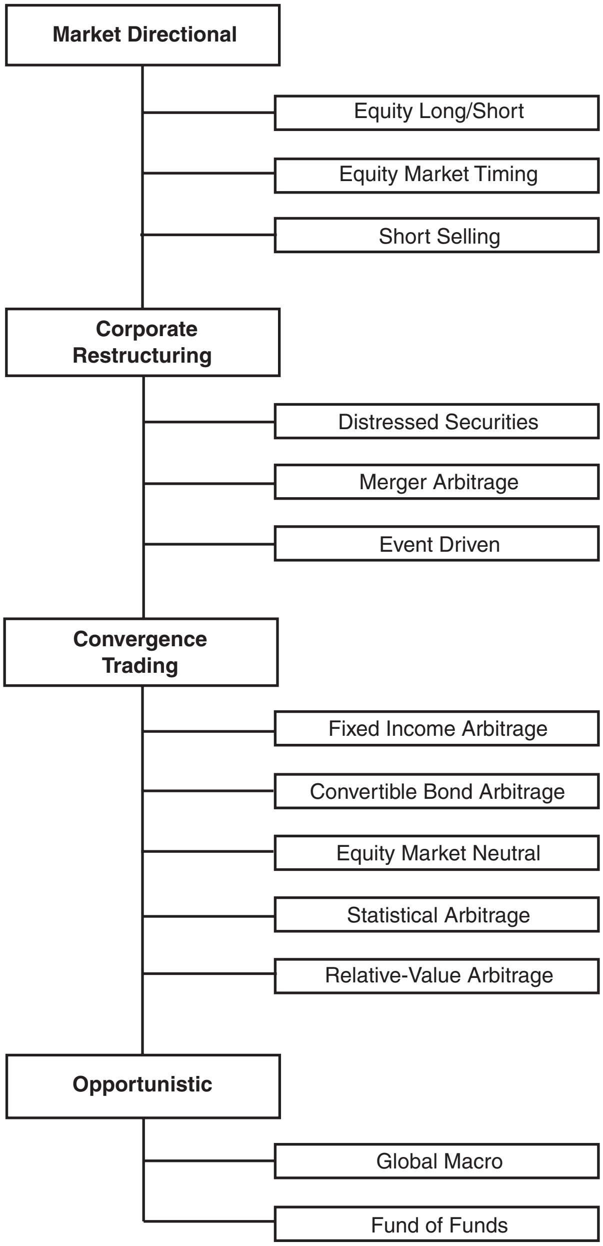


Figure 53.1 Categories of Hedge Funds

example, they may leverage up the hedge fund to go 150% long on stocks that they like while simultaneously shorting 80% of the fund value with stocks that they think will decline in value. The remaining net long market exposure is 70%. Thus, they retain some amount of systematic risk exposure that will be affected by the direction of the stock market.

Corporate restructuring hedge funds take advantage of significant corporate transactions like mergers, acquisitions, or bankruptcies. These funds earn their living by concentrating their portfolios on a handful of companies where it is more important to understand the likelihood that the corporate transaction will be completed than it is to determine whether the corporation is under- or overvalued.

Convergence trading hedge funds are the hedge funds that practice the art of arbitrage. In fact, the specialized subcategories within this bucket typically contain the word *arbitrage* in their description, such as statistical arbitrage, fixed income arbitrage, or convertible arbitrage. In general, these hedge funds make bets that two similar securities but with dissimilar prices will converge to the same value over the investment holding period.

Finally, we have the opportunistic category. We include global macrohedge funds as well as fund of funds (FOF) in this category. These funds are designed to take advantage of whatever opportunities present themselves, hence the term *opportunistic*. For example, FOF often practice tactical asset allocation among the hedge funds contained in the FOF based on the FOF manager's view as to which hedge fund strategies are currently poised to earn the best results. This shifting of the assets around is based on the FOF manager's assessment of the opportunity for each hedge fund contained in the FOF to earn a significant return.

HEDGE FUND STRATEGIES

Hedge funds invest in the same equity and fixed income securities as traditional long-only managers. Therefore, it is not the alternative "assets" in which hedge funds invest that differentiates them from long-only managers, but rather, it is the alternative investment strategies that they pursue. In this section we provide more detail on the types of strategies pursued by hedge fund managers.

Market Direction Hedge Funds

The strategies in this bucket of hedge funds either retain some systematic market exposure associated with the stock market such as equity long/short or are specifically driven by the movements of the stock market such as market timing or short selling.

Equity Long/Short

Equity long/short managers build their portfolios by combining a core group of long stock positions with short sales of stock or stock index options/futures. Their net market exposure of long positions minus short positions tends to have a positive bias. That is, equity long/short managers tend to be long market exposure. The length of their exposure depends on current market conditions. For instance, during the great stock market surge of 1996 to 1999, these managers tended to be mostly long their equity exposure. However, as the stock market turned into a bear market in 2000 to 2002, these managers decreased their market exposure as they sold more stock short or sold stock index options and futures.

The ability to go both long and short in the market is a powerful tool for earning excess returns. The ability to fully implement a strategy not only about stocks and sectors that are expected to increase in value but also stocks and sectors that are expected to decrease in value allows the hedge fund manager to maximize the value of her market insights.

Equity long/short hedge funds essentially come in two flavors: fundamental or quantitative. Fundamental long/short hedge funds conduct traditional economic analysis on a company's business prospects compared to its competitors and the current economic environment. These shops will visit with management, talk with Wall Street analysts, contact customers and competitors and essentially conduct bottom-up analysis. The difference between these hedge funds and long-only managers is that they will short the stocks that they consider to be poor performers and buy those stocks that are expected to outperform the market. In addition, they may leverage their long and short positions.

Fundamental long/short equity hedge funds tend to invest in one economic sector or market segment. For instance, they may specialize in buying and selling Internet companies (sector focus) or buying and selling small market capitalization companies (segment focus).

In contrast, quantitative equity long/short hedge fund managers tend not to be sector or segment specialists. In fact, quite the reverse, quantitative hedge fund managers like to cast as broad a net as possible in their analysis. These managers are often referred to as statistical arbitrageurs because they base their trade selection on the use of quantitative statistics instead of fundamental stock selection.

Market Timers

Market timers, as their name suggests, attempt to time the most propitious moments to be in the market and invest in cash otherwise. More specifically, they attempt to time the market so that they are fully invested during bull markets, and strictly in cash during bear markets.

Unlike equity long/short strategies, market timers use a top-down approach as opposed to a bottom-up approach. Market-timing hedge fund managers are not stock pickers. They analyze fiscal and monetary policy as well as key macroeconomic indicators to determine whether the economy is gathering or running out of steam.

Macroeconomic variables they may analyze are labor productivity, business investment, purchasing managers' surveys, commodity prices, consumer confidence, housing starts, retail sales, industrial production, balance of payments, current account deficits/surpluses, and durable good orders.

They use this macroeconomic data to forecast the expected gross domestic product (GDP) for the next quarter. Forecasting models typically are based on multifactor linear regressions, taking into account whether a variable is a leading or lagging indicator and whether the variable experiences any seasonal effects.

Once market timers have their forecast for the next quarter(s), they position their investment portfolio in the market according to their forecast. Construction of their portfolio is quite simple. They do not need to purchase individual stocks. Instead, they buy or sell stock index futures and options to increase or decrease their exposure to the market as necessary. At all times, contributed capital from investors is kept in short-term, risk-free, interest bearing accounts. Treasury bills are often purchased, which not only yield a current risk-free interest rate, but also can be used as margin for the purchase of stock index futures.

When a market timer's forecast is bullish, he may purchase stock index futures with an economic exposure equivalent to the contributed capital. He may apply leverage by purchasing futures contracts that provide an economic exposure to the stock market greater than that of the underlying capital. However, market timers generally do not borrow investment capital.

When the hedge fund manager is bearish, he will trim his market exposure by selling futures contracts. If he is completely bearish, he will sell all of his stock index futures and call options and just sit on his cash portfolio. Some market timers may be more aggressive and short stock index futures and buy stock index put options to take advantage of bear markets.

In general, though, market timers tend to have long exposure to the market at all times, making them market directional. However, they attempt to trim this exposure when markets appear bearish. This was demonstrated during the bear market years of 2000 to 2002. Consequently, we find that market timers have a similar, but slightly more conservative, risk profile than the stock market.

Short Selling

Short selling hedge funds have the opposite exposure of traditional long-only managers. In that sense, their return distribution should be the mirror image of long-only managers: they make money when the stock market is declining and lose money when the stock market is gaining.

These hedge fund managers may be distinguished from equity long/short managers in that they generally maintain a net short exposure to the stock market. However, short selling hedge funds tend to use some form of market timing. That is, they trim their short positions when the stock market is increasing and go fully short when the stock market is declining. When the stock market is gaining, short sellers maintain that portion of their investment capital not committed to short selling in short-term interest rate–bearing accounts.

Corporate Restructuring Hedge Funds

Many hedge fund articles call these strategies "event driven" or "risk arbitrage," but that does not really describe what is at the heart of each of these type of strategies. The focal point is some form of corporate restructuring such as a merger, acquisition, or bankruptcy. Companies that are undergoing a significant transformation generally provide an opportunity for trading around that event. These strategies are driven by the deal, not by the market.

Distressed Securities

Distressed debt hedge funds invest in the securities of a corporation that is in bankruptcy, or is likely to fall into bankruptcy. Companies can become distressed for any number of reasons such as too much leverage on their balance sheet, poor operating performance, accounting irregularities, or even competitive pressure. Some of these strategies can overlap with private equity strategies that we will discuss in Part Four of this book. The key difference here is that hedge funds are less concerned with the fundamental value of a distressed corporation and, instead, concentrate on trading opportunities surrounding the company's outstanding stock-and-bond securities.

There are many different variations on how to play a distressed situation, but most fall into three categories. In its simplest form, the easiest way to profit from a distressed corporation is to sell its stock short. This requires the hedge fund manager to borrow stock from its prime broker and sell in the marketplace stock that it does not own with the expectation that the hedge fund manager will be able to purchase the stock back at a later date and at a cheaper price as the company continues to spiral downward in its distressed situation. This is nothing more than "sell high and buy low."

However, the short selling of a distressed company exposes the hedge fund manager to significant risk if the company's fortunes should suddenly turn around. Therefore, most hedge fund managers in this space typically use a hedging strategy within a company's capital structure.

A second form of distressed securities investing is called capital structure arbitrage. Consider Company A, which has four levels of outstanding capital: senior secured debt, junior subordinated debt, preferred stock, and common stock. A standard distressed security investment strategy would be to:

1. Buy the senior secured debt and short the junior subordinated debt.
2. Buy the preferred stock and short the common stock.

In a bankruptcy situation, the senior secured debt stands in line in front of the junior subordinated debt for any bankruptcy-determined payouts. The same is true for the preferred stock compared to Company A's common stock. Both the senior secured debt and the preferred stock enjoy a higher standing in the bankruptcy process than either junior debt or common equity. Therefore, when the distressed situation occurs or progresses, senior secured debt should appreciate in value relative to the junior subordinated debt. In addition, there should be an increase in the spread of prices between preferred stock and common stock. When this happens, the hedge fund manager closes out her positions and locks in the profit that occurs from the increase in the spread.

Finally, distressed securities hedge funds can become involved in the bankruptcy process to find significantly undervalued securities. This is where an overlap with private equity firms can occur. To the extent that a distressed securities hedge fund is willing to learn the arcane workings of the bankruptcy process and to sit on creditor committees, significant value can be accrued if a distressed company can restructure and regain its footing. In a similar fashion, hedge fund managers do purchase the securities of a distressed company shortly before it announces its reorganization plan to the bankruptcy court with the expectation that there will be a positive resolution with the company's creditors.

Merger Arbitrage

Merger arbitrage is perhaps the best-known corporate restructuring investment among investors and hedge fund managers. Merger arbitrage generally entails buying the stock of the firm that is to be acquired and selling the stock of the firm that is the acquirer. Merger arbitrage managers seek to capture the price spread between the current market prices of the merger partners and the value of those companies upon the successful completion of the merger.

The stock of the target company usually trades at a discount to the announced merger price. The discount reflects the risk inherent in the deal; other market participants are unwilling to take on the full exposure of the transaction-based risk. Merger arbitrage is then subject to event risk. There is the risk that the two companies will fail to come to terms and call off the deal. There is the risk that another company will enter into the bidding contest, ruining the initial dynamics of the arbitrage. There is finally regulatory risk. Various U.S. and foreign regulatory agencies may not allow the merger to take place for antitrust reasons. Merger arbitrageurs specialize in assessing event risk and building a diversified portfolio to spread out this risk.

Merger arbitrageurs conduct significant research on the companies involved in the merger. They will review current and prior financial statements, SEC electronic data gathering analysis and retrieval (EDGAR) filings, proxy statements, management structures, cost savings from redundant operations, strategic reasons for the merger, regulatory issues, press releases, and competitive position of the combined company within the industries in which it competes. Merger arbitrageurs will calculate the rate of return that is implicit in the current spread and compare it to the event risk associated with the deal. If the spread is sufficient to compensate for the expected event risk, they will execute the arbitrage.

Once again, the term *arbitrage* is used loosely. As discussed earlier, there is plenty of event risk associated with a merger announcement. The profits earned from merger arbitrage are not riskless. Consider the saga of the purchase of MCI Corporation by Verizon Communications. Throughout 2005, Verizon was in a bidding war against Qwest Communications for the purchase of MCI. On February 3, 2005, Qwest announced a $6.3 billion merger offer for MCI. This bid was quickly countered by Verizon on February 10 that matched the $6.3 billion bid established by Qwest. The bidding war raged back and forth for several months before Verizon finally won the day in October of 2005 with an ultimate purchase price of $8.44 billion.

To see the vicissitudes of merger arbitrage at work, we follow both the successful Verizon bid for MCI as well as the unsuccessful bid by Qwest.

Starting with Verizon: at the announcement of its bid for MCI, its stock was trading at $36.00, while MCI was trading at $20. Therefore, the merger arbitrage trade was:

Sell 1,000 shares of Verizon at $36 (short proceeds of $36,000).
Buy 1,000 shares of MCI at $20 (cash outflow of $20,000).

While for the Qwest bid, the trade was:

Sell 1,000 shares of Qwest at \$4.20 (short proceeds of \$4,200).
Buy 1,000 shares of MCI at \$20 (cash outflow of \$20,000).

Throughout the spring and summer of 2005, Qwest and Verizon battled it out for MCI, with Verizon ultimately winning in October 2005. At that time, MCI's stock had increased in value to \$25.50, while Verizon's stock had lost value and was trading at \$30, and finally Qwest was trading unchanged at \$4.20.

Total return for the MCI/Verizon merger arbitrage trade:

Gain on MCI long position:	1,000 × (\$25.50 – \$20)	=	\$5,500
Gain on Verizon short position:	1,000 × (\$36 – \$30)	=	\$6,000
Interest on short rebate:	4% × 1,000 × \$36 × 240/360	=	\$960
Total			\$12,460

The return on invested capital is: \$12,460 ÷ \$20,000 = 62.3%.

If the merger arbitrage manager had applied 50% leverage to this deal and borrowed half of the net outflow, the return would have been (ignoring financing costs):

$$\$12{,}460 \div \$10{,}000 = 124.6\% \text{ Total return}$$

Turning to the MCI/Qwest merger arbitrage trade, the total return was:

Gain on MCI long position	1,000 × (\$25.50 – \$20)	=	\$5,500
Gain on Qwest short position:	1,000 × (\$4.20 – \$4.20)	=	\$0
Gain on short rebate:	4% × 1,000 × \$4.20 × 240/360	=	\$112
Total			\$5,612

The return on invested capital is: \$5,612 ÷ \$20,000 = 28.06%. With 50% leverage the return would be: \$5,612 ÷ \$10,000 = 56.12%.

While both merger arbitrage trades made money, clearly, it made more sense to bet on the Verizon/MCI merger than the Qwest/MCI merger. This is where merger arbitrage managers make their money, by assessing the likelihood of one bid over another. Also, in a situation where there are two bidders for a company, there is a very high probability that there will be a successful merger with one of the bidders. Consequently, many merger arbitrage hedge fund managers will play both bids. This is exactly what happened in the MCI deal—many merger arbitrage managers bet on both the MCI/Verizon deal and the MCI/Qwest deal, expecting that one of the two suitors would be successful in winning the hand of MCI.

Some merger arbitrage managers invest only in announced deals. However, other hedge fund managers will put on positions on the basis of rumor or speculation. The deal risk is much greater with this type of strategy, but so too is the merger spread (the premium that can be captured).

To control for risk, most merger arbitrage hedge fund managers have some risk of loss limit at which they will exit positions. Some hedge fund managers concentrate only on one or two industries, applying their specialized knowledge regarding an economic sector to their advantage. Other merger arbitrage managers maintain a diversified portfolio across several industries to spread out the event risk.

Merger arbitrage is deal driven rather than market driven. Merger arbitrage derives its return from the relative value of the stock prices between two companies as opposed to the status of the current market conditions. Consequently, merger arbitrage returns should not be highly correlated with the general stock market.

Event Driven

Event-driven hedge funds are very similar, in their approach to investing, to distressed securities and merger arbitrage. The only difference is that their mandate is broader than the other two corporate restructuring strategies. Event-driven transactions include mergers and acquisitions, spin-offs, tracking stocks, accounting write-offs, reorganizations, bankruptcies, share buybacks, special dividends, and any other significant market event. Event-driven managers are nondiscriminatory in their transaction selection.

By their very nature, these special events are nonrecurring. Therefore, the financial markets typically do not digest the information associated with these transactions in a timely manner. The financial markets are simply less efficient when it comes to large, isolated transactions. This provides an opportunity for event-driven managers to act quickly and capture a premium in the market. Additionally, most of these events may be subject to certain conditions such as shareholder or regulatory approval. Therefore, there is significant deal risk associated with this strategy for which a savvy hedge fund manager can earn a return premium. The profitability of this type of strategy is dependent on the successful completion of the corporate transaction within the expected time frame.

Convergence Trading Hedge Funds

Hedge fund managers tend to use the term *arbitrage* somewhat loosely. Arbitrage is defined simply as riskless profits. It is the purchase of a security for cash at one price and the immediate resale for cash of the same security at a higher price. Alternatively, it may be defined as the simultaneous purchase of security A for cash at one price and the selling of identical security B for cash at a higher price. In both cases, the arbitrageur has no risk. There is no market risk because the holding of the securities is instantaneous. There is no basis risk because the securities are identical, and there is no credit risk because the transaction is conducted in cash.

Instead of riskless profits, in the hedge fund world, arbitrage is generally used to mean low-risk investments. Instead of the purchase and sale of identical instruments, there is the purchase and sale of similar instruments. The securities also may not be sold for cash, so there may be credit risk during the collection period. Finally, the purchase and sale may not be instantaneous. The arbitrageur may need to hold onto its positions for a period of time, exposing him to market risk.

Fixed Income Arbitrage

Fixed income arbitrage involves purchasing one fixed income security and simultaneously selling a similar fixed income security with the expectation that over the investment holding period, the two security prices will converge to a similar value. Hedge fund managers search continuously for these pricing inefficiencies across all fixed income markets. This is nothing more than buying low and selling high and waiting for the undervalued security to increase in value or the overvalued security to decline in value, or wait for both to occur.

The sale of the second security is done to hedge the underlying market risk contained in the first security. Typically, the two securities are related either mathematically or economically such that they move similarly with respect to market developments. Generally, the difference in pricing between the two securities is small, and this is what the fixed income arbitrageur hopes to gain. By buying and selling two fixed income securities that are tied together, the hedge fund manager hopes to capture a pricing discrepancy that will cause the prices of the two securities to converge over time.

However, because the price discrepancies can be small, the way hedge fund managers add more value is to leverage their portfolio through direct borrowings from their prime broker, or by creating leverage through swaps and other derivative securities. Bottom line: They find pricing anomalies, then "crank up the volume" through leverage.

Fixed income arbitrage does not need to use exotic securities. For example, it can be nothing more than buying and selling U.S. Treasury bonds. In the bond market, the most liquid securities are the *on-the-run* Treasury bonds. These are the most currently issued bonds issued by the U.S. Treasury Department. However, there are other U.S. Treasury bonds outstanding that have very similar characteristics to the on-the-run Treasury bonds. The difference is that *off-the-run* bonds were issued at an earlier date, and are now less liquid than the on-the-run bonds. As a result, price discrepancies occur. The difference in price may be no more than one-half or one quarter of a point ($25) but can increase in times of uncertainty when investor money shifts to the most liquid U.S. Treasury bond. During the Russian bond default crisis, for example, on-the-run U.S. Treasuries were valued as much as $100 more than similar, off-the-run U.S. Treasury bonds of the same maturity.

Nonetheless, when held to maturity, the prices of these two bonds will converge to the same value. Any difference will be eliminated by the time they mature, and any price discrepancy may be captured by the hedge fund manager.

Fixed income arbitrage is not limited to the U.S. Treasury market. It can be used with corporate bonds, municipal bonds, sovereign debt, or mortgage-backed securities.

Another form of fixed income arbitrage involves trading among fixed income securities that are close in maturity. This is a form of yield curve arbitrage. These types of trades are driven by temporary imbalances in the term structure of interest rates.

Still another subset of fixed income arbitrage uses mortgage-backed securities (MBSs). MBSs represent an ownership interest in an underlying pool of individual mortgages loaned by banks and other financial institutions. Therefore, an MBS is a fixed income security with underlying prepayment options. MBS hedge funds seek to capture pricing inefficiencies in the U.S. mortgage-backed market.

MBS arbitrage can be between fixed income markets, such as buying MBS and selling U.S. Treasuries. This investment strategy is designed to capture credit spread inefficiencies between U.S. Treasuries and MBSs. MBSs trade at a credit spread over U.S. Treasuries to reflect the uncertainty of cash flows associated with MBSs compared to the lack of credit risk associated with U.S. Treasury bonds.

During a flight to quality, investors tend to seek out the most liquid markets such as the on-the-run U.S. Treasury market. This may cause credit spreads to temporarily increase beyond what is historically or economically justified. In this case the MBS market will be priced "cheap" to U.S. Treasuries. The arbitrage strategy would be to buy MBS and sell U.S. Treasury, where the interest rate exposure of both instruments is sufficiently similar so as to eliminate most (if not all) of the market risk between the two securities. The expectation is that the credit spread between MBSs and U.S. Treasuries will decline and MBS bonds will increase in value relative to U.S. Treasuries.

MBS arbitrage can be quite sophisticated. MBS hedge fund managers use proprietary models to rank the value of MBS by their option-adjusted spread (OAS). The hedge fund manager evaluates the present value of an MBS by explicitly incorporating assumptions about the probability of prepayment options being exercised. In effect, the hedge fund manager calculates the option-adjusted price of the MBS and compares it to its current market price. The OAS reflects the MBS's average spread over U.S. Treasury bonds of a similar maturity, taking into account the fact that the MBS may be liquidated early from the exercise of the prepayment option by the underlying mortgagors.

The MBSs that have the best OAS compared to U.S. Treasuries are purchased, and then their interest rate exposure is hedged to zero. Interest rate exposure is neutralized using Treasury bonds, options, swaps, futures, and caps. MBS hedge fund managers seek to maintain a duration of zero. This allows them to concentrate on selecting the MBSs that yield the highest OAS.

There are many risks associated with MBS arbitrage. Chief among them are duration, convexity, yield curve rotation, prepayment risk, credit risk, and liquidity risk. Hedging these risks may require the purchase or sale of other MBS products such as interest-only strips and principal-only strips, U.S. Treasuries, interest rate futures, swaps, and options.

What should be noted about fixed income arbitrage strategies is that they do not depend on the direction of the general financial markets. Arbitrageurs seek out pricing inefficiencies between two securities instead of making bets on market direction.

Convertible Bond Arbitrage

Convertible bonds combine elements of both stocks and bonds in one package. A convertible bond is a bond that contains an embedded option to convert the bond into the underlying company's stock.

Convertible arbitrage funds build long positions of convertible bonds and then hedge the equity component of the bond by selling the underlying stock or options on that stock. Equity risk can be hedged by selling the appropriate ratio of stock underlying the convertible option. This hedge ratio is known as the "delta" and is designed to measure the sensitivity of the convertible bond value to movements in the underlying stock.

Convertible bonds that trade at a low premium to their conversion value tend to be more correlated with the movement of the underlying stock. These convertibles then trade more like stock than they do a bond. Consequently, a high hedge ratio, or delta, is required to hedge the equity risk contained in the convertible bond. Convertible bonds that trade at a premium to their conversion value are highly valued for their bondlike protection. Therefore, a lower delta hedge ratio is necessary.

However, convertible bonds that trade at a high conversion act more like fixed income securities and therefore have more interest rate exposure than those with more equity exposure. This risk must be managed by selling interest rate futures, interest rate swaps, or other bonds. Furthermore, it should be noted that the hedging ratios for equity and interest rate risk are not static; they change as the value of the underlying equity changes and as interest rates change. Therefore, the hedge fund manager must continually adjust his hedge ratios to ensure that the arbitrage remains intact.

If this all sounds complicated, it is, but that is how hedge fund managers make money. They use sophisticated option-pricing models and interest rate models to keep track of all of the moving parts associated with convertible bonds. Hedge fund managers make arbitrage profits by identifying pricing discrepancies between the convertible bond and its component parts, and then continually monitoring these component parts for any change in their relationship.

Consider the following example: A hedge fund manager purchases 10 convertible bonds with a par value of $1,000, a coupon of 7.5%, and a market price of $900. The conversion ratio for the bonds is 20. The conversion ratio is based on the current price of the underlying stock, $45, and the current price of the convertible bond. The delta, or hedge, ratio for the bonds is 0.5. Therefore, to hedge the equity exposure in the convertible bond, the hedge fund manager must short the following shares of underlying stock:

$$10 \text{ Bonds} \times 20 \text{ Conversion ratio} \times 0.5 \text{ Hedge ratio} \\ = 100 \text{ Shares of stock}$$

To establish the arbitrage, the hedge fund manager purchases 10 convertible bonds and sells 100 shares of stock. With the equity exposure hedged, the convertible bond is transformed into a traditional fixed income instrument with a 7.5% coupon.

Additionally, the hedge fund manager earns interest on the cash proceeds received from the short sale of stock. This is known as the "short rebate." The cash proceeds remain with the hedge fund manager's prime broker, but the hedge fund manager is entitled to the interest earned on the cash balance from the short sale (a rebate). (The short rebate is negotiated between the hedge fund manager and the prime broker. Typically, large, well-established hedge fund managers receive a larger short rebate.) We assume that the hedge fund manager receives a short rebate of 4.5%. Therefore, if the hedge fund manager holds the convertible arbitrage position for one year, he expects to earn interest not only from his long bond position, but also from his short stock position.

The catch to this arbitrage is that the price of the underlying stock may change as well as the price of the bond. Assume the price of the stock increases to $47 and the price of the convertible bond increases to $920. If the hedge fund manager does not adjust the hedge ratio during the holding period, the total return for this arbitrage will be:

Appreciation of bond price:	10 × ($920 − $900)	=	$200
Appreciation of stock price:	100 × ($45 − $47)	=	−$200
Interest on bonds:	10 × $1,000 × 7.5%	=	$750
Short rebate:	100 × $45 × 4.5%	=	$202.50
Total:			$952.50

If the hedge fund manager paid for the 10 bonds without using any leverage, the holding period return is

$$\$952.50 \div \$9{,}000 = 10.58\%$$

Suppose the underlying stock price declined from $45 to $43, and the convertible bonds declined in value from $900 to $880. The hedge fund manager would then earn:

Depreciation of bond price:	10 × ($880 – $900)	=	−$200
Depreciation of stock price:	100 × ($45 – $43)	=	$200
Interest on bonds:	10 × $1,000 × 7.5%	=	$750
Short rebate:	100 × $45 × 4.5%	=	$202.50
Total			$952.50

What this example demonstrates is that with the proper delta or hedge ratio in place, the convertible arbitrage manager should be insulated from movements in the underlying stock price so that the expected return should be the same regardless of whether the stock price goes up or goes down.

However, suppose that the hedge fund manager purchased the convertible bonds with $4,500 of initial capital and $4,500 of borrowed money. We further assume that the hedge fund manager borrows the additional investment capital from his prime broker at a prime rate of 6%.

Our analysis of the total return is then:

Appreciation of bond price:	10 × ($920 − $900)	=	$200
Appreciation of stock price:	100 × ($47 − $45)	=	−$200
Interest on bonds:	10 × $1,000 × 7.5%	=	$750
Short rebate:	100 × $45 × 4.5%	=	$202.5
Interest on borrowing:	6% × $4,500	=	−$270
Total:			$682.5

And the total return on capital is

$$\$682.5 \div \$4{,}500 = 15.17\%$$

The amount of leverage used in convertible arbitrage will vary with the size of the long positions and the objectives of the portfolio. Yet, in the preceding example, we can see how using a conservative leverage ratio of 2:1 in the purchase of the convertible bonds added almost 500 basis points of return to the strategy and earned a total return equal to twice that of the convertible bond coupon rate.

It is easy to see why hedge fund managers are tempted to use leverage. Hedge fund managers earn incentive fees on every additional basis point of return they earn. Furthermore, even though leverage is a two-edged sword—it can magnify losses as well as gains—hedge fund managers bear no loss if the use of leverage turns against them. In other words, hedge fund managers have everything to gain by applying leverage, but nothing to lose.

Leverage is also inherent in the shorting strategy because the underlying short equity position must be borrowed. Convertible arbitrage leverage can range from two to six times the amount of invested capital. This may seem significant, but it is lower than other forms of arbitrage.

Convertible bonds earn returns for taking on exposure to a number of risks such as (1) liquidity (convertible bonds are typically issued as private securities); (2) credit risk (convertible bonds are usually issued by less than investment grade companies); (3) event risk (the company may be downgraded or declare bankruptcy); (4) interest rate risk (as a bond it is exposed to interest rate risk); (5) negative convexity (most convertible bonds are callable); and (6) model risk (it is complex to model all of the moving parts associated with a convertible bond). These events are magnified only when leverage is applied.

Since convertible bond managers hedge away the equity risk through delta-neutral hedging, we should see little impact from the U.S. stock market. In addition, for undertaking all of the risks listed above, convertible bond arbitrage managers should earn a return premium to U.S. Treasury bonds.

Market Neutral

Market-neutral hedge funds also go long and short the market. The difference is that they maintain integrated portfolios, which are designed to neutralize market risk. This means being neutral to the general stock market as well as having neutral risk exposures across industries. Security selection is all that matters.

Market-neutral hedge fund managers generally apply the rule of one alpha (see Jacobs and Levy, 1995). This means that they build an integrated portfolio designed to produce only one source of alpha. This is distinct from equity long/short managers that build two separate portfolios: one long and one short, with two sources of alpha. The idea of integrated portfolio construction is to neutralize market and industry risk and concentrate purely on stock selection. In other words, there is no beta risk in the portfolio with respect to either the broad stock market or any industry. Only stock selection, or alpha, should remain.

Market-neutral hedge fund managers generally hold equal positions of long and short stock positions. Therefore, the manager is dollar neutral; there is no net exposure to the market either on the long side or on the short side. Additionally, market-neutral managers generally apply no leverage because there is no market exposure to leverage. However, some leverage is always inherent when stocks are borrowed and shorted. Nonetheless, the nature of this strategy is that it does not have credit risk.

Generally, market-neutral managers follow a three-step procedure in their strategy. The first step is to build an initial screen of "investable" stocks. These are stocks traded on the manager's local exchange, with sufficient liquidity so as to be able to enter and exit positions quickly, and with sufficient float so that the stock may be borrowed from the hedge fund manager's prime broker for short positions. Additionally, the hedge fund manager may limit his universe to a capitalization segment of the equity universe such as the midcap range.

Second, the hedge fund manager typically builds factor models. These models are often known as "alpha engines." Their purpose is to find those financial variables that influence stock prices. These are bottom-up models that concentrate solely on corporate financial information as opposed to macroeconomic data. This is the source of the manager's skill—his stock-selection ability.

The last step is portfolio construction. The hedge fund manager will use a computer program to construct his portfolio in such a way that it is neutral to the market as well as across industries. The hedge fund manager may use a commercial "optimizer"—computer software designed to measure exposure to the market and produce a trade list for execution based on a manager's desired exposure to the market—or he may use his own computer algorithms to measure and neutralize risk.

Most market-neutral managers use optimizers to neutralize market and industry exposure. However, more sophisticated optimizers attempt to keep the portfolio neutral to several risk factors. These include size, book to value, price/earnings ratios, and market price to book value ratios. The idea is to have no intended or unintended risk exposures that might compromise the portfolio's neutrality.

We have more to say about transparency in our chapters regarding the selection of hedge fund managers and whether the hedge fund industry should be institutionalized. For now, it is sufficient to point out that black boxes tend to be problematic for investors.

We would expect market-neutral managers to produce returns independent of the general market (they are neutral to the market).

Statistical Arbitrage

A close cousin to equity market-neutral hedge fund managers is statistical arbitrage. The key difference is the amount of quantitative input. While equity market neutral is based more on fundamental research, statistical arbitrage is driven purely by quantitative factor models.

These managers use mathematical analysis to review past company performance in light of several quantitative factors. For instance, these managers may build regression models to determine the impact of market price to book value (price/book ratio) on companies across the universe of stocks as well as different market segments or economic sectors. Or they may analyze changes in dividend yields on stock price performance.

These are linear and quadratic regression equations designed to identify those economic factors that consistently have an impact on share prices. This process is very similar to that discussed with respect to equity long/short hedge fund managers. Indeed, the two strategies are very similar in their stock-selection methods. The difference is that equity long/short managers tend to have a net long exposure to the market while market-neutral managers have no exposure.

Typically, these managers build multifactor models, both linear and quadratic, to identify those economic factors that have a consistent impact on share prices. A key part of building their alpha engines is to apply the quantitative model on prior stock price performance to see if there is any predictive power in determining whether the stock of a particular company will rise or fall. If the model proves successful on historical data, the hedge fund manager will then conduct an "out of sample" test of the model. This involves testing the model on a subset of historical data that was not included in the model-building phase.

If a hedge fund identifies a successful quantitative strategy, it will apply its model mechanically. Buy and sell orders will be generated by the model and submitted to the order desk. In practice, the hedge fund manager will put limits on its model such as the maximum short exposure allowed or the maximum amount of capital that may be committed to any one stock position. In addition, quantitative hedge fund managers usually build in some qualitative oversight to ensure that the model is operating consistently.

Statistical arbitrage programs tend to be labeled *black boxes*. This is a term for sophisticated computer algorithms that lack transparency. The lack of transparency associated with these investment strategies comes in two forms. First, hedge fund managers, by nature, are secretive. They are reluctant to reveal their proprietary trading programs. Second, even if a hedge fund manager were to reveal his proprietary computer algorithms, these algorithms are often so sophisticated and complicated that they are difficult to comprehend.

Note that this strategy does not share in the large up-and-down cycles of the stock market. It earns a steady return, not as great as the stock market, but in excess of U.S. Treasuries. Remember, the goal of this strategy is to neutralize market risk and to profit on small price discrepancies between stocks in the same industry or sector. Consistent profits are the key; large bets are avoided.

Relative-Value Arbitrage

Relative-value arbitrage might be better named the *smorgasbord* of arbitrage. This is because relative-value hedge fund managers are catholic in their investment strategies; they invest across the universe of arbitrage strategies. The best known of these managers was Long-Term Capital Management. Once the story of LTCM unfolded, it was clear that their trading strategies involved merger arbitrage, fixed income arbitrage, volatility arbitrage, stub trading, and convertible arbitrage.

In general, the strategy of relative value managers is to invest in spread trades: the simultaneous purchase of one security and the sale of another when the economic relationship between the two securities (the "spread") has become mispriced. The mispricing may be based on historical averages or mathematical equations. In either case, the relative arbitrage manager purchases the security that is "cheap" and sells the security that is "rich." It is called relative-value arbitrage because the cheapness or richness of a security is determined *relative* to a second security. Consequently, relative-value managers do not take directional bets on the financial markets. Instead, they take focused bets on the pricing relationship between two securities.

Relative-value managers attempt to remove the influence of the financial markets from their investment strategies. This is made easy by the fact that they simultaneously buy and sell similar securities. Therefore, the market risk embedded in each security should cancel out. Any residual risk can be neutralized through the use of options or futures. What is left is pure security selection: the purchase of those securities that are relatively cheap and the sale of those securities that are relatively rich. Relative-value managers earn a profit when the spread between the two securities returns to normal. They then unwind their positions and collect their profit.

We have already discussed fixed income arbitrage, convertible arbitrage and statistical arbitrage. Two other popular forms of relative-value arbitrage are stub trading and volatility arbitrage.

Stub trading is an equity-based strategy. Frequently, companies acquire a majority stake in another company, but their stock price does not fully reflect their interest in the acquired company. As an example, consider Company A, whose stock is trading at $50. Company A owns a majority stake in Company B, whose remaining outstanding stock, or stub, is trading at $40. The value of Company A should be the combination of its own operations, estimated at $45 a share, plus its majority stake in Company B's operations, estimated at $8 a share. Therefore, Company A's share price is undervalued relative to the value that Company B should contribute to Company A's share price. The share price of Company A should be about $53, but instead, it is trading at $50. The investment strategy would be to purchase Company A's stock and sell the appropriate ratio of Company B's stock.

Let us assume that Company A's ownership in Company B contributes to 20% of Company A's consolidated operating income. Therefore, the operations of Company B should contribute one fifth to Company A's share price. A proper hedging ratio would be four shares of Company A's stock to one share of Company B's stock.

The arbitrage strategy is:

Buy four shares of Company A stock at 4 × $50 = $200
Sell one share of Company B stock at 1 × $40 = $40

The relative-value manager is now long Company A stock and hedged against the fluctuation of Company B's stock. Let us assume that over three months, the share price of Company B increases to $42 a share, the value of Company A's operations remains constant at $45, but now the shares of Company A correctly reflect the contribution of Company B's operations. The value of the position will be:

Value of Company A's operations:	4 × $45	=	$180
Value of Company B's operations	4 × $42 × 20%	=	$33.6
Loss on short of Company B stock	1 × ($40 – $42)	=	−$2
Short rebate on Company B stock	1 × $40 × 4.5% × 3/12	=	$0.45
Total:			$212.05

The initial invested capital was $200 for a gain of $12.05 or 6.02% over three months. Suppose the stock of Company B had declined to $30, but Company B's operations were properly valued in Company A's share price. The position value would be:

Value of Company A's operations:	4 × $45	=	$180
Value of Company B's operations:	4 × $30 × 20%	=	$24
Gain on short of Company B's stock:	1 × ($40 – $30)	=	$10
Short rebate on Company B's stock:	1 × $40 × 4.5% × 3/12	=	$0.045
Total:		=	$214.45

The initial invested capital was $200 for a gain of $14.45 or 7.22% over three months. Stub trading is not arbitrage. Although the value of Company B's stock has been hedged, the hedge fund manager must still hold its position in Company A's stock until the market recognizes its proper value.

Volatility arbitrage involves options and warrant trading. Option prices contain an *implied* number for volatility. That is, it is possible to observe the market price of an option and back out the value of volatility implied in the current price using various option pricing models. The arbitrageur can then compare options on the same underlying stock to determine if the volatility implied by their prices are the same.

The implied volatility derived from option pricing models should represent the expected volatility of the underlying stock that will be realized over the life of the option. Therefore, two options on the same underlying stock should have the same implied volatility. If they do not, an arbitrage opportunity may be available. Additionally, if the implied volatility is significantly different from the historical volatility of the underlying stock, then relative-value arbitrageurs expect the implied volatility will revert back to its historical average.

Volatility arbitrage generally is applied in one of two models. The first is a mean reversion model. This model compares the implied volatility from current option prices to the historical volatility of the underlying security with the expectation that the volatility reflected in the current option price will revert to its historical average and the option price will adjust accordingly.

A second volatility arbitrage model applies a statistical technique called generalized autoregressive conditional heteroskedasticity (GARCH). GARCH models use prior data points of realized volatility to forecast future volatility. The GARCH forecast is then compared to the volatility implied in current option prices.

Both models are designed to allow hedge fund managers to determine which options are priced "cheap" versus "rich." Once again, relative-value managers sell those options that are rich based on the implied volatility *relative* to the historical volatility and buy those options with cheap volatility relative to historical volatility.

Opportunistic Hedge Fund Strategies

Along the lines of the *smorgasbord* comment for relative-value hedge funds, these strategies have the broadest mandate across the financial, commodity, and futures markets. These all-encompassing mandates can lead to specific bets on currencies or stocks as well as a well-diversified portfolio.

Global Macro

As their name implies, global macro hedge funds take a macroeconomic approach on a global basis in their investment strategy. These are top-down managers who invest opportunistically across financial markets, currencies, national borders, and commodities. They take large positions depending on the hedge fund manager's forecast of changes in interest rates, currency movements, monetary policies, and macroeconomic indicators.

Global macro managers have the broadest investment universe. They are not limited by market segment or industry sector, nor by geographic region, financial market, or currency. Global macro also may invest in commodities.

In fact, a fund of global macro hedge funds offers the greatest diversification of investment strategies.

Global macro hedge funds tend to have large amounts of investor capital. This is necessary to execute their macroeconomic strategies. In addition, they may apply leverage to increase the size of their macro bets. As a result, global macro hedge funds tend to receive the greatest attention and publicity in the financial markets.

The best known of these hedge funds was the Quantum Hedge Fund managed by George Soros. It is well documented that this fund made significant gains in 1992 by betting that the British pound would devalue (which it did). This fund was also accused of contributing to the "Asian Contagion" in the fall of 1997 when the government of Thailand devalued its currency, the baht, triggering a domino effect in currency movements throughout Southeast Asia.

In recent times, however, global macro hedge funds have fallen on hard times. One reason is that many global macrohedge funds were hurt by the Russian bond default in August 1998 and the bursting of the technology bubble in March 2000. These two events caused large losses for the global macro hedge funds.

A second reason, as indicated above, is that global macro hedge funds had the broadest investment mandate of any hedge fund strategy. The ability to invest widely across currencies, commodities, financial markets, geographic borders, and time zones is a two-edged sword. On the one hand, it allows global macro hedge funds the widest universe in which to implement their strategies. On the other hand, it lacks focus. As more institutional investors have moved into the hedge fund marketplace, they have demanded greater investment focus as opposed to free investment rein.

Fund of Funds

Finally, we come to hedge fund of funds. These are hedge fund managers that invest their capital in other hedge funds. These managers practice tactical asset allocation; reallocating capital across hedge fund strategies when they believe that certain hedge fund strategies will do better than others. For example, during the bear market of 2000 to 2002, short-selling strategies performed the best of all hedge fund categories. Not surprisingly, fund of fund managers allocated a significant portion of their portfolios to short sellers during the recent bear market. Other strategies that are popular in fund of funds are global macro, fixed income arbitrage, convertible arbitrage, statistical arbitrage, equity long/short, and event driven.

One drawback on fund of funds is the double layer of fees. Investors in hedge fund of funds typically pay a management fee plus profit-sharing fees to the hedge fund of funds managers in addition to the management and incentive fees that must be absorbed from the underlying hedge fund managers. This double layer of fees makes it difficult for fund of fund managers to outperform some of the more aggressive individual hedge fund strategies. However, the trade-off is better risk control from a diversified portfolio.

SHOULD HEDGE FUNDS BE PART OF AN INVESTMENT PROGRAM?

A considerable amount of research has been dedicated to examining the return potential of several hedge fund styles. Additionally, a number of studies have considered hedge funds within a portfolio context, that is, hedge funds blended with other asset classes.

The body of research on hedge funds demonstrates two key qualifications for hedge funds. First, over the time period of 1989 to 2000, the returns to hedge funds were positive. The highest returns were achieved by global macro hedge funds, and the lowest returns were achieved by short selling hedge funds. Not all categories of hedge funds beat the Standard and Poor's (S&P) 500. However, in many cases, the volatility associated with hedge fund returns was lower than that of the S&P 500, resulting in higher Sharpe ratios.

Second, the empirical research demonstrates that hedge funds provide good diversification benefits. In other words, hedge funds do, in fact, hedge other financial assets. Correlation coefficients with the S&P 500 range from -0.7 for short selling hedge funds to 0.83 for opportunistic hedge funds investing in the U.S. markets. The less-than-perfect positive correlation with financial assets indicates that hedge funds can expand the efficient frontier for asset managers.

In summary, the recent research on hedge funds indicates consistent, positive performance with low correlation with traditional asset classes. The conclusion is that hedge funds can expand the investment opportunity set for investors, offering both return enhancement as well as diversification benefits.

IS HEDGE FUND PERFORMANCE PERSISTENT?

This is the age-old question with respect to all asset managers, not just hedge funds: Can the manager repeat her good performance? This issue, though, is particularly acute for the hedge fund marketplace for two reasons. First, hedge fund managers often claim that the source of their returns is "skill based" rather than dependent on general financial market conditions. Second, hedge fund managers tend to have shorter track records than traditional money managers.

Unfortunately, the evidence regarding hedge fund performance persistence is mixed. The few empirical studies that have addressed this issue have provided inconclusive evidence whether hedge fund managers can produce enduring results. Part of the reason for the mixed results is the short track records of most hedge fund managers. A three-year or five-year track record is too short a period of time to be able to estimate an accurate expected return or risk associated with that manager.

In addition, the skill-based claim of hedge fund managers makes it more difficult to assess their performance relative to a benchmark. Without a benchmark index for comparison, it is difficult to determine whether a hedge fund manager has outperformed or underperformed her performance "bogey." As a result, the persistence of hedge fund manager performance will remain an open issue until manager databases with longer performance track records can be developed.

A HEDGE FUND INVESTMENT STRATEGY

The preceding discussion demonstrates that hedge funds can expand the investment opportunity set for investors. The question now becomes: What is to be accomplished by the hedge fund investment program? The strategy may be simply a search for an additional source of return. Conversely, it may be for risk management purposes. Whatever its purpose, an investment plan for hedge funds may consider one of three strategies. Hedge funds may be selected on an opportunistic basis, as a hedge fund of funds, or as an absolute-return strategy. A fourth possible strategy is a joint venture where an investor provides seed capital and investment capital for a new hedge fund manager. The investor receives professional hedge fund management plus a "piece of the action."

Opportunistic Hedge Fund Investing

The term *hedge fund* can be misleading. Hedge funds do not necessarily have to hedge an investment portfolio. Rather, they can be used to expand the investment opportunity set. This is the opportunistic nature of hedge funds—they can provide an investor with new investment opportunities that she cannot otherwise obtain through traditional long-only investments.

There are several ways hedge funds can be opportunistic. First, many hedge fund managers can add value to an existing investment portfolio through specialization in a sector or in a market strategy. These managers do not contribute portable alpha. Instead, they contribute above market returns through the application of superior skill or knowledge to a narrow market or strategy.

Consider a portfolio manager whose particular expertise is the biotechnology industry. She has followed this industry for years and has developed a superior information set to identify winners and losers. On the long-only side, the manager purchases those stocks that she believes will increase in value and avoids those biotech stocks she believes will decline in value. However, this strategy does not utilize her superior information set to its fullest advantage. The ability to go both long and short biotech stocks in a hedge fund is the only way to maximize the value of the manager's information set. Therefore, a biotech hedge fund provides a new opportunity: the ability to extract value on both the long side and the short side of the biotech market.

The goal of this strategy is to identify the best managers in a specific economic sector or specific market segment that complements the existing investment portfolio. These managers are used to enhance the risk and return profile of an existing portfolio, rather than hedge it.

Opportunistic hedge funds tend to have a benchmark. Take the example of the biotech long/short hedge fund. An appropriate benchmark would be the AMEX Biotech Index that contains 17 biotechnology companies. Alternatively, if the investor believed that the biotech sector will outperform the general stock market, she could use a broad-based stock index such as the S&P 500 for the benchmark. The point is that opportunistic hedge funds are not absolute-return vehicles (discussed later). Their performance can be measured relative to a benchmark.

As another example, most institutional investors have a broad equity portfolio. This portfolio may include an index fund, external value and growth managers, and possibly, private equity investments. However, along the spectrum of this equity portfolio, there may be gaps in its investment lineup. For instance, many hedge funds combine late-stage private investments with public securities. These hybrid funds are a natural extension of an institution's investment portfolio because they bridge the gap between private equity and index funds. Therefore, a new opportunity is identified: the ability to blend private equity and public securities within one investment strategy.

Alternative "assets" are really alternative investment strategies, and these alternative strategies are used to expand the investment opportunity set rather than hedge it. In summary, hedge funds may be selected not necessarily to reduce the risk of an existing investment portfolio, but instead, to complement its risk and return profile. Opportunistic investing is designed to select hedge fund managers that can enhance certain portions of a broader portfolio.

Another way to consider opportunistic hedge fund investments is that they are finished products because their investment strategy or market segment complements an institutional investor's existing asset allocation. In other words, these hybrid funds can plug the gaps of an existing portfolio. No further work is necessary on the part of the institution because the investment opportunity set has been expanded by the addition of the hybrid product. These "gaps" may be in domestic equity, fixed income, or international investments. Additionally, because opportunistic hedge funds are finished products, it makes it easier to establish performance benchmarks.

Constructing an opportunistic portfolio of hedge funds will depend on the constraints under which such a program operates. For example, if an investor's hedge fund program is not limited in scope or style, then diversification across a broad range of hedge fund styles would be appropriate. If, however, the hedge fund program is limited in scope to, for instance, expanding the equity investment opportunity set, the choices will be less diversified across strategies. Table 53.1 demonstrates these two choices.

Hedge Fund of Funds

A *hedge fund of funds* is an investment in a group of hedge funds, from 5 to more than 20. The purpose of a hedge fund of funds is to reduce the idiosyncratic risk of any one hedge fund manager. In other words, there is safety in numbers. This is simply the modern portfolio theory (MPT) applied

Table 53.1 Implementing an Opportunistic Hedge Fund Strategy

Diversified Hedge Fund Portfolio	Equity-Based Hedge Fund Portfolio
Equity long/short	Equity long/short
Short selling	Short selling
Market neutral	Market neutral
Merger arbitrage	Merger arbitrage
Event driven	Event driven
Convertible arbitrage	Convertible arbitrage
Global macro	
Fixed income arbitrage	
Relative-value arbitrage	
Market timers	

to the hedge fund marketplace. Diversification is one of the founding principles of MPT, and it is as applicable to hedge funds as it is to stocks and bonds.

Absolute Return

Hedge funds are often described as *absolute-return* products. This term comes from the skill-based nature of the industry. Hedge fund managers generally claim that their investment returns are derived from their skill at security selection rather than that of broad asset classes. This is due to the fact that most hedge fund managers build concentrated portfolios of relatively few investment positions and do not attempt to track a stock or bond index. The work of Fung and Hsieh (1997) shows that hedge funds generate a return distribution that is very different from mutual funds.

Further, given the generally unregulated waters in which hedge fund managers operate, they have greater flexibility in their trading style and execution than traditional long-only managers. This flexibility provides a greater probability that a hedge fund manager will reach his return targets. As a result, hedge funds have often been described as absolute-return vehicles that target a specific annual return regardless of what performance might be found among market indices. In other words, hedge fund managers target an absolute return rather than determine their performance relative to an index.

All traditional long-only managers are benchmarked to some passive index. The nature of benchmarking is such that it forces the manager to focus on his benchmark and his tracking error associated with that benchmark. This focus on benchmarking leads traditional active managers to commit a large portion their portfolios to tracking their benchmark. The necessity to consider the impact of every trade on the portfolio's tracking error relative to its assigned benchmark reduces the flexibility of the investment manager.

In addition, long-only active managers are constrained in their ability to short securities. They may only "go short" a security up to its weight in the benchmark index. If the security is only a small part of the index, the manager's efforts to short the stock will be further constrained. The inability to short a security beyond its benchmark weight deprives an active manager of a significant amount of the mispricing in the marketplace. Furthermore, not only are long-only managers unable to take advantage of overpriced securities, but they also cannot fully take advantage of underpriced securities because they cannot generate the necessary short positions to balance the overweights with respect to underpriced securities.

The flexibility of hedge fund managers allows them to go both long and short without benchmark constraints. This allows them to set a target rate of return or an "absolute return."

Specific parameters must be set for an absolute-return program. These parameters will direct how the hedge fund program is constructed and operated and should include risk and return targets as well as the type of hedge fund strategies that may be selected. Absolute-return parameters should operate at two levels: that of the individual hedge fund manager and for the overall hedge fund program. The investor sets target return ranges for each hedge fund manager but sets a specific target return level for the absolute return program. The parameters for the individual managers may be different than that for the program. For example, acceptable levels of volatility for individual hedge fund managers may be greater than that for the program.

The program parameters for the hedge fund managers may be based on such factors as volatility, expected return, types of instruments traded, leverage, and historical drawdown. Other qualitative factors may be included such as length of track record, periodic liquidity, minimum investment, and assets under management. Liquidity is particularly important because an investor needs to know with certainty her time-frame for cashing out of an absolute-return program if hedge fund returns turn sour.

Table 53.2 demonstrates an absolute-return program strategy. Notice that the return for the portfolio has a specific target rate of 15%, while for the individual hedge funds the return range is 10% to 25%. Also, the absolute-return portfolio has a target level for risk and drawdowns, while for the individual hedge funds, a range is acceptable.

However, certain parameters are synchronized. Liquidity, for instance, must be the same for both the

Table 53.2 An Absolute-Return Strategy

Absolute-Return Portfolio	Individual Hedge Fund Managers
Target return: 15%	Expected return: 10% to 25%
Target risk: 7%	Target risk: 5% to 15%
Largest acceptable draw-down: 10%	Largest drawdown: 10% to 20%
Liquidity: Semiannual	Liquidity: Semiannual
Hedge fund style: equity based	Hedge fund style: equity L/S, market neutral, merger arbitrage, short selling, event driven, convertible arbitrage
Length of track record: 3 years	Minimum track record: 3 years

absolute-return portfolio and that of the individual hedge fund managers. The reason is that a range of liquidity is not acceptable if the investor wishes to liquidate her portfolio. She must be able to cash out of each hedge fund within the same time-frame as that established for the portfolio.

SELECTING A HEDGE FUND MANAGER

The hedge fund industry is still relatively new. Most of the academic research on hedge funds was conducted during the 1990s. As a result, for most hedge fund managers, a two- to three-year track record is considered long term. In fact, Park, Brown, and Goetzmann (2001) find that the attrition rate in the hedge fund industry is about 15% per year and that the half-life for hedge funds is about 2.5 years. Liang (2001) documents an attrition rate of 8.54% per year for hedge funds. Weisman and Abernathy (2000) indicate that relying on a hedge fund manager's past performance history can lead to disappointing investment results. Consequently, performance history, while useful, cannot be relied upon solely in selecting a hedge fund manager.

Beyond performance numbers, there are three fundamental questions that every hedge fund manager should answer during the initial screening process. The answers to these three questions are critical to understanding the nature of the hedge fund manager's investment program. The three questions are:

1. What is the investment objective of the hedge fund?
2. What is the investment process of the hedge fund manager?
3. What makes the hedge fund manager so smart?

A hedge fund manager should have a clear and concise statement of its investment objective. Second, the hedge fund manager should identify its investment process. For instance, is it quantitatively or qualitatively based? Last, the hedge fund manager must demonstrate that he or she is smarter than other money managers.

The questions presented are threshold issues. These questions are screening tools designed to reduce an initial universe of hedge fund managers down to a select pool of potential investments. They are not, however, a substitute for a thorough due diligence review.

Investment Objective

The question of a hedge fund manager's investment objective can be broken down into three questions:

1. In which markets does the hedge fund manager invest?
2. What is the hedge fund manager's general investment strategy?
3. What is the hedge fund manager's benchmark, if any?

Although these questions may seem straightforward, they are often surprisingly difficult to answer. Consider the following language from a hedge fund disclosure document:

> The principal objective of the Fund is capital appreciation, primarily through the purchase and sale of securities, commodities and other financial instruments including without limitation, stocks, bonds, notes, debentures, and bills issued by corporations, municipalities, sovereign nations or other entities; options, rights, warrants, convertible securities, exchangeable securities, synthetic and/or structured convertible or exchangeable products, participation interests, investment contracts, mortgages, mortgage and asset-backed securities, real estate and interests therein; currencies, other futures, commodity options, forward contracts, money market instruments, bank notes, bank guarantees, letters of credit, other forms of bank obligations; other swaps and other derivative instruments; limited partnership interests and other limited partnership securities or instruments; and contracts relating to the foregoing; in each case whether now existing or created in the future.

Let's analyze the above statement in light of our three investment objective questions.

Question 1: In which markets does the hedge fund manager invest? Answer: In every market known to exist.

By listing every possible financial, commodity, or investment contract currently in existence (or to exist in the future), the hedge fund manager has covered all options, but has left the investor uninformed. Unfortunately, the unlimited nature of the hedge fund manager's potential investment universe does not help to narrow the scope of the manager's investment objective.

Question 2: What is the hedge fund manager's general strategy? Answer: Capital appreciation.

This answer too, is uninformative. Rarely does any investor invest in a hedge fund for capital *depreciation*. Generally, hedge funds are not used as tax shelters. Furthermore, many institutional investors are tax exempt so that taxes are not a consideration. Capital appreciation is assumed for most investments, including hedge funds. The preceding language is far too general to be informative.

Question 3: What is the manager's benchmark, if any? Answer: There is no effective benchmark. The manager's investment universe is so widespread as to make any benchmark useless.

Unfortunately, the preceding disclosure language, while very detailed, discloses very little. It does cover all of the manager's legal bases, but it does not inform the investor.

Where does this manager fall within the hedge fund spectrum? The very broad nature of this hedge fund's investment objective places it in the global macro category. Its investment universe is far too broad to be an arbitrage fund. By the same token, its strategy is too expansive to be considered an equity long/short program. Its only appropriate category is global macro.

By contrast, consider the following language from a second hedge fund disclosure document:

> The Fund's investment objective is to make investments in public securities that generate a long-term return in excess of that generated by the overall U.S. public

> equity market while reducing the market risk of the portfolio through selective short positions.

This one sentence answers all three investment objective questions. First, the manager identifies that it invests in the U.S. public equity market. Second, the manager discloses that it uses a long/short investment strategy. Finally, the manager states that its objective is to outperform the overall U.S. equity market. Therefore, a suitable benchmark might be the S&P 500, the Russell 1000, or a sector index.

This hedge fund is clearly identified as an equity long/short strategy. Its primary purpose is to take on market risk, not credit risk.

In summary, long-winded disclosure statements are not necessary. A well-thought-out investment strategy can be summarized in one sentence.

Investment Process

Most investors prefer a well-defined investment process that describes how an investment manager makes its investments. The articulation and documentation of the process can be just as important as the investment results generated by the process. Consider the following language from another hedge fund disclosure document:

> The manager makes extensive use of computer technology in both the formulation and execution of many investment decisions. Buy and sell decisions will, in many cases, be made and executed algorithmically according to quantitative trading strategies embodied in analytical computer software running the manager's computer facilities or on other computers used to support the Fund's trading activities.

This is a "black box." A black box is the algorithmic extension of the hedge fund manager's brain power. Computer algorithms are developed to quantify the manager's skill or investment insight.

For black box managers, the black box itself is the investment process. It is not that the black boxes are bad investments. In fact, the hedge fund research indicates that proprietary quantitative trading strategies can be quite successful. Rather, the issue is whether good performance results justify the lack of a clear investment process.

Black box programs tend to be used in arbitrage or relative-value hedge fund programs. Hedge fund managers use quantitative computer algorithms to seek out pricing discrepancies between similar securities or investment contracts. They then sell the investment that appears to be "expensive" and buy the investment that appears to be "cheap." The very nature of arbitrage programs is to minimize market risk. Leverage is then applied to extract the most value from their small net exposure to market risk.

A black box is just one example of process versus investment results. The hedge fund industry considers itself to be "skill based." However, it is very difficult to translate manager skill into a process. This is particularly true when the performance of the hedge fund is dependent on the skill of a specific individual.

Let's consider another, well-publicized skill-based investment process. In the spring of 2000, the hedge funds headed by George Soros stumbled, leading to the departure of Stanley Druckenmiller, the chief investment strategist for Soros Fund Management. The *Wall Street Journal* (May 1, 2000, p. C1) documented the concentrated skill-based investment style of this hedge fund group:

> For years, [Soros Fund Management] fostered an entrepreneurial culture, with a cadre of employees battling wits to persuade Mr. Druckenmiller to invest.
>
> "[Mr. Druckenmiller] didn't scream, but he could be very tough. It could be three days or three weeks of battling it out until he's convinced, or you're defeated."

The preceding statement does not describe an investment process. It is a description of an individual. The hedge fund manager's investment analysis and decision making is concentrated in one person. This is a pure example of "skill-based" investing. There is no discernible process. Instead, all information is filtered through the brain of one individual. In essence, the institutional investor must trust the judgment of one person.

Mr. Druckenmiller compiled an exceptional track record as the manager of the Soros Quantum Fund. However, the concentration of decision-making authority is not an economic risk, it is a process risk.

Investors should accept economic risk but not process risk. Soros Fund Management is a well-known global macro hedge fund manager. The fundamental risks of an investment in a global macro fund are credit risk and market risk.

Investors are generally unwilling to bear risks that are not fundamental to their tactical and strategic asset allocations. Process risk is not a fundamental risk. It is an idiosyncratic risk of the hedge fund manager's structure and operations.

Generally, process risk is not a risk that investors wish to bear. Nor is it a risk for which they expect to be compensated. Furthermore, how would an investor go about pricing the process risk of a hedge fund manager? It can't be quantified, and it can't be calibrated. Therefore, there is no way to tell whether an institutional investor is being properly compensated for this risk. For example, Park and Staum (1998) demonstrate that idiosyncratic process risks can largely be eliminated through a diversified fund of funds program. They indicate that a portfolio of 15 to 20 hedge funds can eliminate much of the idiosyncratic risk associated with hedge fund investments.

Process risk also raises the ancillary issue of lack of transparency. Skill-based investing usually is opaque. Are the decisions of the key individual quantitatively based? Qualitatively based? There is no way to really tell. This is similar to the problems discussed earlier with respect to black boxes.

To summarize, process risk cannot be quantified and it is not a risk that investors are willing to bear. Process risk also raises issues of transparency. Investors want clarity and definition, not opaqueness and amorphousness.

What Makes the Hedge Fund Manager so Smart?

Before investing money with a hedge fund manager, an investor must determine one of the following. The hedge fund manager must be able to demonstrate that he or she is smarter than the next manager. One way to be smarter than another hedge fund manager is to have superior skill in filtering information. That is, the hedge fund manager must be able to look at the same information set as another manager but be able to glean more investment insight from that data set.

Alternatively, if the hedge fund manager is not smarter than the next manager, he must demonstrate that he has a better information set; his competitive advantage is not filtering information but gathering it. To be successful, a hedge fund manager must demonstrate one or both of these competitive advantages.

Generally speaking, quantitative, computer-driven managers satisfy the first criterion. That is, hedge fund managers that run computer models access the same information set as everyone else, but have better (smarter) algorithms to extract more value per information unit than the next manager. These managers tend to be relative-value managers.

Relative-value managers extract value by simultaneously comparing the prices of two securities and buying and selling accordingly. This information is available to all investors in the marketplace. However, it is the relative-value managers that are able to process the information quickly enough to capture mispricings in the market. These arbitrage strategies expose an investor to credit risk.

Alternatively, hedge fund managers that confine themselves to a particular market segment or sector generally satisfy the second criterion. They have a larger information set that allows them to gain a competitive edge in their chosen market. Their advantage is a proprietary information set accumulated over time rather than a proprietary data filtering system.

Consider the following statement from a hedge fund disclosure document:

> The Adviser hopes to achieve consistently high returns by focusing on small and midcap companies in the biotechnology market.

The competitive advantage of this type of manager is his or her knowledge not only about a particular economic sector (biotechnology), but also, about a particular market segment of that sector (small and midcap). This type of manger tends to take more market risk exposure than credit risk exposure and generally applies equity long/short programs.

Identifying the competitive advantage of the hedge fund manager is the key to determining whether the hedge fund manager can sustain performance results. We indicated earlier that the issue of performance persistence is undecided.

Therefore, an investor cannot rely on historical hedge fund performance data as a means of selecting good managers from bad managers. Furthermore, every hedge fund disclosure document contains some variation of the following language:

> Past performance is no indication of future results.

Essentially, this statement directs the investor to ignore the hedge fund manager's performance history.

To assess the likelihood of performance persistence, the investor must then determine whether the hedge fund manager is an information gatherer or an information filterer. Consider the following language from a hedge fund disclosure document:

> The General Partner will utilize its industry expertise, contacts, and databases developed over the past 11 years to identify ___ company investment ideas outside traditional sources and will analyze these investment opportunities using, among other techniques, many aspects of its proven methodology in determining value.

This hedge fund manager has a superior information set that has been developed over 11 years. He is an information gatherer. This manager applies an equity long/short program within a specific market sector.

Finally, consider the following disclosure language from a merger arbitrage hedge fund manager:

> [The] research group [is] staffed by experienced M&A [merger and acquisition] lawyers with detailed knowledge of deal lifecycle, with extensive experience with corporate law of multiple U.S. states, U.S. and foreign securities laws regarding proxy contests, and antitrust laws (both of the United States and EU), and who have made relevant filings before regulators and have closed a wide variety of M&A transactions.

This hedge fund manager is an information filterer. His expertise is sifting through the outstanding legal and regulatory issues associated with a merger and determining the likelihood that the deal will be completed.

To summarize, a good lesson is that successful hedge fund managers know the exact nature of their competitive advantage, and how to exploit it.

SUMMARY

This chapter was intended to provide an overview of the hedge fund market; it was not intended to draw any conclusions about the value of hedge funds as an investment vehicle. There are two key points that the reader should take from this chapter.

First, the hedge fund strategies discussed in this chapter invest in the same securities as traditional long-only managers. Hedge fund managers use the same securities as long-only managers. However, the distinguishing feature of hedge fund managers is the strategies in which they employ those securities. Therefore, hedge fund managers do not employ alternative assets but, rather, alternative strategies.

Second, there are many different hedge fund strategies. Which is best for the investor? That will really depend on the strategic approach that the investor wishes to take. Some investors may be more focused on equity-based strategies. For them, equity long/short funds or market timing might be appropriate. For an investor with a fixed income bias, convertible arbitrage, fixed income arbitrage, or relative-value arbitrage may be appropriate. Suffice it to say that there is sufficient variety in the hedge fund marketplace to suit most investors.

The number of hedge fund strategies in the financial markets sometimes seem like the many colors of a rainbow. Indeed, hedge fund managers excel at finding new ways to extract value from the financial markets, whether it be a form of arbitrage, large bets on the movements of the global asset classes, or taking advantage of corporate restructurings. Keep in the mind that within the four broad categories of hedge fund strategies can be many different substrategies that try to exploit a security pricing inefficiency, an informational advantage, or a misalignment of values across global asset classes. Hedge fund strategies are often complex and can be quite risky. However, the industry has evolved from the days of Long-Term Capital Management. Strong risk control and prudent risk taking are more the norm for hedge fund managers today. Their reputation for "swinging for the fences" is long gone. The goal for most hedge fund managers today is strong risk control with persistent and predictable returns.

REFERENCES

Anson, M. J. P. (2006). *Handbook of Alternative Assets*, 2nd edition. Hoboken, N.J.: John Wiley & Sons.

Jacobs, B. and Levy, K. (1995). The law of one alpha. *Journal of Portfolio Management* 21, Summer: 78–79.

Fung, W. and Hsieh, D. A. (1997). Empirical characteristics of dynamic trading strategies: The case of hedge funds. *Review of Financial Studies* 10, Summer: 275–302.

L'habitant, F-S. (2004). *Hedge Funds: Quantitative Insights*. West Sussex, England: John Wiley & Sons.

Liang, B. (2001). Hedge fund performance: 1990–1999. *Financial Analysts Journal* 57, January/February: 11–18.

Nicholas, J. G. (2000). *Market Neutral Investing*. Princeton, N.J.: Bloomberg Press.

Park, J., Brown, S. J., and Goetzmann, W. (2001). Careers and survival: Competition and risk in the hedge fund and CTA industry. *Journal of Finance* 56, 5: 1869–1886.

Park, J. and Staum, J. (1998). Fund of funds diversification: How much is enough? *Journal of Alternative Investments* 1, Winter: 39–42.

Weisman, A., and Abernathy, J. (2000). The dangers of historical hedge fund data. In L. Rahl (ed), *Risk Budgeting* (pp. 65–81), London, Risk Books.

CHAPTER 54

Introduction to Venture Capital

MARK J.P. ANSON, PhD, JD, CPA, CFA, CAIA
President and Executive Director of Nuveen Investment Services

Abstract: The private equity sector purchases the private stock or equity-linked securities of nonpublic companies that are expected to go public or provides the capital for public companies (or their divisions) that may wish to go private. The key component in either case is the private nature of the securities purchased. Private equity, by definition, is not publicly traded. Therefore, investments in private equity are illiquid. Investors in this marketplace must be prepared to invest for the long haul—investment horizons may be as long as 5 to 10 years. *Private equity* is a generic term that encompasses four distinct strategies in the market for private investing. First, there is venture capital, the financing of start-up companies. Second, there are leveraged buyouts (LBOs) where public companies repurchase all of their outstanding shares and turn themselves into private companies. Third, there is mezzanine financing, a hybrid of private debt and equity financing. Finally, there is distressed debt investing. These are private equity investments in established (as opposed to start-up) but troubled companies.

Keywords: venture capitalist, limited partnership, limited liability company, investment restrictions, management fees, profit sharing, business plan, intellectual property, management team, exit plan, J curve, angel investor, seed capital, early stage, late stage, mezzanine stage, initial public offering

Private equity is as old as Columbus's journey to America. Queen Isabella of Spain sold her jewelry to finance Columbus's small fleet of ships in return for whatever spoils Columbus could find in the New World. The risks were great, but the potential rewards were even greater. This in a nutshell summarizes the private equity market: a large risk of failure but the potential for outstanding gains.

More generally, private equity provides the long-term equity base of a company that is not listed on any exchange and therefore cannot raise capital via the public stock market. Private equity provides the working capital that is used to help private companies grow and succeed. It is a long-term investment process that requires patient due diligence and hands-on monitoring.

In this chapter, we focus on the best known of the private equity categories: venture capital. Venture capital is the supply of equity financing to start-up companies that do not have a sufficient track record to attract investment capital from traditional sources (e.g., the public markets or lending institutions). Entrepreneurs that develop business plans require investment capital to implement those plans. However, these start-up ventures often lack tangible assets that can be used as collateral for a loan. In addition, start-up companies are unlikely to produce positive earnings for several years. Negative cash flows are another reason why banks and other lending institutions as well as the public stock market are unwilling to provide capital to support the business plan.

It is in this uncertain space where nascent companies are born that venture capitalists operate. Venture capitalists finance these high-risk, illiquid, and unproven ideas by purchasing senior equity stakes while the firms are still privately held. The ultimate goal is to make a buck. Venture capitalists are willing to underwrite new ventures with untested products and bear the risk of no liquidity only if they can expect a reasonable return for their efforts. Often, venture capitalists set expected target rates of return of 33% or more to support the risks they bear. Successful start-up companies funded by venture capital money include Cisco Systems, Cray Research, Microsoft, and Genentech.

We begin with the role of a venture capitalist in a start-up company raising a venture capital fund. Next, we review the heart of the venture capital industry—the business plan. We then review the current structure of the industry. This is followed by a review of the different stages of venture capital financing.

ROLE OF A VENTURE CAPITALIST

Venture capitalists have two roles within the industry. Raising money from investors is just the first part. The second is to invest that capital with start-up companies.

Venture capitalists are not passive investors. Once they invest in a company, they take an active role either in an advisory capacity or as a director on the board of the company. They monitor the progress of the company, implement incentive plans for the entrepreneurs and management, and establish financial goals for the company.

Besides providing management insight, venture capitalists usually have the right to hire and fire key managers, including the original entrepreneur. They also provide access to consultants, accountants, lawyers, investment bankers, and most importantly, other business that might purchase the start-up company's product.

In this section we focus on the relationship between the venture capitalist and his investors. In the next section we consider the process by which a venture capitalist selects investments.

Relationship of the Venture Capitalist to Her Investors

Before a venture capitalist can invest money with start-up ventures, she must go through a period of fund raising with outside investors. Most venture capital funds are structured as limited partnerships, where the venture capitalist is the general partner and the investors are limited partners. Each venture capital fund first goes through a period of fund raising before it begins to invest the capital raised from the limited partners.

The venture capitalist, or her company, is the general partner of the venture capital fund. All other investors are limited partners. As the general partner, the venture capitalist has full operating authority to manage the fund as she pleases, subject to restrictions placed in the covenants of the fund's documents.

As the venture capital industry grew and matured through the 1980s and 1990s, sophisticated investors such as pension funds, endowments, foundations, and high-net-worth individuals began to demand that contractual provisions be placed in the documents and subscription agreements that establish and govern a private equity fund. These covenants ensure that the venture capitalist sticks to her knitting and operates in the best interest of the limited partners who have invested in the venture capital fund.

These protective covenants can be broken down into three broad classes of investor protections: (1) covenants relating to the overall management of the fund; (2) covenants that relate to the activities of the general partners; and (3) covenants that determine what constitutes a permissible investment (see Lerner, 2000).

Restrictions on the Management of the Venture Capital Fund

Typically, the most important covenant is the size of an investment by the venture capital fund in any one start-up venture. This is typically expressed as a percentage of the capital committed to the venture capital fund. The purpose is to ensure that the venture capitalist does not bet the fund on any single investment. In any venture capital fund, there will be start-up ventures that fail to generate a return. This is expected. By diversifying across several venture investments, this risk is mitigated.

Other covenants may include a restriction on the use of debt or leverage by the venture capitalist. Venture capital investments are risky enough without the venture capitalist's gearing up the fund through borrowing.

In addition, there may be a restriction on coinvestments with prior or future funds controlled by the venture capitalist. If a venture capitalist has made a poor investment in a prior fund, the investors in the current fund do not want the venture capitalist to throw more good money after bad. Last, there is usually a covenant regarding the distribution of profits. It is optimal for investors to receive the profits as they accrue. Furthermore, distributed profits reduce the amount of committed capital in the venture fund, which in turn reduces the fees paid to the venture capitalist. It is in the venture capitalist's economic interest to hold onto profits, while investors prefer to have them distributed as they accrue.

Restrictions on the Activities of the General Partner

Primary among these is a limit on the amount of private investments the venture capitalist can make in any of the firms funded by the venture capital fund. If the venture capitalist makes private investments on her own in a select group of companies, these companies may receive more attention than the remaining portfolio of companies contained in the venture fund.

In addition, general partners are often limited in their ability to sell their general partnership interest in the venture fund to a third party. Such a sale would likely reduce the general partner's incentive to monitor and produce an effective exit strategy for the venture fund's portfolio companies.

Two other covenants are related to keeping the venture capitalist's eye on the ball. The first is a restriction on the amount of future fund raising. Fund raising is time consuming and distracting—less time is spent managing the investments of the fund. Also, the limited partners typically demand that the general partner spend substantially all of his time on managing the investments of the fund—outside interests are limited or restricted.

Restrictions on the Type of Investments

Generally, these covenants serve to keep the venture capitalist focused on investing in those companies, industries, and transactions where she has the greatest experience. So, for instance, there may be restrictions or prohibitions on investing in leveraged buyouts, other venture capital funds, foreign securities, or companies and industries outside the realm of the venture capitalist's expertise.

Venture Capital Fees

Venture capitalists earn fees two ways: management fees and a percentage of the profits earned by the venture fund. Management fees can range anywhere from 1% to 3.5%, with most venture capital funds in the 2% to 2.5% range. Management fees are used to compensate the venture capitalist while she looks for attractive investment opportunities for the venture fund.

A key point is that the management fee is assessed on the amount of committed capital, not invested capital. Consider the following example: The venture capitalist raises $100 million in committed capital for her venture fund. The management fee is 2.5%. To date, only $50 million dollars of the raised capital has been invested. The annual management fee that the venture capitalist collects is $2.5 million—2.5% × $100 million—even though not all of the capital has been invested. Investors pay the management fee on the amount of capital they have agreed to commit to the venture fund whether or not that capital has actually been invested.

Consider the implications of this fee arrangement. The venture capitalist collects a management fee from the moment that an investor signs a subscription agreement to invest capital in the venture fund—even though no capital has actually been contributed by the limited partners yet. Furthermore, the venture capitalist then has a call option to demand—according to the subscription agreement—that the investors contribute capital when the venture capitalist finds an appropriate investment for the fund. This is a great deal for the venture capitalist—she is paid a large fee to have a call option on the limited partners' capital. Not a bad business model. We will see later that this has some keen implications for leveraged buyout funds.

The second part of the remuneration for a venture capitalist is the profit-sharing or incentive fees. This is really where the venture capitalist makes her money. Incentive fees provide the venture capitalist with a share of the profits generated by the venture fund. The typical incentive fee is 20%, but the better-known venture capital funds can charge up to 35%. That is, the best venture capitalists can claim one-third of the profits generated by the venture fund.

The incentive fees for venture capital funds are a free option. If the venture capitalist generates profits for the venture fund, she can collect a share of these profits. If the venture fund loses money, the venture capitalist does not collect an incentive fee. This option has significant value to the venture capitalist. Furthermore, valued within an option context, venture capital profit-sharing fees provide some interesting incentives to the venture capitalist.

For example, one way to increase the value of a call option is to increase the volatility of the underlying asset. This means that the venture capitalist is encouraged to make riskier investments with the pool of capital in the venture fund to maximize the value of his incentive fee. This increased risk may run counter to the desires of the limited partners to maintain a less risky profile. It is also fascinating to realize that this incentive fee is costless to the venture capitalist—she does not pay any price for the receipt of this option. Indeed, the venture capitalist gets paid a management fee in addition to this free call option on the profits of the venture fund. As we noted previously, this is not a bad business model for the venture capitalist.

Fortunately, there is a check and balance on incentive fees in the venture capital world. Most, if not all, venture capital limited partnership agreements include some restrictive covenants on when incentive fees may be paid to the venture capitalist. There are three primary covenants that are used.

First, most venture capital partnership agreements include a clawback provision. A clawback covenant allows the limited partners to clawback previously paid incentive fees to the venture capitalist if, at the end/liquidation of the venture fund, the limited partners are still out of pocket some costs or lost capital investment. This prevents the venture capitalist from making money if the limited partners do not earn a profit.

Second, there is often an escrow agreement where a portion of the venture capitalist incentive fees are held in a segregated escrow account until the fund is liquidated.

Again this ensures that the venture capitalist does not walk away with any profit unless the limited partners also earn a profit. If a profit is earned by every limited partner, the escrow proceeds are released to the venture capitalist.

Finally, there is often a prohibition on the distribution of profit-sharing fees to the venture capitalist until all committed capital is paid back to the limited partners. In other words, the limited partners must first be paid back their invested capital before profits may be shared in the venture fund. Sometimes this covenant also includes that all management fees must be recouped by the limited partners before the venture capitalist can collect his incentive fees.

Just as a side observation, it is interesting to note that these types of profit-sharing covenants are not used in hedge fund limited partnership agreements.

THE BUSINESS PLAN

The venture capitalist has two constituencies: investors on the one hand, and start-up portfolio companies on the other. In the prior section we discussed the relationship between the venture capitalist and her investors. In this section we discuss how a venture capitalist selects her investments for the venture fund.

The most important document upon which a venture capitalist will base her decision to invest in a start-up company is the business plan. The business plan must be comprehensive, coherent, and internally consistent. It must clearly state the business strategy, identify the niche that the new company will fill, and describe the resources needed to fill that niche.

The business plan also reflects the start-up management team's ability to develop and present an intelligent and strategic plan of action. The business plan not only describes the business opportunity but also gives the venture capitalist an insight to the viability of the management team.

Last, the business plan must be realistic. One part of every business plan is the assumptions about revenue growth, cash-burn rate, additional rounds of capital injection, and expected date of profitability and/or initial public offering (IPO) status. The financial goals stated in the business plan must be achievable. Additionally, financial milestones identified in the business plan can become important conditions for the vesting of management equity, the release of deferred investment commitments, and the control of the board of directors.

In this section we review the key elements of a business plan for a start-up venture. This is the heart and soul of the venture capital industry—it is where new ideas are born and capital is committed.

Executive Summary

The executive summary is the opening statement of any business plan. In this short synopsis, it must be clear what is the unique selling point of the start-up venture. Is it a new product, distribution channel, manufacturing process, chip design, or consumer service? Whatever it is, it must be spelled out clearly for a nontechnical person to understand (see the British Venture Capital Association, 2004).

The executive summary should quickly summarize the eight main parts of the business plan:

1. The market
2. The product/service
3. Intellectual property rights
4. The management team
5. Operations and prior operating history
6. Financial projections
7. Amount of financing
8. Exit opportunities

We next discuss briefly each part of the business plan.

The Market

The key issue here is whether there is a viable commercial opportunity for the start-up venture. The first question is whether there is an existing market already. If the answer is yes, this is both good and bad. It is good because the commercial opportunity has already been demonstrated by someone else. It is bad because someone else has already developed a product or service to meet the existing demand.

This raises the issue of competition. Virtually every new product already has some competition at the outset. It is most unlikely that the product or service is so revolutionary that there is no form of competition. Even if the start-up venture is first to market, there must be an explanation on how this gap in the market is currently being filled with existing (but deficient) solutions.

An existing product makes a prima facie case for market demand, but then the start-up venture must describe how its product/service improves upon the existing market solution. Furthermore, if there is an existing product, the start-up venture should make a direct product comparison including price, quality, length of warranty, ease of use, product distribution, and target audience.

In addition to a review of the competition, the start-up venture must describe its market plan. The marketing plan must include three elements: pricing, product distribution, and promotion.

Pricing is clear enough. If the product is first to market, it can command a price premium. Furthermore, in today's electronic markets, prices erode rapidly. The start-up venture must describe its initial margins, but also how those margins will be affected as technology advances are made.

Product distribution is simply a way to describe how the start-up venture will get its product to the market. Will it use wholesalers, retailers, the Internet, or direct sales? Is a sales force needed? Is a 24-hour help desk required? Also, different distribution channels may require different pricing. For example, wholesalers will need price discounts to be able to make a profit when they sell to retailers. Conversely, the start-up company may wish to offer a discount to those that order the product directly from the start-up venture.

Finally, the start-up venture must describe its promotion strategy. A discussion of trade shows, the Internet, mass media, and tie-ins to other products should be described. The start-up venture should indicate whether its product should be marketed to a targeted audience or whether it has mass appeal. The cost of promotional materials and events must also be evaluated as part of the business plan.

The Product or Service

A description of the product or service should be done along every dimension that establishes the start-up venture's unique selling point. Furthermore, this discussion must be done in plain English without the psychobabble or jargon that normally creeps into the explanation of technology products.

In fact, the key part of this section of the business plan is to cement the unique selling point of the product or service. Is it new to the market, available at a lower price, constructed with better quality, constructed in a shorter time frame, provided with better customer service, smaller in size, easier to operate, and so on? Each of these points can provide a competitive advantage on which to build a new product or service.

One-shot, single products are a concern for a venture capitalist. The upside will be inevitably limited as competition is drawn into the market. Therefore, business plans that address a second generation of products are generally preferred.

Intellectual Property Rights

The third essential part of the business plan is a discussion of intellectual property rights. Most of the industries where venture capital has flowed in recent years are technology related, such as computer software, telecom, biotech, and semiconductors.

Most start-ups in the technology and other growth sectors base their business opportunity on the claim to proprietary technology. It is very important that a start-up's claim and rights to that intellectual property be absolute. Any intellectual property owned by the company must be clearly and unequivocally assigned to the company by third parties (usually the entrepreneur and management team). A structure where the entrepreneur still owns the intellectual property but licenses it to the start-up company are disfavored by venture capitalists because license agreements can expire or be terminated, leaving the venture capitalist with a shell of a start-up company.

Generally, before a venture capitalist invests with a start-up company, it will conduct patent and trademark searches, seek the opinion of a patent counsel, and possibly ask third parties to confidentially evaluate the technology owned by the start-up company.

Additionally, the venture capitalist may ask key employees to sign noncompetition agreements, where they agree not to start another company or join another company operating in the same sector as the start-up for a reasonable period of time. Key employees may also be asked to sign nondisclosure agreements because protecting a start-up company's proprietary technology is an essential element to success.

The Start-up Management Team

Venture capitalists invest in ideas and people. Once the venture capitalist has reviewed the start-up venture's unique selling point, she will turn to the management team. Ideally, the management team should have complementary skill sets: marketing, technology, finance, and operations. Every management team has gaps. The business plan must carefully address how these gaps will be filled.

The venture capitalist will closely review the resumes of every member of the management team. Academic backgrounds, professional work history, and references will all be checked. Most important to the venture capitalist will be the professional background of the management team. In particular, a management team that has successfully brought a previous start-up company to the IPO stage will be viewed most favorably.

In general, a great management team with a good business plan is viewed more favorably than a good management team with a great business plan. The best business plan in the world can still fail from inability to execute. Thus, a management team that has demonstrated a previous ability to follow and execute a business plan gets a greater chance of success than an unproven management team with a great business opportunity.

However, this is where a venture capitalist can add value. Recognizing a great business opportunity but a weak management team, the venture capitalist can bring his or her expertise to the start-up company as well as bring in other, more seasoned management professionals. While this often creates some friction with the original entrepreneur, the ultimate goal is to make money. Egos often succumb when there is money to be made.

In addition to filling in the gaps of the management team, the venture capitalist will need to round out the board of directors of the start-up venture. One seat on the board will be filled by a member of the venture capitalist's own team. However, other directors may be added to fill in some of the gaps found among the management team. These gaps might include distribution expertise. In addition, the venture capitalist may ask an executive from an established company to sit on the board of the start-up to provide contacts within the industry when the start-up is ready to look for a strategic buyer. In addition, a seasoned board member from a successful company can lend credibility to a start-up venture when it decides to go public (see case study on CacheFlow/Blue Coat in Anson [2006]).

Last, the management team will need a seasoned chief financial officer (CFO). This will be the person primarily responsible for bringing the start-up company public. The CFO will work with the investment bankers to establish the price of the company's stock at the IPO. Since the IPO is often the exit strategy for the venture capitalist as well as some of the founders and key employees, it is critical that the CFO have IPO experience.

Operations and Prior Operating History

The operations section of the business plan discusses how the product will be built or the service delivered. This will include a discussion of production facilities, labor requirements, raw materials, tax incentives, regulatory approvals, and shipping.

In addition, if a prototype has not yet been developed, then the business plan must lay out a timeline for its production as well as its cost. Cost of production must be discussed because this will feed into the gross margin discussion as part of the financial projections (discussed next).

Last, barriers to entry should be described. While there might be a higher cost of production at the outset, it will also prevent competition from entering the market later.

Venture capitalists are not always the first investors in a start-up company. In fact, they may be the third source of financing for a company. Many start-up companies begin by seeking capital from friends, family members, and business associates. Next, they may seek a so-called "angel investor": a wealthy private individual or an institution that invests capital with the company but does not take an active role in managing or directing the strategy of the company. Then come the venture capitalists.

As a result, a start-up company may already have a prior history before presenting its business plan to a venture capitalist. At this stage, venture capitalists ensure that the start-up company does not have any unusual history such as a prior bankruptcy or failure.

The venture capitalist will also closely review the equity stakes that have been previously provided to family, friends, business associates, and angel investors. These equity stakes should be clearly identified in the business plan and any unusual provisions must be discussed. Equity interests can include common stock, preferred stock, convertible securities, rights, warrants, and stock options. There must still be sufficient equity and upside potential for the venture capitalist to invest. Finally, all prior security issues must be properly documented and must comply with applicable securities laws.

The venture capitalist will also check the company's articles of incorporation to determine whether it is in good legal standing in the state of incorporation. Furthermore, the venture capitalist will examine the company's bylaws, and the minutes of any shareholder and board of directors meetings. The minutes of the meetings can indicate whether the company has a clear sense of direction or whether it is mired in indecision.

Financial Projections

In light of the discussion on operations and cost of projections, this information leads right into the financial projections. A comprehensive set of financial statements are required including income statement, balance sheet, and cash flow projections. These projections must be realistic but at the same time, entice the venture capitalist that there is a sufficient return to be earned to warrant the investment of capital.

First, the income statement must show in which year a breakeven point will be achieved. Most business plans show a profit being turned by the third year after initial financing. The income statement should include realistic sales forecasts, allowances for discounts, clear numbers for the cost of goods sold, and reasonable estimates of marketing and other overhead costs. Gross margins and net margins must meet the return requirements of the venture capitalist.

The balance sheet is important to determine at what point debt and other forms of financing should be added to the capital structure of the start-up venture. Also, the balance sheet should reflect the receivables received from the sale of the product as well as reasonable assumptions about the timing and collection of those receivables.

Finally, the cash flow statement provides the venture capitalist with a realistic burn rate on the cash on hand. Initially, all firms require infusions of capital to fund their working capital. However, at some point in time, the start-up venture must become self-financing such that its operating and expansion needs can draw from the money raised from the sale of its products.

For all of these financial projections, different scenarios must be included. What happens if a new competitor comes to the market quickly or the economy experiences a period of recessionary growth? Generally, the forecasts should include a base case of sales growth, a pessimistic case, and an optimistic case.

Amount of Financing

This section of the business plan gets down to brass tacks: how much money is the start-up venture requesting? This ties in neatly from the financial projections. As part of the assessment of cash flows, the start-up company needs to estimate its burn rate. The burn rate is simply the rate at which the start-up venture uses cash on a monthly basis. The amount of financing requested must be equal to the burn rate over the time horizon expected by the start-up venture.

Exit Plan

Eventually, the venture capitalist must liquidate her investment in the start-up company to realize a gain for herself and her investors. When a venture capitalist reviews a business plan, she will keep in mind the timing and probability of an exit strategy.

An exit strategy is another way the venture capitalist can add value beyond providing start-up financing. Venture capitalists often have many contacts with established operating companies. An established company may be willing to acquire the start-up company for its technology as part of a strategic expansion of its product line. Alternatively, venture capitalists maintain close ties with investment bankers. These bankers will be necessary if the start-up company decides to seek an IPO. In addition, a venture capitalist may ask other venture capitalists to invest in the start-up company. This helps to spread the risk as well as provide additional sources of contacts with operating companies and investment bankers.

Venture capitalists almost always invest in the convertible preferred stock of the start-up company. There may be several rounds (or series) of financing of preferred stock before a start-up company goes public. Convertible preferred shares are the accepted manner of investment because these shares carry a priority over common stock in terms of dividends, voting rights, and liquidation preferences. Furthermore, venture capitalists have the option to convert their shares to common stock to enjoy the benefits of an IPO.

Other investment structures used by venture capitalists include convertible notes or debentures that provide for the conversion of the principal amount of the note or bond into either common or preferred shares at the option of the venture capitalist. Convertible notes and debentures may also be converted upon the occurrence of an event such as a merger, acquisition, or IPO. Venture capitalists may also be granted warrants to purchase the common equity of the start-up company as well as stock rights in the event of an IPO.

Other exit strategies used by venture capitalists are redemption rights and put options. Usually, these strategies are used as part of a company reorganization. Redemption rights and put options are generally not favored because they do not provide as large a rate of return as an acquisition or IPO. These strategies are often used as a last resort when there are no other viable alternatives. Redemption rights and put options are usually negotiated at the time the venture capitalist makes an investment in the start-up company (often called the registration rights agreement).

Usually, venture capitalists require no less than the minimum return provided for in the liquidation preference of a preferred stock investment. Alternatively, the redemption rights or put option might be established by a common stock equivalent value that is usually determined by an investment banking appraisal. Last redemption rights or put option values may be based on a multiple of sales or earnings. Some redemption rights take the highest of all three valuation methods: the liquidation preference, the appraisal value, or the earnings/sales multiple.

In sum, there are many issues a venture capitalist must sort through before funding a start-up company. These issues range from identifying the business opportunity to sorting through legal and regulatory issues. Along the way, the venture capital must assess the quality of the management team, prior capital infusions, status of proprietary technology, operating history (if any) of the company, and timing and likelihood of an exit strategy.

CURRENT STRUCTURE OF THE VENTURE CAPITAL INDUSTRY

The structure of the venture capital industry has changed dramatically over the past 20 years. We focus on three major changes: sources of venture capital financing, venture capital investment vehicles, and specialization within the industry.

Sources and Uses of Venture Capital Financing

The structure of the venture capital marketplace has changed considerably since 1985. What is most notable is the leading sources of venture capital financing. For example, over the period 1985 to 1990, the leading source of venture capital financing was pension funds. This came as a result of the revisions to the prudent person standard for pension fund investing in 1979. Over the 1985 to 1990 period, pension funds accounted for almost 70% of venture capital funding. Endowments and intermediaries, on the other hand, were a smaller source of venture capital funds. Also, in 1985 to 1990, government agencies accounted for about 11% of the total source of venture capital funds (see Lipin, 2000).

By 2005, the landscape of venture capital financing had changed considerably. Pension funds account for only about 50% of the source of venture capital funds. Government agencies supplied almost no money to venture capital in 2005, squeezed out by private sources. The federal and state governments no longer need to support the venture capital industry. Virtually all money comes from institutional and other investors willing to take the risk of start-up companies in return for sizeable gains.

To replace the decline of pension funds and government agencies, three new sources of venture capital funds have grown over the last 15 years: endowments and foundations, intermediaries, and individuals. Endowments, with their perpetual investment horizons, are natural investors for private equity. Also, as the wealth of the United States has grown, wealthy individuals have allocated a greater share of their wealth to venture capital investments. Finally, intermediaries such as private equity fund of funds, hedge funds, crossover funds, and interval funds have entered the venture capital market.

Venture Capital Investment Vehicles

As the interest for venture capital investments has increased, venture capitalists have responded with new vehicles for venture financing. These include limited partnerships, limited liability companies, corporate venture funds, and venture capital fund of funds.

Limited Partnerships

The predominant form of venture capital investing in the United States is the limited partnership. Venture capitalists operate either as "3(c)(1)" or "3(c)(7)" funds to avoid registration as an investment company under the Investment Company Act of 1940. As a limited partnership, all income and capital gains flow through the partnership to the limited partner investors. The partnership itself is not taxed. The appeal of the limited partnership vehicle has increased since 1996 with the "check the box" provision of the U.S. tax code.

Previously, limited partnerships had to meet several tests to determine if their predominant operating characteristics resembled more a partnership than a corporation. Such characteristics included, for instance, a limited term of existence. Failure to qualify as a limited partnership would mean double taxation for the investment fund—first, at the fund level, and second, at the investor level.

This changed with the U.S. Internal Revenue Service's decision to let entities simply decide their own tax status by checking a box on their annual tax form as to whether they wished to be taxed as a corporation or as a partnership. "Checking the box" greatly encouraged investment funds to establish themselves as a limited partnership.

Limited partnerships are generally formed with an expected life of 10 years with an option to extend the limited partnership for another 1 to 5 years. The limited partnership is managed by a general partner who has day-to-day responsibility for managing the venture capital fund's investments as well as general liability for any lawsuits that may be brought against the fund. Limited partners, as their name implies, have only a limited (investor) role in the partnership. They do not partake in the management of the fund, and they do not bear any liability beyond their committed capital.

All partners in the fund will commit to a specific investment amount at the formation of the limited partnership. However, the limited partners do not contribute money to the fund until it is called down or "taken down" by the general partner. Usually, the general partner will give one to two months' notice of when it intends to make additional capital calls on the limited partners. Capital calls are made when the general partner has found a start-up company in which to invest. The general partner can make capital calls up to the amount of the limited partners' initial commitments.

An important element of limited partnership venture funds is that the general partner/venture capitalist has also committed investment capital to the fund. This assures the limited partners of an alignment of interests with the venture capitalist. Typically, limited partnership agreements specify a percentage or dollar amount of capital that the general partner must commit to the partnership.

Limited Liability Companies

Another financing vehicle in the venture capital industry is the limited liability company (LLC). Similar to a limited partnership, all items of net income or loss as well as capital gains are passed through to the shareholders in the LLC. Also, like a limited partnership, an LLC must adhere to the safe harbors of the Investment Company Act of 1940. In addition, LLCs usually have a life of 10 years with possible options to extend for another 1 to 5 years.

The managing director of an LLC acts like the general partner of a limited partnership. She has management responsibility for the LLC including the decision to invest in start-up companies the committed capital of the LLC's shareholders. The managing director of the LLC might itself be another LLC or a corporation. The same is true for limited partnerships: The general partner need not be an individual; it can be a legal entity like a corporation.

In sum, LLCs and limited partnerships accomplish the same goal—the pooling of investor capital into a central fund from which to make venture capital investments. The choice is dependent on the type of investor sought. If the venture capitalist wishes to raise funds from a large number of passive and relatively uninformed investors, the limited partnership vehicle is the preferred status. However, if the venture capitalist intends to raise capital from a small group of knowledgeable investors, the LLC is preferred.

The reason is twofold: First, LLCs usually have more specific shareholder rights and privileges. These privileges are best utilized with a small group of well-informed investors. Second, an LLC structure provides shareholders with control over the sale of additional shares in the LLC to new shareholders. This provides the shareholders with more power with respect to the twin issues of increasing the LLC's pool of committed capital and from whom that capital will be committed.

Corporate Venture Capital Funds

With the explosive growth of technology companies in the late 1990s, many of these companies found themselves with large cash balances. Microsoft, for example, had current assets (cash, cash equivalents, and receivables) of over \$48 billion, and generated a free cash flow of over \$15 billion in 2005. Microsoft and other companies need to invest this cash to earn an appropriate rate of return for their investors.

A corporate venture capital fund is an ideal use for a portion of a company's cash. First, venture capital financing is consistent with Microsoft's own past; it was funded with venture capital over 20 years ago. Second, Microsoft can provide its own technological expertise to help a start-up company. Finally, the start-up company can provide new technology and cost savings to Microsoft. In a way, financing start-up companies allows Microsoft to "think outside of the box" without committing or diverting its own personnel to the task.

Corporate venture capital funds are typically formed only with the parent company's capital; outside investors are not allowed to join. In addition to Microsoft, other corporate venture funds include Xerox Venture Capital, Hewlett-Packard Company. Corporate Investments, Intel Capital, and Amoco Venture Capital. Investments in start-up companies are a way for large public companies to supplement their research-and-development budgets. In addition to accessing to new technology, corporate venture capital funds also gain the ability to generate new products, identify new or diminishing industries, acquire a stake in a future potential competitor, derive attractive returns for excess cash balances, and learn the dynamics of a new marketplace.

Perhaps the best reason for corporate venture capital funds is to gain a window on new technology. Consider the case of Supercomputer Systems of Wisconsin. Steve Chen, the former CEO of Cray Research, left Cray to start his own supercomputer company. Cray Research is a supercomputer company that was itself a spin-off from Control Data Corporation, which in turn was an outgrowth of Sperry Corporation. When Chen founded his new company, IBM was one of his first investors, even though IBM had shifted its focus from large mainframe computers to laptop computers, personal computers, and service contracts (see Schilit, 1998).

Another example is Intel Capital, Intel Corporation's venture capital subsidiary. The goal of Intel Capital is to develop a strategic investment program that focuses on making equity investments and acquisitions to grow the Internet economy, including the infrastructure, content, and services in support of Intel's main business, which is providing computer chips to power personal and laptop computers. To further this goal, Intel Capital has provided venture capital financing to companies like Peregrine Semiconductor Corporation, a start-up technology company that designs, manufactures, and markets high-speed communications integrated circuits for the broadband fiber, wireless, and satellite communications markets.

Since its founding in 1991, Intel Capital has invested more than $4 billion in approximately 1,000 companies in more than 30 countries. Of this 1,000, 160 portfolio companies have been acquired and another 150 have gone public on exchanges around the world—a combined success rate of 31% for start-up ventures. Intel Capital's program is sufficiently mature now that Intel has five separate funds from which to seed start-up ventures.

There are, however, several potential pitfalls to a corporate venture capital program. These may include conflicting goals between the venture capital subsidiary and the corporate parent. In addition, the 5- to 10-year investment horizon for most venture capital investments may be a longer horizon than the parent company's short-term profit requirements. Furthermore, a funded start-up company may be unwilling to be acquired by the parent company. Still, the benefits from corporate venture capital programs appear to outweigh these potential problems.

Another pitfall of corporate venture capital funds is the risk of loss. Just as every venture capitalist experiences losses in her portfolio of companies, so too will the corporate venture capitalist. This can translate into significant losses for the parent company.

Take the case of Dell Computers. Dell took a charge of $200 million in the second quarter of 2001 as a result of losses from Dell Ventures, the company's venture capital fund. Additionally, in June 2001, Dell reported that its investment portfolio had declined in value by more than $1 billion (see Menn, 2001).

Eventually, Dell decided to exit the venture capital business altogether. It sold the remainder of its venture capital portfolio to Lake Street Capital, a San Francisco private equity firm, for $100 million in 2005.

Intel Corporation reported in 2001 that its technology portfolio had declined more than $7 billion in value. For example, in the second quarter of 2000, Intel reported a $2.1 billion gain from the sale of its venture capital investments. Gains from Intel's technology portfolio helped to keep its earnings growth intact. Conversely, in the second quarter of 2001, Intel reported only a $3 million gain from the sale of its investments from its venture capital subsidiary (see Antonelli, 2001 and Menn, 1998).

However, where Dell did not succeed, Intel has recovered and has rebuilt its venture capital portfolio. In 2005, Intel had over $1 billion of venture capital investments on its financial statements.

Perhaps the most extreme case of nonperforming corporate venture capital investments is that of Comdisco Inc. Comdisco sought bankruptcy protection in July 2001 after making $3 billion in loans to start-up companies that were unable to repay most of the money. The company wrote off $100 million in loans made by its Comdisco Ventures unit, which leases computer equipment to start-up companies. In addition, Comdisco also took a $206 million reserve against earnings from investments in those ventures (see St. Onge, 2001).

Venture Capital Fund of Funds

A venture capital fund of funds is a venture pool of capital that, instead of investing directly in start-up companies, invests in other venture capital funds. The venture capital fund of funds is a relatively new phenomenon in the venture capital industry. The general partner of a fund of funds does not select start-up companies in which to invest. Instead, she selects the best venture capitalists with the expectation that they will find appropriate start-up companies to fund.

A venture capital fund of funds offers several advantages to investors. First, the investor receives broad exposure to a diverse range of venture capitalists and, in turn, a wide range of start-up investing. Second, the investor receives the expertise of the fund of funds manager in selecting the best venture capitalists with whom to invest money. Finally, a fund of funds may have better access to popular, well-funded venture capitalists whose funds may be closed to individual investors. In return for these benefits, investors pay a management fee (and, in some cases, an incentive fee) to the fund of funds manager. The management fee can range from 0.5% to 2% of the net assets managed.

Fund of fund investing also offers benefits to the venture capitalists. First, the venture capitalist receives one large investment (from the venture fund of funds) instead of several small investments. This makes fund raising and investor administration more efficient. Second, the venture capitalist interfaces with an experienced fund of funds manager instead of several (potentially inexperienced) investors.

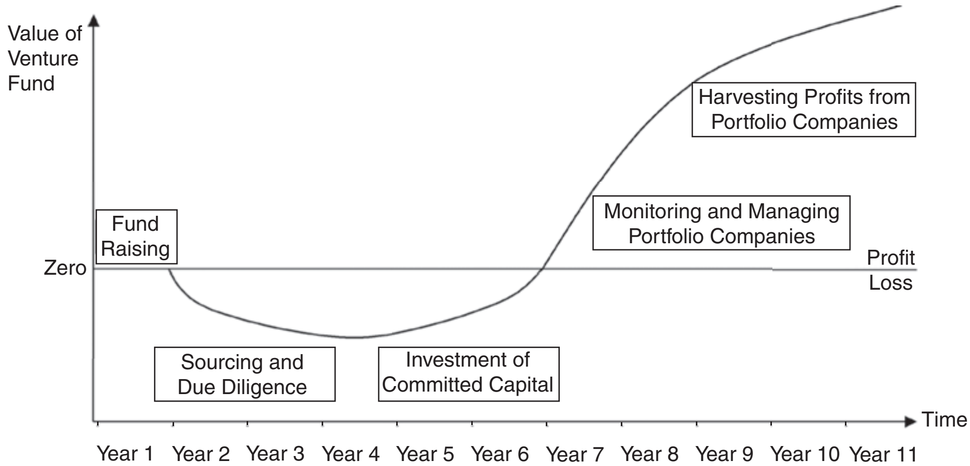


Figure 54.1 The Life Cycle of a Venture Capital Fund

Life Cycle of a Venture Capital Fund

A venture capital fund is a long-term investment. Typically, investors' capital is locked up for a minimum of 10 years—the standard term of a venture capital limited partnership. During this long investment period, a venture capital fund will normally go through five stages of development.

The first stage is the fund-raising stage where the venture capital firm raises capital from outside investors. Capital is committed—not collected. This is an important distinction noted above. Investors sign a legal agreement (typically a subscription) that legally binds them to make cash investments in the venture capital fund up to a certain amount. This is the committed, but not yet drawn, capital. The venture capital firm/general partner will also post a sizeable amount of committed capital.

Fund raising normally takes six months to a year. However, the more successful venture funds such as Kleiner, Perkins, Caufield, and Byers typically fund raise in just two to three months.

The second stage consists of sourcing investments, reading business plans, preparing intense due diligence on start-up companies, and determining the unique selling point of each start-up company. This period begins the moment the fund is closed to investors and normally takes up the first five years of the venture fund's existence.

During stage two, no profits are generated by the venture capital fund. In fact, quite the reverse: The venture capital fund generates losses because the venture capitalist continues to draw annual management fees (which can be up to 3.5% a year on the total committed capital). These fees generate a loss until the venture capitalist begins to extract value from the investments of the venture fund.

Stage three is the investment of capital. During this stage, the venture capitalist determines how much capital to commit to each start-up company, at what level of financing, and in what form of investment (convertible preferred shares, convertible debentures, etc.). At this stage the venture capitalist will also present capital calls to the investors in her venture fund to draw on the capital of the limited partners. Note that no cash flow is generated yet; the venture fund is still in a deficit.

Stage four begins after the funds have been invested and lasts almost to the end of the term of the venture capital fund. During this time the venture capitalist works with the portfolio companies in which the venture capital fund has invested. The venture capitalist may help to improve the management team, establish distribution channels for the new product, refine the prototype product to generate the greatest sales, and generally position the start-up company for an eventual public offering or sale to a strategic buyer. During this time period, the venture capitalist will begin to generate profits for the venture fund and its limited partner investors. These profits will initially offset the previously collected management fees until a positive net asset value is established for the venture fund.

The last stage of the venture capital fund is its windup and liquidation. At this point, all committed capital has been invested and now the venture capitalist is in the harvesting stage. Each portfolio company is either sold to a strategic buyer, brought to the public markets in an IPO, or liquidated through a Chapter 7 bankruptcy liquidation process. Profits are distributed to the limited partners and the general partner/venture capitalist now collects her incentive/profit-sharing fees.

These stages of a venture capital firm lead to what is known as the "J curve effect." Figure 54.1 demonstrates the J curve. We can see that during the early life of the venture capital fund, it generates negative revenues (losses), but eventually, profits are harvested from successful companies and these cash flows overcome the initial losses to generate a net profit for the fund. Clearly, given the initial losses that pile up during the first four to five years of a venture capital fund, this type of investing is only for patient, long-term investors.

Specialization within the Venture Capital Industry

Like any industry that grows and matures, expansion and maturity lead to specialization. The trend toward specialization in the venture capital industry exists on several levels: by industry, geography, stage of financing, and "special situations." Specialization is the natural by-product of two factors. First, the enormous amount of capital flowing into venture capital funds has encouraged venture capitalists to distinguish themselves from other funds by narrowing their investment focus. Second, the development of many new technologies over the past decade has encouraged venture capitalists to specialize in order to invest most profitably.

Specialization by Industry

Specialization by entrepreneurs is another reason why venture capitalists have tailored their investment domain. Just as entrepreneurs have become more focused in their start-up companies, venture capitalists have followed suit. The biotechnology industry is a good example.

The biotech industry was born on October 14, 1980, when the stock of Genentech, Inc. went public. On that day, the stock price went from $39 to $85 and a new industry was born. Today, Genentech is a Fortune 500 company with a market capitalization of $28 billion. Other successful biotech start-ups include Cetus Corporation, Biogen, Inc., Amgen Corporation, and Centacor, Inc.

The biotech paradigm has changed since the days of Genentech. Genentech was founded on the science of gene mapping and splicing to cure diseases. However, initially it did not have a specific product target. Instead, it was concerned with developing its gene-mapping technology without a specific product to market.

Compare this situation to that of Applied Microbiology, Inc. of New York. It has focused on two products with the financial support of Merck and Pfizer, two large pharmaceuticals (see Schilit, 1997). One of its products is an antibacterial agent to fight gum disease contained in a mouthwash to be marketed by Pfizer.

Specialized start-up biotech firms have led to specialized venture capital firms. For example, Domain Associates of Princeton, New Jersey, focuses on funding new technology in molecular engineering. However, specialization is not unique to the biotech industry. Other examples include Communication Ventures of Menlo Park, California. This venture firm provides financing primarily for start-up companies in the telecommunications industry. Another example is American Health Capital Ventures of Brentwood, Tennessee, that specializes in funding new health care companies.

Specialization by Geography

With the boom in technology companies in Silicon Valley, Los Angeles, and Seattle, it is not surprising to find that many California-based venture capital firms concentrate their investments on the west coast of the United States. Not only are there plenty of investment opportunities in this region, it is also easier for the venture capital firms to monitor their investments locally. The same is true for other technology centers in New York, Boston, and Texas.

As another example, consider Marquette Ventures based in Chicago. This venture capital company invests primarily with start-up companies in the Midwest. Although it has provided venture capital financing to companies outside of this region, its predominant investment pattern is with companies located in the midwestern states (see Schilit, 1997). Similarly, the Massey Birch venture capital firm of Nashville, Tennessee, has provided venture financing to a number of companies in its hometown of Nashville as well as other companies throughout the southeastern states.

Regional specialization has the advantage of easier monitoring of invested capital. Also, larger venture capital firms may overlook viable start-up opportunities located in more remote sections of the United States. Regional venture capitalists step in to fill this niche.

The downside of regional specialization is twofold. First, regional concentration may not provide sufficient diversification to a venture capital portfolio. Second, a start-up company in a less exposed geographic region may have greater difficulty in attracting additional rounds of venture capital financing. This may limit the start-up company's growth potential as well as exit opportunities for the regional venture capitalist.

Special Situation Venture Capital

In any industry, there are always failures. Not every start-up company makes it to the IPO stage. However, this opens another specialized niche in the venture capital industry: the turnaround venture deal. Turnaround deals are as risky as seed financing because the start-up company may be facing pressure from creditors. The turnaround venture capitalist exists because mainstream venture capitalists may not be sufficiently well versed in restructuring a turnaround situation.

Consider the following example. (A similar example is in Schilit [1997].) A start-up company is owned 50% by early and midstage venture capitalists and 50% by the founder. Product delays and poor management have resulted in $10 million in corporate assets and $15 million in liabilities. The company has a negative net worth and is technically bankrupt.

The turnaround venture capitalist offers the founder/entrepreneur of the company $1 million for his 50% ownership plus a job as an executive of the company. The turnaround venture capitalist then offers the start-up company's creditors 50 cents for every one dollar of claims. The total of $8.5 million might come from a $1 million contribution from the turnaround venture capitalist and $7.5 million in bank loans secured by the $10 million in assets. Therefore, for $1 million the turnaround venture capitalist receives 50% of the start-up company and restores it to a positive net worth.

The founder of the company is happy because he receives $1 million for a bankrupt company plus he remains as an executive. The other venture capitalists are also happy because now they will be dealing with another venture specialist, plus the company has been restored to financial health. With some additional hard work the company may proceed on to an IPO. The creditors, however, will not be as pleased, but may make the deal anyway because 50 cents on the dollar may be more than they could expect to receive through a formal liquidation procedure.

An example of such a turnaround specialist is Reprise Capital Corporation of Garden City, New Jersey. In 1997, this company raised $25 million for turnaround venture capital deals.

In summary, the growth of the venture capital industry has created the need for venture capital specialists. The range of new business opportunities is now so diverse that it is simply not possible for a single venture capital firm to stay on top of all opportunities in all industries. Therefore, by necessity, venture capitalists have narrowed their investment domain to concentrate on certain niches within the start-up universe. Specialization also leads to differentiation, which allows venture capitalists to distinguish themselves from other investment funds.

STAGE OF FINANCING

While some venture capital firms classify themselves by geography or industry, by far the most distinguishing characteristic of venture capital firms is the stage of financing. Some venture capitalists provide first-stage or "seed capital," while others wait to invest in companies that are further along in their development. Still other venture capital firms come in at the final round of financing before the IPO. A different level of due diligence is required at each level of financing because the start-up venture has achieved another milestone on its way to success. In all, there are five discrete stages of venture capital financing: angel investing, seed capital, first-stage capital, late-stage/expansion capital, and mezzanine financing. We discuss each of these separately below.

Angel Investing

Angel investors often come from "F & F": friends and family. (Sometimes, venture capitalists include a third "F" for fools.) At this stage of the new venture, typically there is a lone entrepreneur who has just an idea—possibly sketched out at the kitchen table or in the garage. There is no formal business plan, no management team, no product, no market analysis—just an idea.

In addition to family and friends, angel investors can also be wealthy individuals who "dabble" in start-up companies. This level of financing is typically done without a private placement memorandum or subscription agreement. It may be as informal as a "cocktail napkin" agreement. Yet without the angel investor, many ideas would wither on the vine before reaching more traditional venture capitalists.

At this stage of financing, the task of the entrepreneur is to begin the development of a prototype product or service. In addition, the entrepreneur begins the draft of his business plan, assesses the market potential, and may even begin to assemble some key management team members. No marketing or product testing is done at this stage.

The amount of financing at this stage is very small—$50,000 to $500,000. Any more than that would strain family, friends, and other angels. The funds are used primarily to flush out the concept to the point where an intelligent business plan can be constructed.

Seed Capital

Seed capital is the first stage where venture capital firms invest their capital. At this stage, a business plan is completed and presented to a venture capital firm. Some parts of the management team have been assembled at this point, a market analysis has been completed, and other points of the business plan as discussed previously in this chapter are addressed by the entrepreneur and his small team. Financing is provided to complete the product development and, possibly, to begin initial marketing of the prototype to potential customers. This phase of financing usually raises $1 to $5 million.

At this stage of financing, a prototype may have been developed and the testing of a product with customers may have begun. This is often referred to as "beta testing," and is the process where a prototype product is sent to potential customers free of charge to get their input into the viability, design, and user friendliness of the product.

Very little revenue has been generated at this stage, and the company is definitely not profitable. Venture capitalists invest in this stage based on their due diligence of the management team, their own market analysis of the demand for the product, the viability of getting the product to the market while there is still time and not another competitor, the additional management team members that will need to be added, and the likely timing for additional rounds of capital from the same venture capital firm or from other venture capital funds.

Examples of seed financing companies are Technology Venture Investors of Menlo Park, California; Advanced Technology Ventures of Boston; and Onsent, located in Silicon Valley (Schilit, 1997). Seed capital venture capitalists tend to be smaller firms because large venture capital firms cannot afford to spend the endless hours with an entrepreneur for a small investment that usually is no greater than $1 to $2 million.

Early-Stage Venture Capital

At this point the start-up company should have a viable product that has been beta tested. Alpha testing may have already begun. This is the testing of the second-generation prototype with potential end users. Typically, a price is charged for the product or a fee for the service. Revenues are being generated and the product/service has now demonstrated commercial viability. Early-stage venture capital financing is usually $2 million or more.

Early-stage financing is typically used to build out the commercial scale manufacturing services. The product is no longer being produced out of the entrepreneur's garage or out of some vacant space above a grocery store. The company is now a going concern with an initial, if not complete, management team. At this stage, there will be at least one venture capitalist sitting on the board of directors of the company.

The goal of the start-up venture is to achieve market penetration with its product. Some of this will have already been accomplished with the beta and alpha testing of the product. However, additional marketing must now be completed. In addition, distribution channels should be identified by now and the product should be established in these channels. Reaching a breakeven point is the financial goal.

Late-Stage/Expansion Venture Capital

At this point, the start-up company may have generated its first profitable quarter, or be just at the breakeven point. Commercial viability is now established. Cash flow management is critical at this stage, as the company is not yet at the level where its cash flows can self-sustain its own growth.

Late-stage/expansion capital fills this void. This level of venture capital financing is used to help the start-up company get through its cash crunch. The additional capital is used to tap into the distribution channels, establish call centers, expand the manufacturing facilities, and attract the additional management and operational talent necessary to the make the start-up company a longer-term success. Because this capital comes in to allow the company to expand, financing needs are typically greater than for seed and early stage. Amounts may be in the $5 million to $15 million range.

At this stage, the start-up venture enjoys the growing pains of all successful companies. It may need additional working capital because it has focused on product development and sales, but now finds itself with a huge backload of accounts receivable from customers on which it must now collect. Inevitably, start-up companies are very good at getting the product out of the door but very poor at collecting receivables and turning sales into cold, hard cash.

Again, this is where expansion capital can help. Late-stage venture financing helps the successful start-up get through its initial cash crunch. Eventually, the receivables will be collected and sufficient internal cash will be generated to make the start-up company a self-sustaining force. Until then, one more round of financing may be needed.

Mezzanine Stage

Mezzanine venture capital is the last stage before a start-up company goes public or is sold to a strategic buyer. At this point, a second-generation product may already be in production if not distribution. The management team is together and solid, and the company is working on managing its cash flow better. Manufacturing facilities are established, and the company may already be thinking about penetrating international markets. Amounts vary depending on how long the bridge financing is meant to last but generally is in the range of $5 to $15 million.

The financing at this stage is considered "bridge" or mezzanine financing to keep the company from running out of cash until the IPO or strategic sale. The start-up company may still have a large inventory of uncollected accounts receivable that need to be financed in the short term. Profits are being recorded, but accounts receivable are growing at the same rate of sales.

Mezzanine financing may be in the form of convertible debt. In addition, the company may have sufficient revenue and earning power that traditional bank debt may be added at this stage. This means that the start-up company may have to clean up its balance sheet as well as its statement of cash flows. Commercial viability is more than just generating sales, it also requires turning accounts receivable into actual dollars.

J Curve for a Start-up Company

Figure 54.2 presents the J curve for a start-up company. Similar to the J curve for a venture capital fund, the initial years of a start-up company generate a loss. Money is spent turning an idea into a prototype product and from there beta testing the product with potential customers. Little or no revenue is generated during this time. It is not until the product goes into alpha testing that revenues may be generated and the start-up becomes a viable concern.

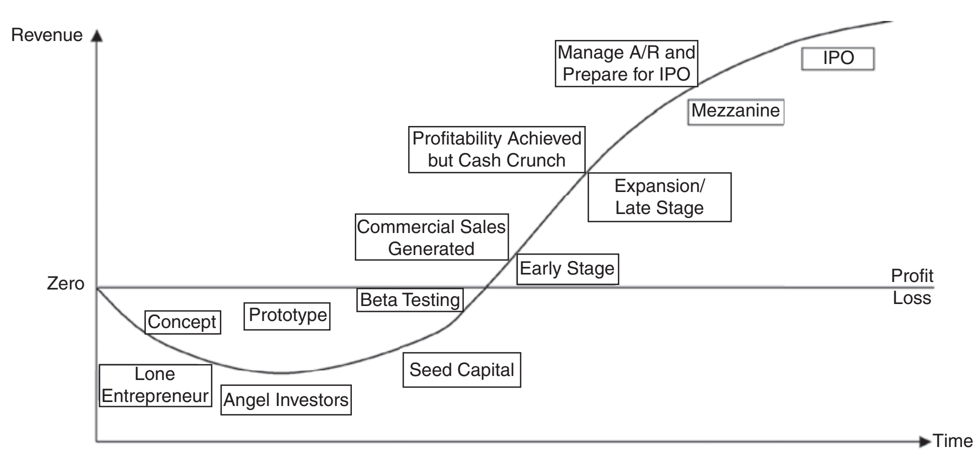


Figure 54.2 The Life Cycle of a Start-up Company

Once a critical mass is generated—where sales are turned into profits and accounts receivable is turned into cash—then it becomes a matter of timing until the start-up company achieves a public offering. Additional rounds of financing may be needed to get the company to its IPO nirvana. At this point, commercial viability is established, but managing the cash crunch becomes critical.

SUMMARY

Venture capital investing is a natural part of the equity cycle. Every company has to start someplace with some amount of capital to finance its initial operations. Long before a successful company reaches traditional investors in an IPO, venture capitalists are hard at work providing strategy, financing, and direction to start-up companies. Without the acorn of venture capital, most start-up companies would wither on the vine and their products would not come to fruition.

Also, venture capital plays a role in economic Darwinism. Because venture capitalists are rewarded only for those technologies and ideas that will have the greatest economic impact on society, they prune out the weak ideas and technology. Only the strong ideas survive—and these are the ideas, products, services, and technologies that will reward the venture capitalist the greatest while serving society to the largest extent possible.

REFERENCES

Anson, M. J. P. (2006). *Handbook of Alternative Assets*, 2nd edition. Hoboken, N.J.: John Wiley & Sons.

Antonelli, C. (2001). Chipmaker's profit plunges 94%; Intel still beats analysts' forecasts. *Bloomberg News*, July 18.

British Venture Capital Association. (2004). A guide to private equity. White paper, October.

Lerner, J. (2000). *Venture Capital and Private Equity*. New York: John Wiley & Sons.

Lipin, S. (2000). Venture capitalists "R" us. *Wall Street Journal*, February 22: C1.

Menn, J. (2001). Tech giants lose big on start-up ventures. *Los Angeles Times*, June 11.

Schilit, K. W. (1997). The nature of venture capital investments. *Journal of Private Equity*, Winter: 59–75.

Schilit, K. W. (1998). Structure of the venture capital industry. *Journal of Private Equity*, Spring: 60–67.

St. Onge, J. (2001). Comdisco seeks bankruptcy protection from creditors. *Bloomberg News*, July 16.

CHAPTER 55

Assessing Hedge Fund Investment Risk in Common Hedge Fund Strategies

ELLEN J. RACHLIN
Managing Director–Portfolio & Risk Manager, Mariner Investment Group, Inc.

Abstract: A key task for the professional hedge fund investor is to look for exposures which are vulnerable to losses during times of increased market volatility or large market dislocations. The professional hedge fund investor should assess continually the market environment for changes in volatility and liquidity so that portfolio risks can be anticipated. Securities present differently in a portfolio's risk metrics under different volatility and liquidity scenarios. This chapter examines many of the key portfolio risks which may be present in the major hedge fund strategies: government bond arbitrage, mortgage-backed security arbitrage, corporate bond arbitrage, emerging markets, capital structure arbitrage, distressed securities, merger arbitrage, long/short equity, multistrategy, market-neutral equity, and convertible bond arbitrage.

Keywords: government bond arbitrage, mortgage-backed security arbitrage, corporate bond arbitrage, emerging markets, capital structure arbitrage, distressed securities, merger arbitrage, long/short equity, multistrategy, market neutral equity, convertible bond arbitrage, leverage, volatility, liquidity, position concentration, tail risk

Professional hedge fund investor, including fund of funds managers, consultants, and others serving in an advisory role must not only understand the risks that portfolio managers intend to take while constructing their portfolios but the risks they might be unintentionally assuming. The professional hedge fund investor must ascertain if each portfolio manager is capable of responding effectively, both defensively and opportunistically, to an abrupt change in the market environment. The professional hedge fund investor must use analytical tools and market experience, based on a thorough understanding of the markets and modern risk software, to make these judgments.

Professional hedge fund investor must also understand the language of risk: its quantitative measurements of aggregated exposures. They must also be a student of the markets to understand what can go wrong as market volatility changes and consequently, asset valuation change. They must remain current not only on the markets but also on new market instruments and how these instruments can present market risks. They must understand the appropriate leverage for each strategy. The professional hedge fund investor must also be aware of the attitude of lenders of credit to the capital markets. Lenders can change their risk appetites fairly quickly. It is also important to understand and remain current on the portfolio manager's settlement and accounting procedures. New market instruments such as derivatives can present difficult challenges for accounting, settlement and risk software applications.

A key task for the professional hedge fund investor is to look for exposures which are vulnerable to losses during times of increased market volatility or large market dislocations. The professional hedge fund investor should assess continually the market environment for changes in volatility and liquidity so that portfolio risks can be anticipated. The portfolio manager may fail to take into consideration certain risks when assessing the relative value of a security. Securities present differently in a portfolio's risk metrics under different volatility and liquidity scenarios. Failure to anticipate these changes on the part of the portfolio manager can result in bad judgments regarding leverage and position sizing as well as their willingness or ability to take a loss on a position.

These assessments go well beyond looking at a hedge fund's performance record.

FIXED INCOME STRATEGIES

Fixed income arbitrage portfolio managers seek to profit from temporary mispricings in specific bonds. Fixed-income arbitrage portfolio managers will determine if the current market price of a given bond differs from their own valuation of the bond. Fixed income arbitrage has many disciplines. These disciplines focus on government-issued debt and privately issued debt. Strategies that focus on government-issued debt are called interest rate arbitrage strategies. Strategies which focus on privately issued debt are called credit arbitrage strategies.

Government Bond Arbitrage

Government bond arbitrage hedge funds exploit temporary pricing anomalies in cash and derivative securities within the global government bond markets. The strategies used may include yield curve arbitrage, basis, volatility trading, cross-currency, and asset swaps trades.

The most important assessment a professional hedge fund investor makes when considering a government bond arbitrage portfolio is whether the portfolio manager has macro bets on interest rates or relative value trades with little directional interest rate exposure. The portfolio which is comprised of macro bets on interest rates will be highly correlated to directional changes in rates of the underlying countries in the portfolio. Abrupt changes in profit and loss (P&L) will be associated in changes in interest rate policy or anticipated changes in interest rate policy. They can be easy or difficult to identify in advance. The macro portfolio may not be overtly long or short government bonds, but it may have certain positions whose success is dependent upon changes in interest rate levels. Macro trades include ones which have large curve exposures. Some examples include a 2-year versus 10-year trade; a 2-year, 5-year, 10-year butterfly trade; or an unevenly weighted Eurodollar future trade. By contrast, relative value trades are expected to be largely uncorrelated to interest rate movements. They also present with tamer P&L swings for the portfolio. The most common relative value trades are swap spread trades, basis trades, on the run versus off the run trades, and Eurodollar trades which do not take significant curve exposure.

Government bond arbitrage generally contains the highest leverage among all hedge fund strategies. The true relative value portfolio will contain spreads trades that seek to profit from movements of just a few basis points. The collateral posted is often considered on a par with cash. An example is U.S. government bonds arbitrage. Commensurate with such small anticipated moves, high leverage must be utilized. But for each substrategy of government bond arbitrage there is an appropriate amount of leverage. Basis trades (cash against futures) or an on-the-run versus off-the-run trade with less than three months of curve exposure, command larger amounts of leverage than does a curve trade. The appropriate leverage factor in basis trades or on-the-run versus off-the-run trades can easily exceed in excess of 30 times levered before 99% value at risk (VaR) approaches 5% of net asset value. VaR describes the potential for portfolio impairment. For example, a 99% value at risk means that on any given day, the portfolio has a 99% probability of not losing more than its stated VaR amount for stated standard deviation. Swap spreads should be less levered. Curve exposure greater than six months and positions involving swaptions (which have volatility exposure) should be significantly less levered.

If the portfolio contains interest rate options, the professional hedge fund investor must monitor the frequency and amount by which the portfolio manager is a net receiver of option premium. A key question when assessing the risk profile of a portfolio manager is the dollar vega for a 1% increase in index interest rate volatility. If it is a negative number, the manager is short volatility. The larger the negative dollar vega, the larger is the potential loss for an increase in volatility. Even if the portfolio is not a net receiver of premium, it is important to know if and where on the volatility surface the manager is receiving premium. Volatility spikes can strike different tenors differently, resulting in uncovered tail risk exposure. *Tail risk* increases if the portfolio is a net receiver of option premium and the amount of tail risk increases as the dollar amount of the premium received increases.

Options have asymmetric risks. The asymmetry changes for each new price level of the underlying asset. The asymmetry is adverse containing potential tail risk to the manager that sells options. The asymmetry is positive and contains a built in positive convex hedge to the manager that purchases options. This positive asymmetry costs money in the form of premium paid. *Volatility* cannot go below zero and can increase to spectacular levels under the right circumstances. In short, significant market movements in the interest rate markets can be expected to generate losses for the portfolio that receives option premium.

Leverage in option books is important to assess. One way to view this embedded leverage is to notionalize the option reference amount of the underlying government bonds. (This delta amount will change for any movement in the markets.) There are other useful measures such as the metric lambda which is a metric that measures how the option price varies as reference security's price varies. This result is the leverage embedded in an option position. This measurement is particularly useful in uncovering the embedded leverage in out of the money options where the embedded leverage can be quite high. That is to say, small changes in rates can create large percentage changes in option premium. The potential for portfolio impairment depends on the leverage in the option book.

Another type of potential tail risk exposure is in positions which consistently have the potential to generate losses in times of market dislocation. In times of global political or economic turmoil, AAA rated sovereign debt can be expected to out perform lower rated sovereign debt and bank credit instruments. The later exposure is most frequently represented by interest rate swaps in the government bond markets. Interest rate swaps involve parties exchanging London Interbank Offered Rate (LIBOR) rates for fixed rate sovereign debt rates at specified semiannual payment dates. Because default is always a possibility, the asymmetric bet favors the buyer of the AAA sovereign debt over the buyer of the less than AAA sovereign debt and over the receiver of LIBOR. Professional hedge fund investor should segregate the government bond arbitrage book between AAA credits and other credit. Then they should total up the exposures to ascertain whether or not the portfolio manager is net long the more highly rated debt in the portfolio.

Because government bond arbitrage is a highly levered strategy, it is susceptible to short squeeze situations on a widely shorted government bonds. Widely shorted bonds or ones with large open short interest can be so scarce, they are not deliverable. Even issues with in excess of 20 billion can be susceptible to squeezes. It may take regulatory intervention to alleviate these situations, but, sometimes, not before such shorts are costly to the portfolio manager. A measure of the portfolio manager's risk discipline is whether or not or after how long they pay up to cover or buy back the short position that is being squeezed.

Most government bond arbitrage portfolio managers will supply investors with VaR reports. However, it is useful to be aware of how the portfolio net asset value (NAV) alters during significant moves in interest rates such as 50 or 100 basis points. These large P&L moves help reveal adverse asymmetric risks in the manager's portfolio. This is particularly useful to those managers trading rate volatility.

Mortgage-Backed Security Arbitrage

Mortgage-backed security arbitrage hedge funds predominately invest in the residential mortgage and interest rate markets. The portfolio manager will typically employ hedging techniques to hedge out the adverse investment risks inherent in the prepayment feature embedded in mortgage securities. Mortgage debt includes mortgages pooled by financial institutions and government-sponsored enterprises. Portfolio managers can invest in a variety of styles including coupon arbitrage, basis, synthetic coupons and credit trading.

There are many sub-strategies in mortgage-backed security arbitrage. The most significant initial assessment to make is whether or not the portfolio has credit exposure. This exposure is broadly defined as securities which have a credit rating of below AA. Securities below AA have an increased probability of default over AA or AAA mortgage securities.

Mortgage securities are among the most complex financial securities. The embedded market exposures are difficult to detect unless one has access to the calculations of sophisticated risk systems which model the prospective cash flows. These securities have embedded optionality through their prepayment features and, hence potentially, widely divergent income streams. It is essential to look at mortgage portfolios through the lens of good risk software. The most important stress tests for a noncredit portfolio include changes in the yield curve, interest rate, prepayments, basis (mortgage versus Treasury positions) and volatility. A long mortgage portfolio contains an implied short embedded option exposure. The long mortgage security holder is short the prepayment option. Risk software will measure just how much exposure in dollar terms the portfolio can be expected to loose for changes in interest rate volatility. Yield curve, interest rate, basis and prepayments stress tests are paramount exposures to understand in a mortgage portfolio. In a given portfolio, the professional hedge fund investor may simply decide there is too much exposure for the current anticipated return.

The most relevant stress tests for mortgage credit portfolios are default sensitivity tests. For pools of mortgages which represent the lowest rated lenders one must evaluate a number of other metrics such as delinquency rates and days delinquent. All portfolio stress results should be relied upon for assessing tail risk as the mortgage markets can go through long periods of time without major disruptions. Volatility of returns in any given year may well disguise tail risk potential.

It is important to have quantitative risk software which is reliable and capable of evaluating all positions in the portfolio. It is also important to ensure that the prepayment model applied to the software is widely acceptable. Proprietary prepayment models must be vetted by knowledgeable professionals. The wrong prepay assumptions can be misleading about the tail risk in a portfolio. Reliable stress results give the professional hedge fund investor parameters with which to judge the potential for tail risk. Specifically, the professional hedge fund investor can assess the likelihood of a given move in the underlying markets and use the software results to understand the exposures they are assuming if they invest.

The underlying risk in a mortgage-backed security arbitrage portfolio is dependent upon the degree to which the portfolio is hedged. Beyond the aggregated position results as calculated by sophisticated risk software, the professional hedge fund investor should consider viewing the portfolio by its arbitrage substrategy security groupings. For example, how much exposure there is to basis, synthetic mortgages (e.g., interest-only and principal-only mortgage strips), coupon rolls, and so on? This qualitative or empirical approach allows the professional hedge fund investor to assess potential portfolio impairment during adverse moves in the underlying interest rate markets. For example, a prospective flight to quality in the interest rate markets can be assessed separately in the context of each substrategy exposure.

In most strategies, a view into positions is more revealing than top line stress results. However, because mortgage securities contain embedded options with many possible price paths on a forward-looking basis, a positions-level view is less helpful than with other strategies. Risk software must be relied on to reprice cash flows as underlying markets move and to reveal portfolio exposures.

While the primary stresses that should be evaluated are the portfolio's sensitivity to interest rate shifts, curve shifts, prepayment speed increases, volatility increases and mortgage versus U.S. Treasury spreads widening, equally important risk factors include the portfolio's duration, convexity and leverage. All these metrics are best understood when stated as a percentage of NAV impairment.

The professional hedge fund investor should be prepared to make assumptions on near term market volatility and on the prospects for a significant or an extreme rate or curve move in order to assess the risk versus reward contained in the mortgage-backed security arbitrage portfolio. Of course, the portfolio manager can alter the hedges or dynamically alter the included trades. The professional hedge fund investor when assessing portfolio risk must also consider these possibilities. In general, the professional hedge fund investor should be concerned if a reasonable set of stress events or stress events that are likely to occur in tandem could impair the portfolio beyond their loss tolerance.

Corporate Bond Arbitrage

Corporate bond arbitrage hedge funds trade investment grade and noninvestment-grade credit on a relative value basis. The strategies employed include basis (cash bonds against credit default swaps), long/short, pairs and structured credit trading.

When evaluating a corporate bond arbitrage portfolio manager's trading style, it is important to ascertain whether they trade in a relative value or in a macro style. The hedged portfolio manager will not have a significant difference between the notional long and short books regarding average spread, rating and industry descriptions. Let's consider two extreme examples: a basis trading portfolio and a portfolio that is long high-yield names and short AAA investment-grade names. In the former example, the basis trader could be long a corporate bonds and a holder of credit default swap protection. The differences in average spread will be limited and the rating and industry descriptions of the trading book will be identical between the long and short positions. The portfolio consisting of high-yield longs and AAA investment grade bond shorts will have large discrepancies in ratings, spread, and, perhaps, industries between the long and short positions. At times the long and short positions could perform in an uncorrelated fashion. Over long periods of time, the basis book can be expected to be the more conservative one, demonstrating fewer drawdowns in performance numbers and lower volatility. Either portfolio is not necessarily a good or bad investment. However, the styles are very different and the portfolio manager should be relied upon to demonstrate an understanding of the differences.

Corporate cash bond/derivative trades can present similar risks as government cash bond/futures trades. Assumptions on deliverability can prove to be incorrect at or well before settlement. If enough leverage is applied, small miscalculations can generate significant P&L swings. These changes occur due to prolonged supply and demand technicals which usually affect the cash markets more often. One such dislocation in the investment grade corporate bond markets could be caused by a corporate restructuring action which alters the value of an affected firm's debt. The bonds could become targets for short sellers and the bonds may become difficult to borrow. This would cause volatility in the basis trades. Two similar investment grade securities can present with large P&L changes. The professional hedge fund investor has to be sensitive to the possible risks a long and short book of unrelated credits could be subject to.

Corporate cash bond/derivative trades also in the lower rated credits are more susceptible than higher rated debt to jump to default risk. This risk is the uneven trajectory in prices which the cash and derivatives can take when a credit's rating migration is headed dramatically lower. This can present with prolonged dislocations and expensive borrows on short cash positions. This type of risk is difficult to forecast or protect against other than by limiting the potential harm it can do to the portfolio's value.

As can be expected, *leverage* plays a role in the corporate bond arbitrage portfolio. Even the more conservative basis trading style portfolio can become unstable from a volatility point of view if enough leverage is applied. A basis book at 20 times leveraged could become more risky than a long/short book at two times leveraged with large spread and rating discrepancies between the long and short positions. Each substrategy has an appropriate range of leverage for a given level of underlying volatility in the markets. They can be arguably rank ordered lowest to highest: unrelated long/short credit, net long credit exposure on a beta-adjusted basis, structured credit, pairs trades, and basis trades. The lowest levels of leverage should be applied to the seemingly unrelated long and short book or those with embedded leverage, as is often the case with structure credit securities. The highest leverage should be applied to trading books of securities which are highly similar.

Leverage should be assessed in context of the volatility of the underlying markets and to the extent the underlying markets' volatility can be understood. The more leveraged a position, the greater the P&L volatility. The professional hedge fund investor must not only evaluate current P&L swings within the portfolio but also anticipate prospective P&L swings. The professional hedge fund investor should consider the history of the markets as well as the current credit cycle when anticipating prospective volatility.

While net credit exposure in a portfolio usually presents with greater P&L volatility than pairs or basis trades, net credit exposure should be further evaluated. On a beta-adjusted basis, the long and short book differences in spread, curve location, credit rating, and industry should be considered. Most portfolios will not be fully hedged and will have net long credit exposure which correlates highly with a corporate bond index. Net credit exposure should be determined in context of a beta matching of the long and short books which should also factor in duration, particularly, in investment grade portfolios. A notional matched book of long and shorts should be considered to have net credit exposure if spread, curve, rating and/or industry differentials exist. The professional hedge fund investor considers these differences as containing larger amounts of investment risk the larger the net duration and spread exposures.

Single-name issuer concentration or trade-size concentration is an important feature to examine in a portfolio. One usually cannot predict consistently single-name credit problems, but through diversification one can limit the potential portfolio impairment a single credit problem can cause. At times, unexpected industry problems may emerge or expected problems may not emerge. Therefore, industry concentration limits are also prudent. For example, one country can lose its competitive advantage in a given industry. Suppliers to that industry can be negatively impacted in sympathy. The same is true for sectors. Diversification should be assessed in context of the percentage of NAV, not percentage to total portfolio value. Leverage plays a key role in the cost of an uncontrollable position. The larger the leverage in the portfolio strategy, the more likely one position can impair the entire portfolio in adverse market circumstances. If an industry problem occurs, the portfolio will be impaired faster if industry diversification limits are not placed and monitored as part of the risk process.

Another type of concentration risk is single-issue concentration risk. Prudent ownership percentages of a single issue are larger for longs than shorts. Liquidity is vitally important to the portfolio manager. It allows the portfolio manager to enter and exit positions at optimally chosen times. Large single issue concentration levels may afford the portfolio manager pricing power but not the ability to exit a position, which for a number of reasons, should be more important to the professional hedge fund investor. The other danger of large position concentration is that it can distract the portfolio manager from other profitable trades. Ten to 15% ownership of an issue for a long position is probably about the maximum one should assume allows for optimal trading flexibility. Shorts, which can be subject to squeezes, should probably be limited to 5% to 10% of a given issue. If a squeeze ensues, a small short position can be covered more easily than a large one. At times, a large single-issue concentrated short cannot be covered for a very long and costly period of time.

Corporate bond arbitrage can include structured security trades. Structured securities are constructed of multiple asset classes which include corporate bonds both cash and derivatives, asset-backed securities including mortgages and bank loans, among other securities. The portfolio risks are dependent on which tranche is owned in a given structured security. Ratings are a secondary consideration to tranche location. Lower tranches include the equity, preference shares or mezzanine. The principal risk in these tranches is default risk. (At the issue date, each structured credit security is overcollateralized which means that the market value of the pooled assets in the security exceeds the face value of the structured security.) Defaults beyond the over collateralized amount will impair the equity tranches first. The equity, preference shares, and mezzanine tranches are comparatively small in size versus the highest-rated tranche. In a sense these lower tranches contain high degrees of embedded leverage. That is to say, small amounts of capital assume the risks for a relatively large amount of securities. The leverage is implied but can be inferred through modeling of default scenarios and cash flow payout terms. If a portfolio contains deeply subordinated tranches of structured notes, the professional hedge fund investor needs to view leverage in a more circumspect way that by notional market value. It should be grossed up to reflect the leverage relative to a higher tranche to put the leverage in a risk equivalent terms.

Lower tranches are commensurately rewarded with higher coupons and the equity tranche is priced for the highest internal rate of return. These yields are dependent on supply and demand for the underlying collateral. The yields may be cheap on a relative basis but expensive on a historical basis. Ratings must be viewed in context of other structured transactions, not in context of corporate bonds. Professional hedge fund investors should be aware of subordinated tranches in a portfolio and be current on the possible impairment per new defaults in the pools of assets. In general, aging deals increasingly suffer collateral impairment, that is, the prospect of default increases as the deal ages. Structured product securities often have a trustee whose duty is to measure the statistical qualities of the collateral to determine compliance with the offering documents regarding securities diversity among other guidelines to preserve the rating. The professional hedge investor should occasionally request a sample of the trustees' reports to validate the performance of the individual investments. The structured credit markets include many derivative products. These products reference single-asset-backed security deals, aggregated loan pools, or tranches of these loan pools. These derivatives are often more liquid than the cash markets, however, issues about volatility, liquidity, and leverage have to be monitored and prudently traded.

While documents and rating evaluations can remain standardized for long periods of time, changes occur which affect the investor. It is important for the professional hedge fund investor to remain current with document and ratings changes.

Emerging Markets

Emerging markets credit hedge funds invest in the debt of economically developing countries. These countries tend to have lower per capita income and smaller stock and bond markets than developed countries' markets. Emerging market countries tend to use hard and local currency issued securities to access the publicly traded markets.

Some of the underlying risks in this strategy are sourced from the sovereign investment environment which, at times, can include local policy shifts, currency devaluations, ineffective central bank policy to control economic growth and inflation, volatile credit spreads and unreliable liquidity.

Emerging market debt over long periods of time can have significant ratings migrations, and be susceptible to principal impairment and currency devaluations. Consequently, it is very difficult, over long periods of time, to not incur losses due to these reasons. The same risk rules apply to this strategy as highlighted in the section on government and corporate bond arbitrage. The prudently managed emerging market debt portfolio should employ lower levels of leverage than one constructed of investment grade sovereign debt, even in periods of prosperity and tight spreads. Within emerging markets portfolios, those with local currency debt should employ even less leverage than those portfolios constructed in hard currency, if all other factors are constant. As with government bond arbitrage, the professional hedge fund investor should expect lower levels of leverage when large amounts of yield curve or outright country exposures are present over portfolios with long versus short exposures which cover shorter maturity differentials on the same yield curve.

The carefully monitored and actively managed emerging markets hedge fund can present tremendous opportunity due to the potential volatility and significant yield curve movements. From a risk perspective the monitoring and hedging must be well timed and reactive.

EVENT-DRIVEN STRATEGIES

Event-driven hedge funds are designed to profit from anticipated corporate events which will alter a given companies' debt and equity valuations. The event-driven portfolio manager will determine the likely change in debt and equity valuations and position their portfolio to reflect the anticipated outcome. The event-driven portfolio manager will have an expectation regarding the timing of the corporate event. The event is considered to be a catalyst for revaluing debt and equity prices. Event driven trading has many disciplines. These disciplines focus on companies experiencing financial distress such as capital structure arbitrage and distressed trading or on companies expected to experience a change in the corporate structure such as merger arbitrage.

Capital Structure Arbitrage

Capital structure arbitrage hedge funds take offsetting long and short positions within the capital structure of a given company. The long and short positions may be debt or equity securities of a company that is experiencing financial difficulty. The arbitrageur buys and sells securities within the capital structure in accordance with their interpretation of recovery values. Capital structure arbitrage trades are constructed to profit from relative mispricings within the capital structure of a given company.

The capital structure arbitrageur in addition to assessing the recovery values of the securities of a given company must also assess the relative amounts to position long and short. The market movements of the different parts of the capital structure of credit-impaired companies are unpredictable and often path dependent with large relative price fluctuations. Because the price paths of the securities are highly sensitive to changing default probabilities, prospective valuations and, hence, price expectations can be widely divergent. Not only miscalculating the likely outcome of a company can be costly but also miscalculating the hedge can exacerbate potential losses. The price trajectories of companies heading into default can occur more rapidly than recovering ones. Consequently, for a given weighting, a professional hedge fund investor should evaluate capital structure trades which are long senior debt and short junior debt or equity as less risky than ones which are short senior debt and long junior debt or equity. Of course, the truth of this statement is dependent upon the sizing of the trade. The proper sizing is difficult to assess other than at the extreme where one can usually presume that securities more junior in payment priority are more volatile and should be sized smaller relative to securities more senior in payment priority. A portfolio which is sized counter to this expectation should be viewed as having outright long or short exposure to the financial outcome of the company.

The portfolio manager will typically have a catalyst event in mind when positioning each trade. They size their trades in accordance with their perception of the eventual outcome. The capital structure arbitrageur will position defensively through long and short positions to profit from the expected outcome of a catalyst event. For example, if they presume a company will recover, they may buy the securities that will improve the most and hedge with a security that will improve less. The risk-averse capital structure arbitrageur will position in a way that is counter to this assertion but will just weight the trade in a bullish or long bias manner.

This strategy lends itself to containing net credit exposure, either long or short for even the most risk-averse arbitrageur. The professional hedge fund investor should view the net exposure in light of whether that exposure seems balanced on a volatility basis. The appropriate weighting and net exposure are assessments, not necessarily a measurement. Relative volatility changes and presents a challenge to the professional hedge fund investor to assess the weightings of these trades. At the extreme, where the capital structure arbitrageur is quite long, the risk assessment is easier. But often, this is not the case. But the professional hedge fund investor should find sufficient evidence and logic offered by the capital structure arbitrageur that volatilities are monitored, hedges, dynamically adjusted, all with fundamental rationale.

Capital structure arbitrage's success is dependent on the ability of the portfolio manager to borrow securities. Often, the securities that the manager wishes to borrow are scarce. The prudent manager will carefully manage their borrowing relationships in a way which ensures maximum access to the securities they wish to borrow. Dialogues with various lending institutions are important. The professional hedge fund investor should monitor the frequency with which a portfolio manager has its short positions called by a lending institution, forcing the capital structure arbitrageur to close out a position.

As with most strategies, the professional hedge fund investor should be concerned about the diversification of a capital structure arbitrageur's portfolio. Because the portfolio is subject to unsystematic risk, it does not necessarily follow that systematic factors may not trigger the unsystematic risk. For example, if corporate default rates are on the rise, many troubled companies may be affected. Or if interest rates are rising, the cost of borrowing may prove too onerous for troubled companies. A high level of diversification will help protect the portfolio's value.

Distressed Securities

Distressed securities portfolio managers invest in the debt and equity of companies which are highly leveraged and liquidity impaired. These companies require legal action or restructuring to improve asset value and financial stability. This strategy generally employs low leverage and has a longer investment time horizon than most other hedge fund strategies. Usually, the investment opportunities occur due to potentially favorable restructuring, recapitalizations or fundamental improvements in the companies' circumstances.

Risk evaluations of distressed portfolios do not lend themselves to a heavy reliance on quantitative risk analytical tools. Distressed securities have unsystematic risk. The valuations of these securities are dependent upon structural changes in a given company's outlook. Changes by management and in the business environment may be particular or unique to individual companies. Changes in the underlying markets as represented by the broad market indices are least likely to impact companies that are experiencing financial distress. Therefore, the securities within a distressed portfolio must be evaluated individually for the risks that the portfolio might assume.

One key factor the professional hedge fund investor will consider is default characteristics. First, is the company in default already? If so, the professional hedge fund investor will assess the presumed recovery rates or what percentage of par the debt holder is likely to recover. In the case of defaulted debt, the recovery value is presumed to be the current price. If the defaulted debt does not trade frequently, a very high valuation should be a signal that the manager may be marking the book unrealistically. For each company that is stressed but not in default, the portfolio manager should be not only aware of the probability of default through a Merton-style model but also be able to provide an estimate of a possible recovery value as well. Recovery values as assigned by each portfolio manager should be justifiable in terms of each credit's applicable assets.

In general, distressed securities are the most vulnerable to market illiquidity. (At times, however, large companies with fairly deep capital structures can defy this rule.) This strategy has little room for leverage. Under times of great market stress, a leveraged distressed portfolio's survival could be compromised. The portfolio without leverage would be able to wait through some period of difficulty, surviving the daily position markdowns as long as it did not suffer withdrawals. Workout times for these companies can exceed the hedge fund's liquidity. It is important for the professional hedge fund investor to assess whether the hedge fund's liquidity is consistent with the workout times of the distressed companies in the portfolio.

Distressed portfolios are often subject to binary risk. That is to say, the companies' survival depends upon certain outcomes regarding board votes, financings or regulatory intervention to name a few. These outcomes create risk in the portfolio and sharp price swings. Therefore, as with most strategies, sizing becomes very important. The portfolio with fewer positions will be riskiest and can be expected to experience sharp price swings.

Another feature of distressed investing is that the portfolio manager may become involved in creditor committees. This work requires a great deal of time and focus. This time needs to be well managed, as it is time spent away from managing the portfolio. The professional hedge fund investor must focus on how the portfolio manages this aspect of their time. There are real and implied costs to this process including legal liability.

Merger Arbitrage

Merger arbitrage is the investment discipline of buying target companies and often, selling their acquirers. The practice may involve announced deals or proposals. The types of transactions may include exchange offers, cash tenders, stock for stock, leveraged buyouts, among others. The principal risk is that deal terms change adversely or the deal breaks.

Merger arbitrage deals have varying degrees of risk depending on the certainty with which a particular deal will close. Deals in which the target and acquirer have agreed on terms are the lowest risk deals. Unforeseen circumstances are generally the only risk factor to the arbitrager. Deal types with greater risk and whose outcome is dependent upon regulatory approval such as anti-trust, activist hedge fund financing, speculation that a higher bid will materialize, hostile takeovers, among other situations present higher risk to the arbitrageur. Some merger arbitrage portfolio managers will become actively involved in merger terms to facilitate a desired outcome for their investments. Their success is dependent upon their persuasiveness with the board of directors of the target company. Consequently, deal spreads or projected returns on deals can range from barely over LIBOR to thousands of basis points over LIBOR. The professional hedge fund investor should be aware of the spreads of the deals in which their invested merger arbitrage funds invest.

Risk arbitrage portfolios can have very different risk profiles depending on the types of deals that are included in the portfolio. However, one risk remains the same threat level to any risk arbitrage portfolio regardless of the type of deals that are included in the portfolio. Even the most assured low-risk deal can be the victim of an external shock in the broader equity markets which can threaten the portfolio NAV. If a portfolio is not adequately diversified, the portfolio is at greater risk for potential impairment. Beyond proper diversification which may mean no position is greater than 7% of NAV, the portfolio manager should demonstrate awareness of the portfolio's downside value during multiple deal breaks. This is defined as predeal prices for both the target and acquiring companies.

VaR or equity stresses are not as useful as a measurement of merger arbitrage portfolio risks as a semi-qualitative assessment of the portfolio's vulnerability during deal breaks. The merger arbitrage portfolio is likely hedged and risk software might not calculate the risk in some deal types accurately. It is useful to the professional hedge fund investor to view the portfolio as if all deals broke or fell through. This repriced value is a return to preproposal status of both the target and acquiring companies and a very useful measure.

EQUITY STRATEGIES

Equity portfolio managers seek to build a portfolio which profits from temporary mispricings in equity valuations as represented by stock prices. The portfolio manager will determine if the current market price of a given equity differs from their own valuation. There are various equity strategies including ones with a global, country or sector focus. Additionally, long/short equity hedge funds may have a predominantly quantitative or qualitative investment process. Some hedge funds may be market neutral, which generally means that the portfolio manager will try to minimize market beta exposure. Convertible bond arbitrage is a hybrid debt/equity strategy and is often part of equity-based multistrategy funds.

Long/Short Equity

Long/short equity hedge funds hold long and short positions in equities. The portfolio may contain longer time horizon investments as well as shorter time horizon investments. The portfolios tend to be net long, decreasing net exposure and leverage in times of market downturns. Long/short equity portfolios reflect investment and risk decisions on sector, market cap, and net exposures.

Traditionally, at investment banks and banks, proprietary capital was allocated more typically to fixed income and currency markets traders and less so to equity traders. The equity divisions of investment banks focused on distribution. Equity proprietary capital was more commonly allocated to risk arbitrage or special situation investments over the style of trading which is encompassed in the long/short equity hedge fund strategy. It is common for long/short portfolio managers to have had prior careers as highly regarded analysts or total return managers. Long/short portfolio managers have widely divergent approaches to risk management than do other hedge fund disciplines.

The professional hedge fund investor should evaluate how the long/short equity hedge fund measures and monitors their exposures. The recovery rates in equities can be presumed to be zero. Therefore, the portfolio manager must manage their risk carefully in order to maintain the portfolio's viability and success. Important topics to monitor include cap skew between longs and shorts, sector and industry exposures. The larger the imbalance between the long and short books for these classifications, the greater the potential for volatility in returns. In order for the portfolio manager to prudently manage these risks, they must monitor these exposures.

Diversification is similarly important. Positions which are larger than 7% of the NAV can generally be considered to be concentrated positions. Position limits for longs should be greater than those for shorts. Shorts have negatively asymmetric risk and can be more difficult to finance than long positions. A size limit protects somewhat the total portfolio value at risk during a short squeeze. Inattention to the possibility of short squeezes can be harmful to the portfolio.

The professional hedge fund investor should be familiar with the stop loss rules practiced by the portfolio manager. The advantage of stop loss limits is the imposition of nonemotional, nonanalytical protection of the portfolio's value. Stop losses may mean the end or partial end of a particular open position but they also do not prevent the portfolio manager from recommitting to the same trade. Large losses have a tendency to cloud the judgment of the portfolio manager.

Portfolio managers will often say that at a certain loss level the losing trade will come under greater scrutiny. However, if the original thesis is intact, they remain with the trade. This practice presents greater loss potential to the portfolio than a liquidation or partial liquidation stop loss rule. The riskiest practice is the so-called double-down practice which calls for the trader to take advantage of the lower price of a current long or higher price in the case of a current short by increasing the position in share amount terms. The double-down practice is the riskiest practice of all, as the portfolio manager has increased exposure to a trade which is going the wrong way for the portfolio.

Another consideration that the professional hedge fund investor should be attuned to is that the long/short portfolio manager applies less leverage when trading the equities of smaller cap, distressed and lower-rated sovereign countries. For example, the professional hedge fund investor should expect that the leverage limits applied to German equities be greater than emerging market equities. As well, shorting low price dollar stocks presents adverse asymmetrical risks for the portfolio.

Many long/short equity portfolio managers are required to file position reports with the Securities and Exchange Commission (SEC). These reports are readily available and should be reviewed by the professional hedge fund investor on a continuous basis. When reviewing the SEC reports, the professional hedge fund investor should affirm that the reported positions are consistent with printed material provided by the portfolio manager.

Emerging market equity funds should be monitored for the same practices but with lower leverage thresholds than are tolerated in G-7 government arbitrage and long/short equity portfolios. The professional hedge fund manager should also be aware of small-cap equities in the emerging market portfolio. These stocks can be more easily manipulated in terms of their valuations which not only could misrepresent the portfolio's true value but also become the subject of regulatory scrutiny.

Multistrategy

Multistrategy hedge funds encompass multiple strategies weighted among any of the strategies covered in this chapter.

A risk evaluation of this strategy should begin with the separation of the portfolio into each of the substrategies and applying the risk monitoring techniques covered throughout this chapter. Only then should the professional hedge fund investor evaluate the portfolio in an aggregated perspective. Specifically, the professional hedge fund investor should evaluate whether the portfolio contains considerable overlap in themes or exposures which present concentration risk.

In addition to the strategies covered in this chapter, the multistrategy fund may include special situation investments or private investment in public entities (PIPEs). These substrategies, while rarely involving leverage, present unique risks to the portfolio. Often assets held in the portfolio which are classified as special situation or PIPEs are less liquid or illiquid. Often these investments are side-pocketed with different liquidity terms for the investor than the rest of the portfolio. In times of financial difficulty, these investments may have severely restricted liquidity and, hence, are difficult to price or value. Perhaps the portfolio manager is drawn to these investment opportunities because they present long-term gain possibilities not readily available in the liquid markets. If these investments are included in the portfolio, the professional hedge fund investor must pay careful attention to the basis for the portfolio manager's valuation. These investments can be the catalyst for a manager to enforce gate restrictions on fund withdrawals.

The professional hedge fund investor should evaluate the effectiveness of centralized risk controls within the multistrategy hedge fund for the sake of the unique risks multistrategy funds present. The professional hedge fund investor needs to evaluate the allocation method used to deploy risk capital and leverage among the investment strategies.

Equity Market Neutral

The equity market neutral portfolio can encompass qualitative and/or quantitative investment styles. Market neutral may be market value neutral or beta neutral. Usually, the market neutral portfolio will be more protected from market volatility than a long long/short hedge fund with a net long. Many equity market neutral hedge funds utilize sophisticated computer models and electronic trading systems working in tandem to remove net market beta or market directional exposure.

Quantitative and qualitative market-neutral portfolios may have non-market-neutral elements. It is important to figure out what the manager's tendency is toward exposure in the book. Important topics to monitor include cap skew, sector and industry exposure, and concentration between longs and shorts. The larger these imbalances between the long and short books for these classifications, the greater the potential return volatility. In order for the portfolio manager to prudently manage these risks, they must monitor these exposures. The professional hedge fund investor should be aware of these skews. This awareness will allow one to anticipate portfolio valuations changes in response to broad market movements. If the strategy is an option-based one, the professional hedge fund investor should follow the same guidelines regarding net premium exposure as detailed in the section on government bond arbitrage. The risk principal of being short vega or a net receiver of premium applies to both debt and equity options.

Academically a truly market neutral portfolio is one which is beta neutral at all times. Because market values and betas change constantly, market neutrality in a portfolio is almost impossible to achieve. Even if were to be achievable, it may not be the best strategy for creating alpha which is often the goal of hedge fund investing.

Convertible Arbitrage

A convertible bond arbitrage position generally involves the purchase of a bond which is convertible into equity of the same company at a specified conversion rate and the short sale of the same equity. Convertible arbitrage presents the trader with the opportunity to take advantage of pricing inefficiencies between a bond and its reference equity. The timing of the trade may be catalyst driven or due to the trader's belief in the relative mispricing of the convertible bond and its equity. The trade can be asset swapped which eliminates much of the credit spread risk in the convertible bond arbitrage position. In this case the purpose of the trade is to profit on volatility changes in the embedded option to convert into equity. Convertible bond arbitrage positions which do not involve asset swaps are generally plays on credit as well as volatility and equity performance.

Assets that are good candidates for arbitrage should be relied upon to covary in some predictable way. If they deviate, it can be assumed the deviation will be for a brief period of time. One of the assumptions that an arbitrager must have is that they not only can position an arbitrage but that they can exit or unwind the trade when they want to as well, within some price range. In order for that assumption to hold, there must be an active, nonhomogenous group of buyers and sellers. Historically, convertible bondholders are more likely to be leveraged investors than the holders of other types of debt. The risk that this fact presents is that the arbitrageur will not have a counterparty to unwind the arbitrage with during time of financial stress. The more homogenous the holders, the more likely all holders will assess the value of a given convertible bond similarly. Or more directly stated, all the holders of a given bond could head for the exit door at the same time with little prospect of a buyer emerging. Nonetheless, the professional hedge fund investor should review the issue percentage ownership for each convertible bond in the portfolio. A high percentage per issue held is roughly and arguably in the range of 15% to 20%. The professional hedge fund investor should also try to assess the other holders of the issues of the portfolio. Concentration of ownership per issue by the arbitrageur of the portfolio being assessed is a significant liquidity risk factor. As stated above, if the other holders are hedge funds or leveraged investors, the portfolio risk increases.

Convertible bonds are less senior than other bonds issued by a given corporation. Because of the convert feature, convertible bonds trade with lower yields than other bonds in the capital structure. As the price at which the convertible bondholder can convert to equity becomes significantly out of the money, the convertible bond will decline in price or its yield will rise in line with the corporation's other debt obligations, reflecting the convertible bond's lower payment priority. Finally, one might assume that the market will place a recovery value on the convertible bond that is in line with the payment priority of other debt which is senior to the convertible bond. However, there must be active market participants willing to perceive the convertible bond's value and bid for it. The convertible bond holder may find that the out of the money convertible bond's yield will not only rise but it may rise above any level the arbitrageur envisioned before other, nonconvertible bond arbitrageur investors materialize with bids. In times of market stress such as 1998, convertible bonds lacked bidders. A convertible bond arbitrageur will face, in times of stress, the market's disregard for the implied relationship between convertible bonds and equity. The correlation between them can deteriorate and break down. The portfolio can be severely impaired. At this point the bond will trade increasingly with a credit predominant feature. *Liquidity* risk is a significant risk that convertible bond arbitrage portfolio managers assume.

If the principal activity of the convertible bond arbitrageur is to buy cheap volatility, the professional hedge fund investor should assess the price of convertible bond volatility not only in historical context of implied volatility in the embedded option in the convertible bond but in context of the historical volatility of the reference equity. The professional hedge fund investor should be aware of this deviation and the portfolio manager's sensitivity to it. If the deviation between these volatilities is great enough and the portfolio manager has not in fact bought the cheaper one, the portfolio may not perform well as equity options will attract the marginal buyer looking for the cheapest price for equity volatility. If the convertible bond arbitrageur's volatility portfolio consists of overpriced volatility, liquidity risk rises.

Historically, there have been several well-publicized failures in convertible bond arbitrage hedge funds due to the use of matrix pricing models to determine the funds' net asset value. Investors need to be aware of this history and pay special attention to how convertible bonds are priced.

SUMMARY

The professional hedge fund investor must understand the strategies in which they invest and the potential risks of each strategy. Many of these risks are similar across hedge fund strategies and may only appear to be different. The risks of a corporate bond arbitrage portfolio that is long high-yield debt and short investment-grade debt or of a long/short equity portfolio which is long small-cap stocks and short large cap stocks are similar with regard to the dissimilarity of the short hedge position to the long position. The professional hedge fund investor should look for the differences between short and long positions in a portfolio and know that one of the key risks to the portfolio's value may lay within this difference. The professional hedge fund investor develops a keen sense for not only this issue but for issues on leverage, liquidity, volatility, and position concentration, among other risk factors that could impair the value of their investments.

ACKNOWLEDGMENTS

The author is especially grateful to Dennis Winter for his insights and guidance. The author would also like to thank Ed Cleaver and Barry Campbell for their comments.

REFERENCES

Ackermann, C., McEnally, R., and Ravenscraft, D. (1999). The performance of hedge funds: Risk, return, and incentives. *Journal of Finance* 54, 3: 833–874.

Fung, W., and Hsieh, D. A. (2001). The risk in hedge fund strategies: theory and evidence from trend followers. *Review of Financial Studies* 14, 2: 313–341.

Fabozzi, F. J., and Manning, Steven V. (ed.) (2005). *Securities Finance: Securities Lending and Repurchase Agreements*. Hoboken, NJ: John Wiley & Sons.

Morris, S., and Shin, H. S. (1999). Risk management with interdependent choice. *Oxford Review of Economic Policy* 15: 52–62.

Fabozzi, F. J. (ed.). (2004). *Short Selling: Strategies, Risks and Rewards*. Hoboken, NJ: John Wiley & Sons.

Horowitz, R. (2004). *Hedge Fund Risk Fundamentals: Solving the Risk Management and Transparency Challenge*. New York: Bloomberg Press.

Parker, V. (2005). *Managing Hedge Fund Risk: Strategies and Insights from Investors, Counterparties, Hedge Funds and Regulators*, 2nd edition. London: Risk Books.

Gregoriou, G. N., Karavas, V. N., and Rouah, F. (2003). *Hedge Funds: Strategies, Risk Assessment, and Return*. Maryland: Beard Books.

CHAPTER 56

Diversify a Portfolio with Tangible Commodities

HENRY G. JARECKI, MD
Chairman, The Falconwood Corporation

TERRENCE F. MARTELL, PhD
Saxe Distinguished Professor of Finance, Zicklin School of Business, Baruch College/CUNY

Abstract: Tangible commodities represent an asset class separate from stocks and bonds. Historical returns of a diversified portfolio of commodities are as high as those of equities and are not correlated with them. Thus, the inclusion of a modest allocation of commodities to a typical stock and bond portfolio provides a diversification benefit, reducing the overall risk of that portfolio without reducing the returns. The most efficient way to get consistent exposure to commodities for diversification is by an investment in a long-only commodity fund.

Keywords: diversification, asset allocation, tangible commodities, futures, volatility, risk-adjusted return, investment portfolio, mean-variance optimization, asset class, efficient frontier, correlation, Tangible Asset Program® (TAP®)

Diversifying to reduce the risk of ruin is not a new concept. "Don't put all your eggs in one basket" is age-old wisdom. What lies behind this adage? And how do commodities relate to diversification?

This chapter will help you understand why commodities, when added to a stock and bond portfolio in reasonable quantities, tend to reduce the overall risk of the portfolio without reducing the overall return on the portfolio. The chapter will also explain what a long-only commodity fund is and why long-only commodity funds are the most appropriate way to gain commodity exposure.

THE BENEFITS OF TANGIBLE COMMODITIES

Investors have traditionally owned stocks for appreciation and bonds for stability, but the idea of having commodities in one's *investment portfolio* is relatively new. Only gold, of all the commodities, has had a long history as an investment. Other *tangible commodities* such as cattle, copper, and cotton were not considered for investment purposes until the early 1990s.

Before that, theoreticians, investment institutions, and their advisers—who had been introduced to financial *futures* trading on the Chicago Mercantile Exchange in the early 1970s—knew little about tangible commodities, and the little that they thought they knew was that commodities were far more volatile than stocks and that most commodity investors went broke.

It was probably true that most commodity investors lost money, but this was not due to any inherent danger but to the fact that commodity futures can be bought on so little margin that the average investor's exposure is 10 or even 20 times as great as the limited amount of margin money he or she deposits. As a result, even small price changes lead to large fluctuations in profit and loss; and it is this observation that has led people to believe that the dollar values of the commodities themselves were very volatile. This is not true, however. When theoreticians actually measured commodity price *volatility,* they found that tangible commodities were less volatile than equities.

This fact, coupled with tangible commodities' lack of *correlation* with both equities and bonds, prompted theoreticians to look at them as a unique *asset class* that could provide a substantial *diversification* benefit when added to an investment portfolio.

Traditionally, the commodity market was composed of "hedgers," who used the market to reduce their risk, and "speculators," who tried to profit from price movement. As a result of the application of *asset allocation* principles to commodities, we have seen the development of a third major market participant, the commodity investor. Unlike the speculator who makes his or her own buy and sell decisions or is guided to them by his broker, the commodity investor typically employs a third-party fund manager to make the decisions.

As a consequence of the rise in energy and metals prices, of the prospect for further increases due to the long-term economic growth of the developing Chinese and Indian economies, and of the fear of U.S. inflation induced by expanding budget deficits, more and more people—and institutions as well—began investing in commodities. Many of them have come to include a diversified basket of tangible commodities in their portfolios.

They believe, and we agree, that it makes sense to put commodities into an investment portfolio because:

- Their historical returns and volatilities are similar to those of equities.
- Their returns are not correlated with those of equities or of bonds.
- Over almost any long period since 1960, they improve the *risk-adjusted return* and lower the risk of a large loss of most "equities-plus-bonds" mixes to which they are added.
- They provide protection against political crises and natural disasters.
- They provide significant protection against inflation.

We will review each of these characteristics of tangible commodities, using the 20-year returns of an actual commodity investment portfolio called TAP® as a proxy for commodities. TAP (for Tangible Asset Program®) was created in 1987 by Henry Jarecki as the commodity portion of a diversified investment portfolio, in accordance with the asset allocation principles of Markowitz portfolio theory.

The Historical Returns of Commodities Are Similar to Those of Equities

TAP's returns since its inception in January 1987 through March of 2007 are comparable to those of equities. TAP's commodities have provided a return of 11.6%; the S&P 500 11.7%. More important, as Figure 56.1 reports, commodities have had a lower standard deviation—12.6% compared to 15.2% for equities—despite the fact that there are nearly 28 times as many objects in the S&P 500 index as in the TAP commodities portfolio. (The more different objects there are in any pool, the lower its volatility is likely to be. The 500-member S&P should logically have more internal offsets and thus a lower volatility than the 18-member pool of TAP's exchange-traded commodities.) The higher return and lower volatility significantly enhanced TAP's risk-adjusted return. TAP's Sharpe ratio, a measure of return/risk utility, was 0.54—about as high as the Sharpe ratio of bonds—and significantly higher than the 0.45 Sharpe ratio of equities over the same period.

The TAP portfolio has the lowest average volatility (12.6%) of any object displayed in Figure 56.1, followed by gold and feeder cattle, and only then by the DJIA (15.2%). The figure also shows that individual commodities are no more volatile than individual equities. The average commodity volatility (28.6%) is marginally higher than the average volatility of the 30 blue-chip stocks (27.6%).

Returns of Commodities Are Not Correlated with Those of Equities and Bonds

Commodities help to diversify a portfolio because their returns are not correlated with those of other asset classes. Indeed, they have had a slightly negative correlation with U.S. equities (−0.18), foreign equities (−0.11), and U.S. bonds (−0.13). The lack of correlation between asset classes that have positive returns reduces the volatility (standard deviation) of returns and produces improved risk-adjusted returns. Commodities' lack of correlation with equities and bonds also confers downside protection during events that have a negative impact on stock and bond returns, such as natural disasters, geopolitical uncertainty, or times of rapidly increasing energy prices.

The low correlation of commodities with equities can perhaps be explained by thinking of equities as *anticipatory assets* in the sense that investors value them based on their future cash flows. Commodities, however, are priced based on supply and demand; their prices reflect an *equilibrium* based on production and consumption rates and the adequacy of inventories. Are people living well, buying a lot of things, and thus making rarer and more expensive the commodities of which the things are made? Or are they retrenching and not buying things? Where current inventory and future production appear inadequate to meet current and future needs, current prices rise to balance

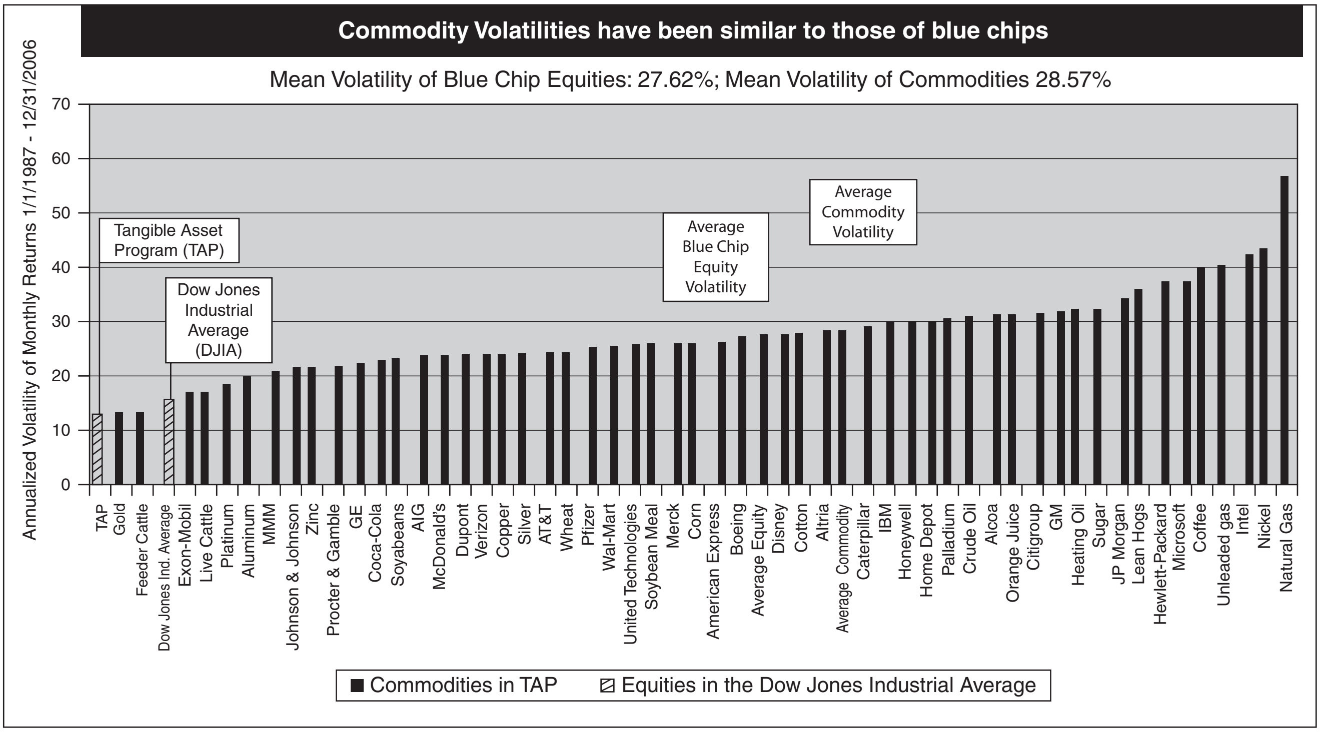


Figure 56.1 Comparison of Commodity and Equity Volatility: January 1987 to December 2006
Source: Gresham Investment Management LLC.

supply and demand and to induce producers to expand future production by allocating investment capital.

Commodities can for these reasons be a source of return across market cycles. As Figure 56.2 shows, commodity futures have historically achieved positive returns not only when stocks and bonds went up, but also when they went down.

There is a sound economic reason for the countercyclical character of stock prices and commodity prices. It relates to the business cycle. Equity returns will, according to this view, be highest at the start of an economic expansion when expectations of future cash flows are running high, whereas commodity returns are highest near the end of an economic expansion when consumption rates are high, inventories are depleted, and new production has not yet come on stream. Returns of commodities and equities thus have somewhat alternating price cycles.

Add Commodities to Your Portfolio for Increased Risk-Adjusted Return

Asset allocators are concerned with improving the overall return/risk profile of their portfolios and with decreasing the possibility of large losses. But, ideally, they would prefer not to sacrifice returns, and that is where commodities come into play. From 1987 onward (when TAP was started), an allocation to a diversified basket of long-only tangible commodities has helped to increase a portfolio's total return while decreasing its risk—even if, in order to buy commodities, one had lowered one's equity allocations. The hypothetical benefit of adding commodities to a traditional equities and bonds portfolio is shown in Figure 56.3.

If we look at the *efficient frontier* of portfolios consisting only of domestic equities and domestic bonds, we can see that, over the 17-year period from 1987 to 2005, adding commodities in various amounts (5, 10, and 20% in Figure 56.3) improved the return/risk profile of whatever equity/bond mix one chooses (40–60, 50–50, or 60–40). In fact, each incremental addition of commodities to the original equity-bond mix defines a more favorable efficient frontier.

Commodities Can Protect a Portfolio during Periods of Uncertainty

As noted above, commodities can protect overall portfolio returns because they have historically shown little or no correlation to the downward moves of stock and bond markets, and they tend to gain in value during events that have a negative impact on stock and bond returns.

During political crises and natural calamities, the correlations of most financial assets—with the possible exception of U.S. Treasury securities—tend to increase. Their prices fall together as investors bail out of assets they perceive as more risky. Commodities, however, offer protection against losses during such times. In addition, large commodity price movements are more often increases than decreases. These relationships are so well known among professionals that it is sometimes said that commodities "fall up."

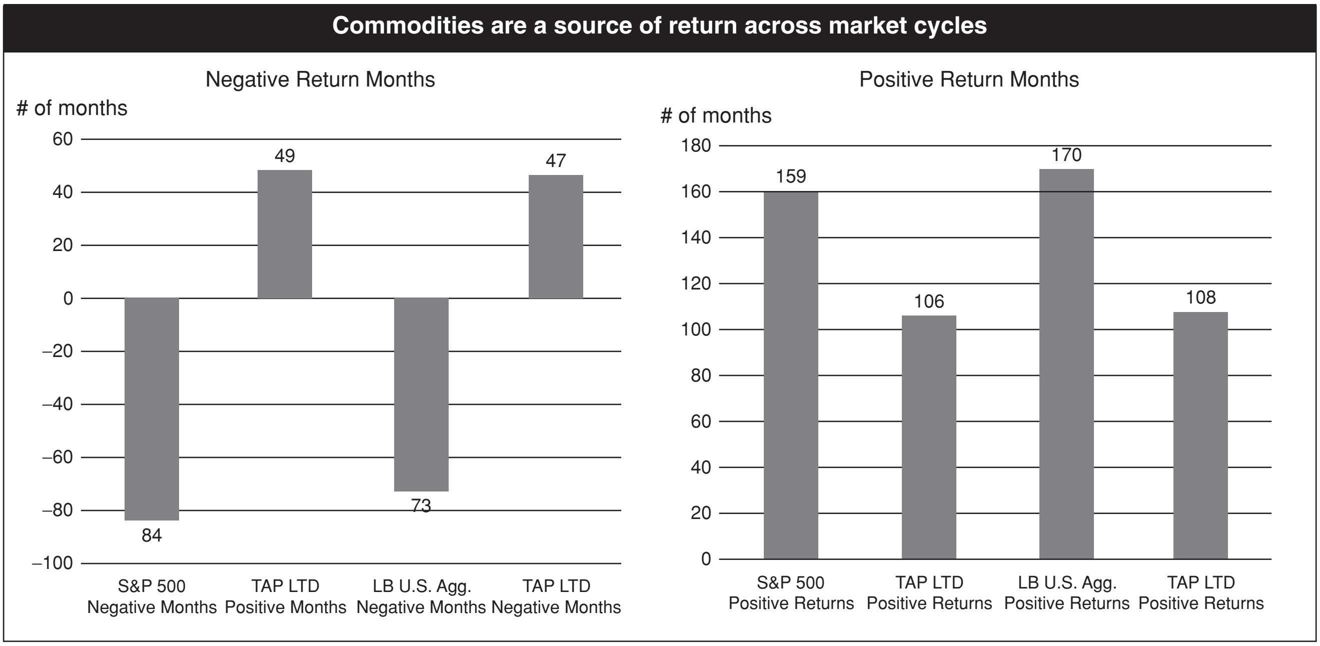


Figure 56.2 Comparison of Returns: January 1987 to March 2007
Source: Gresham Investment Management LLC.

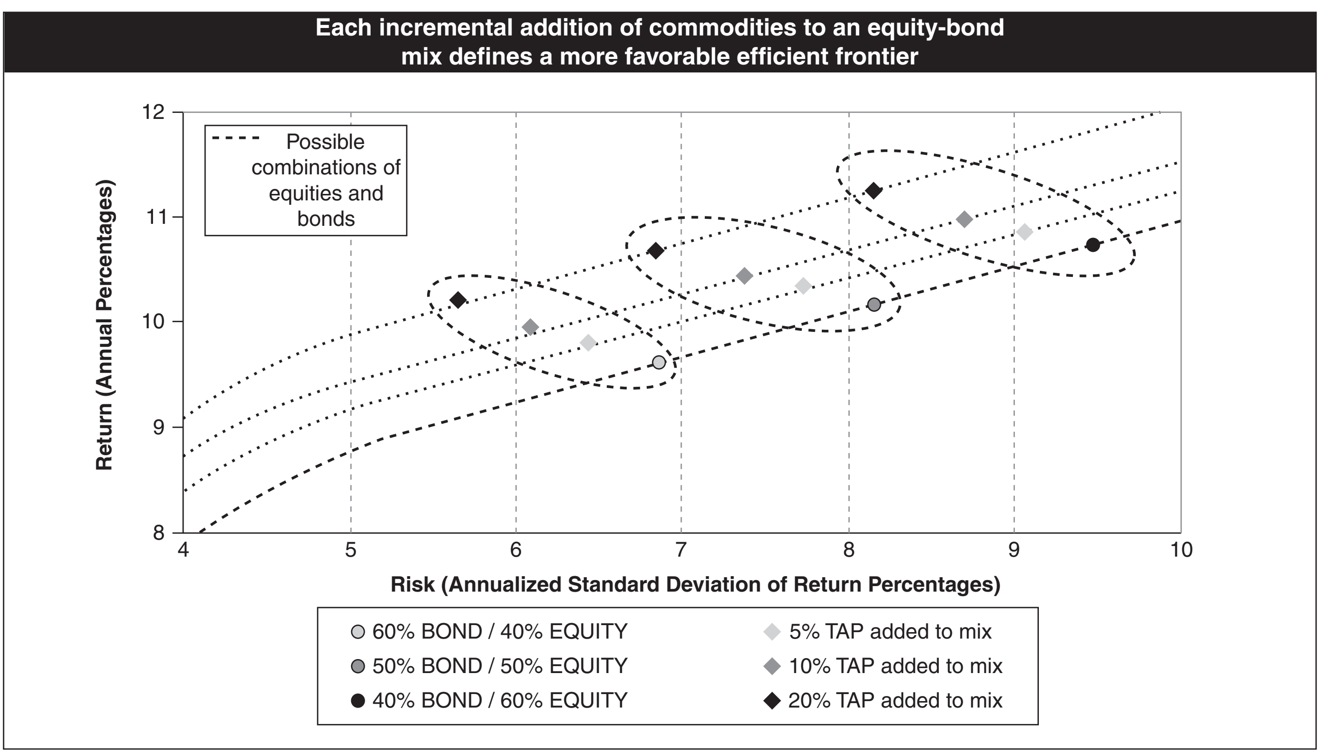


Figure 56.3 Balanced Portfolio Efficient Frontier with and without Commodities: 1987 to 2005
Source: Gresham Investment Management LLC.

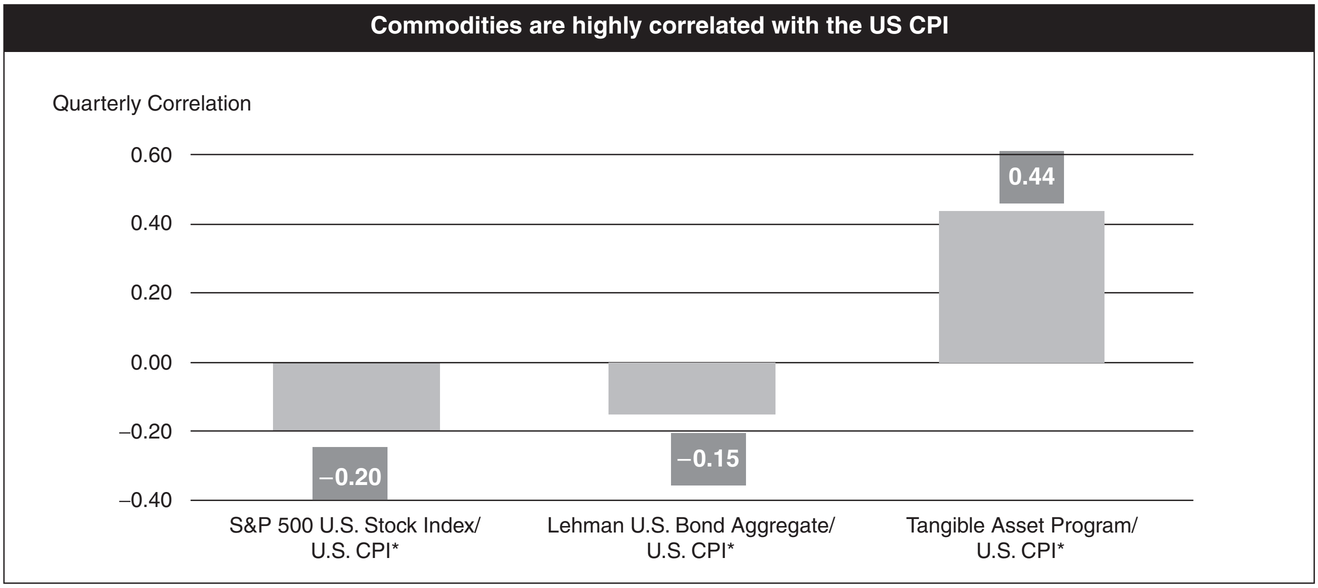


Figure 56.4 Rolling Quarterly Return Correlation Comparison: January 1987 to March 2007

As Figure 56.4 reports, unlike equities and bonds, which are negatively correlated with the U.S. Consumer Price Index, commodities are positively correlated with the CPI (0.44). This is not surprising since commodity prices are a main determinant of the cost of living and thus of inflation. On a shorter-term basis, commodities, denominated in an investor's home currency, tend to perform well when that currency weakens.

HOW TO GET COMMODITY EXPOSURE

What is the best way to get diversified commodity exposure for an investment portfolio? Buying a collection of physical commodities is impractical for most investors because of the difficulties of purchase, transportation, storage, spoilage and shrinkage, and the inability to rebalance. There are only four practical ways to go about obtaining such exposure: investing in natural resource shares or funds, managed futures funds, nondiversified commodity sector funds, and diversified commodity funds and indices.

- **Natural resource shares or funds**. Some investors think they have commodity exposure because they own natural resource stocks. This does not work as well as one might expect, for the value of such stocks is affected by factors other than commodity prices, such as industry competition, management's successful strategy execution, and production-related hedging, all of which affect performance and thus dividends and share prices. There is often little correlation between the price of a commodity producer's shares and the price of the commodity itself. For example, over the period 1990 through 2006, the correlation between American Stock Exchange Oil Company index and the NYMEX Crude Oil contract was 0.31.
- **Managed futures funds**. Some investors hope to obtain exposure to commodities by buying a managed futures (or Commodity Trading Advisor's) fund. While such a fund may or may not provide satisfactory returns, it is not an effective way to get commodity exposure. Most CTA programs involve financial futures strategies, not tangible commodities. And, while CTA returns may not be correlated to equities, they are also not commodity class returns. Even if a CTA includes tangible commodities in a fund, the program may use substantial leverage and may also go short, so an investor will have far more or far less commodity exposure than might have been expected. Finally, a CTA's fees may or may not be reasonable from a trading perspective, but they are high if your goal is adding asset class exposure in commodities.

  Some investors have looked to a managed futures fund to supplement a core investment in a diversified commodity fund, thus using the opportunities available to a talented money manager to gain alpha. This "core-plus-satellite" strategy may make sense for those who understand the risks inherent in active management programs.
- *Nondiversified commodity sector funds.* Exchange-traded funds (ETF's) *and nonliquid structures*: Included in this category are a plethora of old and new investment vehicles such as oil and gas leases, mineral rights, timber, alternative energy plays, single-commodity and single-sector ETFs, funds with a very small number of commodities, and similar structures. Some of these are liquid; some require the investor to give the manager an extended notice period of the desire to withdraw funds (that is, long lock-up periods). By definition, they are all less diversified and therefore can be expected to have a higher volatility than a diversified fund, though they may meet some investors' needs. Because they are not as diversified, they do not provide efficient exposure to the commodity asset class.

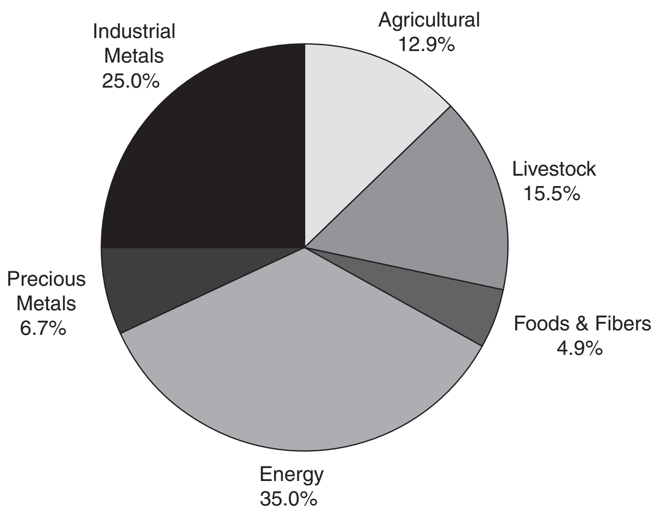
- *Diversified commodity funds and investable indices.* The most efficient way to get commodity class exposure is to invest in a rules-based, long-only fund product or private account that invests in liquid tangible futures contracts. Fees will be substantially less than for an actively managed fund, and managers will typically not demand a profit share for providing asset class exposure alone.

The Tangible Asset Program is an example of a long-only commodity fund. The TAP methodology consists of a set of rules for choosing the commodities and the amounts of each one (their relative weights), two factors that are updated annually. The rules are straightforward and transparent and reflect the experience gained from maintaining TAP since 1987. Some of the more important "rules" for TAP include:

- To make it possible to buy and sell the commodities without being subject to high transaction costs, TAP is limited to exchange-traded commodity futures contracts.
- TAP's managers calculate the dollar value traded of all futures contracts (as a measure of liquidity) and average that number with double the global production value of the underlying commodities (reflecting that real commodity prices are determined in the streets and stores and not on exchange floors).
- TAP's managers then select the top three commodities in each of six separate groups.
- There is no leverage. If the model calls for a $1 million investment in corn, TAP would go long (buy) corn futures contracts with a face value of $1 million, depositing the margin (approximately 5% of the value of the commodity position or $50,000) and placing the remainder in short-term, low-risk money market instruments like U.S. Treasury bills.
- The weight of each commodity and of each group is limited to ensure that the portfolio is truly diversified and not just a tail being wagged by some dominant commodity's dog.
- Finally, each commodity's weight is rebalanced whenever price changes cause the initial weighting to get out of line. This results in reducing positions when prices move up and increasing positions when prices move down.

The 2007 composition of TAP is shown in Figure 56.5. There are a total of 18 physical commodities in the portfolio, with no single commodity dominating.

A Diversified Pool of Tangible Commodities



2007 TAP Commodity Weights

Group	Commodity	Weight (%)
Energy		**35.00%**
	Crude oil	19.00%
	Natural gas	5.70%
	Heating oil	5.60%
	Unleaded gas	4.70%
Industrial Metals		**25.00%**
	Copper	11.10%
	Aluminum	7.70%
	Zinc	3.20%
	Nickel	3.00%
Livestock		**15.50%**
	Live cattle	8.70%
	Lean hogs	4.60%
	Feeder cattle	2.20%
Agricultural		**12.90%**
	Soybeans	3.50%
	Wheat	4.40%
	Soybean meal	1.30%
	Corn	3.70%
Precious Metals		**6.70%**
	Gold	4.70%
	Silver	1.20%
	Platinum	0.60%
	Palladium	0.20%
Foods & Fibers		**4.90%**
	Sugar	2.20%
	Cotton	1.20%
	Coffee	1.10%
	Orange juice	0.40%

Figure 56.5 TAP Is a Rules-Based, Long-Only, Diversified, Multicommodity Strategy that Was Started in 1987
Source: Gresham Investment Management LLC.

Since the pioneering work in defining the TAP methodology, interest has grown in using tangible commodities for diversifying an investment portfolio, as evidenced by the later creation of the Goldman Sachs Commodity Index (GSCI) and Dow Jones-AIG commodity indices. In recent years there has been significant interest in funds such as TAP, and there is now well in excess of $100 billion invested in such funds.

Can a long-only commodity fund be too diversified? A fund with 30 or 35 commodities is certainly diverse and would be expected to offer the return-smoothing benefits that flow from diversification. However, such a fund would be subject to two other potential problems: a lack of liquidity in the futures contracts the fund must buy and roll; and a lack of replicability/transparency if the futures contracts are traded in foreign time zones and are settled in currencies other than the U.S. dollar. Since tangible futures contracts must be rolled to a forward month prior to delivery, trading volumes must be adequate to support the periodic rolling, or the fund will be penalized with less robust returns. Since there are at most 20 to 25 liquid tangible commodities traded on U.S. and U.K. exchanges that are settled in U.S. dollars, one should be wary of a fund that includes futures contracts or exchanges that may not support rolling.

SUMMARY

In this chapter, we have reviewed the case for the inclusion of commodities in a well-diversified investment portfolio. The key components of the arguments are that diversification reduces risk and that the risk-return characteristics of a portfolio of commodities are sufficiently different from other assets so as to comprise a separate asset class.

A diversified commodity portfolio has returns similar to equity returns but with less risk. Moreover—and this is key—the correlation between those returns are low. Thus, the inclusion of a modest allocation of commodities to a typical stock-and-bond portfolio will reduce the overall risk of that portfolio without reducing the returns.

We also discussed how to implement a commodities investment strategy. We argue that the most efficient way to gain exposure to commodities is through an investment in a long-only commodity fund such as TAP. Long-only commodity funds, generally speaking, have lower fees, an important consideration in any investment decision. They have clear rules that govern how they construct their portfolios. We caution against investing in a long-only commodity fund with too many different commodities (concern about liquidity of the individual markets) or which allocates too much of the fund to one commodity or commodity sector (less diversification).

With these caveats, we believe the emergence of diversified commodity funds represents a major opportunity for investors to diversify their investment portfolios and thereby reduce portfolio volatility and risk of ruin.

REFERENCES

Erb, C. B., and Harvey, C. R. (2005). The tactical and strategic value of commodity futures. *Financial Analysts Journal* 62, 2: 69–79.

Gorton, G., and Rouwenhorst, K. G. (2004). Facts and fantasies about commodity futures. *Financial Analysts Journal* 62, 2: 47–68.

Lamle, H., and Martell, T. F. (2005). A new era for commodity investments. *Journal of Financial Transformation* 15, December: 120–125.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance* 7, 1: 77–91.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New Haven, CT: Cowles Foundation Yale University.

Sharpe, W. F. (1994). The Sharpe ratio. *Journal of Portfolio Management* 21, Fall: 49–58.

CHAPTER 57

The Fundamentals of Commodity Investments

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

ROLAND FÜSS, PhD
Professor of Finance, Union Investment Endowed Chair of Asset Management, European Business School (EBS), International University-Schloss Reichartshausen

DIETER G. KAISER, PhD
Director Alternative Investments, Feri Institutional Advisors GmbH and Research Fellow, Centre for Practical Quantitative Finance, Frankfurt School of Finance and Management

Abstract: According to the academic literature, investments in commodity markets are considered an effective way for investors to diversify traditional portfolios. The diversification benefits of commodities come from their low (and sometimes even negative) correlation with equity and bond markets, as well as from their high positive correlation with inflation. Therefore, during times of price increases, commodities as real assets can function as effective inflation hedges. Moreover, the low correlation with stocks and bonds remains even in downward-trending markets (that is, during phases when it is needed most). However, because commodities can be characterized as a heterogeneous asset class, commodity sector risk and return profiles can vary quite significantly, and may even move in opposite directions. In addition, the complexity of commodity investments can be revealed when considering the different ways investors can obtain exposure to this asset class. Commodity stocks, commodity funds, commodity futures, and futures indices all provide specific advantages and specific disadvantages.

Keywords: commodities, hedger, speculator, arbitrageur, basis trade, soft commodities, hard commodities, convenience yield, capital assets, store of value assets, consumable assets, transferable assets, natural resource companies

The twenty-first century has seen a renaissance of sorts for *commodities*, due to their high returns and subsequent increased demand from institutional investors such as pension funds and traditional portfolio managers. Compared to foreign exchange or equity markets, there is almost no way for central banks to intervene in commodity markets. Because the production side reacts very sluggishly to market distortions, short-term supply and demand shocks are compensated for only by price movements. These inherent asset class volatilities are the main reason many investors have historically refrained from investing in commodities, despite the valuable diversification benefits they can add to traditional security portfolios because of their low correlations with bonds and stocks.

In this chapter, we discuss the fundamentals of commodity investments by describing the market participants, the commodity subsectors, the commodity exchanges, and the different kinds of commodity investment vehicles available to investors. We subsequently investigate the risk and return characteristics of commodity futures using commodities futures indices. We then provide an empirical analysis of portfolio allocation of traditional security portfolios, explicitly including commodity futures.

MARKET PARTICIPANTS

Futures market participants are classified into hedgers, speculators (traders), and arbitrageurs. Commodity producers pass on the price risk that results from highly volatile and difficult to forecast commodity futures markets to speculators, and therefore pay a premium. Commodity producers have a distinct interest in hedging the price of their product in advance (a short hedge).

For example, consider the situation in the classic agricultural market. Farmers face a weather-dependent, volatile supply that is met by a relatively stable demand. Contrary to the maintenance cost for cattle breeding or the purchase cost of seed, the selling price is generally known only upon completion.

We see the opposite in the manufacturing industry: As the manufacturing industry hedges increasing commodity prices (a long hedge), the contrarian position to the commodity producers' short positions is taken. Airline companies, for example, often appear as long hedgers to guard against increasing fuel prices, the underlying in which the airline companies are short. If an existing or expected cash position is compensated for via an opposite future, the market participant is classified as a *hedger*. Hence, for the commodity producer, there is a fixed net profit; for the commodity manufacturer, there is a fixed purchase price.

Speculators represent the largest group in the futures markets. Their main task is to provide liquidity on the one hand, while balancing the long and short hedges on the other hand. Contrary to the commodity producers or the manufacturing industry, which try to avoid susceptibility to unfavorable price developments, the intention of speculators is to take a distinct market position and speculate for a price change. To make a profit, speculators deliberately take on risk by betting on rising or falling prices. As opposed to hedging, speculation is subject to both huge gains and huge losses, since speculators do not hold compensating cash positions.

The third and smallest group of market participants are the *arbitrageurs*, who try to take advantage of time- or location-based price differences in commodity futures markets, or between spot and futures markets, in order to generate riskless profits. Clearly, this group also intends to make profits, but their trading activity does not involve taking risky positions. Moreover, they use economic and financial data to detect existing price differences with respect to time and location. If these price differences exceed interlocal or intertemporal transfer costs like shipping, interest rates, warehouse costs, or insurance costs at the spot market, riskless profits can be realized. Consequently, price differences among the markets are adjusted, price relationships among the markets are restored, and arbitrageurs guarantee market balancing.

In the case of cash and carry arbitrage, the resale price of today's leveraged spot position is simultaneously set by selling the commodity futures. This short futures position implies an unconditional commitment to purchase the underlying at maturity. At maturity of the futures, the specified commodities are tendered against the maturing short futures. If the profit from the spot trade of the physical commodity exceeds the value of the futures plus the cost of debt financing, the arbitrageur will realize a profit from what is known as a *basis trade*.

COMMODITY SECTORS

Investments in international commodity markets differ greatly from other investments in several important ways. First, commodities are real assets—primarily consumption and not investment goods. They have an intrinsic value, and provide utility by use in industrial manufacturing or in consumption. Furthermore, supply is limited because in any given period, commodities have only a limited availability. For example, renewable commodities like grains can be produced virtually without limitation. However, their yearly harvest is strictly limited. In addition, the supply of certain commodities shows a strong seasonal component. While metals can be mined almost all year, agricultural commodities like soybeans depend on the harvesting cycle.

Another important aspect of commodities as an asset class is heterogeneity. The quality of commodities is not standardized; every commodity has its own specific properties. A common way to classify them is to distinguish between soft and hard commodities. *Hard commodities* are products from the energy, precious metals, and industrial metals sectors. *Soft commodities* are usually weather-dependent, perishable commodities from agricultural sector serving for consumptional purposes, such as grains, soybeans, or livestock such as cattle or hogs. Figure 57.1 shows the classification of commodity sectors.

Storability and availability (or renewability) are also important features of commodities. Since storability plays a decisive role in pricing, we distinguish between storable and nonstorable commodities. A commodity is said to have a high degree of storability if it is nonperishable and the costs of storage remain low with respect to its total value. Industrial metals such as aluminum or copper are prime examples: They fulfill both criteria to a high degree. In contrast, livestock is storable to only a limited degree, as it must be continuously fed and housed at current costs, and is only profitable in a specific phase of its lifecycle.

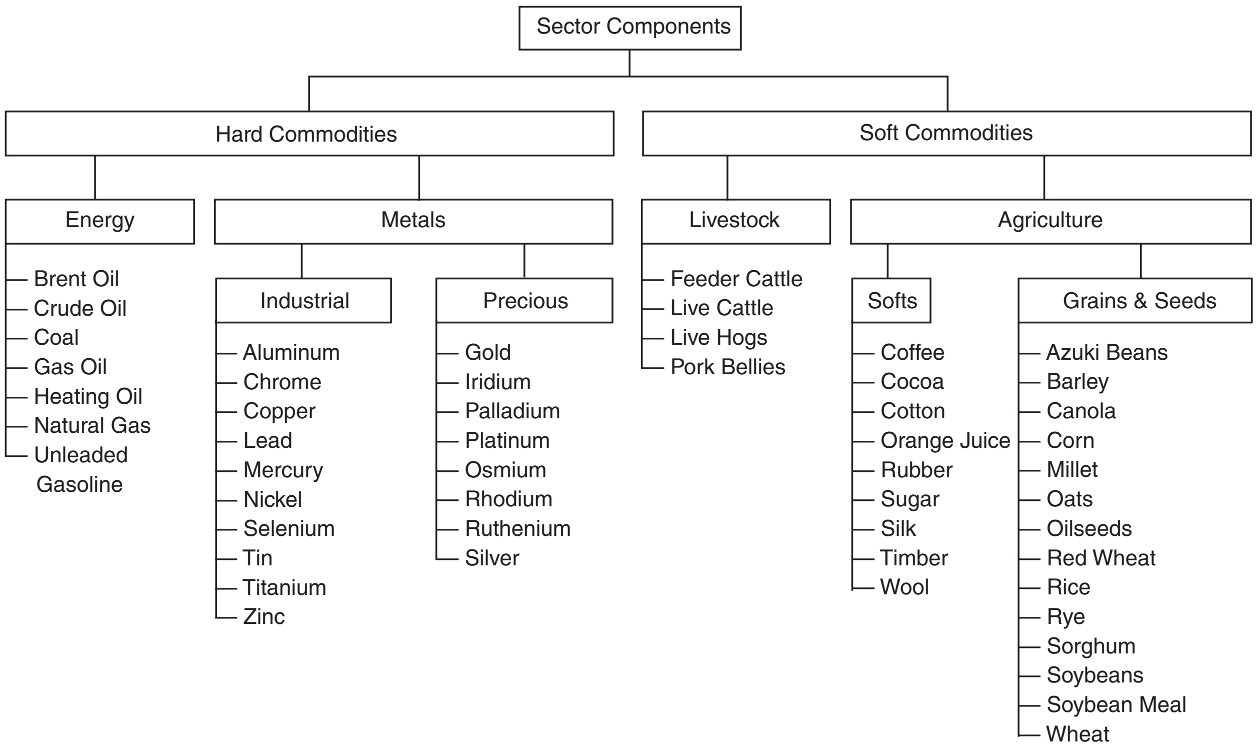


Figure 57.1 Classification of Commodity Sectors

Commodities such as silver, gold, crude oil, and aluminum are nonrenewable. The supply of nonrenewable commodities depends on the ability of producers to mine raw material in both sufficient quantity and quality.

The availability of commodity manufacturing capacities also influences supply. For some metals (excluding precious metals) and crude oil, the discovery and exploration of new reserves of raw materials is still an important issue. For given supply, the price of nonrenewable resources depends strongly on current investor demand, while the price of renewable resources depends more on estimated future production costs. (The events following Hurricane Katrina in 2005 clearly illustrated the insufficiency of the refinery capacities for crude oil and natural gas. Declining investment in this sector over the years has led to a bottleneck. The absence of investment in the industrial metals sector is also an issue for the supply side.)

The monetary benefit from holding a commodity physically instead of being long the respective futures is called the *convenience yield*. The convenience yield reflects market participants' expectations regarding a possible future scarcity of a short-term nonrenewable commodity.

COMMODITIES AS AN ASSET CLASS OF THEIR OWN

There is a broad consensus among academics and practitioners that commodities compared to other alternative assets can be considered—in a portfolio context—as an asset class of their own. (In reality, most alternative investments like hedge funds or private equity are not an asset class of their own, but are considered alternative investment strategies within an existing asset class.) By definition, an asset class consists of similar assets that show a homogeneous risk/return profile (a high internal correlation), and a heterogeneous risk/return profile toward other asset classes (a low external correlation). The key properties are common value drivers, and not necessarily common price patterns. This is based on the idea that a separate asset class contains a unique risk premium that cannot be replicated by combining other asset classes (see Scherer, 2005). Furthermore, it is generally required that the long-term returns and liquidity from an asset class are significant to justify an allocation.

To describe existing asset classes, Greer (1997) explains the decomposition into so-called super classes: *capital assets*, *store of value assets*, and *consumable or transferable assets*. Continuous performance is a characteristic of capital assets. Equity capital like stocks provides a continuous stream of dividend payments, while fixed income guarantees regular interest payments in the absence of the default of the obligor.

Common to all capital assets is that their valuation follows the net present value method by discounting expected future cash flows. In contrast, real estate as an asset class has a hybrid classification. On the one hand, real estate can be classified as a capital asset because it promises a continuous rental stream and has market value. On the other hand, some features of real estate assets can justify their classification as store of value assets (for example, if the real estate is used for the owner's own purpose). Such store of value assets cannot be consumed, nor do they generate income; classic examples are foreign exchange, art, and antiquities.

Commodities belong to the third super class—consumable or transferable (C/T) assets. In contrast to stocks and bonds, physical commodities like energy, grains, or livestock, do not generate continuous cash flows, but rather have an economic value. Grains, for example, can be consumed or used as input goods; crude oil is manufactured into a variety of products. This difference is what makes commodities a unique asset class.

Hence, it is obvious that commodity prices cannot be determined by discounting future cash flows. Thus, interest rates have only a minor influence on the value of commodities. Moreover, commodity prices are the result of the interaction between supply and demand on specific markets (see Scott, 1994). In this context, it is not surprising that the capital asset pricing model (CAPM) cannot adequately explain commodity futures returns. As we have noted, commodities are not capital assets (see Erb and Harvey, 2006; Bodie and Rosanksy, 1980).

The line between the super classes is blurred in the case of gold. On the one hand, gold as a commodity is used in such things as electrical circuitry because of its excellent conductivity. On the other hand, gold as a store of value asset is a precious metal and is used for investment, similarly to currencies. The rising demand of commodities since the stock market downturn in 2002 clearly demonstrates this characteristic. Because gold can be leased, Anson has even classified it as a capital asset. (Precious metals like gold or platinum can generate a lucrative stream of income by being leased at market leasing rates. See Anson [2006].)

Another specific criterion that differentiates commodities from capital assets is that commodities are denominated worldwide in U.S. dollars. Furthermore, the value of a specific commodity is determined through global rather than regional supply and demand. In comparison, equity markets reflect the respective economic development within a country or a region.

Prospects for Commodity Market Participation

In general, there are several ways to participate in commodity markets via a number of different kinds of financial instruments. The most important are (1) direct investment in the physical good, (2) indirect investment in stocks of natural resource companies, (3) commodity mutual funds, (4) an investment in commodity futures, and/or (5) an investment in structured products on commodity futures indices.

Buying the Physical Good

First, it seems obvious to invest directly in commodities by purchasing the physical goods at the spot market. However, the immediate or within two days delivery is frequently not practical for investors. According to Geman (2005), precious metals like gold, silver, or platinum are an exception, as they do not have high current costs and do not require storage capacity. However, a portfolio consisting solely of precious metals would not be a sufficiently diversified portfolio for investors to hold.

Commodity Stocks

An investment in commodity stocks (*natural resource companies*), which generate a majority of their profits by buying and selling physical commodities, may conceivably be considered an alternative investment strategy. In general, the term "commodity stock" cannot be clearly differentiated. It consists of listed companies that are related to commodities (that is, those that explore, mine, refine, manufacture, trade, or supply commodities to other companies). Such an indirect investment in commodities (e.g., the purchase of petrochemical stocks) is only an insufficient substitute for a direct investment. By investing in such stocks, investors do not receive direct exposure to commodities because listed natural resource companies all have their own characteristics and inherent risks, and take action in order not to be exposed too strongly to their commodity product by hedging appropriately.

Georgiev (2005) shows that these sector-specific stocks are only slightly correlated with commodity prices, and hence prices of commodity stocks do not completely reflect the performance of the underlying market. This is because stocks reflect other price-relevant factors such as the strategic position of the company, management quality, capital structure (the debt/equity ratio), the expectations and ratings of company and profit growth, risk sensitivity, as well as information transparency and information credibility. (For example, consider the poor information policy of Shell in the matter of the Brent Spar oil platform in 1995, which led to a massive stock price decline.)

Stock markets also show quick and more sensible reactions to expected developments that can impact company value. Hence, other causes of independent price discovery exist that differ from a pure commodity investment. Moreover, there may be temporary market disequilibriums, especially for stocks with low free float where few buy and sell transactions can already cause major price reactions. Finally, natural resource companies are subject to operational risk caused by human or technical failure, internal regulations, or external events. This means that when investing in a company listed on the stock exchange, both the associated market risk as well as any idiosyncratic risk must be considered carefully. Also note that the majority of large oil and energy companies hedge the risk associated with buying and selling oil products in order to smooth yearly profits.

However, the risk of commodity stocks is not completely reflected in the price volatility. First, particularly in the energy and metal sectors, there is the paradox that companies threaten their own business fundamentals by extracting exhaustible resources. On the one hand, long-term decreasing total reserves mean rising prices and a positive prospective for investors and commodity producers. On the other hand, commodity producers will suffer when resources are depleted.

Second, there is always the risk of a total loss if prices decrease below total production costs and the extraction

Table 57.1 Commodity Futures Indices

	Reuters/Jefferies Commodity Research Bureau (RJ/CRB)	Goldman Sachs Commodity Index (GSCI)	Dow Jones/AIG Commodity Index (DJ-AIGCI)
Introduced in	2005	1991	1998
Historical data available since	1982	1970	1991
# of commodities	19	24	19
Weighting scheme	Within a graduated system of four groups, based on liquidity and economic relevance	Rolling 5-year average of world production	Liquidity data, in conjunction with dollar-weighted production from the past 5 years
Rebalancing frequency	Monthly	Yearly	Yearly
Allocation restrictions	None	None	33% Maximum per sector; 2% market minimum per commodity
Relevant futures price on which the index calculation is based	Next futures contract/delivery month	Next month with sufficient liquidity	Next futures contract/delivery month
Roll period	4 Days	5 Days	5 Days
Calculation method	Arithmetic	Arithmetic	Arithmetic

of a commodity is stopped. By constructing an index consisting of commodity stocks, Gorton and Rouwenhorst (2006) show empirically that observed return correlations with commodity futures are even lower than those with the S&P 500. Furthermore, the commodity stock index exhibits lower historical returns than a direct commodity investment. For example, the returns of European oil companies covary strongly with EuroStoxx, but less with oil price returns. Exceptions are gold and silver stocks, whose beta to the domestic stock index is smaller than the beta to the gold and silver price.

Commodity Funds

In contrast to an investment in commodity stocks, one can actively invest in commodity funds, realizing an adequate diversification benefit with moderate transaction costs. Commodity funds differ in terms of management style, allocation strategy, geographic and temporal investment horizon in the denominated currency, and investment behavior. It is also important for investors to distinguish between active and passive funds (that is, index tracking funds). Commodity stock indices (e.g., the MSCI World Materials, the FTSE World Mining, the HSBC Global Mining, the Morgan Stanley Commodity Related Index, the FTSE World Oil, and Gas or the FTSE Goldmines) and commodity futures indices can be used to benchmark actively managed commodity funds. Commodity trading advisors (CTAs) also present an alternative to actively managed investment products. Today, there are also about 450 hedge funds with energy- and commodity-related trading strategies.

Commodity Futures Indices

Nowadays, investors can choose from an increasing number of investible commodity futures indices as a passive form of investing in commodities (see Table 57.1). Commodities have an exceptional position among alternative investments because they provide investible indices for a broad universe of commodity sectors. According to Doyle, Hill, and Jack (2007), between US $55 billion and $60 billion were invested in the Goldman Sachs Commodity Index (GSCI) in March 2007, and another US $15 billion was linked to the Dow Jones-AIG Commodity Index. Estimates for December 2006 state that about US $90 billion of invested capital from pension and mutual funds are invested in commodity-based indices or products. (In 2001, the total invested capital in the GSCI was between $4 billion and $5 billion. At the beginning of 2007, Standard & Poor's acquired the GSCI Commodity Index, which was subsequently renamed the S&P GSCI Commodity Index.)

For the majority of investors, an index-oriented investment represents the most reasonable way to obtain exposure to commodities or an individual commodity sector. Such an investment can be done cost-effectively using the following two types of financial products:

- Exchange-traded funds (ETFs) on commodity indices, and
- Commodity index certificates closely tied to commodity indices.

Index funds have the advantage of being relatively easy to trade and reasonably priced. Another advantage of funds over certificates is the nonexisting credit risk of the issuer. Because ETFs represent special assets, investor deposits are safe even if the investment company goes bankrupt.

Certificates constitute legal obligations that can be quickly and fairly cheaply issued by banks. In the case of commodity index certificates, the issuing institution invests in futures markets and rolls the futures contracts for a fee. The term of a certificate is normally restricted to a fixed date (e.g., rainbow certificates, whose underlyings are different subindices or asset classes, or discount and bonus certificates). But there are also open-end certificates.

However, because the indices, like the commodities themselves, are denominated in U.S. dollars, investors are exposed to currency risk. Quanto certificates, discount certificates with a currency hedge, can be used to mitigate this risk.

The main disadvantage of index certificates is that they often use excess return indices as the underlying instrument. These indices do not consider all the return components, in contrast to total return indices, which may lead to lower returns during periods of high interest rates. Investing in a low performance excess return index compared to a total return index can nevertheless be an advantage because the latter bears little or no initial costs and no yearly management fees. Hence, for investors with short-term investment horizons, certificates on excess return indices with lower returns can be a smart choice during periods of low interest rates.

Another disadvantage of index-based commodity investments is that due to their construction, they can only consider short-term futures contracts. Commodity funds not linked to commodity indices, however, can freely determine their optimal term by investing directly in commodity futures contracts. And similarly to purchasing rainbow certificates on different asset classes, there is also the possibility of purchasing commodity funds that do not invest exclusively in commodity indices, but also include commodity stocks to a certain extent.

Commodity Futures

In addition to options and other derivatives, commodity products are based primarily on futures contracts. A futures contract is a mutual binding agreement between two parties to deliver or accept and pay (or undertake a cash settlement): (1) a qualitative explicitly determined underlying (in this case commodities), (2) in a certain quantity, (3) at a fixed date, and (4) at a fixed, already at conclusion of the contract determined price. Futures can be described as mutually binding, exchange-traded "unconditional" forward contracts, since the conclusion of a futures contract leads to a legally binding accomplishment in the future if there is no compensating contrary transaction. (In contrast, in the case of conditional forward contracts such as options, the option holder has no obligation to exercise his option right, and can thus abandon the option at maturity.)

Contract sizes in the commodity market are standardized. The smallest tradable unit represents a contract, and the smallest possible price change of a futures is called a tick. The value of the minimum price change is the U.S. dollar and cent–denominated tick, multiplied by the contract size (also known as the point value) of the commodity. It is common practice to deposit a margin for every futures contract. The amount is determined by the exchange, but it is usually between 2% and 10% of the contract. (However, futures commission merchants may charge higher margins than the exchanges.) However, the margin changes according to the price and volatility of the contract.

In this context, we also distinguish between the initial margin, the minimum deposit required to invest in a futures contract, and the maintenance margin, the minimum deposit required to be on account at the exchange as long as the futures position is held. If the capital deposit on the account falls to or below the value of the maintenance margin due to price variations, the broker issues a margin call to recoup the initial value of the clients' capital. If an investor does not want to increase the margin, he can also close part of or the entire position, and accept a loss. For collateral in terms of the initial margin, investors in futures receive interest income from money market interest.

Generally, for commodity futures, there are two forms of settlement: delivery of the commodity at maturity, which happens in about 2% of the cases, and closing the futures position (that is, buying or selling the same amount of contracts before maturity). Daily price limits are a specific characteristic of commodity futures markets. As noted by Edwards and Neftci (1998), they were established to allow the market to stabilize during times of extreme movements (e.g., a cooling-off phase). Hence, daily price limits, again determined by the exchange, represent the maximum possible increase or decrease of a commodity price from the settlement price of the preceding trading day. In the case of limit up (limit down), the sellers (buyers) are outnumbered by buyers (sellers) who are willing to buy (sell) at the upper (lower) price limit. At this price limit, there may still be trading activity, but it may not exceed (limit up) or fall short of (limit down) the price limit.

The following are the contract specifications published regularly by the futures exchanges:

- **The type and quality of the futures underlying**: The type of commodity, abbreviation, and futures exchange.
- **The contract size**: The amount and units of the underlying asset per futures contract.
- **Price determination**: The formal notation of futures prices at the futures exchange.
- **Trading hours**.
- **The tick**: the minimum permissible price fluctuation.
- **The currency** in which the futures contract is quoted.
- **The daily price limit**.
- **The last trading date**.
- **Delivery regulations** (e.g., delivery month, type of settlement).

Investors in commodity futures can profit from price movements of the underlying commodity without having to fulfill the logistical or storage requirements connected with a direct purchase. However, this is possible only if the position is closed before maturity. The advantages of futures investments lie especially in the tremendous flexibility and leveraged nature of the futures position due to the low capital requirements. Thus, a shift of an existing futures position is possible at any time, even in the short term. By holding long or short positions, investors can profit from rising and falling markets. Furthermore, the futures markets are characterized by a high degree of liquidity and low transaction costs.

Despite the numerous advantages of an active investment in commodity futures, it is not always advisable for a private investor to take futures positions in such volatile commodities. Even if diversification by a large number of different futures contracts were guaranteed, the investor would still face the problem of maintaining an exposure to commodity prices without the liability of physical delivery of the underlying contract. This requires continuously closing existing futures positions and reestablishing new positions by opening more futures contracts. This is referred to as rolling of futures contracts, and it may be quite costly depending on the forward curve of the futures market. An active, indirect investment in commodities can be achieved by purchasing futures contracts and closing them prior to maturity. In order to keep an exposure to commodities, investors must buy another futures contract with a later maturity date (this is called "rolling" and must be repeated before each maturity date).

In addition, falling futures prices may constantly trigger margin calls (although margins can be withdrawn if the futures prices increase). Overall, however, compared to traditional assets, managing futures positions requires a great deal of time and effort. It is also possible to invest in commodity swaps and forwards. These instruments, however, are of minor liquidity since they are tailor-made for individual investors. Furthermore, these derivatives are not traded at the exchange, and commodity investment strategies of individual investors cannot be publicly observed.

COMMODITY EXCHANGES

The trading of commodity futures takes place at specialized exchanges that function as public marketplaces, where commodities are purchased and sold at a fixed price for a fixed delivery date. Commodity futures exchanges are mostly structured as membership associations, and operate for the benefit of their members. Transactions must be made as standardized futures contracts by a broker who is also a member of the exchange. Only members are allowed to trade. Membership in commodity exchanges is restricted to individuals who often act in the name of investment banks, brokers, or producers.The main task of a commodity exchange is to provide an organized marketplace with uniform rules and standardized contracts.

The first commodity exchange was founded by Japanese farmers trading rice futures contracts in Osaka. In the United States, the Chicago Board of Trade, founded in 1848, was the first institution. Even today, most commodities are still traded there. According to Geman (2005), in the United States most futures exchanges still function as open outcry trading systems, although many exchanges around the world operate on an electronic platform. The British London Metal Exchange was founded in 1877.

Energy futures trading, however, only began with the foundation of the International Petroleum Exchange (IPE) in London in 1980. Since 2005, the IPE operates under the name ICE Futures. Trading of WTI crude oil at the New York Mercantile Exchange (NYMEX) began in 1983; trading of Brent crude oil began in 1988. In terms of traded volume, the Chicago Mercantile Exchange (CME), founded in 1998, is the world's most important futures exchange. There are about 30 commodity exchanges worldwide; the most important are listed in Table 57.2. Based on traded volume, the majority of commodity futures trading takes place in the United States, the United Kingdom, Japan, and China.

RISK AND PERFORMANCE CHARACTERISTICS

Based on their historical return, risk, and correlation performance, commodity investments have an advantage over traditional assets, but exhibit some similarities to stocks. Kaplan and Lummer (1997), for example, conclude in their empirical investigation that commodities show an equity-like performance over the long run. This finding is also supported by many other studies such as Greer (2000), who concludes that the performance of unleveraged commodity indices from 1970 to 1999 was on average positive, and comparable to equities with regard to return and volatility.

Bodie and Rosansky (1980) analyze an equally weighted commodity futures portfolio between 1949 and 1976, and Gorton and Rouwenhorst (2006) between 1959 and 2004. Both studies confirm equity-like returns for commodities. In addition, during the high inflation period of the 1970s, commodities had the highest real returns by far of all asset classes. Gorton and Rouwenhorst also found differences with traditional assets. They show that commodity returns exhibit positive skewness, in contrast to stocks,

Table 57.2 Major Commodity Exchanges

Exchange Name	Abbreviation	Country	Traded Futures	Web site
Chicago Board of Trade	CBOT	U.S.	Agricultural products and oil	cbot.com
Chicago Mercantile Exchange	CME	U.S.	Agricultural products and livestock	cme.com
New York Mercantile Exchange	NYMEX	U.S.	Energy and metals	nymex.com
Intercontinental Exchange	ICE	GB	Energy	theice.com
London Metal Exchange	LME	GB	Metals	lme.co.uk
Winnipeg Commodity Exchange	WCE	Canada	Agricultural products	wce.ca
Tokyo Commodity Exchange	TOCOM	Japan	Energy and metals	tocom.or.jp
Shanghai Metal Exchange	SHME	China	Metals	shme.com
Dalian Commodity Exchange	DCE	China	Agricultural products and oil	dce.com.cn
Brazilian Mercantile and Futures Exchange	BM&F	Brazil	Agricultural products	bmf.com.br
Risk Management Exchange	RMX	Germany	Agricultural products and livestock	wtb-hannover.de
National Commodity and Derivatives Exchange	CDEX	India	Agricultural products and metals	ncdex.com

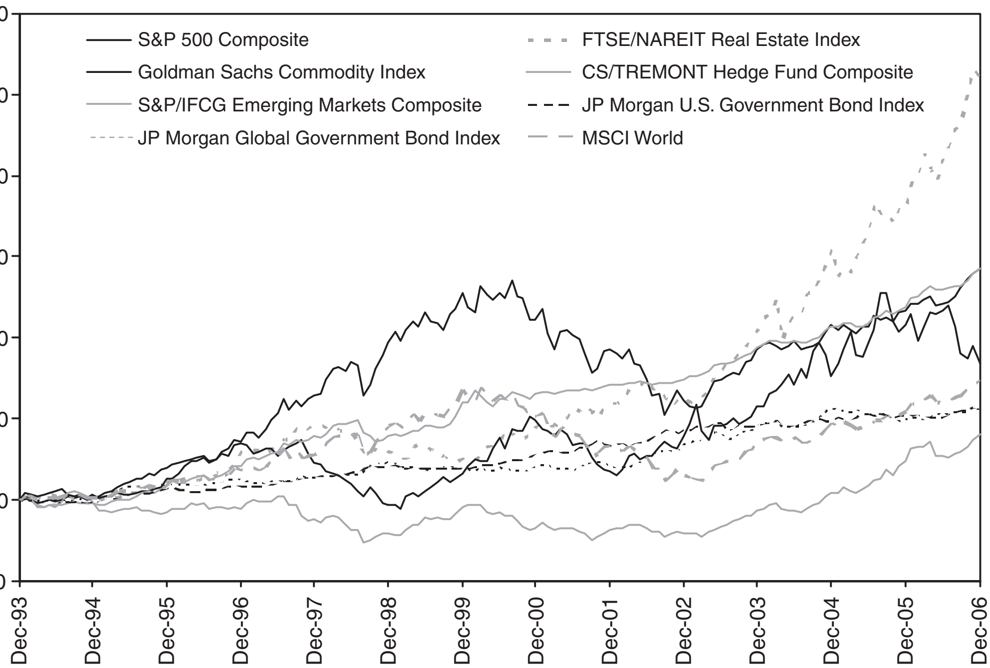


Figure 57.2 Performance of the Goldman Sachs Commodity Index Compared to Other Financial Assets

which have negative skewness and thus include higher downside risk.

Figure 57.2 shows the performance of both traditional and alternative assets starting with a reference basis of 100 in December 1993. After consolidating in 2006, the GSCI, which is heavily invested in energy, currently shows very strong performance, along with indirect real estate and hedge funds. In contrast, equity investments in emerging markets show the smallest price increases.

During the January 1994–December 2006 period, commodities had an annualized arithmetic return of 9.64% (and geometric annualized mean of 7.61%) with a volatility of 20.25% (see Table 57.3). The high variability can be explained by the GSCI's large share in energy. The energy sector currently represents over 70% of the total index (as at end of 2006), and is itself composed of 40% crude oil, which has experienced extreme volatility over the last few years. Thus, compared to other observed asset classes, commodities have a high average volatility. Note, however, that the downside risk of the S&P 500 Composite, the S&P/IFCG Emerging Markets, and the FTSE/NAREIT Real Estate Index are higher because of their negative skewness; commodities possess positive skewness.

The most beneficial investment in terms of the Sharpe ratio is the CS/Tremont Hedge Fund Index. However, hedge fund investors also face high excess kurtosis. When considering only return and volatility, an indirect investment in real estate also seems less favorable due to negative skewness and positive excess kurtosis. Furthermore, the poor performance of emerging market equities seen in Figure 57.2 is also confirmed by the descriptive statistics, especially considering the geometric mean as well as the exorbitant volatility.

As mentioned above, commodities serve an important diversification function in asset allocation due to their long-term low correlation with stocks, bonds, real estate, hedge funds, and, to a lesser extent, their absolute performance characteristics. According to Greer (2000), commodity indices have a negative correlation with stocks and bonds and a positive correlation with the inflation rate, especially unexpected changes in inflation. There are, however, significant differences among the individual commodity sectors: Energy, metals, livestock, and sugar show the best inflation hedging potential. Greer also finds very high correlation coefficients among different kinds of commodity sectors.

According to Kat and Oomen (2007), commodity futures and traditional assets like stocks and bonds are uncorrelated. In specific phases, the correlation admittedly increases—therefore not all commodities are useful for portfolio diversification in every market phase. However, even in down markets, commodities as a group do not lose their diversification potential. According to Anson (2006), there are three reasons for low or negative correlations between commodities and stocks/bonds. First, inflation has a positive effect on commodity prices, but a negative effect on equity and bond markets. Second, investor expectations in commodity markets are different from those in equity and bond markets. Finally, a trade-off between capital return and commodity return exists in industrial production.

Table 57.4 shows the return correlation structure between the total return indices of various asset classes. As can be seen, correlation is significant only at the 5% level between commodities and hedge funds, which turn out to be relatively low at 0.167. This can be traced back to the commodity trading advisers (CTAs) and managed futures

Table 57.3 Annualized Average Monthly Return and Volatility (January 1994–December 2006)

	Average annual			Minimum Annualized	Maximum Annualized			
	Arithmetic Return	Geometric Return	Standard Deviation	Arithmetic Return	Arithmetic Return	Skewness	Excess Kurtosis	Sharpe Ratio
GSCI Composite	9.64%	7.61%	20.25%	−14.41%	16.88%	0.063	0.024	0.281
S&P 500 Composite	11.43%	10.41%	14.27%	−14.46%	9.78%	−0.622	0.838	0.524
MSCI World	7.91%	7.00%	13.43%	−13.45%	8.91%	−0.658	0.890	0.294
Emerging Markets	6.76%	4.58%	20.62%	−25.56%	12.37%	−0.765	1.877	0.136
Hedge Funds Comp.	10.71%	10.42%	7.66%	−7.55%	8.53%	0.099	2.465	0.882
Real Estate Index	14.99%	14.14%	13.04%	−14.58%	10.39%	−0.510	1.472	0.846
JPM U.S. Govt. Bonds	5.91%	5.80%	4.65%	−4.68%	3.71%	−0.509	1.084	0.421
JPM Global Bonds	5.98%	5.79%	6.23%	−4.30%	5.65%	0.320	0.336	0.325
T-Bill Rate	3.96%	3.95%	0.49%	0.07%	0.53%	−0.644	−1.049	—

funds included in the CS/Tremont Hedge Fund Composite Index.

However, the return correlation between the money market and the commodity market is negative. Hence, the results of several academic studies are confirmed for our sample period: Commodities show a high diversification potential in traditional and alternative security portfolios. Chong and Miffre (2006) also find that conditional correlations between commodity futures and the S&P 500 decrease during times of down markets, that is, exactly when market risk increases and diversification is strongly needed. The conditional correlations between commodities and fixed income, on the other hand, increase during times of increased bond volatility.

PORTFOLIO OPTIMIZATION WITH COMMODITIES

In this section, we analyze whether an allocation in commodities yields any diversification benefits for a portfolio consisting of U.S. and global stocks, fixed income, and a riskless asset represented by the Treasury bill rate (that is, whether the efficient frontier shifts into the upper left corner in the expected return–standard deviation diagram). According to Markowitz (1952), these efficient portfolios (efficient in the sense that no others exhibit a superior risk/return combination) are located on the borderline formed by the set of all portfolios between the minimum variance (MVP) and the maximum return portfolio (MaxEP).

Table 57.4 Correlation Matrix

	GSCI Commodity Index	S&P 500 Composite	MSCI World	S&P/IFCG Emerging Markets	CS/Tremont Hedge Fund Comp	FTSE/NAREIT Real Estate	J.P. Morgan U.S. Govt. Bonds	J.P. Morgan Global Govt. Bonds	U.S. Treasury Bill Rate
GSCI Commodity Index	1								
S&P 500 Composite	0.003	1							
MSCI World	0.068	0.937[b]	1						
S&P/IFCG Emerging Markets	0.136	0.643[b]	0.724[b]	1					
CS/Tremont Hedge Fund Comp.	0.167[a]	0.487[b]	0.493[b]	0.503[b]	1				
FTSE/NAREIT Real Estate	0.005	0.299[b]	0.314[b]	0.350[b]	0.223[b]	1			
J.P. Morgan U.S. Govt. Bonds	0.079	−0.098	−0.159[a]	−0.216[b]	0.098	0.032	1		
J.P. Morgan Global Govt. Bonds	0.156	−0.016	0.064	−0.069	−0.050	0.118	0.597[b]	1	
U.S. Treasury Bill Rate	−0.063	0.084	0.008	−0.180[a]	0.102	−0.066	0.105	−0.084	1

Note: [a] and [b] denote significance of the correlation coefficient at the 95% and 99% confidence levels, respectively.

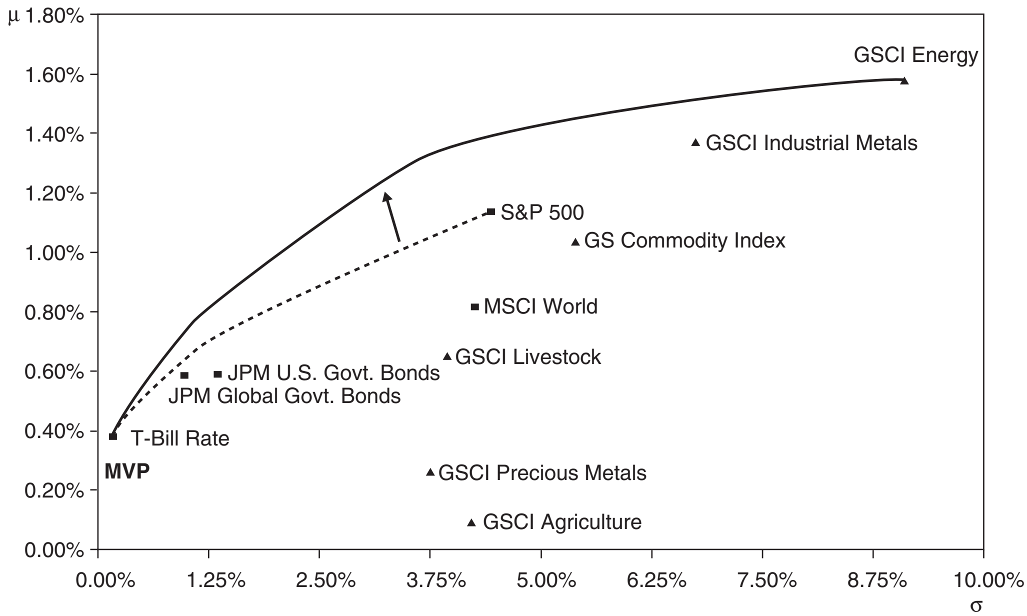


Figure 57.3 Expected Return-Standard Deviation (μ-σ) Portfolio Optimization (Monthly Returns in %)

Figure 57.3 shows how portfolio efficiency can be improved by including commodities in a traditional portfolio, thus rotating the efficient frontier counterclockwise around the MVP (the Treasury bill rate). The upward shift of the efficient frontier also provides higher risk-adjusted returns. The efficient frontier of the traditional portfolio is limited by a 98% investment in Treasury bills for the MVP, and 100% in the S&P 500 for the MaxEP.

Starting from the MVP and incorporating individual commodity sectors, the share of global bonds initially increases to 69% (see Figure 57.4). Subsequently, the proportions of the energy and industrial metals sectors increase continuously, together with the share of U.S. stocks. At a monthly return level of about 1%, livestock is represented with a share of about 4% to 5%. However, agricultural and precious metals are excluded entirely from the allocation. At a monthly return level of about 1.4%, the portfolio only consists of an allocation in the S&P 500 (28%), the energy sector (37%), and the industrial metals sector (35%).

Thus, with an increasing return level, the proportion of commodities in the portfolio expands as the allocation in U.S. stocks increases. It is remarkable that the GSCI

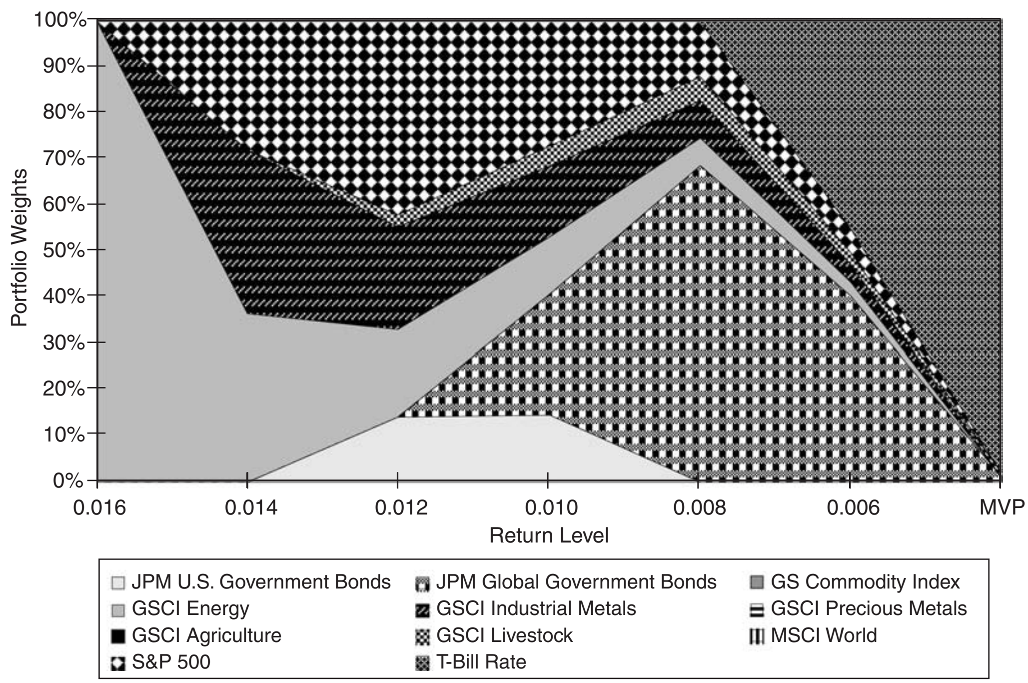


Figure 57.4 μ-σ-Portfolio Allocation (Monthly Returns in %)

Composite is not included in any allocations. It seems advisable to invest directly in the respective individual subsectors.

SUMMARY

An allocation to commodities offers not only a hedge against inflation, but also effective diversification because of their low correlation with traditional asset classes. Over the long run, commodity investments show equity-like returns, accompanied by lower volatility and lower shortfall risk. These advantages hold for passive investment in commodity futures indices, which are considered indicators of commodity market price movements. However, the futures indices of individual providers differ with regard to sector weights, index construction, and calculation method—hence there are tremendous variations in risk/return characteristics.

REFERENCES

Adams, Z., Füss, R., and Kaiser, D. G. (2008). Macroeconomic determinants of commodity futures returns. In F. J. Fabozzi, R. Füss, and D. G. Kaiser (eds.), *The Handbook of Commodity Investing* (pp. 87–112). Hoboken, NJ: John Wiley & Sons.

Akey, R. P. (2005). Commodities: A case for active management. *Journal of Alternative Investments* 8, 2: 8–30.

Anson, M. J. P. (2006). *The Handbook of Alternative Assets*, 2nd edition. Hoboken, NJ: John Wiley & Sons.

Bodie, Z., and Rosansky, V. I. (1980). Risk and return in commodity futures. *Financial Analysts Journal* 35, 3: 27–39.

Chong, J., and Miffre, J. (2006). Conditional risk premia and correlations in commodity futures markets. Working paper.

Doyle, E., Hill, J., and Jack, I. (2007). Growth in commodity investment: Risks and challenges for commodity. Financial Services Authority, working paper.

Edwards, F. R., and Neftci, S. (1998). Extreme price movements and margin levels in futures markets. *Journal of Futures Markets* 8: 639–655.

Erb, C., and Harvey, C. R. (2006). The tactical and strategic value of commodity futures. *Financial Analysts Journal* 62, 2: 69–97.

Füss, R., Hoppe, C., and Kaiser, D. G. (2008). Review of commodity futures performance benchmarks. In F. J. Fabozzi, R. Füss, and D. G. Kaiser (eds.), *The Handbook of Commodity Investing* (pp. 169–202). Hoboken, NJ: John Wiley & Sons.

Geman, H. (2005). *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy*. Chichester: John Wiley & Sons.

Georgiev, G. (2005). Benefits of commodity investment. Working paper.

Gorton, G., and Rouwenhorst, K. G. (2006). Facts and fantasies about commodity futures. *Financial Analysts Journal* 62, 2: 47–68.

Greer, R. J. (1997). What is an asset class, anyway? *Journal of Portfolio Management* 23, 2: 86–91.

Greer, R. J. (2000). The nature of commodity index returns. *Journal of Alternative Investments*, Summer: 45–52.

Kaplan, P. D., and Lummer, S. L. (1997). GSCI collateralized futures as a hedging and diversification tool for institutional portfolios: An update. Working paper.

Kat, H. M., and Oomen, R. C. (2007). What every investor should know about commodities, Part II: Multivariate return analysis. *Journal of Investment Management*, Third Quarter: 1–25.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance* 7, 1: 77–91.

Scott, J. H. (1994). Managing asset classes. *Financial Analysts Journal*, January/February: 62–69.

Scherer, B. (2005). Commodities as an asset class: Testing for mean variance spanning under arbitrary constraints. *Deutsche Bank—An Investors' Guide to Commodities* (pp. 35–42). New York.

Till, H. (2001). Taking full advantage of the statistical properties of commodity investments. *Journal of Alternative Investments* 4, 1: 63–66.

CHAPTER 58

Art Finance

RACHEL A. J. CAMPBELL, PhD
Assistant Professor of Finance, Rotterdam School of Management, Erasmus University & Maastricht University, The Netherlands

Abstract: Art finance is a rapidly developing area of international finance. Although direct investment in art is not new, structured solutions to investing in art indirectly mean that art is being considered an alternative asset class. A number of art funds have recently been launched, and a number of smaller boutique funds actively invest and trade in art, purely for financial gain. With the increasing amount of money pouring into the art market, banks are becoming increasingly interested in using art as collateral. Understanding how art is priced, as well as the risk and return characteristics of art is fundamental to portfolio management in all areas of art finance, when considering art as an alternative asset and the development of art banking and art finance.

Keywords: art banking, alternative investment, art funds, art indices, art pricing, risk and return

Artworks have been trading in auction markets and private dealer markets for centuries, so the concept of trading art is not a new phenomenon. Trading portfolios of artworks for pure financial gain and using art as an asset with an underlying value from which to obtain finance is a much more recent phenomenon.

Typically, turnover in the market is dominated by fine art paintings, making up 75% of the number of art sales; drawings and watercolors and sculptures represent a further 10% each, with prints and photography taking the smaller 5% final share of the market. (See *Artprice* for an annual report on these figures.) We focus in this chapter on the market for fine art, and exclude the market for decorative art. The majority of the world's art is traded in New York, attracting 50% of global art auction sales, with London taking 25%. These two trading hubs dominate the global auction market, with 75% of the global market. France and China take a smaller 5% share each, with the other 20% split between other smaller markets. The two major auction houses are Christie's and Sotheby's, who dominate the market, commanding the lion's share of auction sales, with each house taking almost 40% of all auction sales. Figures on dealer markets are illusive, with prices attained by private dealers difficult to obtain. It is thought that the dealer market represents between 40% and 60% of the total art market. With growing economic prosperity in China and India, the major auction houses have opened auction houses at the local level, with trade in indigenous art currently booming. Both repatriation and a growing interest by Western collectors and investors for Asian art are creating enormous demand in this area.

The art market is changing dramatically, and growing rapidly with continued increase in the demand for contemporary art; other sectors of the art market show continued steady growth. (See Goodwin [2007] for an overview of the global art market.) Art finance is a relatively new area of international finance, which covers both the concept of art investment, as well as using art as an underlying asset for finance.

ART MARKET INDICES

The vogue for investing in fine art has received a boost from the availability of greater information on art prices. Demand is growing, resulting in record prices being reached, and buyers are able to trace previous prices of artworks through a number of data providers. Databases, indices, and market reports are now essential analytical tools with which art investors can assess financial performance.

The use of a variety of indices for various art markets enables us to get a good impression of the returns which fine artworks have made historically and the amount of risk associated with these art prices. This in turn is useful for analyzing art as an investment, for looking at the performance of *art funds,* and in giving an indication of art's performance in a diversified portfolio.

The Mei Moses index, Art Market Research, and art price indices are the three most widely quoted indicators of art market performance. All are reliant on data from sales obtained from the two main auction houses. As mentioned earlier in the chapter, auction results alone provide an incomplete picture of the market performance because they are only a portion of the whole market. The dealer market is largely ignored due to this absence of obtainable data. Although there is some disagreement as to the percentage of the market that dealers comprise, it cannot be denied that dealers have a significant, albeit unquantifiable, impact on the art market. The absence of dealers' transactions from the *art indices* may have a bearing on the rate of return indicated by the indices. This is due to the fact that dealers may buy at lower prices but sell at higher prices, thereby reducing the art investors' rate of return.

Historically, moderate returns have been made financially from investing in art. The return made by art can be split into a financial return and a nonfinancial return, which comes in the form of aesthetic value from holding the artwork. The presence of a nonfinancial return means that the return made by art, when looking at it from a purely financial perspective, tends to be only moderate, especially when compared with other alternative asset classes with comparable levels of risk. It could be argued that the aesthetic value earned by the holder of the artwork is not valued or included in the financial gain and thus not compensated for financially by the level of risk held.

There are four main methodologies for producing art price indices: geometric means, average prices, repeat sales regressions (RSR), and hedonic regressions. Chanel, Gerard-Varet, and Ginsburgh's (1996) study indicates that over long periods the respective methodologies are closely correlated. Issues regarding the various index pricing methodologies are extremely well highlighted in a recent paper by Ginsburgh, Mei, and Moses (2007), which specifically compares hedonic to repeat sales regression.

Hedonic valuation takes into account the characteristics of the artworks. An examination of the subject matter, size, medium, provenance, and condition of the artwork, as well as the artist's popularity, will all materially contribute to the financial value an artwork is given. While many of these are necessarily objective inquiries, it is ultimately the subjective opinion of the purchaser that will be determinative of the price paid. Therefore, unlike stocks and bonds, the price of an artwork comprises an unquantifiable element: taste. For a collector, taste will play an important role in determining whether an artwork is bought.

Ashenfelter and Graddy (2003) provide an excellent survey of average returns estimated from art price data, currently in the academic literature. We have extended the exhibit with a few additional studies; see Table 58.1, which provides estimates of the levels of *risk and return* over various periods.

These indices show that historically, average real returns for fine art are moderate. Returns are above inflation and tend to be greater than for government bonds, but less than for equities.

The survey of art pricing methodologies in Table 58.1 tends to indicate that the repeat sales methodology provides slightly higher estimates of average returns than the other methodologies for similar time periods. For example, Anderson (1974) provides RSR and hedonic price indices for the periods 1780–1970 and 1780–1960 and Chanel, Gerard-Varet, and Ginsburgh (1996) for the period 1855–1969. It is of interest to observe the long-run trend in the market, and to note that there have been periods in which art returns have been substantially higher than average.

There has been a general upward trend in art prices in the market. Figure 58.1 shows the performance of a $1,000 investment in the art market over the period 1976–2003. This is purely theoretical, since trading such an index is not presently possible.

We see that the repeat sales estimates provide a significantly greater estimate of average return over the period than the average prices from Art Market Research. Caution should be urged in using too high an estimate for past historical average returns in forecasting expected returns, due to this upward bias in using the repeat sale methodology. Also, the indices provided do not take account of transaction costs, which can represent a significant fraction.

With the collection and availability of art price data, the concept of art market finance is beginning to flourish. Many of the larger banks offer *art banking* services, and the opportunity to lend against art as collateral. With the proliferation of repeat sales indices, which give a good estimate of the financial return for various artworks over time, there is a booming interest in considering art as an *alternative investment.*

ART AS AN ALTERNATIVE ASSET CLASS

The continued search for alternative asset classes, which exhibit historically low correlations with the more traditional asset classes, render the art market as an attractive avenue to reaping benefits from portfolio diversification. Although notoriously illiquid at times, the art market appears to offer investors an alternative asset class that is only slightly correlated with most other asset classes. The highest correlation exhibited over a 30-year period was with commodities, still less than 10% (see Campbell, 2008). Others sources of risk—above all, liquidity risk—mean that it is difficult to assess the true risk/return tradeoff in this market; however, with the greater amount of data available on the prices at which artworks are sold,

Table 58.1 Estimated Fine Art Market Performance, Seventeenth through Twenty-first Century (as reported by various academic papers by period of study)

Author (Year Published)	Sample	Period	Method	Nominal Return	Real Return	Standard Deviation
Baumol (1986)	Paintings in general	1652–1961	RSR		0.60%	
Frey and Pommerehne (1989)	Paintings in general	1635–1949	RSR		1.40%	
		1653–1987	RSR		1.50%	5.00%
		1950–1987	RSR		1.70%	
Buelens and Ginsburgh (1992)	Paintings in general	1700–1961	Hedonic		0.91%	
	Paintings in general	1780–1970	RSR	3.70%	3.00%*	
Goetzmann (1993)	Paintings in general	1716–1986	RSR	3.20%	2.00%*	5.65%
		1850–1986	RSR	6.20%	3.80%	6.50%
		1900–1986	RSR	17.50%	13.3%	5.19%
Anderson (1974)	Paintings in general	1780–1960	Hedonic	3.30%	2.60%*	
		1780–1970	RSR	3.70%	3.00%*	
Chanel, Gerard-Varet, and Ginsburgh (1996)	Paintings in general	1855–1969	Hedonic		4.90%	
		1855–1969	RSR		5.00%	
Mei and Moses (2002)	American, impressionists and old masters	1875–1999	RSR		4.90%	4.28%
		1900–1986	RSR		5.20%	3.72%
		1900–1999	RSR		5.20%	3.55%
		1950–1999	RSR		8.20%	2.13%
		1977–1991	RSR		7.80%	2.11%
Goetzmann (1996)	Paintings in general	1907–1977	RSR		5.00%	
Fase (1996)	Nineteenth century	1946–1966		11.00%	7.50%	
		1972–1992		10.60%	1.10%	
Stein (1977)	Paintings in general	1946–1968	Geometric Mean	10.47%		
Barre, Docclo, and Ginsburgh (1996)	Great impressionist	1962–1991	Hedonic	12.0%	5.00%*	
	Other impressionist	1962–1991	Hedonic	8.00%	1.00%*	
Czujack (1997)	Picasso paintings	1966–1994	Hedonic		8.30%	
Deutschman (1991)	Old masters	1971–1991		12.30%	6.04%	
Angnello and Pierce (1996)	Nineteenth Century U.S.	1971–1992		9.30%	3.25%	
Campbell (2005)	Paintings in general	1976–2004	Average prices	5.73%	1.44%	8.27%
	U.S. paintings	1976–2004	Average prices	7.94%	3.66%	8.73%
Pesando (1993)	Modern prints	1977–1992	RSR		1.51%	19.94%
Pesando and Shum (1996)	Picasso prints	1977–1992	RSR	12.00%	2.10%	23.38%
Frey and Serna (1990)	Old masters	1981–1988	Hedonic	10.59%	3.20%	
Candela and Scorcu (1997)	Modern contemporary paintings	1983–1994		3.89%	0.21%	

*Real returns estimated additionally by Ashenfelter and Graddy (2003).

financial analysis and research into the art market is becoming widespread.

Although returns are only moderate, especially when including transaction costs, the benefits arising from investing in art is that the price of art appears to be only slightly correlated with other financial asset classes. This renders a small investment in art as beneficial addition to an investment portfolio because for the same amount of return less risk in encountered. That is to say, the low correlation is highly desirable from a diversification perspective, which reduces the overall volatility—and thus the risk of the investment portfolio. Even though art is highly volatile alone, if held in conjunction with stocks and bonds in a portfolio, then the investor is able to obtain higher expected returns than otherwise for a given level of risk.

Moreover, diversification benefits can be gained within the art market itself. Holding a broad portfolio of artworks from a variety of styles (old masters, impressionists, modern, and contemporary, to name a few), and across countries and artists, can reduce the overall risk of the art portfolio considerably. It is unclear just how many artworks are required to completely reduce the unsystematic risk from art prices, so that the art market index risk and return profile is met—indeed, it is likely to be more than a stock portfolio. However, with indices from various styles and countries available (see Art Market Research), correlation statistics between the various "art industries" offer attractive diversification benefits to the art investor.

Investment skill lies in interpreting the available information, assessing whether the risk/return ratio is acceptable, and in deciding whether the investment is appropriate to an existing portfolio. In the art market this is extremely subjective. Taste adds an additional, unquantifiable element to art investment even after market analysis has been undertaken. Considering art as a direct

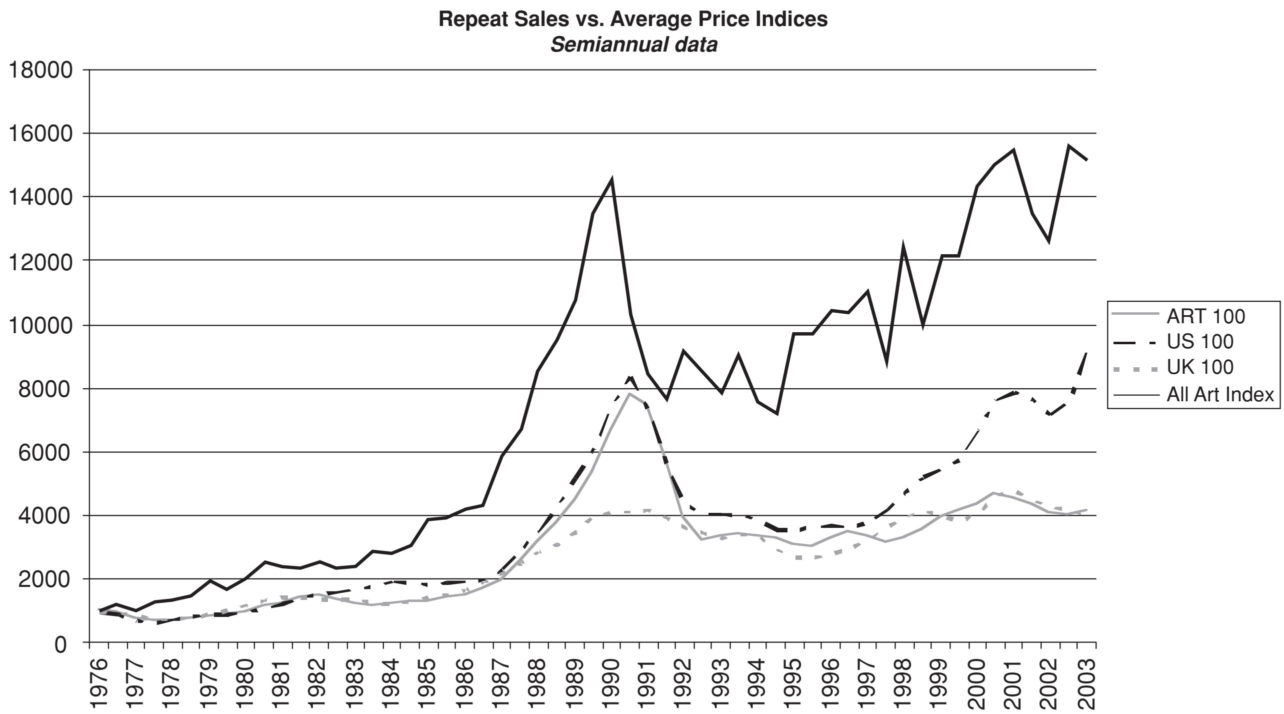


Figure 58.1 Art Price Indices, 1976–2003
Repeat sales All Art Index versus the average price indices from Art Market Research for the general art market (Art 100), a basket of U.S. artists (U.S. 100) and a basket of British artists (U.K. 100).

investment presents a risky investment opportunity, although purchasing according to personal taste means that the aesthetic benefit is also received, which could potentially outweigh any financial benefit or loss incurred.

Market anomalies and inefficiencies can lead to much higher realized returns. With almost no regulation, market efficiencies in the art market are abundant. Many possibilities exist to trade legally on "insider information," gain from pricing discrepancies among markets, and to a great extent make the market.

Market participants can behave in a rather "irrational" manner, with paintings being bought as a status symbol, to exemplify one's wealth. Behavior at auction has been observed, resulting in an overexuberance with participants tending to overpay for objects, with prices hitting the hammer above what buyers had thought they would have been prepared to pay. This aversion to lose in a bidding process, and other such emotive reasons, appear to have a large impact on the market than in other financial asset classes.

There are a number of standard market anomalies that have been observed in the market for art. See Ashenfelter and Graddy (2003) for a detailed overview of some of these irregularities, which have been documented in the academic literature. For example, the law of one price does not consistently hold, higher prices paid for burned artworks (artworks not having sold previously at auction and bought in), declining prices paid for similar objects. These behavioral aspects of market participants can lead to inefficiencies in the art market, and thus for the ability of less emotive art funds (who are dispassionate about the art and interested only in the financial return) to reap higher returns than a benchmark index. Moreover, the lack of liquidity in the market means that buyers with no liquidity constraints can pick up artworks at relatively undervalued prices simply by providing immediate liquidity. Often quoted in the art market are the forced sales arising from the three D's—debt, divorce, and death—which often require sales to be transacted at a faster pace, driving down the price received for such forced sales.

Art may also be attractive to investors seeking a diversified portfolio. First of all, a well-diversified portfolio of art needs to be held to reduce the risk of individual artists falling from fashion. Second, the low correlation between art and other financial assets means that art may form part of an optimal portfolio allocation, with investors holding a variety of assets such as equities, bonds, real estate, and art as well as cash. Estimating the correlation statistics among a variety of asset classes, we find art and equities are approximately 5% correlated over the past 30 years. The correlation between other asset classes is also low, the highest being between art and commodity futures, and even then only 10% correlated.

ART FUNDS

The irregularities of the market have led to the emergence of a number of art funds in recent years, whose proprietary strategy is in trading art for a speculative profit. The additional attraction is the low correlation that art prices appear to have with other financial asset classes, although the lack of liquidity during art market downturns must not be underestimated, and with fewer sales occurring during these periods, keeping correlation statistics down.

The first major fund to launch was the London-based Fine Art Fund in 2003. These types of funds act more like private equity funds, alternatively taking on strategies common to hedge funds, the newest of which is the Art Trading Fund, currently raising capital and run by a former hedge fund manager. Banks are also moving into the arena, with a failed attempt by ABN Amro for a fund-of-funds, and the recent launch of Societe General Alternative Asset Management with a fund out of Paris. With expert knowledge being a crucial factor in the success of these types of funds, only a handful of funds have been successful in attracting capital, with those having the dedicated resources and capabilities of being able to find and exploit inefficiencies in the market. This is due to the fact that dealers tend to buy at lower prices, and are able to sell at relatively higher prices, reducing the investor's rate of return even more. Current art funds total less than $100 million in capital, representing only a marginal fraction of the current trade in fine art and unlikely at this stage to have any major impact on controlling prices in the art market.

However, the large investors in the art market are influential in price movements. As we have seen, asymmetric information between investors and managers is greater in the art market than in the market for most other financial assets. In the art market, information is imperfect, with participants not necessarily as well informed about the quality, resale value, price, and availability of substitutes. Unlike other financial markets, private art dealers, art funds, and other "influential" investors are able to "make the market." Art funds, with their greater weight, can influence the demand for art by promoting particular artworks and artists. In this regard, the art market differs from traditional investment strategies: As pricing anomalies disappear rapidly in other financial markets, it appears that the art market is able to tolerate their existence for greater periods of time.

Certainly, these inefficiencies present opportunities for exploitation and profit, but conversely represent a danger for the uninformed investor. It is the extent to which these inefficiencies and anomalies exist in the art market that determines the positive abnormal returns that can be made by, but can also lead to loses by, the novice investor. Naturally, this position would be sustainable only in the short term. If more funds enter the marketplace, there will be less room for abnormal profits to be made, although the required skills and knowledge to be an art fund manager mean the entry level is high, with many promising funds having fallen at the first gate while trying to raise enough capital to launch. Until then, the inefficient nature of the market means that if artworks are chosen wisely, attractive returns may be made from alternative investment vehicles specializing in art and art mutual funds in the foreseeable future.

Developments in structured products around art have enabled art investment to become more mainstream and accessible to investors worldwide. Many funds have a high entry level to typically closed funds; however, there is a move toward making the funds more accessible to all investors with shares as low as €2500 for a share in Germany's ARTESTATE, an art fund specializing in German and U.S. contemporary art. Arguably, this would bring the notion of art investment to all sectors of society and, consequently, awaken a greater interest in the market and for up-and-coming artists. The tax breaks in the market amplify the attractiveness of art investment as an alternative investment.

It is not just art funds and banks who are offering structured products around artworks. The Artist Pension Fund provides a novel way of providing retirement provision for artists through a scheme of collective investment into a few chosen artworks over time. It is likely that this type of fund will grow in future, providing a secure retirement provision to an industry with otherwise highly disparate income streams.

SUMMARY

Art finance is a new area of international finance which is flourishing primarily to a booming art market and the prevalence of art price data. Art banking is a steadily growing area with many banks offering art banking services. Art is subsequently being considered an alternative asset class, with investment in art purely for financial gain. Although art is an asset class in its financial infancy, from an investment perspective, the market is developing at a rapid pace, and exciting opportunities exist for informed investors to make attractive returns. Despite empirical difficulties associated with pricing artworks, and estimating the true underlying amount of risk, particularly liquidity risk, for an estimated level of historical risk, investors and investment institutions are carefully looking at art as a prospective investment. A number of funds have securitized the purchase of art so that private individuals can own shares in a fund dedicated to fine art investment.

The broader appeal of art as an investment strategy lies in the low correlation with other asset classes. Mean-variance portfolio optimization also shows this using moderate returns for art, even after accounting for the high transaction costs prevalent in the art market.

With the number of art funds on offer steadily increasing it is possible that in the future, a market for a mutual fund of international art funds will develop for investors who seek a truly diversified investment into art, with potentially greater liquidity in the market through greater trading. It is unlikely that art investment will ever become mainstream. Also questionable is whether the art market could cope with such a large investment strategy into the art market, as would be common by institutional investors in other asset classes. What is likely is that a niche alternative investment market is created, with a limited number of funds that have a highly specialized knowledge and level of expertise. The profitable art funds are likely to be

hard to replicate, and market efficiencies are likely to stay abound. Indeed, successful art investment is likely to be an art in itself.

REFERENCES

Anderson, R. C. (1974). Paintings as an investment. *Economic Inquiry* 12, 1: 13–26.

Ashenfelter, O. (1989). How auctions work for wine and art. *Journal of Economic Perspectives* 3, 3: 23–36.

Ashenfelter, O., and Graddy, K. (2003). Auctions and the price of art. *Journal of Economic Literature* 41, 3: 763–788.

Baumol, W. J. (1986). Unnatural value: Or art investment as floating crap game. *American Economic Review* 76, 2: 10–14.

Campbell, R. (2008). Art as a financial investment. *Journal of Alternative Investments* 10 (Spring): 64–81.

Chanel, O., Gerard-Varet, L. A., and Ginsburgh, V. (1996). The relevance of hedonic price indices. *Journal of Cultural Economics* 20, 1: 1–24.

Frey, B. R., and Eichenberger, R. (1995). On the return of art investment return analyses. *Journal of Cultural Economics* 19: 207–220.

Genesove, D., and Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics* 116, 4: 1233–1260.

Ginsburgh, V., Mei, J., and Moses, M. (2007). On the computation of price indices. In V. Ginsburgh and D. Throsby (eds.), *Handbook of Economics Art and Culture* (pp. 948–979), Amsterdam: Elsevier.

Goodwin, J. D. (2007). *A Guide to International Art Markets—Collecting and Investing*. London: Kogan Page.

Grampp, W. (1989). *Pricing the Priceless—Art, Artists and Economics*. New York: Basic Books.

Kochugovindan, S. (2005). *Barclays Capital Equity Gilt Study*. London: Barclays Bank.

Mei, J., and Moses, M. (2002). Art as an investment and the underperformance of masterpieces. *American Economic Review* 92, 1: 1656–1668.

Mei, J., and Moses, M. (2005). Vested interest and biased price estimates: Evidence from an auction market. *Journal of Finance* 60, 5: 2409–2435.

Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance* 53, 5: 1775–1798.

Stein, J. P. (1977). The monetary appreciation of paintings. *Journal of Political Economy* 85, 5: 1021–1035.

CHAPTER 59

Investing in Life Settlements

ANTHONY F. L. PECORE
Research Analyst, Franklin Templeton Investments

Abstract: Life insurance is an asset and can be sold into a willing market. For years, life insurance companies served as the only market, paying cash surrender value. In recent years, the life-settlement market has developed to give competitive market pricing, rather than cash surrender value, to policyholders for their unneeded insurance. Development of a risk framework for life settlements requires an understanding of the mortality risk and expected cash flows.

Keywords: life settlement, life expectancy (LE), mortality risk

Life settlements are a new asset class for the investor. We believe life settlements possess investment characteristics that will be very attractive to those in search of investment-grade credit quality, long duration, assets priced at attractive spreads, with low correlation to other capital markets. The purpose of the chapter is to introduce the life settlement cash flow to asset managers, survey the driving valuation factors, and provide a first-order analytical framework for risk management.

INVESTMENT CHARACTERISTICS

After a policy owner has made the economic decision that their life insurance policy is no longer needed, they may seek to sell the policy or lapse the policy and accept the cash surrender value, if any, from the issuing insurance company. If the policy owner sells the policy to an investor, the investor will pay all future premiums and receive the policy's benefit upon the demise of the insured. Consider a policy with a \$100 benefit, with premiums of \$3 per year and a cash surrender value of \$0. In this example, the investor would pay \$25, \$25 more than the insurance company would offer, for the policy and would have to pay the \$3 per year premium to keep the policy in force. From the investor's perspective, both the acquisition cost and premiums are negative cash flows. If the insured matured 10 years after the sale, the investor would receive \$100. This cash flow would have an internal rate of return (IRR) of roughly 7.5%. Of course, we don't know in advance how long the insured will live, but we can make assumptions based on actuarial and medical data. This cash flow uncertainty is not unusual for participants in the bond market. Many bond market investments contain cash flows where there is uncertainty in either the timing or value of a given cash flow. For example, mortgage-backed securities (MBSs) and credit default swaps (CDSs) possess these cash flow uncertainties, yet are large and mature markets.

In this section, we describe the investment characteristics of life settlement. They include:

- Attractive IRRs and longer duration
- Low correlation to other asset classes
- High credit quality

Higher Expected IRR/Longer Duration

The numbers used in the deterministic cash flow example were typical of those in the life settlement market place: policies are being purchased from individuals with life expectancies (LEs) typically greater than eight years. In its simplest form, there is no coupon payment in the life-settlement transaction; rather, there is capital appreciation supported by regular premium payments. The large benefit payment that is made at the insured's maturity provides for the life settlement's long duration. The discount at which policies can be purchased provides for the higher expected IRRs, which should compensate the investor for the asset's lack of liquidity and regulatory risk.

Low Correlation to Other Asset Classes

Life settlement cash flows consist of an outflow of premium payments followed by a benefit inflow. In the simplest form, both the premium payments and the benefit are known quantities. What is not specifically known are the timing of the benefit payout and the duration of premium payments. The timing of the cash flows is dependent on the remaining life span of the insured. Thus, the ultimate return of the life settlement is not correlated to economic cycles, but rather to the *life expectancy* of the insured.

High Credit Quality

The fundamental credit risk to the investor in life settlements is that the issuing life insurance company fails to pay the benefit. Three points help to protect the investor:

1. Through guaranty funds many states protect up to $300,000 in benefits if a carrier is in default when policy claim is filed.
2. Insurance claims sit at the top of an insurance company's capital structure, and must be paid before any form of loan.
3. Life settlements can be selected from only investment-grade companies.

Purchasing policies from lower-rated insurers is possible, and the investor may be rewarded for assuming the higher risk.

VALUATION

The fair market value of a life-settlement policy is the expected net present value of the policy face, less the expected net present value of the premiums required to keep the policy in force. The probabilities required to calculate expectations on a single life of both policy face and premium cash flows will be developed and described in actuarial terms. With the probability of a cash flow occurring tied to the life status of the underlying insured, both expectations and variances of cash flows can be calculated. We show how the value of a policy increases in time. A portfolio of policies will then be considered and expressions developed to quantify risk in terms familiar to financial management. Throughout, we will be concerned with the question of how much cash flow to expect in any given year given today's information.

While today's actuarial theory incorporates the many insights of, and could be entirely cast in the framework of, modern probability theory, a working knowledge of actuarial math can be built from the life-table concept. The earliest actuarial studies for life insurance and annuity purposes began with the development of the life table. The following overview of actuarial methods is primarily drawn from Gerber (1997) and Chiang (1983).

In a life table, a cohort of a suitably large population was established, and at the end of the year, the remaining number of survivors was marked. (Sir Edmund Halley, of comet fame, is credited with constructing the first life table in 1693.) A life table following a cohort of 10,000 initially 75-year-old females with the same initial health and demographic conditions is given in Table 59.1. The table is truncated at year 100. Column t marks the beginning of the interval year. Column *Age*, x is the age of remaining survivors at time t. Column l_x is the number of survivors at time t. Column d_x is the number of people who survived to age x, then passed within the following year.

The quantities d and l are integer values counting the number of survivors and matured. The quantities p and q are ratios or probabilities of survival or death and range between 0 and 1. Before deriving the remaining columns, note that of the initial 10,000 individuals, only 697 survive to age 100, and that half of the original cohort has passed by approximately 13.5 years.

Table 59.1 Life Table of 10,000 Initially 75-Year-Old Females

t	Age, x	l_x	d_x	q_x	p_x	$_tp_{75}$	$_tp_{75}\,q_{75+t}$
0	75	10000	238	0.02375	0.97625	1.00000	0.02375
1	76	9763	255	0.02610	0.97390	0.97625	0.02548
2	77	9508	273	0.02869	0.97131	0.95077	0.02728
3	78	9235	291	0.03155	0.96845	0.92349	0.02914
4	79	8944	310	0.03464	0.96536	0.89436	0.03098
5	80	8634	329	0.03808	0.96192	0.86338	0.03288
6	81	8305	356	0.04285	0.95715	0.83050	0.03559
7	82	7949	383	0.04824	0.95176	0.79491	0.03835
8	83	7566	405	0.05357	0.94643	0.75656	0.04053
9	84	7160	426	0.05945	0.94055	0.71604	0.04257
10	85	6735	445	0.06609	0.93391	0.67347	0.04451
11	86	6290	453	0.07200	0.92800	0.62896	0.04528
12	87	5837	474	0.08113	0.91887	0.58367	0.04735
13	88	5363	486	0.09064	0.90936	0.53632	0.04861
14	89	4877	491	0.10076	0.89924	0.48771	0.04914
15	90	4386	482	0.10994	0.89006	0.43857	0.04822
16	91	3904	445	0.11402	0.88598	0.39035	0.04451
17	92	3458	425	0.12283	0.87717	0.34584	0.04248
18	93	3034	413	0.13628	0.86372	0.30336	0.04134
19	94	2620	402	0.15346	0.84654	0.26202	0.04021
20	95	2218	388	0.17488	0.82512	0.22181	0.03879
21	96	1830	357	0.19508	0.80492	0.18302	0.03570
22	97	1473	318	0.21588	0.78412	0.14732	0.03180
23	98	1155	252	0.21788	0.78212	0.11551	0.02517
24	99	903	206	0.22837	0.77163	0.09035	0.02063
25	100	697	171	0.24585	0.75415	0.06971	0.01714

Note: Table derived from the ultimate table for 75-year-old females drawn from the Female Non-Smoking 2001 Valuation Basic Table (commonly referred to as the VBT2001 tables).

The number of lives in the cohort aged x is l_x. The number of lives at the end of t years later is l_{x+t}. The ratio of those initially aged surviving t years and is the expected

$$\frac{l_{x+t}}{l_x} = {}_t p_x \tag{59.1}$$

The number of maturities through t years is $l_x - l_{x+t}$. The ratio of maturities through t years to the initial number of lives is x

$$\frac{l_x - l_{x+t}}{l_x} = 1 - {}_t p_x = {}_t q_x, \tag{59.2}$$

and can be expressed in terms of the percentage of survivors.

When the time span t is limited to 1 year the pre-subscript t is dropped. The notation p_x is read "the percentage of lives aged x, surviving one year is p_x." The notation q_x is read "the percentage of lives aged x, maturing within one year is q_x." Defining the number of maturities within one year as $d_x = l_x - l_{x+1}$, the percentage of those initially aged x maturing within one year can also be written

$$q_x = \frac{d_x}{l_x}, \quad \text{similarly} \quad p_x = \frac{l_{x+1}}{l_x}. \tag{59.3}$$

The relation $p_x + q_x = 1$ is immediate. The probability of an individual aged x being alive or dead within one year is 100%. The probability q_x is the quantity given in actuarial tables.

The probability of an individual initially aged x living for t years then maturing in the interval $(t, t+1)$ is the conditional probability

$$\begin{aligned} &\Pr\{\textit{matures in } (t, t+1) | (x) \textit{ survives to age } x+t\} \\ &= \frac{l_{x+t}}{l_x} \times \frac{d_{x+t}}{l_{x+t}} = \frac{d_{x+t}}{l_x}, \end{aligned} \tag{59.4}$$

or using the actuarial notation,

$$\Pr\{\textit{matures in } (t, t+1) | (x) \textit{ survives to age } x+t\} = {}_t p_x \, q_{x+t}. \tag{59.5}$$

Note that

$$\sum_{t=0}^{\infty} \frac{d_{x+t}}{l_x} = \sum_{t=0}^{\infty} {}_t p_x \, q_{x+t} = 1 \tag{59.6}$$

which expresses the sum of all maturities is equal to the original number of lives in the cohort, or the probability of maturing in some year in the future is 100%.

The probability of an individual initially aged x living for t years can also be expressed as the product

$${}_t p_x = \prod_{k=0}^{t-1} p_{x+k} = \prod_{k=0}^{t-1} (1 - q_{x+k}), \tag{59.7}$$

and is read as the probability of an individual aged x living to $x+1$ and then living from $x+1$ to $x+2$ and so on. Equation (59.7) is particularly useful given that only q_x are provided in actuarial tables.

Referring to Table 59.1, the number of survivors at the beginning of year $t = 3$ is $l_{78} = 9235$, these survivors are all 78-years-old. At the end of year 3, $t = 4$, $l_{79} = 8944$. Accordingly, 310 individuals, who began the year at age

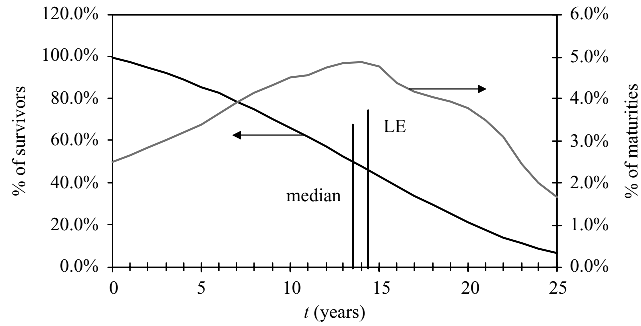


Figure 59.1 Survivor Curve and Mortality Distribution from Table 59.1

78 passed before their 79th birthday and $d_{78} = 310$. The percentage of 78-year-olds who passed compared to those who entered the year alive is $q_{78} = 3.155\%$, and with a sufficiently large population q_{78} is the probability that a 78-year-old non-smoking female will pass before her 79th birthday.

The probability that a 75-year-old non-smoking female will live to 78, then pass before her 79th birthday can be calculated with either (59.4) or (59.5).

$$\begin{aligned} \frac{d_{x+t}}{l_x} &= \frac{d_{78}}{l_{75}} = 2.914\% \text{ or } {}_t p_x q_{x+t} = {}_3 p_{75} \, q_{78} \\ &= 0.92349 \times 0.03155 = 2.914\% \end{aligned}$$

Both methods give the same result, but q_x is again typically given in actuarial tables, and working directly with them is preferred.

The life expectancy of an individual aged x is

$$e_x = \left(\sum_{t=1}^{\infty} {}_t p_x \, q_{x+t} \, t \right) + \frac{1}{2} \tag{59.8}$$

The $\frac{1}{2}$ represents the fraction of the final year of life that lived. The life expectancy for the cohort in Table 59.1 is 13.64 years. In this case, the life expectancy is approximately one year longer than the median life, or point in time at which only $\frac{1}{2}$ of the cohort survives.

Figure 59.1 shows how the cohort is expected to evolve over time. The survivor curve begins with 100% of the cohort alive and declines as individuals mature. The time at which 50% of the cohort has matured is the median remaining life. The LE is slightly longer than the median remaining life.

Variance Estimates

The life table is a binomial process. A 78-year-old female has a $1 - q_{78}$ chance of living to her 79th birthday and a q_{78} of not. The variance for the estimates in any given year x is

$$\text{var}(q_x) = \text{var}(p_x) = \frac{1}{l_x} p_x q_x \tag{59.9}$$

The variance of the survivor estimate t years from today is

$$\text{var}({}_t p_x) = \frac{1}{l_x} {}_t p_x (1 - {}_t p_x) \tag{59.10}$$

The variance in the estimate of those expected to survive t years, then pass within the next year, is

$$\mathrm{var}({}_t p_x q_{x+t}) = \frac{1}{l_x} {}_t p_x q_{x+t}(1 - {}_t p_x q_{x+t}) \tag{59.11}$$

In each case, l_x is the number of independent trials, or in the context of the life table, the number of individuals in the individual cohort.

ELEMENTARY LIFE INSURANCE

First, consider the whole life contract, which provides a payment or death benefit of one unit at the end of the year of death. Using the preceding cohort described in Table 59.1, what is the net present value (NPV) of the whole life policy benefit? Using a discount factor of v, where r is the discount rate,

$$\nu = (1 + r)^{-1} \tag{59.12}$$

a maturity in year T, received at time $t = T + 1$ has the NPV of v^{T+1}. While the amount payment is known, the time of payment, T, is random. Define Z as a random variable

$$Z = v^{T+1}$$

The expected NPV of Z is the probability-weighted cash flow of benefits is

$$E[Z] = A_x = \sum_{t=0}^{\infty} {}_t p_x \quad q_{x+t} \quad v^{t+1} \tag{59.13}$$

The term A_x is known as the net single premium where a payment of A_x would purchase a paid up whole-life policy requiring no further premiums.

The NPV of a term policy's benefits follows from (59.13) where benefits after year n are zero, resulting in a truncated series:

$$A^1_{x:n} = \sum_{t=0}^{n-1} {}_t p_x \quad q_{x+t} \quad v^{t+1} \tag{59.14}$$

The variance of Z is

$$\begin{aligned}\sigma^2_{A_x} &= \mathrm{var}(Z) = E(Z^2) - E(Z)^2 \\ &= \left(\sum_{t=0}^{\infty} {}_t p_x \quad q_{x+t} \quad v^{2(t+1)}\right) - A_x^2 \end{aligned} \tag{59.15}$$

Elementary Life Annuity

A whole-life annuity-due pays one unit as long as the beneficiary is alive. Payments are issued at the beginning of each year the contract is in force at time $t = 0, 1, 2, 3\ldots$ The NPV of this cash flow is

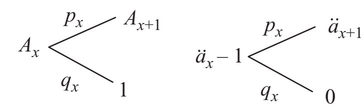
$$E(Y) = \ddot{a}_x = \sum_{t=0}^{\infty} {}_t p_x v^t \tag{59.16}$$

Paying one unit rather than receiving while the insured is alive would match the cash flow of a level premium. The variance is of the whole-life annuity is

$$\sigma^2_{\ddot{a}_x} = \mathrm{var}(Y) = \left(\frac{1+r}{r}\right)^2 \mathrm{var}(Z) \tag{59.17}$$

Forward Values

Values for the quantities A_x and $\ddot{a}_x$, one period forward can be quickly seen from a binomial tree construction:



$$A_x = v p_x A_{x+1} + v q_x \quad \ddot{a}_x = 1 + v p_x \ddot{a}_{x+1} \tag{59.18}$$

Today's value A_x can be understood as the discounted value of the one period forward value A_{x+1} assuming survival plus the discounted benefit value assuming maturity. Today's value $\ddot{a}_x$ is the discounted value of the one period forward $\ddot{a}_{x+1}$ plus an immediate payment of one unit.

Elementary-Level Premium

The annual level premium Π required to keep a whole life policy with benefit F in force is found by using (59.13) and (59.16).

$$F A_x = \Pi \ddot{a}_x \text{ which gives } \Pi = F\frac{A_x}{\ddot{a}_x} \tag{59.19}$$

With the proper level premium, the NPV of the whole life benefit should be equal to the NPV a whole life annuity. Dividing the premium Π by the benefit F results in the *premium to face ratio (pfr)* or π_x

$$\pi_x = \frac{A_x}{\ddot{a}_x} \tag{59.20}$$

This premium is unloaded; it does not include expense loadings, which include acquisition costs, collection expenses, administration fees, or risk loadings for larger benefits. Typical values for π in life settlement market range from 3% to 7%.

The unloaded minimum cost of insurance for a one-year term paid, in the event of death, at the end of the term with a premium paid in advance is expressed as a *pfr* and is simply q_x.

Substandard Health

The analysis developed so far has assumed a base case survivor curve and reflects the survival rates for a large population sharing major factors such as age, gender, and smoking status. The life table allows the calculation of q_x, which is expressed either as the ratio of maturities in a given year to those alive at the beginning of the year, or the probability that a living individual will mature within one year.

An individual health status will likely vary from the average of its cohort. If an individual is very ill, compared to their cohort, they may have four times the chance of maturing within a year. However, a healthy individual with evidence of superior cardiovascular health may have 0.9 times the chance of maturing within a year when compared to their base cohort. Some individuals may experience higher *mortality risk* while exposed to the hazard of some ailment such as cancer, but return to standard health once cured.

The mortality risk relative to the cohort base is expressed with the multiplier m_k which can vary with time. The Multiple Medical Impairment Study gives mortality multipliers, based on experience from 1962 to 1977 for various impairment groups. Though the study is dated it provides a framework and example baseline from which to consider multiple medical impairments.

Depending on the medical impairment, the multiplier can decay in time. One reason is that an individual may be diagnosed with a condition, which, if survived the first few years, would return to the base risk. A second reason would be that as a person ages with impaired health, their cohort is likely to experience the same impairment. An example of the former is an individual who experiences higher mortality risk while exposed to the hazard such as cancer, but returns to standard health once cured. Cardiovascular disease is an example of the later.

Equations (59.5) and (59.7) are combined to give and expression for the probability of maturity in a given year with multiplier *m:*

$$ {}_t\hat{p}_x \hat{q}_{x+t} = \left[\prod_{k=0}^{t-1} [1 - \min(m_k q_{x+k}, 1)] \right] \min(m_t q_{x+t}, 1) \tag{59.21}$$

The *min*() function ensures that the quantity $m_t q_{x+t}$ never exceeds 1, as the probability of maturity in a given year can never exceed 100%.

Figure 59.2 shows the mortality distributions 70 years with notable health risk, and an 80-year-old with minor health concerns. For simplicity, the multipliers are constant 350% and 130%, respectively. The distributions are

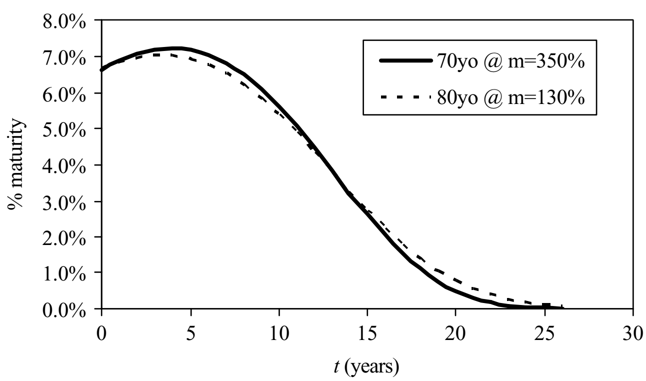


Figure 59.2 Mortality Distributions for a 70-Year Old with Notable Health Impairment and an 80-Year Old with Minor Health Concerns

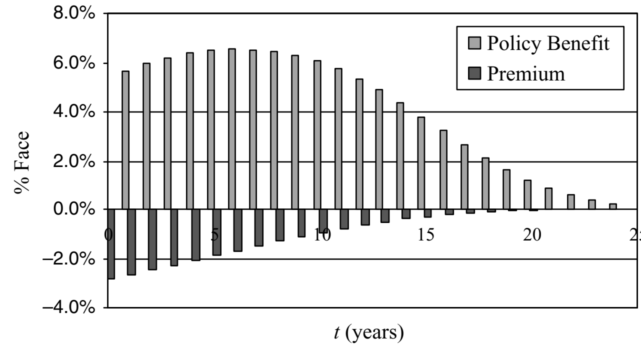


Figure 59.3 Illustrative Cash Flows for Cohort Of 80-Year Old Nonsmoking Females with Level Premiums Of 3% of Face

nearly equivalent in shape as are the median and life expectancy. Note that the life expectancy for the 70-year-old is more sensitive to a change in multiplier m than the life expectancy of the 80-year-old.

Insurance industry underwriting convention adjusts for substandard health by underwriting a policy at a table rating of 0,1,2,3... The table rating *tr*, relates to the multiplier m by

$$m = 1 + \frac{1}{4} tr.$$

While the table rating at issue can often be found stated within the policy, the underlying actuarial tables that are used in the policy construction are not public knowledge.

Elementary Life-Settlement Valuation

The net present value (NPV) of the policy benefit less the NPV of the premium stream is the value of the life settlement. Using the actuarial quantities A_x and $\ddot{a}_x$ developed (59.13) and (59.16), where the effect of substandard mortality is expressed through the multiplier m, the value of the life settlement is

$$P = F\left[A_x(m) - \pi \ddot{a}_x(m)\right] \text{ or } p = A_x(m) - \pi \ddot{a}_x(m), \tag{59.22}$$

where P is the policy value or price and F is the policy face, π is the premium to face ratio. The discount rate or IRR of the life settlement is embedded in the quantities $A_x(m)$ and $\ddot{a}_x(m)$. The policy value expressed as a percentage of face is p. Illustrative cash flows are shown in Figure 59.3. The NPV of these cash flows gives the policy value.

Equation (59.19) assumes a level premium and level policy face, and no loan against the policy, but can be easily generalized to account for such variations. Using the expression for forward values in (59.18), we can project the forward values of the life settlement. Figure 59.4 shows how the price of the policy is expected to increase at future dates.

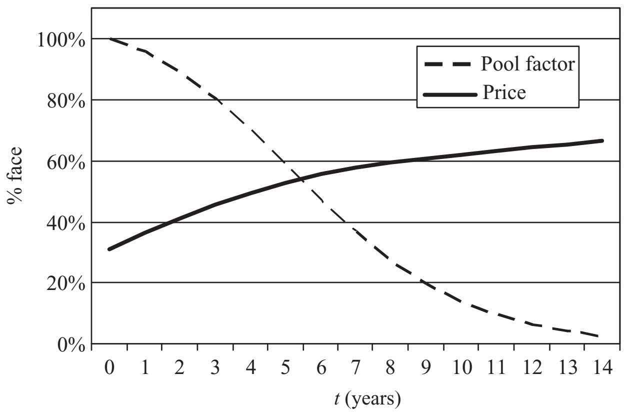


Figure 59.4 Forward Values for Policy Value

PORTFOLIO OF LIFE SETTLEMENTS

The expected net present value of the policy face less the expected NPV of the premiums required to keep the policy in force has been established in (59.22). We will now combine $N_{policies}$ together into a portfolio. Price and other attributes are averaged together using policy face weighting. Assuming each individual life in the portfolio is uncorrelated, a variance estimate for the portfolio price is given. Having paid a known price, this variance can be used to estimate the more useful quantity, the variance of the expected IRR do to actuarial variance.

In dollar terms, the value of the i^{th} policy in a portfolio of life settlements is

$$P_i = F_i \left\lfloor A_{x_i}(m) - \pi \ddot{a}_{x_i}(m) \right\rfloor$$

Summing all policies

$$\sum_{i=1}^{N_{policies}} P_i = \sum_{i=1}^{N_{policies}} F_i A_{x_i}(m) - \sum_{i=1}^{N_{policies}} F_i \pi_i \ddot{a}_{x_i}(m)$$

Dividing by the total face amount in the portfolio gives the face weighted value Φ

$$\Phi = \sum_{i=1}^{N_{policies}} \omega_i \left[A_{xi}(m) - \pi_i \ddot{a}_{x_i}(m) \right] \tag{59.23}$$

where

$$\omega_i = \frac{F_i}{\bar{F}} \text{ and } \bar{F} = \sum_{i=1}^{N_{policies}} F_i$$

The variance of the portfolio price is

$$\text{var}(\Phi) = \sum_{i=1}^{N_{policies}} \omega_i^2 \, \text{var}\left[A_{xi}(m) - \pi_i \ddot{a}_{x_i}(m) \right] \tag{59.24}$$

where

$$\text{var}\left[A_x(m) - \pi \ddot{a}_x(m)\right] = \sigma_{A_x}^2 - 2\pi \rho_{A_x,\ddot{a}_x} \sigma_{A_x} \sigma_{\ddot{a}_x} + \pi^2 \sigma_{\ddot{a}_x}^2 \tag{59.25}$$

and σ_{A_x} and $\sigma_{\ddot{a}_x}$ are defined in (59.15) and (59.17) and $\rho_{A_x,\ddot{a}_x} = -1$;

Table 59.2 Calculated Variance Using Monte Carlo (MC) Simulation

$N_{policies}$	1	5	10	50	100
MC:$\sigma_{A_x}^2$	0.051440	0.010299	0.005164	0.001027	0.000514
MC:$\sigma_{\ddot{a}_x}^2$	4.481010	0.897156	0.449853	0.089423	0.044782
MC: σ_{Φ}^2	0.090537	0.018127	0.009089	0.001807	0.000905

Table 59.3 Analytical Variance for Various Number of Policies

$N_{policies}$	1	5	10	50	100
Eq. (59.15): $\sigma_{A_x}^2$	0.051466	0.010293	0.005147	0.001029	0.000515
Eq. (59.17): $\sigma_{\ddot{a}_x}^2$	4.483290	0.896658	0.448329	0.089666	0.044833
Eq. (59.26): σ_{Φ}^2	0.090583	0.018117	0.009058	0.001812	0.000906

Now consider a portfolio with N independent lives from the same cohort, where each policy has the same face. The face weighting for each policy is $\omega_i = 1/N_{policies}$, and because each life is from the same cohort, the variance is also the same for each policy. The variance of the portfolio is

$$\begin{aligned} \text{var}(\Phi) &= \left(\frac{\text{var}\left[A_x(m) - \pi \ddot{a}_x(m)\right]}{N_{policies}^2} \right) \sum_{i=1}^{N_{policies}} 1 \\ &= \frac{\text{var}\left[A_x(m) - \pi \ddot{a}_x(m)\right]}{N_{polcies}} \end{aligned} \tag{59.26}$$

To confirm the analytical expressions for variance in (59.15), (59.17), and (59.26), we calculate the variance of a portfolio of $N_{policies}$ using Monte Carlo simulation (see Table 59.2) and compare to the analytical values.

Comparing Table 59.2 to Table 59.3 shows agreement between calculated and analytical value. As demonstrated, calculating σ_{Φ}^2 is straightforward enough, though calculating the analytical variances requires as much effort as implementing a Monte Carlo simulation. Fortunately, the distribution of Φ converges to a normal distribution rapidly as $N_{policies}$ increases as shown in Figure 59.5. This useful attribute will allow us to estimate the variance in r as a function of $N_{policies}$.

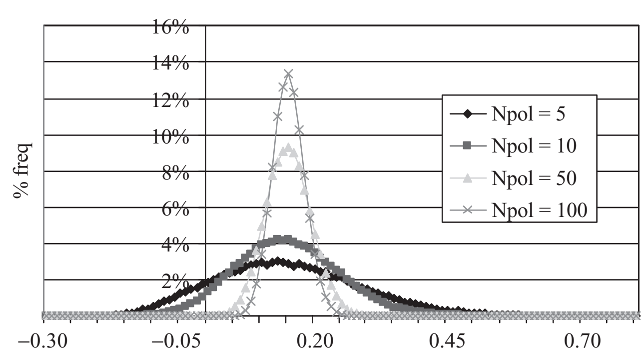


Figure 59.5 Distributions of Φ for Increasing Number of Policies

Table 59.4 Comparison of Standard Deviation in r Using Monte Carlo (MC) Simulation and Analytical Approximation

$N_{policies}$	5	10	50	100	300	500
MC:σ_r	0.17879	0.09353	0.03063	0.02026	0.01168	0.00889
Eq. (59.27): σ_r	0.08748	0.06186	0.02766	0.01956	0.01129	0.00875
err	−51.1%	−33.9%	−9.7%	−3.5%	−3.3%	−1.6%

Assuming a portfolio of policies was assembled using a discount rate of $\bar{r}$ for a price-to face-ratio of $\bar{\Phi}$, what is the expected variance of r? We can connect the standard deviation of Φ, σ_Φ with σ_r via

$$\sigma_\Phi \approx \left|\frac{\partial \Phi}{\partial r}\right| \sigma_r \tag{59.27}$$

with

$$\frac{\partial \Phi}{\partial r} = \sum_{i=1}^{N_{policies}} \omega_i \frac{\partial}{\partial r}\left[A_{xi}(m) - \pi_i \ddot{a}_{x_i}(m)\right] \tag{59.28}$$

where

$$\frac{\partial}{\partial r} A_x = -v \sum_{t=0}^{\infty} (t+1)_t\, p_x\, q_{x+t}\, v^{t+1} \quad \text{and}$$

$$\frac{\partial}{\partial r} \ddot{a}_x = -v \sum_{t=0}^{\infty} t_t\; p_x\, v^t \tag{59.29}$$

The standard deviation σ_r can be approximated directly from actuarial information and (59.27)–(59.29). To confirm this result, we calculate the standard deviation of r, σ_r, for a portfolio of $N_{policies}$ from the same cohort priced at $\bar{\Phi}$ using Monte Carlo simulation and compare to the analytical approximation.

The results in Table 59.4 show converging agreement between the two measures of σ_r with increasing $N_{policies}$. The error measure between the Monte Carlo results and the analytical is

$$err = \frac{\left[(\sigma)_{AnalyticApprox} - (\sigma)_{MC}\right]}{(\sigma_r)_{MC}}$$

The resulting distributions from the Monte Carlo distribution in r are shown in Figure 59.6. As expected, the

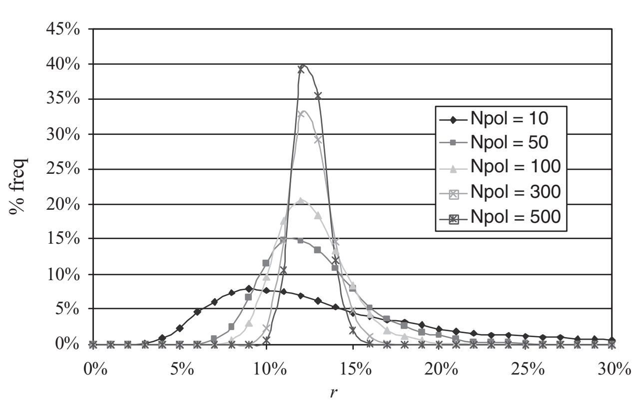


Figure 59.6 Distributions of r for Increasing Number of Policies

distribution tends toward the normal distribution as the number of policies increase.

SUMMARY

A life settlement is the sale of a life insurance policy, while the insured is still alive, from the policy's original owner to a potentially unrelated third party. If the policy were allowed to lapse, the owner would be essentially selling the policy for cash surrender value. In recent years, the life-settlement market has developed to give competitive market pricing, which may exceed a policy's cash surrender value, to policyholders for their unneeded insurance.

The lifespan of the insured, which is arguably uncorrelated to other assets, is the primary factor driving the life settlement cash flow. The life expectancy of the underlying insured is typically greater than eight years which accounts for the life settlement's long duration. Because policies sit atop of the capital structure of high credit quality insurance carriers, the risk of credit default is minimal.

Standard actuarial methods to develop a valuation methodology for single policies and a portfolio of policies can be employed. We show how the valuation methodology can be used to calculate forward price values of the portfolio. Assuming a binomial process, an estimate for the variance of the portfolio's expected internal rate of return is given and compared to Monte Carlo simulations.

REFERENCES

Black, K. Jr., and Skipper, H. D. Jr., (2000). *Life and Health Insurance*. Upper Saddle River, NJ: Pearson Education.

Booth, P. M., Chadburn, R., Cooper, D., Haberman, S., and James, D. (1998), *Modern Actuarial Theory and Practice.* Boca Raton, FL: Chapman & Hall/CRC.

Chiang, C. L. (1983). *Life Table and Its Applications*, Malabar, FL: Krieger Publishing Co.

Daykin, C. D., Pentikainen, T., and Pesonen, M. (1993). *Practical Risk Theory for Actuaries* Boca Raton, FL: Chapman & Hall/CRC.

Deshpande, J. V., Purohit, S. G. (2005). *Life Time Data: Statistical Models and Methods.* Singapore: World Scientific Publishing Co.

Gerber, H. U. (1997). *Life Insurance Mathematics*, 3rd edition. New York: Springer-Verlag.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. (1998). *Loss Models: From Data to Decisions*. New York: Wiley-Interscience.

Luenberger, D. G. (1998). *Investment Science*. New York: Oxford University Press.

Multiple Medical Impairment Study (1998). Staff MIB, SOA AAIM HOLUA-IHOU Mortality & Morbidity Liaison Committee, Center for Medico-Actuarial Statistics of MIB Inc.

PART 6

Investment Companies, ETFs, and Life Insurance Products

CHAPTER 60

Investment Companies

FRANK J. JONES, PhD
Professor, Accounting and Finance Department, San Jose State University and Chairman of the Investment Committee, Salient Wealth Management, LLC

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

Abstract: Investment companies include open-end mutual funds, closed-end funds, and unit trusts. Shares in investment companies are sold to the public and the proceeds invested in a diversified portfolio of securities. The value of a share of an investment company is called its net asset value. The two types of costs borne by investors in mutual funds are the shareholder sales charge or loads and the annual fund operating expense. Two major advantages of the indirect ownership of securities by investing in mutual funds are (1) risk reduction through diversification, and (2) reduced cost of contracting and processing information because an investor purchases the services of a presumably skilled financial advise at less cost than if the investor directly and individually negotiated with such an adviser. There is a wide-range of investment companies that invest in different asset classes and with different investment objectives.

Keywords: investment companies, open-end funds, mutual funds, closed-end funds, net asset value (NAV), shareholder fee, sales charge, load, commission, expense ratio, unit trust, front-end load, operating expense, expense ratio, management fee, family of funds, investment adviser, exchange-traded funds (ETFs), Morningstar, Lipper

Investment companies are entities that sell shares to the public and invest the proceeds in a diversified portfolio of securities. Each share sold represents a proportional interest in the portfolio of securities managed by the investment company on behalf of its shareholders. The type of securities purchased depends on the company's investment objective.

TYPES OF INVESTMENT COMPANIES

There are three types of investment companies: open-end funds, closed-end funds, and unit trusts.

Open-End Funds (Mutual Funds)

Open-end funds, commonly referred to simply as *mutual funds*, are portfolios of securities, mainly stocks, bonds, and money market instruments. There are several important aspects of mutual funds. First, investors in mutual funds own a pro rata share of the overall portfolio. Second, the investment manager of the mutual fund manages the portfolio, that is, buys some securities and sells others (this characteristic is unlike unit investment trusts, discussed later).

Third, the value or price of each share of the portfolio, called the *net asset value* (*NAV*), equals the market value of the portfolio minus the liabilities of the mutual fund divided by the number of shares owned by the mutual fund investors. That is,

$$\text{NAV} = \frac{\text{Market value of portfolio} - \text{Liabilities}}{\text{Number of shares outstanding}}$$

For example, suppose that a mutual fund with 10 million shares outstanding has a portfolio with a market value of \$215 million and liabilities of \$15 million. The NAV is

$$\text{NAV} = \frac{\$215{,}000{,}000 - \$15{,}000{,}000}{10{,}000{,}000} = \$20$$

Fourth, the NAV or price of the fund is determined only once each day, at the close of the day. For example, the NAV for a stock mutual fund is determined from the closing stock prices for the day. Business publications provide the NAV each day in their mutual fund tables. The published NAVs are the closing NAVs.

Fifth, and very importantly, all new investments into the fund or withdrawals from the fund during a day are priced at the closing NAV (investments after the end of the day or on a non–business day are priced at the next day's closing NAV).

The total number of shares in the fund increases if there are more investments than withdrawals during the day, and vice versa. This is the reason such a fund is called an "open-end" fund. For example, assume that at the beginning of a day a mutual fund portfolio has a value of \$1 million, there are no liabilities, and there are 10,000 shares outstanding. Thus, the NAV of the fund is \$100. Assume that during the day \$5,000 is deposited into the fund, \$1,000 is withdrawn, and the prices of all the securities in the portfolio remain constant. This means that 50 shares were issued for the \$5,000 deposited (since each share is \$100) and 10 shares redeemed for \$1,000 (again, since each share is \$100). The net number of new shares issued is then 40. Therefore, at the end of the day there will be 10,040 shares and the total value of the fund will be \$1,004,000. The NAV will remain at \$100.

If, instead, the prices of the securities in the portfolio change, both the total size of the portfolio and, therefore, the NAV will change. In the previous example, assume that during the day the value of the portfolio doubles to \$2 million. Since deposits and withdrawals are priced at the end-of-day NAV, which is now \$200 after the doubling of the portfolio's value, the \$5,000 deposit will be credited with 25 shares (\$5,000/\$200) and the \$1,000 withdrawn will reduce the number of shares by 5 shares (\$1,000/\$200). Thus, at the end of the day there will be 10,020 shares (25 – 5) in the fund with an NAV of \$200, and the value of the fund will be \$2,004,000. (Note that 10,020 shares × \$200 NAV equals \$2,004,000, the portfolio value.)

Overall, the NAV of a mutual fund will increase or decrease due to an increase or decrease in the prices of the securities in the portfolio, respectively. The number of shares in the fund will increase or decrease due to the net deposits into or withdrawals from the fund, respectively. And the total value of the fund will increase or decrease for both reasons.

Closed-End Funds

The shares of a *closed-end fund* are very similar to the shares of common stock of a corporation. The new shares of a closed-end fund are initially issued by an underwriter for the fund. And after the new issue, the number of shares remains constant. This is the reason such a fund is called a "closed-end" fund. After the initial issue, there are no sales or purchases of fund shares by the fund company as there are for open-end funds. The shares are traded on a secondary market, either on an exchange or in the over-the-counter market.

Investors can buy shares either at the time of the initial issue (as discussed below), or thereafter in the secondary market. Shares are sold only on the secondary market. The price of the shares of a closed-end fund are determined by the supply and demand in the market in which these funds are traded. Thus, investors who transact closed-end fund shares must pay a brokerage commission at the time of purchase and at the time of sale.

The NAV of closed-end funds is calculated in the same way as for open-end funds. However, the price of a share in a closed-end fund is determined by supply and demand, so the price can fall below or rise above the net asset value per share. Shares selling below NAV are said to be "trading at a discount," while shares trading above NAV are "trading at a premium." Newspapers list quotations of the prices of these shares under the heading "Closed-End Funds." Some sources also list the NAV and the discount or premium of the shares.

Consequently, there are two important differences between open-end funds and closed-end funds. First, the number of shares of an open-end fund varies because the fund sponsor will sell new shares to investors and buy existing shares from shareholders. Second, by doing so, the share price is always the NAV of the fund. In contrast, closed-end funds have a constant number of shares outstanding because the fund sponsor does not redeem shares and sell new shares to investors (except at the time of a new underwriting). Thus, the price of the fund shares will be determined by supply and demand in the market and may be above or below NAV, as discussed above.

Although the divergence of the price from NAV is often puzzling, in some cases the reasons for the premium or discount are easily understood. For example, a share's price may be below the NAV because the fund has a large built-in tax liability and investors are discounting the share's price for that future tax liability. (We'll discuss this tax liability issue later in this chapter.) A fund's leverage and resulting risk may be another reason for the share's price trading below NAV. A fund's shares may trade at a premium to the NAV because the fund offers relatively cheap access to, and professional management of, stocks in another country about which information is not readily available to or transactions are difficult or expensive for small investors.

Under the Investment Company Act of 1940, closed-end funds are capitalized only once. They make an initial IPO (initial public offering) and then their shares are traded on the secondary market, just like any corporate stock, as discussed earlier. The number of shares is fixed at the IPO; closed-end funds cannot issue more shares. In fact, many closed-end funds become leveraged to raise more funds without issuing more shares.

An important feature of closed-end funds is that the initial investors bear the substantial cost of underwriting the issuance of the funds' shares. The proceeds that the managers of the fund have to invest equals the total paid by initial buyers of the shares minus all costs of issuance. These costs, which average around 7.5% of the total amount paid for the issue, normally include selling fees or commissions paid to the retail brokerage firms that distribute them to the public. The high commissions are strong incentives for retail brokers to recommend these shares to their retail customers, and also for investors to avoid buying these shares on their initial offering.

Exchange-traded funds (ETFs) pose a threat to both mutual funds and closed-end funds. ETFs, which are the subject of Chapter 61 in Volume I, are essentially hybrid closed-end vehicles, which trade on exchanges but which typically trade very close to NAV.

Since closed-end funds are traded like stocks, the cost to any investor of buying or selling a closed-end fund is the same as that of a stock. The obvious charge is the stock broker's commission. The bid/offer spread of the market on which the stock is traded is also a cost.

Unit Trusts

A *unit trust* is similar to a closed-end fund in that the number of unit certificates is fixed. Unit trusts typically invest in bonds. They differ in several ways from both mutual funds and closed-end funds that specialize in bonds. First, there is no active trading of the bonds in the portfolio of the unit trust. Once the unit trust is assembled by the sponsor (usually a brokerage firm or bond underwriter) and turned over to a trustee, the trustee holds all the bonds until they are redeemed by the issuer. Typically, the only time the trustee can sell an issue in the portfolio is if there is a dramatic decline in the issuer's credit quality. As a result, the cost of operating the trust will be considerably less than costs incurred by either a mutual fund or a closed-end fund. Second, unit trusts have a fixed termination date, while mutual funds and closed-end funds do not. (There are, however, exceptions. Target term closed-end funds have a fixed termination date.) Third, unlike the mutual fund and closed-end fund investor, the unit trust investor knows that the portfolio consists of a specific portfolio of bonds and has no concern that the trustee will alter the portfolio. While unit trusts are common in Europe, they are not common in the United States.

All unit trusts charge a sales commission. The initial sales charge for a unit trust ranges from 3.5% to 5.5%. In addition to these costs, there is the cost incurred by the sponsor to purchase the bonds for the trust that an investor indirectly pays. That is, when the brokerage firm or bond-underwriting firm assembles the unit trust, the price of each bond to the trust also includes the dealer's spread. There is also often a commission if the units are sold.

In the remainder this chapter of our primary focus chapter is on open-end (mutual) funds.

FUND SALES CHARGES AND ANNUAL OPERATING EXPENSES

There are two types of costs borne by investors in mutual funds. The first is the *shareholder fee*, usually called the *sales charge* or *load*. For securities transactions, this charge is called a *commission*. This cost is a "one-time" charge debited to the investor for a specific transaction, such as a purchase, redemption or exchange. The type of charge is related to the way the fund is sold or distributed. The second cost is the annual fund operating expense, usually called the *expense ratio*, which covers the funds' expenses, the largest of which is for investment management. This charge is imposed annually. This cost occurs on all funds and for all types of distribution. We discuss each cost next.

Sales Charges or Loads

Sales charges on mutual funds are related to their method of distribution. The current menu of sales charges and distribution mechanisms has evolved significantly and is now much more diverse than it was a decade ago. To understand the current diversity and the evolution of distribution mechanisms, consider initially the circumstances of a decade ago. At that time, there were two basic methods of distribution, two types of sales charges, and the type of the distribution was directly related to the type of sales charge.

The two types of distribution were sales-force and direct. Sales-force distribution occurred via an intermediary, that is via an agent, a stockbroker, insurance agent, or other entity who provided investment advice and incentive to the client, actively "made the sale," and provided subsequent service. This distribution approach is active, that is the fund is typically sold, not bought.

The other approach is direct (from the fund company to the investor), whereby there is no intermediary or salesperson to actively approach the client, provide investment advice and service, or make the sale. Rather, the client approaches the mutual fund company, most likely by a toll-free telephone number, in response to media advertisements or general information, and opens the account. Little or no investment counsel or service is provided either initially or subsequently. With respect to the mutual fund sale, this is a passive approach, although these mutual funds may be quite active in their advertising and other marketing activities. Funds provided by the direct approach are bought, not sold.

There is a quid pro quo, however, for the service provided in the sales-force distribution method. The quid pro quo is a sales charge borne by the customer and paid to the agent. The sales charge for the agent-distributed fund is called a load. The traditional type of load is called a front-end load, since the load is deducted initially or "up front." That is, the load is deducted from the amount invested by the client and paid to the agent/distributor. The remainder is the net amount invested in the fund in the client's name. For example, if the load on the mutual fund is 5% and the investor invests $100, the $5 load is paid to the agent and the remaining $95 is the net amount invested in the mutual fund at NAV. Importantly, only $95, not $100, is invested in the fund. The fund is, thus, said to be "purchased above NAV" (that is, the investor pays $100 for $95 of the fund). The $5 load compensates the sales agent for the investment advice and service provided to the client by the agent. The load to the client, of course, represents income to the agent.

Let's contrast this with directly placed mutual funds. There is no sales agent and, therefore, there is no need for a sales charge. Funds with no sales charges are called no-load mutual funds. In this case, if the client provides $100 to the mutual fund, $100 is invested in the fund in the client's name. This approach to buying the fund is called buying the fund "at NAV," that is, the whole amount provided by the investor is invested in the fund.

Previously, many observers speculated that load funds would become obsolete and no-load funds would dominate because of the sales charge. Increasingly financially sophisticated individuals, the reasoning went, would make their own investment decisions and not need to compensate agents for their advice and service. But the actual trend has been quite different.

Why has the trend not been away from the more costly agent distributed funds as many expected? There are two reasons. First, many investors have remained dependent on the investment counsel and service, and perhaps more importantly, the initiative of the sales agent. Second, sales-force distributed funds have shown considerable ingenuity and flexibility in imposing sales charges, which both compensate the distributors and are acceptable to the clients. Among the adaptations of the front end sales load are back-end loads and level loads. While the front-end load is imposed at the time of the purchase of the fund, the back-end load is imposed at the time fund shares are sold or redeemed. Level loads are imposed uniformly each year. These two alternative methods both provide ways to compensate the agent. However, unlike with the front-end load, both of these distribution mechanisms permit the client to buy a fund at NAV—that is, not have any of their initial investment debited as a sales charge before it is invested in their account.

The most common type of back-end load currently is the contingent deferred sales charge (CDSC). This approach imposes a gradually declining load on withdrawal. For example, a common "3,3,2,2,1,1,0" CDSC approach imposes a 3% load on the amount withdrawn within one year, 3% within the second year, 2% within the third year, and so on. There is no sales charge for withdrawals after the sixth year. Thus, the sales charge is postponed or deferred, and it is contingent upon how long the investment is held.

The third type of load is neither a front-end load at the time of investment nor a (gradually declining) back-end load at the time of withdrawal, but a constant load each year (e.g., a 1% load every year). This approach is called a level load. Most mutual fund families are strictly either no-load (direct) or load (sales-force).

Many load type mutual fund families often offer their funds with all three types of loads—that is, front-end loads (usually called "A shares"); back-end loads (often called "B shares"); and level loads (often called "C shares"). These families permit the distributor and its client to select the type of load they prefer. [See O'Neal (1999).] These different types of load shares are called share "classes." A recent type of share class is "F shares." F shares have no front, level or back loads. In this way they are like C shares. But F shares have considerably lower annual expenses than C shares, as will be seen below. F shares are designed for financial planners who charge annual fees (called fee-based financial planners) rather than sales charges such as commissions or loads. F shares of a fund family may only be sold by investment dealers and their representatives which have an arrangement with the fund family.

According to the National Association of Securities Dealers (NASD), the maximum allowable sales charge is 8.5%, although most funds impose lower charges.

The sales charge for a fund applies to most, even very small, investments (although there is typically a minimum initial investment). For large investments, however, the sales charge may be reduced. For example, a fund with a 4.5% front-end load may reduce this load to 3.0% for investments over $1 million. At some level of investment the front-end load will be 0%. There may be in addition further reductions in the sales charge at greater investments. The amount of investment needed to obtain a reduction in the sales charge is called a breakpoint—the breakpoint is $1 million in this example. There are also mechanisms whereby the total amount of the investment necessary to qualify for the breakpoint does not need to be invested up front, but only over time (according to a "letter of intent" signed by the investor). (See Inro, Jaing, Ho, and Lee, 1999.) Fund returns are calculated without subtracting sales charges since different individual investors have different sales charges (e.g., may have different breakpoints).

The sales charge is, in effect, paid by the client to the distributor. How does the fund family, typically called the sponsor or manufacturer of the fund, cover its costs and make a profit? This is the topic of the second type of "cost" to the investor, the fund annual operating expense.

Annual Operating Expenses (Expense Ratio)

The *operating expense*, also called the *expense ratio*, is debited annually from the investor's fund balance by the fund sponsor. The three main categories of annual operating expenses are the management fee, distribution fee, and other expenses.

The *management fee*, also called the investment advisory fee, is the fee charged by the *investment adviser* for managing a fund's portfolio. If the investment adviser is part of a company separate from the fund sponsor, some or all of this investment advisory fee is passed on to the investment adviser by the fund sponsor. In this case, the fund manager is called a subadviser. The management fee varies by the type of fund, specifically by the risk of the asset class of the fund. For example, the management fee as well as the risk may increase from money market funds to bond funds, to U.S. growth stock funds, to emerging market stock funds, as illustrated by examples to come.

In 1980, the Securities and Exchange Commission (SEC) approved the imposition of a fixed annual fee, called the l2b-1 fee, which is, in general, intended to cover distribution costs, also including continuing agent compensation and manufacturer marketing and advertising expenses. Such l2b-1 fees are now imposed by many mutual funds. By law, 12b-1 fees cannot exceed 1% of the fund's assets per year. The 12b-1 fee may also include a service fee of up to 0.25% of assets per year to compensate sales professionals for providing services or maintaining shareholder accounts. The major rationale for the component of the l2b-1 fee which accrues to the selling agent is to provide an incentive to selling agents to continue to service their accounts after having received a transaction-based fee such as a front-end load. As a result, a 12b-1 fee of this type is consistent with sales-force sold, load funds, not with directly sold, no-load funds. The rationale for the component of the 12b-1 fee which accrues to the manufacturer of the fund is to provide incentive and compensate for continuing advertising and marketing costs.

Other expenses include primarily the costs of (1) custody (holding the cash and securities of the fund), (2) the transfer agent (transferring cash and securities among buyers and sellers of securities and the fund distributions, etc.), (3) independent public accountant fees, and (4) directors' fees.

The sum of the annual management fee, the annual distribution fee, and other annual expenses is called the expense ratio or annual operating expense. All the cost information on a fund, including selling charges and annual expenses, are included in the fund prospectus. In addition to the annual operating expenses, the fund prospectus provides the fees which are imposed only at the time of a fund transaction.

As we explained earlier, many agent-distributed funds are provided in different forms, typically the following: (1) A shares: front-end load; (2) B shares: back-end load (contingent deferred sales charge); (3) C shares: level load; and (4) F shares: fee based program. These different forms of the same fund are called share classes. Table 60.1 provides an example of hypothetical sales charges and annual expenses of funds of different classes for an agent distributed stock mutual fund. The sales charge accrues to the sales agent. The management fee accrues to the mutual fund manager. The 12b-1 fee accrues to the sales agent and the fund sponsor. Other expenses, including custody and transfer fees and the fees of managing the fund company, accrue to the fund sponsor to cover expenses. All of these expenses are deducted from fund returns on an annual basis.

Multiple Share Classes

Share classes were first offered in 1989 following the SEC's approval of multiple share class. Initially share classes were used primarily by sales-force funds to offer alternatives to the front-end load as a means of compensating brokers. Later, some of these funds used additional share classes as a means of offering the same fund or portfolio through alternative distribution channels in which some fund expenses varied by channel. Offering new share classes was more efficient and less costly than setting up two separate funds. [See Reid (2000).] By the end of the 1990s, the average long-term sales-force fund offered nearly three share classes. Direct market funds tended to continue to offer only one share class.

ADVANTAGES OF INVESTING IN MUTUAL FUNDS

There are several advantages of the indirect ownership of securities by investing in mutual funds. The first is risk reduction through diversification. By investing in a fund, an investor can obtain broad-based ownership of a sufficient number of securities to reduce portfolio risk. While an individual investor may be able to acquire a broad-based portfolio of securities, the degree of diversification will be limited by the amount available to invest. By investing in an investment company, however, the investor can effectively achieve the benefits of diversification at a lower cost even if the amount of money available to invest is not large.

The second advantage is the reduced cost of contracting and processing information because an investor purchases

Table 60.1 Hypothetical Sales Charges and Annual Expenses of Funds of Different Classes for an Agent Distributed Stock Mutual Fund

	Sales Charge			Annual Operating Expenses			
	Front	Back	Level	Management Fee	Distribution (12b-l Fee)	Other Expenses	Expense Ratio
A	4.5%	0	0%	0.90%	0.25%	0.15%	1.30%
B	0	[a]	0%	0.90%	1.00%	0.15%	2.05%
C	0	0	1%	0.90%	1.00%	0.15%	2.05%
F	0	0	0	0.90%	0.25%	0.15%	1.30%

[a]3%, 3%, 2%, 2%, 1%, 0%.

the services of a presumably skilled financial adviser at less cost than if the investor directly and individually negotiated with such an adviser. The advisory fee is lower because of the larger size of assets managed, as well as the reduced costs of searching for an investment manager and obtaining information about the securities. Also, the costs of transacting in the securities are reduced because a fund is better able to negotiate transactions costs; and custodial fees and record-keeping costs are less for a fund than for an individual investor. For these reasons, there are said to be economies of scale in investment management.

Third, and related to the first two advantages, is the advantage of the professional management of the mutual fund. Fourth is the advantage of liquidity. Mutual funds can be bought or liquidated any day at the closing NAV. Fifth is the advantage of the variety of funds available, in general, and even in one particular funds family, as discussed later.

Finally, money market funds and some other types of funds provide payment services by allowing investors to write checks drawn on the fund, although this facility may be limited in various ways.

TYPES OF FUNDS BY INVESTMENT OBJECTIVE

Mutual funds have been provided to satisfy the various investment objectives of investors. In general, there are stock funds, bond funds, money market funds, and others. Within each of these categories, there are several subcategories of funds. There are also U.S.-only funds, international funds (no U.S. securities), and global funds (both U.S. and international securities). There are also passive and active funds. Passive (or indexed) funds are designed to replicate an index, such as: the S&P 500 Stock Index; the Lehman Aggregate Bond Index; or the Morgan Stanley Capital International EAFE Index (Europe, Australasia, and the Far East). Active funds, on the other hand, attempt to outperform an index by actively trading the fund portfolio. There are also many other categories of funds, as discussed below. Each fund's objective is stated in its prospectus, as required by the SEC and the "1940 Act," as discussed below.

Stock funds differ by:

- The average market capitalization ("market cap") (large, mid, and small) of the stocks in the portfolio.
- Style (growth, value, and blend).
- Sector—"sector funds" specialize in one particular sector or industry, such as technology, healthcare, or utilities.

With respect to style, stocks with high price-to-book value and price-to-earnings ratios are considered "growth stocks," and stocks with low price-to-book value and price-to-earnings ratios are considered value stocks, although other variables may also be considered. There are also blend stocks with respect to style.

Bond funds differ by the creditworthiness of the issuers of the bonds in the portfolio (e.g., U.S. government and investment-grade and high-yield corporates) and by the maturity (or duration) of the bonds (long, intermediate, and short.) There is also a category of bond funds called municipal bond funds whose interest income is exempt from federal income taxes. Municipal funds may be single state (that is, all the bonds in the portfolio were issued by issuers in the same state) or multistate or "national."

There are also other categories of funds such as asset allocation, hybrid, target date, and balanced funds (all of which hold both stocks and bonds), and convertible bond funds.

There is also a category of money market funds (maturities of one year or less) which provide protection against interest rate fluctuations. These funds may have some degree of credit risk (except for the U.S. government money market category). Many of these funds offer check-writing privileges. In addition to taxable money market funds, there are also tax-exempt municipal money market funds.

Among the other fund offerings are index funds and funds of funds. Index funds, as discussed above, attempt to passively replicate an index. Funds of funds invest in other mutual funds not in individual securities. A fund of funds is a fund that invests in other mutual funds.

Several organizations provide data on mutual funds. The most popular ones are *Morningstar* and *Lipper*. These firms provide data on fund expenses, portfolio managers, fund sizes, and fund holdings. But perhaps most importantly, they provide performance (that is, rate of return) data and rankings among funds based on performance and other factors. To compare fund performance on an "apples to apples" basis, these firms divide mutual funds into several categories which are intended to be fairly homogeneous by investment objective. The categories provided by Morningstar and Lipper are similar but not identical. Many of the categories of these two services are shown and compared in Table 60.2. Thus, the performance of one Morningstar "large-cap blend" fund can be meaningfully compared with another fund in the same category, but not with a "small-cap value" fund. Morningstar's performance ranking system whereby each fund is rated on the basis of return and risk from one star (the worst) to five stars (the best) relative to the other funds in its category is well known.

Mutual fund data are also provided by the Investment Company Institute, the national association for mutual funds.

THE CONCEPT OF A FAMILY OF FUNDS

A concept that revolutionized the fund industry and benefitted many investors is what the mutual fund industry calls a *family of funds*, a group of funds or a complex of funds. That is, many fund management companies offer investors a choice of numerous funds with different investment objectives in the same fund family. In many cases, investors may move their assets from one fund to another within the family at little or no cost, and with only a phone call. Of course, if these funds are in a taxable account, there may be tax consequences to the sale. While the same policies regarding loads and other costs may apply

Table 60.2 Fund Categories: Morningstar versus Lipper

Morningstar		Lipper	
LG	Large Growth	LG	Large-Cap Growth
LV	Large Value	LV	Large-Cap Value
LB	Large Bland	LC	Larga-Cap Core
MG	Mid-Cap Growth	MG	Mid-Cap Growth
MV	Mid-Cap Value	MV	Mid-Cap Value
MB	Mid-Cap Blend	MC	Mid-Cap Core
SG	Small Growth	SG	Small-Cap Growth
SV	Small Value	SV	Small-Cap Value
SB	Small Blend	SC	Small-Cap Core
		XG	Multi-Tap Growth
		XV	Multi-Cap Value
		XC	Multi-Cap Core
MA	Moderate Allocation	BL	Balanced
CA	Conservationai Allocation	MP	Stock/Bond Blend
TA	Target—Date 2004–2014		
TB	Target—Date 2015–2029		
TC	Target—Date 2030+		
DH	Domestsc Hybrid	EI	Equity Income
		SP	S&P 500 Funds
		SQ	Specialty Diversified Equity
ST	Technology	TK	Science & Technology
SU	Utilities	UT	Utility
SH	Health	HB	Health/Biotech
SC	Communication	—	Telecommunications
SF	Financial	—	
SN	Natural Resources	NR	Natural Resources
SP	Precious Metals	AU	Gold Oriented
SR	Reai.Estate	—	Real Estate
BM	Bear Market	—	
LO	Long-Short	—	
—		SQ	Special Equity
—		SE	Sector
FS	Foreign Stock	IL	International Stock (non-U.S.)
WS	World Stock	GL	Global Stock (inc. U.S.)
ES	Europe Stock	EU	European Region
EM	Diversified Emerging Mkt.	EM	Emerging Markets
DP	Diversified Pacific Asia	PR	Pacific Region
PJ	Pacific ex-Japan	—	
JS	Japan Stock	—	
LS	Latin America Stock	LT	Latin American
IH	International Hybrid	—	
CS	Short-Term Bond—General	SB	Short-Term Bond
GS	Short Government	SU	Short-Term U.S. Govt.
CI	Interm.-Term Bond—General	IB	Intermediate Bond
GI	Interm. Government	IG	Intermediate U.S. Govt.
MT	Mortgage	MT	Mortgage
CL	Long-Term Bond—General	AB	Long-Term Bond
GL	Long Government	LU	Long-Term U.S. Govt.
IP	Inflation-Protected Bond	—	
		GT	General U.S. Taxable
CV	Convertibles	—	
UB	Ultrashort Bond	—	
HY	High-Yield Bond	HC	High-Yield Taxable
MU	Multisector Bond	—	
IB	World Bond	WB	World Bond
EB	Emerging Market Bond		
BL	Bank Loan		
ML	Muni National Long	GM	General Muni Debt
MI	Muni National Interm.	IM	Interm. Muni Debt
MS	Muni National Short	SM	Short-Term Muni Debt
HM	High Yield Muni	HM	High-Yield Muni
SL	Muni Single St. Long	NM	Insured Muni
SI	Muni Single St. Interm.	SS	Single-State Muni
SS	Muni Single St. Short		
MY	Muni New York Long		
MC	Muni California Long		
MN	Muni New York Interm./Sht		
MF	Muni California Interm./Sht		

to all the funds in a family, a management company may have different fee structures for transfers among different funds in its family.

Large fund families usually include money market funds, U.S. bond funds of several types, global stock and bond funds, broadly diversified U.S. stock funds, U.S. stock funds which specialize by market capitalization and style, and stock funds devoted to particular sectors such as healthcare, technology or gold companies. Well-known management companies, such as Vanguard, American Funds, and Fidelity the three largest fund families, sponsor and manage varied types of funds in a family. Fund families may also use external investment advisers (called subadvisors) along with their internal advisers in their fund families.

Fund data provided in newspapers group the various funds according to their families. For example, all the American Funds are listed under the American Fund heading, all the funds of Vanguard are listed under their name, and so on.

TAXATION OF MUTUAL FUNDS

Mutual funds must distribute at least 90% of their net investment income earned (bond coupons and stock dividends) exclusive of realized capital gains or losses to shareholders (along with meeting other criteria) to be considered a regulated investment company (RIC) and, thus, not be required to pay taxes at the fund level prior to distributions to shareholders. Consequently, funds always make these distributions. Taxes, if this criterion is met, are then paid on distributions, only at the investor level, not the fund level. Even though many mutual fund investors choose to reinvest these distributions, the distributions are taxable to the investor, either as ordinary income or capital gains (long term or short term), whichever is relevant.

Capital gains distributions must occur annually, and typically occur late during the calendar year. The capital gains distributions may be either long-term or short-term capital gains, depending on whether the fund held the security for a year or more. Mutual fund investors have no control over the size of these distributions and, as a result, the timing and amount of the taxes paid on their fund holdings is largely out of their control. In particular, withdrawals by some investors may necessitate sales in the fund, which in turn cause realized capital gains and a tax liability to accrue to investors who maintain their holding.

New investors in the fund may assume a tax liability even though they have no gains. That is, all shareholders as of the date of record receive a full year's worth of dividends and capital gains distributions, even if they have owned shares for only one day. This lack of control over capital gains taxes is regarded as a major limitation of mutual funds. In fact, this adverse tax consequence is one of the reasons suggested for a closed-end company's price selling below par value. Also, this adverse tax consequence is one of the reasons for the popularity of exchange-traded funds to be discussed later.

Of course, the investor must also pay ordinary income taxes on distributions of income. Finally, when the fund investors sell the fund, they will have long-term or short-term capital gains, taxes, depending on whether they held the fund for a year or less.

REGULATION OF FUNDS

There are four major laws or Acts which relate either indirectly or directly to mutual funds. The first is the Securities Act of 1933 ("the '33 Act") which provides purchasers of new issues of securities (the "primary market") with information regarding the issuer and, thus, helps prevent fraud. Because open-end investment companies issue new shares on a continuous basis, mutual funds must comply with the '33 Act. The Securities Act of 1934 ("the '34 Act") is concerned with the trading of securities once they have been issued (the "secondary market"), with the regulation of exchanges, and with the regulation of broker-dealers. Mutual fund portfolio managers must comply with the '34 Act in their transactions.

All investment companies with 100 or more shareholders must register with the SEC according to the Investment Company Act of 1940 ("the '40 Act"). The primary purposes of the '40 Act are to reduce investment company selling abuses and to ensure that investors receive sufficient and accurate information. Investment companies must provide periodic financial reports and disclose their investment policies to investors. The '40 Act prohibits changes in the nature of an investment company's fundamental investment policies without the approval of shareholders. This Act also provides some tax advantages for eligible RICs, as indicated below. The purchase and sale of mutual fund shares must meet the requirements of fair dealing that the SEC '40 Act and the NASD (National Association of Securities Dealers), a self-regulatory organization, have established for all securities transactions in the United States.

Finally, the Investment Advisers Act of 1940 specifies the registration requirements and practices of companies and individuals who provide investment advisory services. This Act deals with registered investment advisers (RIAs).

Overall, while an investment company must comply with all aspects of the '40 Act, it is also subject to the '33 Act, the '34 Act, and the Investment Advisers Act of 1940.

The SEC also extended the '34 Act in 1988 to provide protections such that advertisements and claims by mutual funds would not be inaccurate or misleading to investors. New regulations aimed at potential self-dealing were established in the Insider Trading and Securities Fraud Enforcement Act of 1988, which requires mutual fund investment advisers to institute and enforce procedures that reduce the chances of insider trading.

An important feature of the '40 Act exempts any company that qualifies as a "regulated investment company" from taxation on its gains, either from income or capital appreciation, as indicated above. To qualify as an RIC, the fund must distribute to its shareholders 90% of its net income excluding realized capital gains each year. Furthermore, the fund must follow certain rules about the diversification and liquidity of its investments, and the degree of short-term trading and short-term capital gains.

Fees charged by mutual funds are also, as noted previously, subject to regulation. The foundation of this regulatory power is the government's de facto role as arbiter of costs of transactions regarding securities in general. For example, the SEC and the NASD have established rules as part of the overall guide to fair dealing with customers about the markups dealers can charge financial institutions on the sale of financial assets. The SEC set a limit of 8.5% on a fund's load but allows the fund to pass through certain expenses under the 12b-1 rule, as indicated below. On July 1, 1993, the SEC amended the rule to set a maximum of 8.5% on the total of all fees, inclusive of front-end and back-end loads as well as expenses such as advertising.

Some funds charge a 12b-1 fee, as authorized in the '40 Act and created in 1980. At the time mutual funds were losing money and the SEC allowed funds to charge the fees to pay for marketing and distribution expenses to increase fund assets. This was envisioned as a temporary measure. The 12b-1 fee may be divided into two parts. The first component is a distribution fee, which can be used for fund marketing and distribution costs. The maximum distribution fee is 0.75% (of net assets per year). The second is a service fee (or trail commission), which is used to compensate the sales professionals for their ongoing services. The maximum service fee is 0.25%. Thus, the maximum 12b-1 fee is 1%. While no-load funds can have 12b-1 fees, the practice has been that in order to call itself a no-load fund, its 12b-1 fee must be at most 0.25% (all of which would be a distribution fee). In general, the distribution fee component of the 12b-1 fee is used to develop new customers while the service fee is used for servicing existing customers.

A rule called "prospectus simplification" or "Plain English Disclosure" was enacted on October 1, 1998 to improve the readability of the fund prospectus and other fund documents. According to the SEC, prospectuses and other documents were written by lawyers for other lawyers and not for the typical mutual fund investor. This initiative mandated that prospectuses and other document be written in "plain English" for individual investors. Efforts to simplify fund information continues.

Among the recent SEC priorities that have directly affected mutual funds are:

1. Reporting after-tax fund returns. This requires funds to display the pre-liquidation and post-liquidation impact of taxes on one, five, and ten year returns both in the fund's prospectus and in annual reports. Such reporting could increase the popularity of tax–managed funds (funds with a high tax efficiency).
2. More complete reporting of fees, including fees in dollars and cents terms as well as in percentage terms.
3. More accurate and consistent reporting of investment performance.
4. Requiring fund investment practices to be more consistent with the name of a fund to more accurately reflect their investment objectives. The SEC now requires that 80% of a fund's assets be invested in the type of security that its name implies (e.g., healthcare stocks).
5. Disclosing portfolio practices such as "window dressing" (buying or selling stocks at the end of a reporting period to include desired stocks or eliminate undesired stocks from the reports at the end of the period in order to improve the appeared composition of the portfolio), or "portfolio pumping" (buying shares of stocks already held at the end of a reporting period to improve performance during the period).

Among the current SEC priorities are the following:

- **Reviewing 12b-1 fees.** A common view is that 12b-1 fees no longer solve their original intended function and so they should be altered or eliminated.
- **Considering soft dollars again.** Soft dollars are the use of transaction charges to a dealer to pay for fund expenses. Among the considerations are what type of expense can be paid, for example securities research, computer systems or computers and other real assets.
- **A general SEC topic is disclosure.** Disclosure would pertain to the transparency of fund charges and fees including sales charges on various share classes. Disclosure would also pertain to conflicts of interest which would include selling agreements between funds and distributors and other types of revenue sharing. Some perceive that money is being transferred among the participants in the provision of 401(k).

The SEC is trying to specify the relationship between the providers of transactions services, R/Rs (registered representatives) and investment advice, IARs (investment adviser representatives) and their conflicts and obligations. Specifying which providers have a suitability responsibility and which have a fiduciary responsibility is a current SEC concern.

STRUCTURE OF A FUND

A mutual fund organization is structured as follows:

- A board of directors (also called the fund trustees), which represents the shareholders who are the owners of the mutual fund.
- The mutual fund, which is an entity based on the Investment Company Act of 1940.
- An investment adviser, which manages the fund's portfolios and is a registered investment adviser (RIA) according to the Investment Adviser's Act of 1940.
- A distributor or broker/dealer, which is registered under the Securities Act of 1934.
- Other service providers, both external to the fund (the independent public accountant, custodian, and transfer agent) and internal to the fund (marketing, legal, reporting, etc.).

The role of the board of directors is to represent the fund shareholders. The board is composed of both "interested" (or "inside") directors who are affiliated with the investment company (current or previous management) and "independent" (or "outside") directors who have no affiliation with the investment company. Currently, regulations require that more than half of the board be composed of independent directors and that the chairperson can be either an interested or independent director.

The mutual fund enters into a contract with an investment adviser to manage the fund's portfolios. The investment adviser can be an affiliate of a brokerage firm, an insurance company, a bank, an investment management firm, or an unrelated company.

The distributor, which may or may not be affiliated with the mutual fund or investment adviser, is a broker-dealer.

The role of the custodian is to hold the fund assets, segregating them from other accounts to protect the shareholders' interests. The transfer agent processes orders to buy and redeem fund shares, transfers the securities and cash, collects dividends and coupons, and makes distributions. The independent public accountant audits the fund's financial statements.

RECENT CHANGES IN THE MUTUAL FUND INDUSTRY

There have been several significant recent changes in the mutual fund industry in addition to those discussed earlier in this chapter. Next we discuss these changes.

Distribution Channels

As explained earlier in this chapter, at the beginning of the 1990s there were two primary distribution channels, direct sales to investors and sales through brokers. Since then, fund companies and fund distributors developed and expanded sales channels beyond the two traditional channels. By the end of the 1990s, fund companies' use of multiple distribution channels resulted in a blurring of the distinction between direct and sales-force funds that had characterized funds at the beginning of the decade.

Fund companies and distribution companies developed new outlets for selling mutual funds and expanded their traditional sales channels. The changes that occurred are evident in the rising share of sales through third parties and intermediaries. Significant market trends account for these changes. In particular, many funds that had previously marketed only directly turned increasingly toward third parties and intermediaries for distribution.

Like direct-market funds, funds that were traditionally sold through a sales force moved increasingly to nontraditional sources of sales such as employer-sponsored pension plans, banks and life insurance companies, and fee-based advisers during the 1990s.

Below we describe the various nontraditional distribution channels.

Supermarkets

The introduction of the first mutual fund supermarket in 1992 marked the beginning of a significant change in the distribution of direct market funds. Specifically, during 1992, Charles Schwab & Co. introduced its *OneSource* service. With this and other supermarket programs, the organizer of the supermarket offers no-load funds from a number of different mutual fund companies. These supermarkets allow investors to purchase funds from participating companies without investors having to contact each fund company. The organizer of the supermarket also provides the investor with consolidated recordkeeping and a simple account statement.

These services provide a non-transaction-fee (NTF) program to provide access to multiple fund families under one roof and to help service the back-office needs of financial advisers. Through this service, investors can access many mutual fund families through one source and buy all the funds with no transaction fee (that is, no load).

On the one hand, these services make a mutual fund family more accessible to many more investors. On the other hand, they break the direct link between the mutual fund and the investor. According to these services, the mutual fund company does not know the identity of its investors through the supermarkets; only the supermarket, which distributes the funds directly to the investor knows their identity. These supermarkets fit the needs of fee-based financial planners very well. For individual investors and planners as well, supermarkets may offer one-stop shopping including the current "best of the breed."

Wrap Programs

Wrap accounts are managed accounts, typically mutual funds "wrapped" in a service package. The service provided is often asset allocation counsel; that is, advice on the mix of managed funds. Thus, mutual fund wrap programs provide investors with advice and assistance for an asset-based fee rather than the traditional front-end load. Wrap products are currently offered by many fund and nonfund companies. Wrap accounts are not necessarily alternatives to mutual funds, but may be different ways to package the funds.

Traditional direct market funds as well as sales-force funds are marketed through this channel.

Fee-Based Financial Advisers

Fee-based financial advisers are independent financial planners who charge investors a fee, typically as a percentage of assets under management, or an hourly charge. In return, they provide investment advice to their clients by selecting portfolios of mutual funds, ETFs, and securities. While many planners recommend mutual funds to their clients, others recommend portfolios of planner-selected securities.

Variable Annuities

Variable annuities represent another distribution channel. Variable annuities are "mutual funds in an insurance wrapper." Among their insurance features are the tax deferral of investment earnings until they are withdrawn, and higher charges (including a mortality charge for an insurance feature provided). Variable annuities are sold through insurance agents and other distributors as well as directly through some fund companies.

"Mix and Match" (Open Architecture)

Until recently, fund manufacturers distributed only their own funds; fund distributors distributed only one manufacturer's funds; and typically employer defined contribution plans, such as 401(k)s, offered funds from only one distributor. However, the investors' demands for choice and convenience, and also the distributors' need to appear independent and objective, have incented essentially all institutional users of funds and distribution organizations to offer funds from other fund families in addition to their own (that is, if they also manufacture their own funds). In addition, mutual fund supermarkets distribute funds of many fund families with considerable facility and low costs. Many fund families offer funds from other families. When a distributor or distribution system sells the investment products of many mutual fund families, it is referred to as "open architecture."

The balance of power between fund manufacturers and distributors currently significantly favors distribution. That is, in general there are more funds available than distributors to sell them. In the mutual fund business, "distribution is king."

MUTUAL FUNDS VERSUS EXCHANGE-TRADED FUNDS

While mutual funds have become very popular with individual investors during the 1980s and 1990s, they are often criticized for two reasons. First, mutual funds shares are priced at, and can be transacted only at, the end-of-the-day (closing) price. Specifically, transactions (that is, purchases and sales) cannot be made at intraday prices, but only at the end-of-the-day closing prices. The second issue relates to taxes and the investors' control over taxes. As noted earlier in this chapter, withdrawals by some fund shareholders may cause taxable realized capital gains for shareholders who maintain their positions and in some cases even if they have held them for a few days.

During 1993, a new investment vehicle which has many of the same features of mutual funds but responds to these two limitations was introduced. This investment vehicle,

Table 60.3 Mutual Funds versus Exchange-Traded Funds

	Mutual Funds	ETFs
Variety	Wide choice, passive and active portfolios	Choices currently limited to passive indexes.
Taxation	Subject to taxation on dividend and realized capital gains.	Subject to taxation on dividend and realized capital gains.
	May have gains/losses when other investors redeem funds.	No gains/losses when other investors redeem funds.
	May have gains/losses when stocks in index are changed.	May have gains/losses when stocks in index are changed.
Valuation	NAV based on actual stock market prices.	Creations and redemptions at NAV. Secondary market prices may be valued somewhat above or below NAV, but deviation typically small due to arbitrage.
Pricing	End-of-Day	Continuous
Expenses	Low for Index Funds	Low, and in some cases, even lower than for index mutual funds
Transaction Cost	None for no-load funds; sales charge for load funds.	Commission or brokerage
Management Fee	Depends on fund; index funds have a range of management fees.	Depends on fund; tends to be very low on many stock index funds

called exchange-traded funds (ETFs), consists of investment companies that are similar to mutual funds but trade like stocks on an exchange. ETFs are described in more detail in Chapter 61 of Volume I. While they are open-ended, ETFs are, in a sense, similar to closed-end funds which have very small premiums or discounts from their NAV.

ETFs have been based on U.S. and international stock and bond indexes and subindexes. In addition to broad stock indexes, ETFs are also based on style, sector, and industry-oriented indexes.

In an ETF, it is the investment adviser's responsibility to maintain the portfolio such that it replicates the index and the index's return accurately. Because supply and demand determine the secondary market price of these shares, the exchange price may deviate slightly from the value of the portfolio and, as a result, may provide some imprecision in pricing. The deviation will be small, however, because arbitrageurs can create or redeem large blocks of shares on any day at NAV, significantly limiting the deviations.

Along with being able to transact in ETFs at current prices throughout the day comes the flexibility to place limit orders, stop orders, orders to short sell and buy on margin, none of which can be done with open-end mutual funds.

The other major distinction between open-end mutual funds and ETFs relates to taxation. For both open-ended funds and ETFs, dividend income and capital gains realized when the funds or ETFs are transacted are taxable to the investor. However, in addition, when there are redemptions, open-end mutual funds may have to sell securities (if the cash position is not sufficient to fund the redemptions), thus causing a capital gain or loss for those who held their shares, while ETFs do not have to sell portfolio securities since redemptions are effected by an in-kind exchange of the ETF shares for a basket of the underlying portfolio securities—not a taxable event to the investors according to the IRS. Therefore, investors in ETFs are subject to significant capital gains taxes only when they sell their ETF shares (at a price above the original purchase price). However, ETFs do distribute cash dividends and may distribute a limited amount of realized capital gains and these distributions are taxable. Overall, with respect to taxes, ETFs, like index mutual funds, avoid realized capital gains and the taxation thereof due to their low portfolio turnover. But, unlike index mutual funds (or other funds for that matter), they do not cause potentially large capital gains tax liabilities which accrue to those investors who hold their positions in order to meet other shareholder redemptions due to the unique way in which they are redeemed.

The pros and cons of mutual funds and ETFs are summarized in Table 60.3. Table 60.4 considers the tax differences

Table 60.4 Taxes: Mutual Funds versus ETFs

	Mutual Funds	ETFs
Holding/Maintaining		
1. Taxes on Dividend, Income, and Realized Capital Gains	Fully Taxable	Fully Taxable
2. Turnover of Portfolio	Withdrawal by other investors may necessitate portfolio sales and realized capital gains for holder.	Withdrawal by others does not cause portfolio sales and, thus, no realized capital gains for holder.
Disposition		
3. Withdrawal of Investment	Capital gains tax on difference between sales and purchase price.	Capital gains tax on difference between sales and purchase price.
4. Overall	May realize some capital gains due to some portfolio turnover.	Will not realize significant capital gains due to very low portfolio turnover.

in more detail. Overall, the ETFs have the advantages of intraday pricing and tax management, and many, but not all, have lower expenses than their corresponding index mutual funds. However, since open-ended funds are "transacted" through the fund sponsor and ETFs are traded on an exchange, the commissions on each ETF trade may make them unattractive for a strategy that involves several small purchases, as for instance, would result from strategies such as dollar cost averaging or monthly payroll deductions. However, ETFs may provide a viable alternative to mutual funds for many other purposes.

REFERENCES

Inro, D. C., Jaing, C. X., Ho, M. Y. and Lee, W. Y. (1999). Mutual fund performance: Does fund size matter? *Financial Analysts Journal*, May/June: 74–87.

O'Neal, E. S. (1999). Mutual fund share classes and broker incentives. *Financial Analysts Journal*, September/October: 76–87.

Reid, B. (2000). The 1990s: A decade of expansion and changes in the U.S. mutual fund industry. *Perspectives: Investment Company Institute* 6, 3: 1–20.

CHAPTER 61

Exchange-Traded Funds

GARY L. GASTINEAU
Managing Director, ETF Consultants, LLC

Abstract: Exchange-traded funds are the most popular examples of a category of financial instrument that might be characterized as a "portfolio-in-a-single-share". In addition to open-end exchange-traded funds based on the SPDR structure, closed-end funds, HOLDRs, exchange-traded notes and even FOLIOs sometimes compete in the portfolio-as-a-share market. While the products all feature multiple instruments in a single transaction, these products and structures have distinct differences in tax treatment, trading costs and convenience. The open-end exchange-traded fund structure offers unique opportunities for increased shareholder efficiency and the delivery of actively managed portfolios in a tax-efficient format. The genesis of exchange-traded funds was in portfolio or program trading and its cousin, index arbitrage.

Keywords: exchange-traded funds (ETFs), mutual funds, portfolio trading, program trading, SPDRs, HOLDRs, exchange-traded notes (ETNs), exchange of futures for physicals (EFP), TIPs, WEBS, expense ratio, closed-end funds, open-end funds, FOLIOs, creation units, arbitrageur, creation, redemption, tax efficiency, shareholder accounting, Regulated Investment Company (RIC), grantor trust, structured product, special purpose vehicle (SPV), actively managed ETFs

Exchange-traded funds (ETFs) are the most important—and potentially the most versatile—financial instruments introduced since the debut of financial futures a generation ago. We begin this chapter by explaining the origins of ETFs and some of their important features like intra-day trading on a stock exchange, creation and redemption of fund shares "in-kind," and tax efficiency. We also compare the recently popular open-end ETFs to competitive products like closed-end funds, conventional mutual funds, HOLDRs, exchange traded notes (ETN), and Folios in terms of costs, and applications. Advocates of conventional mutual funds, ETFs, and separate stock portfolios (including HOLDRs and Folios) have engaged in extensive discussions about the relative tax-efficiency of their respective approaches to equity portfolio management. (For a discussion of the principal tax-efficiency issues, see Gastineau [2002].)

THE HISTORY AND STRUCTURE OF ETFs AND SOME COMPETITORS

Exchange-traded funds, referred to by friends and foes alike as "ETFs," are outstanding examples of step-by-step evolution of new financial instruments starting with a series of proto-products that led in a natural progression to the current generation of exchange-traded funds and set the stage for products yet to come. (A more detailed discussion appears in Gastineau [2001].)

Portfolio Trading

The basic idea of trading an entire portfolio in a single transaction did not originate with the TIPS or SPDRS, which are the earliest successful examples of the modern portfolio-traded-as-a-share structure. The idea originated with what has come to be known as "portfolio trading" or "program trading." In the late 1970s and early 1980s, program trading was the then revolutionary ability to trade an entire portfolio, often a portfolio consisting of all the S&P 500 stocks, with a single order placed at a major brokerage firm. Some modest advances in electronic order entry technology at the NYSE and the Amex and the availability of large order desks at some major investment banking firms made these early portfolio or program trades possible. At about the same time, the introduction of S&P 500 index futures contracts at the Chicago Mercantile Exchange provided an arbitrage link between the futures contracts and the traded portfolios of stocks. It even became possible, in a trade called an *exchange of futures for physicals* (EFP) to exchange a stock portfolio position, long or short, for a stock index futures position, long or short. The effect of these developments was to make portfolio trading either in cash or futures markets an attractive activity for many trading desks and for many institutional investors.

As a logical consequence of these developments affecting large investors, there arose interest—one might even say insistent demand—for a readily tradable portfolio or basket product for smaller institutions and the individual investor. Before the introduction of "mini" contracts, futures contracts were relatively large in notional size. Even with "mini" contracts, the variation margin requirements for carrying a futures contract are cumbersome and relatively expensive for a small investor. Perhaps even more important, there are many more securities salespeople than futures salespeople. The need for a security—that is, an SEC-regulated portfolio product—that could be used by individual investors was apparent. An important predecessor came from Canada.

Toronto Stock Exchange Index Participations (TIPs)

TIPs were a warehouse receipt-based instrument designed to track the TSE-35 index and a later product tracked the TSE-100 index as well. The TSE-100 product was initially called HIPs. These products traded actively and attracted substantial investment from Canadians and from international indexing investors. TIPs were unique in their expense ratio. The ability of the trustee (State Street Bank) to loan out the stock in the TIPs portfolio and frequent demand for stock loans on shares of large companies in Canada led to what was, in effect, a negative expense ratio at times.

The TIPs were a victim of their own success. They proved costly for the Exchange and for some of its members who were unable to recover their costs from investors. Early in 2000, the Toronto Stock Exchange decided to get out of the portfolio share business and TIPs positions were liquidated or rolled into a Barclays Global Investors (BGI) 60 stock index share at the option of the TIPs holder. The BGI fund was relatively low cost, but not as low cost as the TIPs, so a large fraction of the TIPs shares were liquidated.

Standard & Poor's Depository Receipts (SPDRs)

SPDRs (pronounced "spiders"), developed by the American Stock Exchange (Amex), are the shares of a unit trust which holds an S&P 500 portfolio that, unlike the portfolios of most U.S. unit trusts, can be changed as the index changes. The reason for the selection of the unit trust structure was the Amex's concern for simplicity and costs. A mutual fund must pay the costs of a board of directors, even if the fund is very small. The Amex was uncertain of the demand for SPDRs and did not want to build a more costly infrastructure than was necessary. SPDRs traded reasonably well on the Amex in their earlier years, but only in the late 1990s did SPDRs asset growth become truly exponential. Investors began to look past the somewhat esoteric in-kind share creation and redemption process (used by market makers and large investors to acquire and redeem SPDRs in large blocks) and focused on the investment characteristics and tax efficiency of the SPDRs shares.

Today, the S&P 500 SPDRs have more assets than any other index fund except the Vanguard 500 mutual fund. The SPDRs account for less than one-sixth of ETF assets in the United States. Interestingly, however, from 70% to 90% of traditional U.S. index fund money goes into S&P 500 portfolios. Clearly, the interest in ETFs based on indexes other than the S&P 500 suggests that there is more to ETFs than an alternative to conventional index funds. (For specific analysis of the S&P 500 SPDRs see Elton, Gruber, Comer, and Li [2002].)

World Equity Benchmark Shares (WEBS)—Renamed iShares MSCI Series

The *WEBS*, originally developed by Morgan Stanley, are important for two reasons. First, they are foreign index funds. More precisely, they are U.S.-based funds holding stocks issued by non-U.S.-based firms. Second, they are one of the earliest exchange-traded index products to use a mutual fund as opposed to a unit trust structure. The mutual fund structure has more investment flexibility and there are some other differences in dividend reinvestment and stock lending. We would expect most new funds to use the mutual fund structure.

In addition to WEBS, a variety of additional ETF products are now available. The Mid-Cap SPDRs (a unit trust run by the Bank of New York) actually came before WEBS, and the DIAMONDS (a unit trust based on the Dow Jones Index Industrial Average and run by State Street Bank) and the Nasdaq 100 (a unit trust run by the Bank of New York) were introduced later. The Select Sector SPDRs used a mutual fund structure similar to the WEBS and were introduced in late 1998.

Barclays Global Investors, a major institutional index portfolio manager, launched iShares (mutual fund type ETFs based on a large number of benchmark indexes) in a bid to develop a retail branded family of financial products. The streetTRACKS Funds (another group of mutual fund type ETFs) represent State Street's first solo ETF effort in the United States. State Street is also behind the Hong Kong TraHKers Fund and other funds for investors outside the United States. (For a slightly different perspective of the ETF landscape with more data on individual funds, see Fredman [2001a].) The fund roster continues to grow.

ETFs AND OTHER TRADABLE BASKET PRODUCTS

While most readers think of the fund products described above as ETFs, various financial instruments, each referred to by some of its advocates as an exchange-traded fund, are designed to meet specific portfolio investment needs. In many cases, the needs met are practically identical; in other cases, they are quite different. In spite of some confusion about what the term ETF includes, most observers agree that a range of exchange-traded portfolio basket products compete for investors' dollars.

Our purpose in this section is to introduce the major categories of financial instruments which sometimes have been called "ETFs" or which compete with ETFs. We will appraise the features of each. Our objective is to provide a relatively straightforward comparison of features. The purpose of the comparison is not to suggest that one structure is always superior or that the emphasis should always be on competition between the products. In fact, folio customers have been important users of the fund-type ETFs described in the previous section and of HOLDRs which are described below.

Closed-End Funds

Nuveen Investments began using the term "exchange-traded funds" for its closed-end municipal bond funds traded on the New York and American Stock Exchanges in the very early 1990s, several years before the first SPDRs began trading on the American Stock Exchange. The use of the name "exchange-traded funds" was selected to emphasize the fact that someone buying and selling these municipal bond fund shares enjoyed the investor protections afforded by investment company (fund) regulation and by the auction market on a major securities exchange.

"Open" Exchange-Traded Funds

The SEC requires that references to what we have been calling exchange-traded funds as open-end funds be made only in the context of a comparison with conventional open-end investment companies (mutual funds). We are about to make such a comparison so we will drop the quotes around open, and fully qualify the limits of openness in such funds. Shares in open ETFs are issued and redeemed directly by the fund at their net asset value (NAV) only in creation unit aggregations, typically 50,000 fund shares or multiples of 50,000 shares. The shareholder who wants to buy or sell fewer than 50,000 shares may only buy and sell smaller lots on the secondary market at their current market price. The secondary market participant is dependent on competition among the exchange specialist, other market makers and arbitrageurs to keep the market price of the shares very near the intra-day value of the fund portfolio. The effectiveness of market forces in promoting tight bid asked spreads and fair pricing has been impressive. ETF shares have consistently traded very close to the value of the underlying portfolio in a contemporaneously priced market.

For the typical retail or even institutional investor, purchasing and selling ETF shares is the essence of simplicity. The trading rules and practices are those of the stock market. ETF shares are purchased and sold in the secondary market, much like stocks or shares of closed-end funds, rather than being purchased from the fund and resold to the fund, like conventional mutual fund shares.

Because they are traded like stocks, shares of ETFs can be purchased or sold any time during the trading day, unlike shares of most conventional mutual funds which are sold only at the 4:00 p.m. net asset value (NAV) as determined by the fund and applied to all orders received since the prior day's share trading deadline. While the opportunities for intra-day trading may not be important to every investor, they certainly have appeal to many investors during a period when there is concern about being able to get out of a position before the market close when prices are volatile.

Primary market transactions in ETF shares, that is, trades when shares are bought and redeemed with the fund itself as a party to the trade, consist of in-kind creations and redemptions in large size. There have been occasions when creation and redemption of fund shares has resulted in asset flows of billions of dollars in or out of the SPDR or the Nasdaq 100 Trust in a single day. Exchange specialists, market makers, and arbitrageurs buy ETF shares from the fund by depositing a stock portfolio and a cash balancing component that essentially match the fund in content and are equal in value to, say, 50,000 ETF shares on the day the fund issues the shares. The same large market participants redeem fund shares by tendering them to the fund in 50,000 share multiples and receiving a stock portfolio plus or minus balancing cash equivalent in value to the 50,000 ETF shares redeemed. The discipline of possible creation and redemption at each day's market closing NAV is a critical factor in the maintenance of fund shares at a price very, very close to the value of the fund's underlying portfolio, not just at the close of trading, but intra-day. A proxy for intra-day net asset value per share is disseminated for each ETF throughout the trading day to help investors check the reasonableness of bids and offers on the market. (This proxy value does not have the status of a formal NAV calculation.)

An extremely important feature of the creation and, more particularly, the redemption process is that redemption-in-kind does more than provide an arbitrage mechanism to assure a market price quite close to net asset value. Redemption in kind also reduces the fund's transaction costs and enhances the tax efficiency of the fund. While a conventional mutual fund can require shareholders to take a redemption payment in-kind rather than in cash for large redemptions, most funds are reluctant to do this, and most shareholders hold fund positions considerably smaller than the $250,000 minimum usually required for redemption in-kind. As a consequence, most redemptions of conventional mutual fund shares are for cash, meaning that an equity fund faced with significant shareholder redemptions is required to sell shares of portfolio stocks, frequently shares that have appreciated from their original cost. When gains taken to obtain cash for redemptions are added to gains realized on merger stocks that are removed from the index for a premium over the fund's purchase price, many conventional index funds distribute substantial capital gains to their shareholders, even though the continuing shareholders who pay taxes on these distributions have made no transactions, and the fund, looked at from a longer perspective, has been a net buyer of its index's component securities.

The in-kind redemption process for exchange-traded funds enhances tax efficiency in a simple way. The lowest cost shares of each stock in the portfolio are delivered against redemption requests. In contrast to a conventional fund which would tend to sell its highest cost stocks first, leaving it vulnerable to substantial capital gains realizations when a portfolio company is acquired at a premium and exits the index and the fund, the lowest cost lot of stock in each company in the portfolio is tendered to ETF shareholders redeeming in multiples of 50,000 fund shares. The shares of stock in each company remaining in the portfolio have a relatively higher cost basis, which means that acquired companies generate smaller or no gains when they leave the index and are sold for cash by the fund.

One further feature of the existing exchange-traded funds which causes a degree of misunderstanding and which seems to create an expectation that all ETFs will be extremely low cost funds requires an explanation. First, the existing ETFs are all index funds. Index funds generally have lower management fees than actively-managed funds, whatever their share structure. Second, ETFs enjoy somewhat lower operating costs than their conventional fund counterparts. The principal reasons for lower costs are (1) the opportunity to have a somewhat larger fund because of the popularity of the exchange-traded fund structure, (2) slightly lower transaction costs due to in-kind deposits from and payments to buyers and redeemers in the primary market and, most importantly, (3) the elimination of the transfer agency function—that is, the elimination of shareholder accounting—at the fund level.

As all U.S. ETFs are "book entry only" securities, an exchange-traded fund in the United States has one registered shareholder: the Depository Trust Company (DTC). If you want a share certificate for a SPDR or QQQ position, you are out of luck. Certificates are not available. The only certificate is held by the Depository Trust Company, and the number of shares represented by that certificate is "marked to market" for increases and decreases in shares as creations and redemptions occur.

Shareholder accounting for ETFs is maintained at the investor's brokerage firm, rather than at the fund. This creates no problems for the shareholder, although it does have some significance for the distribution of exchange-traded funds. One of the traditional functions of the mutual fund transfer agent is to keep track of the salesperson responsible for the placement of a particular fund position, so that any ongoing payments based on 12b-1 fees or other marketing charges can be made to the credit of the appropriate salesperson. There is no way for the issuer of an ETF to keep track of salespeople because these fund positions do not carry the record keeping information needed to use the DTC Fund/SERV process. They are, in a word, just like shares of a stock—and a stock with no certificates at that. The elimination of the individual shareholder transfer agency function reduces operating costs by a minimum of five basis points and probably by much more in many cases. ETF expenses tend to reflect the cost savings on this function.

The trading price of an exchange-traded fund share will be subject to a bid-asked spread in the secondary market (although these are very narrow on most products) and a brokerage commission. A simple break-even analysis divides the round-trip trading costs by the daily difference in operating expenses. Anyone planning to retain a reasonably large fund position for more than a short period of time and/or anyone who values the intra-day purchase and sale features of the exchange-traded funds will find the combination of the lower expense ratio and greater flexibility make the ETF share more attractive than a conventional mutual fund share. New delivery systems developed for 401(k) accounts will reduce most small lot ETF trading costs.

Powerful advantages notwithstanding, there are a few disadvantages in the exchange-traded fund format for some investors. An investor cannot be certain of his or her ability to buy or sell shares at a price no worse than net asset value without incurring some part or all of a trading spread and a commission. (The specialized index ETFs introduced after 2004 are not enhanced index funds. The latter track a specified benchmark closely using optimization and other quantitative techniques to improve return and/or reduce risk.) It is the trading spread in the secondary market which covers the costs of insulating the ongoing shareholder from the cost of in-and-out transactions by active traders. These transaction costs in open market ETF trades mean that, even with lower fund expenses, certain small investors will not find ETFs as economical as traditional funds if they are in the habit of making periodic small investments. Since most conventional mutual funds take steps to refuse investments from in-and-out traders if they trade in and out too frequently, the transaction costs associated with ETFs are simply a more equitable allocation of these costs among various fund shareholders. A long-term investor, particularly a taxable long-term investor will benefit greatly from the exchange-traded fund structure because in the long run that investor should enjoy lower fund expenses and a higher after-tax return than he would find in an otherwise comparable conventional fund. This allocation of costs and benefits is ironic given the only significant criticism which has been leveled at exchange-traded funds, that is, that they encourage active trading. In fact, the long-term taxable investor enjoys the greatest benefits from the ETF structure. Even so, the ETF structure has probably reduced the active trader's costs as well, given the obstacles and special redemption fees these traders often incur when they use conventional funds.

As noted, all current open exchange-traded funds are index funds. As time goes by, there will be a wider variety of funds available. The introduction of enhanced index funds and ultimately actively-managed funds seems inevitable. It is in the advance from simple indexation with full replication of the index in the portfolio that the investment management company structure shows its greatest advantages over the open UIT structure because the latter structure does not provide a mechanism for anything beyond full replication of an index. The open-end management investment company structure permits a portfolio to differ from the structure of an index fairly easily if the index structure is not consistent with the diversification requirements that allow the fund to qualify as a regulated investment company (RIC) for tax purposes. The UIT structure provides for replication of an index with limited variations based on rounding share positions and limited timing adjustments of index replicating transactions by advancing or deferring them for a few days.

Alternative portfolio or basket structures differ both from the UIT and the exchange-traded investment management company. These other structures have their own unique features. Foremost among these are Holding Company Depository Receipts (HOLDRs), a structure pioneered by Merrill-Lynch, and Folios, which have been introduced by a number of firms that would otherwise be characterized primarily as deep discount brokers. Both HOLDRs and Folios are unmanaged baskets of securities which may have an initial structure based on an index, a theme, or just a diversification policy. Exchange-traded notes (ETNs) are used for assets and risk-modified positions not easily accommodated in the other structures.

Holding Company Depository Receipts (HOLDRs)

HOLDRs use a grantor trust structure which makes them similar to the open ETFs discussed above in that additional HOLDRs shares can be created and existing HOLDRs can be redeemed. The creation unit aggregation for the open ETF management company structures is typically 50,000 fund shares and the minimum trading unit on the secondary market is a single fund share. In contrast, the creation unit *and* the minimum trading unit in HOLDRs is generally 100 shares. Most brokerage firms will not deal in fractional shares or odd lots of HOLDRs. (DTC does not transfer fractional shares or fractions of the basic trading unit of a security, which is 100 shares in the case of the HOLDRs. However, some firms use trading and accounting systems that accommodate the New York Stock Exchange's Monthly Investment Plan (MIP). MIP was designed to let investors buy odd lots and fractional shares as a start in owning their share of America. Firms which can accommodate fractional share positions (including Foliofn) see the ability to handle fractional shares as a competitive advantage.) An investor can buy and sell HOLDRs in the secondary market or an existing HOLDRs position can be redeemed (exchanged for its specific underlying stocks). A new HOLDRs position can be created by simply depositing the stocks behind the 100-share HOLDRs unit with the Bank of New York. (The stock basket underlying a 100-share HOLDRs unit will initially consist of whole shares of the component stocks. In the event of a merger affecting one of the companies, any cash proceeds will be distributed. The surviving company's whole shares will usually be retained in the HOLDRs basket.)

The creation/redemption fee for HOLDRs will generally be roughly similar in relative magnitude to the comparable fee on investment company ETFs and the pricing principles and arbitrage pricing constraints operate in a similar way. To the extent that one of the stocks in a HOLDRs basket performs poorly and the investor wants to use the loss on that stock to offset gains elsewhere, the HOLDRs can be taken apart and reassembled without affecting the tax status of any shares not sold. The ability to realize a loss on an individual position may give the HOLDRs structure a slight tax advantage over the investment company-based ETFs. On the other hand, unlike the redemption in-kind of the shares of an open ETF, the HOLDRs structure does not permit elimination of a low-cost position in the HOLDRs portfolio without realization of the gain by the investor.

The principal disadvantages of HOLDRs are that they lack the indefinite life of an investment company and there is no provision for adding positions to offset attrition through acquisitions of basket components by other companies. No HOLDRs component that disappears in a cash merger or bankruptcy can be replaced in the HOLDRs basket. If some stocks do well and others do poorly, there is no mechanism for rebalancing positions.

The HOLDRs share one very important characteristic with the index ETFs: It is frequently less costly to trade the basket in the form of HOLDRs than it is to trade the individual shares, particularly for a small- to mid-sized investor who might be trading odd lots in many of the basket components if HOLDRs or ETFs were unavailable. The grantor trust structure of HOLDRs is also used by a few securitized commodity products. The most prominent of these is the StreetTracks Gold Shares (GLD). In contrast to the HOLDRs, these commodity securitization products trade in single share units and are created and redeemed in much larger than 100-share lots.

Folios

In contrast to the other ETF variations and competitors described here, Folios are not standardized products nor are they investment companies or some kind of trust. They are baskets of stocks that can be modified one position at a time or traded with a single order through a brokerage firm. The firms which advocate and provide Folio baskets for trading do provide semi-standardized baskets—in some cases based on indexes, and in other cases based on a simple diversification rule. In practice, however, each investor's implementation of the Folio basket may be slightly different.

Because Folio baskets will not be standardized, Folios cannot be traded like fund shares or like HOLDRs. Each of the stocks in a Folio will trade separately. While the brokerage firm can provide low-cost commissions and even the opportunity to execute trades against its other customer trades at selected times during the day, if the basket does not trade as a standardized basket, the investor will miss some of the transaction cost advantages which traders in standardized basket shares often enjoy.

A tax advantage of Folios over investment companies in certain circumstances is similar to a tax feature of HOLDRs. An investor can sell one position out of a Folio to take a loss and use that loss to offset gains obtained elsewhere—outside the Folio basket. In contrast, a fund taxed as a regulated investment company cannot pass losses through to shareholders. If the fund experiences large losses, an investor can take a loss on the fund shares by selling the share position; but losses on an individual portfolio component are not available to the investor who continues to hold the shares as a passthrough. In a reasonably bullish market environment, the ability of the UIT or management company ETF to modify its portfolio with creations and redemptions without taxable gain realizations will probably be more important to an individual investor than the ability to take specific losses in either HOLDRs or Folios. Other market environments may make the selected loss realization opportunity of the HOLDRs or Folios more valuable.

In contrast to the ETFs' fund structure, there is no "tax-realization-free" mechanism for reducing the impact of a very successful position in either HOLDRs or Folios. In the regulated investment company structures (exchange-traded unit trusts or funds), tax rules would limit the size of any single stock to 25% of the assets of the fund under most circumstances. Reductions in the commitment to a particular position in a regulated investment company with redemptions in-kind might be obtainable without realization of taxable gains. This would not be possible for very successful positions underlying HOLDRs or for components of a Folio. Basket mechanisms that do not offer a way to reduce a large, successful position without capital gains realization force the investor to choose between tax deferral and diversification.

Exchange-Traded Notes (ETNs) and Other Structured Products

Exchange-traded notes and other structured products were introduced long before the earliest products now called ETFs were available to investors in the United States or other major markets. Some publicly traded structured products are based on special purpose vehicles (SPVs) if they require credit enhancement or separation from other entities for credit or regulatory purposes. However, most exchange-traded notes and other structured products are liabilities of major financial institutions and appear on the liability side of corporate balance sheets. The exchange-traded notes closest in structure and function to the open-end ETFs which are the primary focus of this chapter are probably the iPath notes issued by Barclays Bank PLC and marketed through its affiliate, Barclays Global Investor Services.

In contrast to the daily creation and redemption of ETFs, open-end exchange-traded notes are redeemable either weekly or monthly in most cases. The open-end note creation baskets are typically comparable in value to ETF creation or redemption baskets. Depending on the nature of the note and the underlying index portfolio or commodity exposure it provides, an ETN may pay an interest-like payment or be marked to market daily and purchased and sold on the basis of its net asset value. While the expiration dates of exchange-traded notes may be distant, a maturity date more than 30 years in the future is very rare. Not all ETNs have explicit management fees or expenses analogous to the expenses of an exchange-traded fund. Consequently, understanding the economics of an ETN can be a complex exercise. The flexibility of ETN and structured products formats has made these instruments increasingly popular with many investors.

The relatively recent introduction and popularity of open-end ETNs suggests opportunities for considerable growth. The principal constraint on growth is that, in contrast to a fund or a grantor trust where the underlying portfolio is held for the benefit of investors, the value of ETNs and most other structured products is highly dependent on the credit of the issuer. While lower rated financial service firms frequently "rent" the balance sheets of more highly rated firms to issue structured products, credit evaluation is always a significant consideration in any decision to use ETNs or other structured products.

A Side-by-Side Comparison of Tradable Basket Products

Table 61.1 provides an eclectic comparison of the mutual fund-style and UIT-style versions of open exchange-traded funds and conventional mutual funds to the other basket products we have discussed. Most of the items on this comparison table are relatively straightforward and readily understandable from the previous text, but several items do require some discussion. (For a slightly different but useful perspective, see Fredman [2001b].)

In assigning tax-efficiency ratings, we have placed significantly greater value on the redemption in-kind feature of the open ETFs and open UITs than on the separable loss feature available in Folios with no particular change and in HOLDRs through the exchange of the HOLDR for the basket of underlying securities followed by realization of the loss, reestablishment of the position that incurred the loss after the wash sale period is past and reconstitution of the HOLDR—a relatively complex and non-user-friendly process. Open ETNs vary in tax efficiency, but most provide substantial tax deferral.

Closed-end funds are rated higher than conventional mutual funds on tax efficiency because they are

Table 61.1 Basket Product Comparisons

Feature/Product Structure	Open ETFs	Open UITs	Conventional Mutual Funds	Closed ETFs	HOLDRs	Folios	Open ETNs
Creation of shares—primary market	In-kind deposit	In-kind deposit	Cash deposit with fund	IPO	IPO/in-kind deposit	NA	Cash or in-kind deposit
Purchase of shares—secondary market	Open market purchase	Open market purchase	NA	Open market purchase	Open market purchase	Open market purchase	Open market purchase
Sale of shares—secondary market	Open market sale	Open market sale	NA	Open market sale	Open market sale	Open market sale	Open market sale
Redemption of shares—primary market	In-kind redemption	In-kind redemption	Cash redemption	NA	In-kind redemption	NA	Cash or in-kind redemption
Underlying portfolio structure (available today)	Index	Index	Index or managed	Managed	Preset basket	Investor's choice	Index, commodities, option features
Tax structure	RIC	RIC	RIC	RIC	Structure is tax transparent	No structure	Typically, prepaid contracts
Tax-efficiency factors	Redemption in-kind	Redemption in-kind	Cash redemption	Cash redemption	Separable losses	Separable losses	Tax deferral until sale
Investor tax-efficiency rating	1	1	5	4	3	2	1
Effect of structure on shareholder's trading cost	Usually reduces	Usually reduces	Usually reduces	Usually reduces	Usually reduces	No effect except discount brokerage	Usually reduces
Investor's trading cost rating	1	1	3	1	1	3	2
Shareholder attention required	Minimal dividend reinvestment	Minimal dividend reinvestment	Minimal dividend reinvestment	Minimal dividend reinvestment	Dividend reinvestment principal reinvestment tax loss sales and replacements significant	Dividend reinvestment principal reinvestment tax loss sales and replacements significant	Minimal, some structures require reinvestment

Ratings: 1 = best, 5 = worst. See text for discussion.

characterized by a closed portfolio and do not face the forced realization of gains which can come about through cash redemptions in an open-end mutual fund.

The investor's trading cost ratings are based largely on the advantages associated with trading a basket at the share level versus transacting separately in all the securities making up the basket. All of the standardized ETFs are ranked highly because trading in the composite share should be more efficient than trading in the underlying positions separately. It is certainly possible to differentiate among individual products in terms of the cost of trading the product or trading the underlying securities separately, but the difference is more related to the nature of the underlying market and the quality of the market in the basket product than it is on anything systematically related to the product structure. The assets underlying most ETNs are more costly to trade than the average ETF basket. The conventional mutual funds are rated slightly below the exchange-traded products other than the unstructured Folios on the assumption that, on average, a redemption charge or other obstacles to short-term trading will increase an investor's costs of trading. (An investor can do an in-and-out trade in some conventional mutual funds with almost no transaction cost, but many funds will probably not accept a repeat order from that investor.) In any event, the free liquidity mutual funds offer traders is an ongoing trading cost borne by all mutual fund investors. Folios are rated less favorably on trading cost simply because they do not provide any of the advantages associated with trading the other products as portfolios or baskets. Even when the transactions in a Folio are aggregated, each stock is traded separately. None of the Folio providers have reached a size that permits them to match and offset many customer orders to eliminate the bid-asked spread.

HOLDRs and Folios require somewhat greater investor (or manager) attention than the conventional fund or exchange-traded fund products for at least two reasons: First, to the extent that any of the companies in the HOLDRs or Folios are taken over in a cash acquisition, the shares will automatically be turned into cash and the shareholder will have to deal with reinvestment of the principal. Also, both these less structured products provide for their variety of tax-efficiency by permitting tax loss sales of individual securities. Folios, which are marketed principally as a way to take advantage of the automatic diversification a portfolio of stocks provides, require some kind of replacement or re-balancing activity to maintain a useful degree of diversification. With the other products, either a portfolio manager or the process for weighting or re-weighting the index and insuring regulated investment company diversification compliance in the fund will retain a minimal level of diversification without action by the investor or an advisor employed to manage the investor's position.

Improving ETFs

It is appropriate to look at some new ETF features that will improve the performance of these funds for investors. If any fund is going to serve the interest of its shareholders, the portfolio manager needs to implement portfolio changes without revealing the fund's ongoing trading plans. Whether a fund is attempting to replicate an index or to follow an active portfolio selection or allocation process, portfolio composition changes cannot be made efficiently if the market knows what changes a fund will make in its portfolio before the fund completes its trades. A number of recent studies have highlighted an index composition change problem which many of indexing's strong supporters have been aware of for some time: Benchmark indexes like the S&P 500 and the Russell 2000 do not make efficient portfolio templates. Investors in index funds based on popular, transparent indexes are disadvantaged by the fact that anyone who cares will know what changes the fund must make before the fund's portfolio manager can make them. When transparency means that someone can earn an arbitrage profit by frontrunning a fund's trades, transparency is not desirable.

The cost to ongoing shareholders of preannounced portfolio composition changes in index ETFs must be eliminated. The best way to improve index fund performance is to use silent indexes-—indexes that keep portfolio composition changes confidential until after the fund has traded. This requires radically new procedures for the management of indexes and for the management of index funds.

Everyone seems to agree that actively managed funds require confidential treatment of portfolio composition changes until after the fund has traded. Only recently have investors begun to understand the costs that index transparency imposes on index fund investors. Making portfolio changes confidential and efficient requires changes in the ETF structure and the portfolio trading process.

Many individual investors have a stake in being able to make small, periodic purchases or sales in their fund share accounts. The prototypical investor of this type is the 401(k) investor who invests a small amount in his defined contribution retirement plan every payroll period. The mutual fund industry has developed an elaborate framework which permits small orders for a large number of investors to be aggregated and for cash to enter or leave the fund to accommodate small investors at net asset value. There are ways to modify ETF procedures so that these investors, while paying a little more than they have paid in the past to cover the transaction costs of their entry and exit, will still be accommodated at low cost. The snowballing rush to greater transparency in the economics of defined contribution accounts like 401(k) plans will make fund cost and performance comparisons easier—to the advantage of ETFs. The only "problem" that limits the ability of ETFs to deliver this degree of shareholder protection is that the true transaction costs associated with buying and selling shares of an ETF can be difficult for an investor to determine in advance of trading.

One solution to this problem is a new trading process that increases the transparency of ETF transaction costs and, consequently, improves the ETF structural shareholder protection without compromising the ETF "gold" standard whereby investors entering and leaving the fund pay the costs of their entry and exit. In most discussions of actively managed ETFs, there has been appropriate concern expressed for the cost of achieving enough portfolio transparency to facilitate trading in ETFs without subjecting the fund's trades to the front running risk that all of today's index funds experience. The SEC's Concept Release on actively managed ETFs stressed the importance of finding a solution to this problem. It is now apparent that the manager of an actively managed ETF needs to offer no more information on his portfolio composition and portfolio changes than the manager of a conventional mutual fund must publish today. Funds that do not require the full measure of confidentiality available under today's rules for fund asset disclosure can reduce transaction costs for their entering and leaving shareholders and for market makers by providing more frequent disclosure. But more frequent disclosure is not essential. An investment process that requires the maximum permitted portfolio confidentiality can work well inside an actively-managed ETF.

Fund issuers can build on the compelling advantages of exchange-traded funds to offer better and more varied portfolios. New actively managed and improved index funds can offer their shareholders full protection from the cost of entry and exit by other fund shareholders and the tax efficiency that are inherent in the initial generation of SPDR-style exchange-traded funds.

SUMMARY

This chapter describes the relatively short history of exchange-traded funds and their principal competitors. It provides an analysis of the various products and their investment, tax, legal and structural characteristics. The version of the exchange-traded fund that is based on the investment company structure (which it uses in common with conventional mutual funds) shows great promise for active management in competition with mutual funds.

REFERENCES

Elton, E. J., Gruber, M. J., Comer, G., and Li, K. (2002). Where are the bugs? *Journal of Business*, June: 453–472.

Fredman, A. J. (2001a). An investor's guide to analyzing exchange-traded funds. *AAII Journal*, May: 8–13.

Fredman, A. J. (2001b). Sizing up mutual fund relatives: low-cost alternative investing. *AAII Journal*, July: 9–14.

Gastineau, G. L. (2001a). Exchange-traded funds—An introduction. *Journal of Portfolio Management*, Spring: 88–96.

Gastineau, G. L. (2001b). Silence is golden: The importance of stealth in pursuit of the perfect fund index. *Journal of Indexes*, Second Quarter: 8–13.

Gastineau, G. L. (2002). *The Exchange-Traded Funds Manual*. Hoboken, NJ: John Wiley & Sons.

Gastineau, G. L. (2003). Converting actively-managed mutual funds to ETFs. Unpublished manuscript, ETF Consultants LLC.

Gastineau, G. L. (2004a). Single stock futures on exchange-traded funds. *Futures*, January: 38–41.

Gastineau, G. L. (2004b). Protecting fund shareholders from costly share trading. *Financial Analysts Journal*, May/June: 22–32.

Gastineau, G. L. (2004c). An exchange-traded fund or conventional fund—You can't really have it both ways. *Journal of Indexes*, First Quarter: 32–34.

Gastineau, G. L. (2005). The anatomy of tax efficiency. *Journal of Indexes*, May–June: 12–15.

Lazzara, C. J. (2003). Index construction issues for exchange-traded funds. Unpublished manuscript, ETF Consultants LLC.

CHAPTER 62

Investment-Oriented Life Insurance

FRANK J. JONES, PhD
Professor, Accounting and Finance Department, San Jose State University and Chairman of the Investment Committee, Salient Wealth Management, LLC

Abstract: Insurance and investments are distinct concepts. This distinction leads to the development of various insurance and investment products. In practice, however, there is an overlap between some types of insurance products and investment products. This overlap occurs due partially to specific tax advantages provided to investment-oriented life insurance products. The two major types of investment-oriented life insurance are cash value life insurance and annuities.

Keywords: Pure insurance, investment-oriented insurance, cash value life insurance, whole life, variable life, universal life, variable universal life, annuities, variable annuities, fixed annuities, guaranteed investment contract (GIC), participating policies, general account products, separate account products, immediate annuity, deferred annuity, credited rate, flexible-premium deferred annuity (FPDA), single-premium deferred annuity (SPDA)

This chapter begins with an overview of insurance. The remainder of the chapter considers the major types of investment-oriented life insurance, mainly cash value life insurance and annuities.

INSURANCE

Insurance is defined as a contract whereby one party—the insured—substitutes a small certain cost (the insurance premium) for a large uncertain financial loss based on a future contingent event. Thus, there are two parties to an insurance contract, the insured, who pays the premium and receives protection; and the insurer (or insurance company), which collects the premium and provides the protection.

Most types of insurance provide for a prespecified payment from the insurer to the insured if and when the contingent insured event occurs and otherwise have no value. This is called *pure insurance*. Other types of insurance have a "cash value" even if the contingent event does not occur. This is called *investment-oriented insurance*. The two types of investment-oriented insurance are discussed later.

The major types of insurance, in general, are:

- Life
- Health
- Disability
- Property (home and automobile)
- Liability

Other types of insurance include long-term care, business interruption, and workers' compensation.

Of these types, only life insurance has a cash value form in addition to pure insurance. Cash value life insurance is a very important type of investment-oriented life insurance. Therefore, let's consider life insurance in more detail.

According to a pure life insurance contract, the insurer (the life insurance company) pays the beneficiary of the

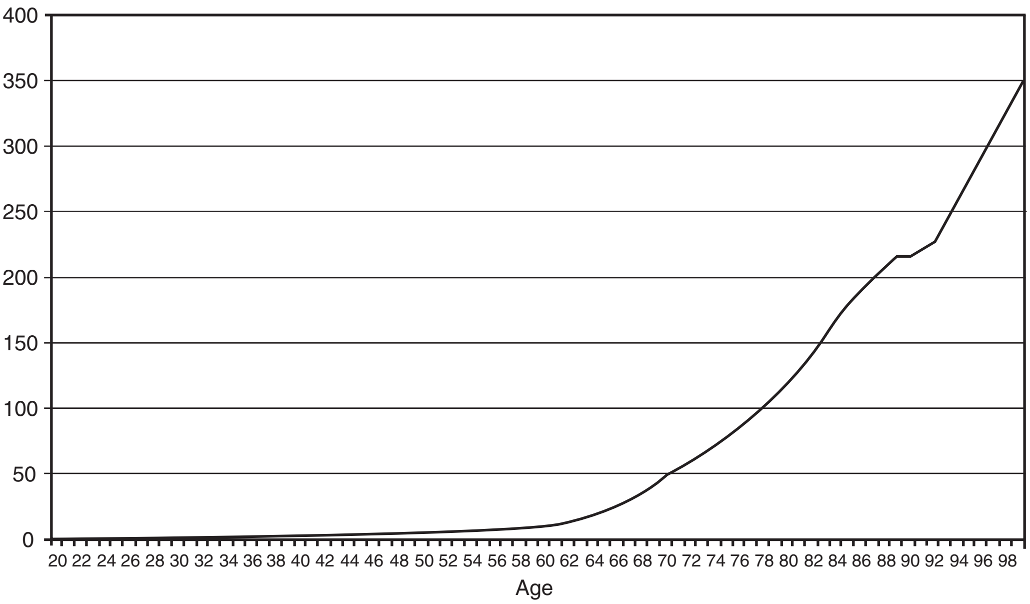


Figure 62.1 Annual Premium for Yearly Renewable Term Policy

contract a fixed amount if the insured dies while the life insurance contract is valid. If the insured does not die while the policy is valid, the insurance contract becomes worthless at its expiration. To provide pure life insurance contracts, the insurance company—specifically its actuaries—calculate the probability of the insured dying during the period the contract is valid. Many variables affect this probability, including physical health, whether the person smokes, and gender. The most important variable, however, is the insured's age. Specifically, the probability of death increases with age. Actuaries estimate this relationship with some degree of precision.

Obviously the insurance premium charged by the insurance company must cover the average amount paid to all insureds, the administrative and distribution costs, and a profit. The cost of pure insurance to the company depends on the probability of the insured dying during the period which increases with age.

Overall, the premium charged to the insured for a pure life insurance contract is shown in Figure 62.1, which is determined by the probability of death (the "cost" of paying the death benefit) plus the distribution and administrative costs plus the profits. Pure life insurance is called term insurance. It is applicable over the term of the policy.

There are three types of term insurance. The most common type is called annual renewable term. According to this type, the insured has the right to renew the coverage every year without new underwriting (that is, without a new medical examination). Premiums, however, change; that is, they increase each year and become very expensive at older ages, as indicated below. A second type of term insurance, much less common, does not have the guaranteed renewability feature of the above.

The third type of term insurance is level-premium term, wherein the premium is constant during the life of the policy. Its level is higher than for annual renewable term early in the policy. However, the premium does not increase with age and is lower than an annual renewable term policy late in the life of the policy. Typically, policies of ten years or more are written on a level-premium term.

For annual renewable term the annual premium increases significantly with age. Traditional whole life policy premiums are much higher than for term insurance, often ten times higher or more.

The costs of non-life types of pure insurance are determined in a similar manner. However, in other types of insurance, factors other than the age of the insured may be the dominant variables. For example, location may be important in home insurance: It costs more to insure against hurricanes in Miami and Galveston than in Chicago and San Francisco. And it costs more for a young male (age is a also a factor here) than a middle-aged female to buy automobile insurance. For both, however, prior driving record is important.

Conceptual Issues in Risk Management

Consider some conceptual issues regarding risk management from the perspective of the insured and the willingness to provide risk coverage from the perspective of the insurer.

From the perspective of the insured, insurance is a mechanism for managing risk. Individuals experience many types of risk and the manner in which they manage the risk depends on the characteristics of the risk. Two important characteristics of the risk are the severity of the risk (the cost) and the frequency of the risk.

There are, in general, also four different ways to manage the risk. Consider specifically these four ways in the context of managing the risk of fire for a house.

- **Avoidance:** Avoid the risk-producing activity. For example, do not build a house in a hot, dry area.
- **Reduction:** Reduce the risk of an activity. For example, build a house in a hot, dry area, but add a sprinkler system.
- **Retain:** Continue the risk-producing activity, but do not insure the risk, that is, self-insure. For example, build a house in a hot, dry area; do not buy insurance; and be prepared to pay for the house yourself if the house burns.
- **Insure:** Engage in an insurance contract on the risk and pay the premium thereon. For example, buy fire insurance on your house. Fire insurance on a house in a hot, dry area will, however, be expensive.

The way in which an individual manages risk will depend on the characteristics of the risk, as summarized in Table 62.1. That is, insurance is most appropriate when the frequency of the insured event is low and the severity is high. Examples of this type of risk might be a serious automobile accident, your house burning down, or the death of a young person. From the perspective of the insurer, the diversification of the risk is important. The essence of insurance is that the financial burden of the losses suffered by a few is shared among many. Suppose it is estimated that in one year, 100 out of 100,000 homeowners will experience losses caused by fire. This is determined based on data assembled about what happened in the past. Instead of those 100 homeowners bearing the entire financial burden of the losses, the burden is shared among the 100,000 homeowners through premiums for homeowners' insurance, which includes protection against fire losses. It is necessary to be able to estimate in advance with reasonable accuracy the aggregate losses that will be suffered by the 100,000 homeowners.

The statistical concept of the "law of large numbers" is relevant. Considering again life insurance, assume that the probability of death during a 12-month period is 20% for a given age. If only one person of this age is insured, either 0% or 100% of the insureds die, and the insurer experiences either a large loss or a small gain (the premium). But if the insurer insures 100 people of this age at an actuarially determined premium, the insurer is likely to have a profit close to the average profit actuarially expected. The law of large numbers says there is more statistical certainty when a large number of insureds (which are diversified) are involved.

Table 62.1 Treatment of Risk by Type of Risk

	Frequency	
Severity	**High**	**Low**
High[a]	Avoidance or reduction (Insurance is very expensive)[b]	Insurance
Low[c]	Retention or reduction	Retention

[a]When the severity of loss is high, retention is not realistic—another technique is needed.
[b]When the frequency of loss is high and the severity is high, insurance is very expensive.
[c]When the severity of the loss is low, insurance is not needed.

Table 62.2 Investment-Oriented Insurance Product

1. Cash Value Life Insurance
   - Whole Life
   - Variable Life
   - Universal Life
   - Variable Universal Life
2. Annuities
   - Variable
   - Fixed
   - GICs

The correlation or independence of the individual events is also important. For example, providing hurricane insurance to 100 houses in Galveston, Texas, does not benefit from the law of large numbers—either all or none of the houses are likely to experience a hurricane.

In the calculation of premiums, insurers estimate the future based on the past. Insurers need to feel comfortable that their estimates will apply to the future. To calculate the loss component of insurance premiums, insurers multiply their estimates of the probability of future losses times the dollar value of the loss.

Investment-Oriented Life Insurance Products

This chapter does not consider any of the pure insurance products. Rather, it considers only various types of investment-oriented life insurance products. Such products are shown in Table 62.2. Each product is discussed in more detail in this chapter.

There is an important distinction in investment-oriented life insurance with respect to whether the insured or the insurance company bears the investment risk, that is, who gains or loses if the investment experience is greater or less than expected. Table 62.3 segregates the products by who bears the investment risk.

The products in the first column are called "general account products" and those in the second column are called "separate account products." The nature of this distinction is discussed later in this chapter.

In all types of pure insurance, the insurer, that is the insurance company, bears the risk of honoring the contract. That is, it is the obligation of the insurer to deliver the exact amount specified in the insurance contracts. But either the insurance company or the insured may bear the risk of underperforming.

Table 62.3 Types of Investment-Oriented Insurance by Risk Bearer

General Account (Insurer Risk)	Separate Account (Insured Risk)
Whole Life Insurance	Variable Life Insurance
Universal Life Insurance	
GICs	Variable Annuities
Fixed Annuities	

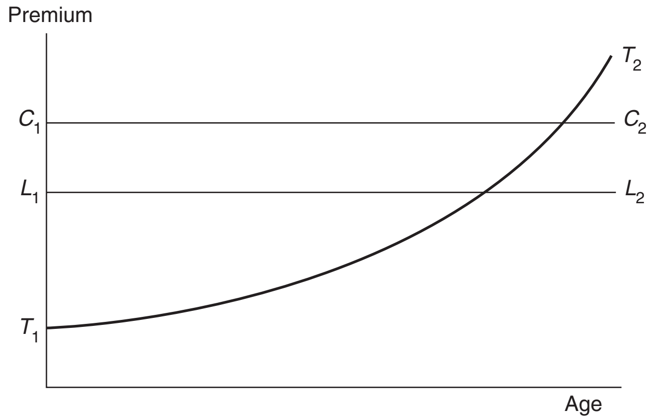


Figure 62.2 Annual Premium for Pure Life Insurance Policy

Cash Value Life Insurance

Consider how *cash value life insurance* relates to the discussion based on Figure 62.1 for pure life insurance. The premium for an annual pure life insurance (term insurance) contract is paid each year for a contract that expires after one year and is shown in Figure 62.1 which is reproduced in Figure 62.2. The annual term insurance premium is denoted by T_1T_2.

This premium, T_1T_2, has two important characteristics:

1. The premium increases each year for a new one-year contract, and increasingly so as age increases.
2. If the insured does not die during the year, the insurance contract expires worthless at the end of the year (and can be replaced by paying a higher premium the next year).

This consideration provides a transition to cash value life insurance. Suppose the insurance company provided pure life insurance for a period much longer than a year, for example the insured's entire life, but charged a constant, called level, premium. In fact, level premium term life insurance is available. In this case, the level premium represents the average premium over the term of the policy. Let L_1L_2 be the level premium of the term life insurance. Such a policy has no cash value.

Second, suppose that the initial (and constant) premium paid is higher than the cost of pure life insurance. This excess premium can be invested and build up cash value during the term of the policy. For example, in Figure 62.2 T_1/T_2 represents the initial cost of annual term insurance, L_1/L_2 the cost of level premium term insurance, and C_1/C_2 the level premium of cash value insurance. The excess amount of premium for cash value insurance over annual term, $C_1 - T_1$ at a young age, in addition to potentially covering the deficit between the cost of pure insurance and cash value insurance at an older age (e.g., $T_2 - C_2$), can be entered into an investment account of the insured. This is the essence of cash value life insurance.

Each year's premium is segregated into two components by the insurance company. The first is the amount needed to pay for the pure insurance, which, as indicated, increases each year. The second goes into the insured's investment account, which is the cash value of the life insurance contract. An investment return is earned on this cash value, which further increases the cash value. The buildup of this cash value and the ability to borrow against it both have tax advantages, as discussed below. Two important observations can be made here.

First, a common marketing or sales advantage attributed to cash value life insurance is that the higher premium paid will "force" the individuals to save, whereas if they did not pay the higher insurance premium, they would use their income for consumption rather than savings. According to this rationale, the higher insurance premium is, thus, forced savings.

Whether or not this first observation has merit, the second observation unequivocally does. The federal government encourages the use of cash value life insurance by providing significant tax advantages. Thus, the second advantage of cash value life insurance is *tax-advantaged* savings.

There are several tax advantages to cash value life insurance. The first and major tax advantage is called "inside buildup." This means that the returns on the investment component of the premium, both income and capital gains, are not subject to taxation (income or capital gains) while held in the insurance contract. Inside buildup is a significant advantage to "saving" via a cash value life insurance policy rather than, for example, saving via a mutual fund.

The second tax advantage of a cash value life insurance policy relates to borrowing against the policy. In general, an amount equal to the cash value of the policy can be borrowed. However, there are some tax implications. The taxation of life insurance is covered in more detail in a following section. In addition to the above, the death benefit, that is the amount paid to the beneficiary of the life insurance contract at the death of the insured, is exempt from income taxes, although it may be subject to estate taxes. This benefit applies both to cash value and pure life insurance.

Term insurance has become much more of a commodity product and, in fact, there are web sites that provide premium quotes for term life insurance for various providers. Cash value life insurance, due to its complexity and multiple features, is not, however, a commodity.

Obviously, the cost of annual term life insurance is much lower than that of whole life insurance, particularly for the young and middle-aged. For example, while there is a wide range of premiums for both term and whole life insurance, for a 35-year-old male, the annual cost of \$500,000 of annual term insurance may be \$400 and the cost of whole life insurance may be \$5,000.

The Nature of Insurance Companies

The nature of an insurance company is quite different than that of a traditional manufacturing company. Consider, for a simple comparison, a bread manufacturing company.

The pricing of bread and the calculation of the profits of a bread manufacturing company are quite simple. The bread manufacturer buys flour and other ingredients, produces the bread with its ovens and bakers, and sells the bread soon thereafter. The costs of the inputs are straightforward (the ovens, of course, must be depreciated) and the revenues are received soon after the costs are incurred. Bread prices may be altered as the costs of the inputs vary. Profits can be measured over short periods of time.

The insurance business is much more complex. Premiums—revenues—are determined initially and may be collected once or over a long period of time. The events that trigger an insurance payout are not only deferred but are also contingent on the occurrence of a specified event, for example death or an automobile accident. Since there is a long and uncertain period between the collection of the premium and the payment of the benefit, the receipts may be invested in the interim and the investment returns represent an important but initially uncertain source of revenue. Insurance company investment practices are not considered in this chapter.

Another important distinction between bread manufacturers and insurance companies is the timing of the claim of the customer on the producing company. The purchaser of a loaf of bread is not concerned about the solvency of the bread manufacturer. The purchaser leaves the store with the bread; that is, the business is "cash and carry."

The purchaser of a life insurance contract, however, has a deferred claim on the life insurance company. This claim may arise decades from the purchase of the life insurance contract. For this reason, the customer is concerned about the long-term solvency of the life insurance company. Rating agencies provide credit ratings on life insurance companies to assist customers in this evaluation. The "claims paying ability," as assessed by these rating agencies, may be an important characteristic to customers in their overall choice of a life insurance company.

In addition, to assure that the insurance company will be able to pay the insurance benefit, if necessary, regulators require that the insurance company retain reserves (in an accounting sense) for the security of future payments. Other accounting complexities are also relevant. Thus, overall, the pricing and measurement of the profits of an insurance company are much more complex than that of a bread manufacturer. And to insure that insurance companies are solvent and pay deferred insurance claims, insurance companies are more regulated than bread manufacturers.

Thus, the fundamental difference between bread manufacturers and life insurance companies is that for bread manufacturers the timing of the costs and revenues is approximately synchronous, while for life insurance companies the timing is potentially very different. There are also significant differences in this regard between annual term insurance and whole life insurance. Companies providing annual term life insurance collect the revenue at the beginning of the year and pay the death benefit by the end of the year, if at all. Companies providing whole life insurance, however, may collect premiums for several years and make a large payment after decades.

Stock and Mutual Insurance Companies

There are two major forms of life insurance companies, stock and mutual. A stock insurance company is similar in structure to any corporation (also called a public company). Shares (of ownership) are owned by independent shareholders and may be traded publicly. The shareholders care only about the performance of their shares, that is the stock appreciation and the dividends over time. Their holding period and, thus, their view may be short term or long term. The insurance policies are simply the products or businesses of the company in which they own shares.

In contrast, mutual insurance companies have no stock and no external owners. Their policyholders are their owners. The owners, that is the policyholders, care primarily or even solely about the performance of their insurance policies, notably the company's ability to eventually pay on the policy and to, in the interim, provide investment returns on the cash value of the policy, if any. Since these payments may occur considerably into the future, the policyholders' view will be long term. Thus, while stock insurance companies have two constituencies, their stockholders and their policyholders, mutual insurance companies only have one, since their policyholders and their owners are the same. Traditionally, the largest insurance companies have been mutual, but recently there have been many demutualizations, that is, conversions by mutual companies to stock companies. Currently several of the largest life insurance companies are stock companies.

The debate on which is the better form of insurance company, stock or mutual, is too involved to be considered in any depth here. However, consider selected comments on this issue. First, consider this issue from the perspective of the policyholder. Mutual holding companies have only one constituency, their policyholder or owner. The liabilities of many types of insurance companies are long term, particularly the writers of whole life insurance. Thus, mutual insurance companies can appropriately have a long time horizon for their strategies and policies. They do not have to make short-term decisions to benefit their shareholders, whose interests are usually short term, via an increase in the stock price or dividend, in a way that might reduce their long-term profitability or the financial strength of the insurance company. In addition, if the insurance company earns a profit, it can pass the profit onto its policyholders via reduced premiums. (Policies that benefit from an increased profitability of the insurance company are called *participating policies*, as discussed later.) These increased profits do not have to accrue to stockholders because there are none.

Finally, mutual insurance companies can adopt a longer time frame in their investments, which will most likely make possible a higher return. Mutual insurance companies, for example, typically hold more common stock in their portfolios than stock companies. However, whereas the long time frame of mutual insurance companies may be construed as an advantage over stock companies, it may also be construed as a disadvantage. Rating agencies and others assert that, due to their longer horizon and their long time frame, mutual insurance companies may be less efficient and have higher expenses than stock companies. Empirically, rating agencies and others assert that mutual insurance companies have typically significantly reduced their expenses shortly before and after converting to stock companies.

Overall, it is argued, mutual insurance companies have such long planning horizons that they may not operate efficiently, particularly with respect to expenses. Stock companies, on the other hand, have very short planning horizons and may operate to the long-term disadvantage of their policyholders to satisfy their stockholders in the short run. Recently, however, mutual insurance companies have become more cost conscientious.

General Account versus Separate Account Products

The general account of an insurance company refers to the overall resources of the life insurance company, mainly its investment portfolio. Products "written by the company itself" are said to have a "general account guarantee," that is, they are a liability of the insurance company. When the rating agencies (Moody's, Standard & Poor's, Fitch) provide a credit rating, these ratings are on products written by or guaranteed by the general account, specifically on the "claims-paying ability" of the company. Typical products written by and guaranteed by the general account are whole life, universal life, and fixed annuities (including GICs). Insurance companies must support the guaranteed performance of their general account products to the extent of their solvency. These are called *general account products*.

Other types of insurance products receive no guarantee from the insurance company's general account, and their performance is based, not on the performance of the insurance company's general account, but solely on the performance of an investment account separate from the general account of the insurance company, often an account selected by the policyholder. These products are called *separate account products*. Variable life insurance and variable annuities are separate account products. The policyholder selects specific investment portfolios to support these separate account products. The performance of the insurance product depends almost solely on the performance of the portfolio selected, adjusted for the fees or expenses of the insuring company (which do depend on the insurance company). The performance of the separate account products, thus, is not affected by the performance of the overall insurance company's general account portfolio.

Most general account insurance products, including whole life insurance, participate in the performance of the company's general account performance. For example, whereas a life insurance company provides the guarantee of a minimum dividend on its whole life policies, the policies' actual dividend may be greater if the investment portfolio performs well. This is called the "interest component" of the dividend. (The other two components of the dividend are the expense and mortality components.) Thus, the performance of the insurance policy participates in the overall company's performance. Such a policy is called a *participating policy*, in this case a participating whole life insurance policy.

In addition, the performance of some general account products may not be affected by the performance of the general account portfolio. For example, disability income insurance policies may be written on a general account, and while their payoff depends on the solvency of the general account, the policy performance (e.g., its premium) may not participate in the investment performance of the insurance companies' general account investment portfolio.

Both stock and mutual insurance companies write both general and separate account products. However, most participating general account products tend to be written in mutual companies.

Overview of Cash Value Whole Life Insurance

The details of cash value *whole life* insurance (CVWLI) are very complex. This section provides a simple overview of CVWLI, partially by contrasting it with term life insurance.

As discussed above, in annual term life insurance, the owner of the policy, typically also the insured, pays an annual premium which reflects the actuarial risk of death during the year. The premium, thus, increases each year. If the insured dies during the year, the death benefit is paid to the insurer's beneficiary. If the insured does not die during the year, the term policy has no value at the end of the year.

The construction and performance of CVWLI and term life insurance are quite different. Primarily, the owner of the CVWLI policy pays a constant premium. This premium on the CVWLI policy is initially much higher than the initial premium on a term policy (the pure insurance cost) because the constant premium must cover not only lower insurance risk early in the policy but also higher insurance risk later in the policy when the insured has a higher age and the annual cost of the pure insurance exceeds the level premium. However, assuming the same interest and mortality assumptions on both products, the CVWLI premium should be lower than the average of the term premium over time. This is because in the early years, the excess of the level CVWLI premium over the term premium can earn interest, which lowers the overall premium needed to fund the policy; and some CVWLI policy holders paying the level premium die in the early years, leaving funds (from the excess of the level premium over the early life insurance cost) available to the remaining policy holders, which can be used to decrease the CVWLI premium.

In the early years of the policy, the excess of the premium over the pure insurance cost is invested by the insurance company in its general account portfolio. In the later years, there is a shortfall in the premiums relative to the pure insurance cost and the previous cash value buildup is used to fund this shortfall. This portfolio generates a return which accrues to the policy owner's cash value. Typically, the insurance company guarantees a minimum increase in cash value, called the guaranteed cash value buildup. The insurance company, however, may provide an amount in excess of the guaranteed cash value buildup based on earnings for participating policies. What happens to this excess? Assume that the insurance company has a mutual structure, that is, it is owned by the policyholders. In this case, with no stockholders, the earnings accrue to the policyholders as dividends.

The arithmetic of the development of the cash value in a life insurance contract follows:

+ Premium
− Cost of Insurance (Mortality) (denoted by *M*)
− Expenses (denoted by *E*)
+ Guaranteed (Minimum) Cash Value Buildup
+ (Participating) Dividend
= Increase (Buildup) in Cash Value

Note that the overall dividend is calculated from the investment income, the cost of paying the death benefit (the mortality expense denoted by *M*), and the expense of running the company (denoted by *E*). The latter two together are called the M&E charges.

If the insurance company is owned by stockholders, some or all of the earnings might go to the stockholders as dividends.

The returns to the insurance company and, therefore, the dividends to the policyholder can increase if: (1) investment returns increase; (2) company expenses decrease; or (3) mortality costs decrease (that is, the life expectancy of the insured increases).

The dividends can be "used" by the policyholder in either of two ways. The first is to decrease the annual premium. In this case, the death benefit remains constant. The second is to increase the death benefit and the cash value of the policy. Such increases are called "*paid up additions*" (PUAs). In this case, the annual premium remains constant. Most policies are written in the second way.

The intended way for the life insurance policy to terminate is for the insured to die and the life insurance company to pay the death benefit to the beneficiary. There are other ways, however. First, the policy can be lapsed (alternatively called forfeited or surrendered). In this case, the owner of the policy withdraws the cash value of the policy and the policy is terminated.

There are also two nonforfeiture options—that is methods whereby an insurance policy for the insured remains. The owner can use the cash value of the policy to buy extended term insurance (the amount and term of the resulting term insurance policy depends on the cash value). In addition, the cash value of the policy can be used to buy a reduced amount of fully paid (that is, no subsequent premiums are due) whole life insurance—this is called reduced paid up.

In addition to the forfeiture option and the two nonforfeiture options of terminating the CVWLI policy, the policy could be left intact and borrowed against. This is called a *policy loan*. An amount equal to the cash value of the policy can be borrowed. There are two effects of the loan on the policy. First, the dividend is paid only on the amount equal to the cash value of the policy minus the loan. Second, the death benefit of the policy paid is the policy death benefit minus the loan.

The taxation of the death benefit payout, a policy lapse, and borrowing against the loan are considered next. For taxation of life insurance, it is important to recall that the insurance premium is paid by the policy owner with after-tax dollars (this is often called the *cost* of the policy). But the cash value is allowed to build up inside the policy with taxes deferred (or usually tax free), often called the return on the policy.

Taxability of Life Insurance

A major attraction of life insurance as an investment product is its taxability. Consider the four major tax advantages of life insurance.

The first tax advantage is that when the death benefit is paid to the beneficiary of the insurance policy, the benefit is free of income tax. If the life insurance policy is properly structured in an estate plan, the benefit is also free of estate taxes.

The second tax advantage is called "inside buildup"—that is, all earnings (interest, dividend, and realized capital gains) are exempt from income and capital gains taxes. Thus, these earnings are tax deferred (and when included in the death benefit become income tax free, and in some cases also estate tax free).

The third relates to the lapse of a policy. When the policy is lapsed, the owner receives the cash value of the policy. The amount taxed is the cash value minus the cost of the policy (the total premiums paid plus the dividends, if paid in cash). That is, the tax basis of the policy is the cost (accumulated premiums) of the policy. The cost, thus, increases the basis and is recovered tax free. (Remember, however, that these costs were paid with after-tax dollars.) And, the remainder was allowed to accumulate without taxation but is taxed at the time of the lapse.

The fourth tax issue relates to borrowing against the policy—that is, a policy loan. The primary tax issue is the distinction between the cost (accumulated premium) and the excess of the policy cash value over the cost (call it the excess). When a policy loan is made, the cost is deemed to be borrowed first (that is, FIFO [first in-first out] accounting is employed). The amount up to the cash value of the policy can be borrowed and not be subject to the ordinary income tax. (An exception to this practice is for a modified endowment contract (MEC). If the loan is outstanding at the time of the policy lapse, the loan is treated on a FIFO basis whereby the cost basis is assumed to be borrowed first and is not taxable, and when the cost basis is exhausted by the loan, the remainder of the loan [up to the cash value of the policy] is taxable.)

Although CVWLI has both insurance and investment characteristics, Congress provided insurance policies tax advantages because of their insurance, not their investment, characteristics. And Congress does not wish to apply these insurance-directed tax benefits to primarily investment products. In this regard, in the past some activities related to borrowing against insurance policies were considered abuses by Congress and tax law changes were made to moderate these activities. These abuses originated with a product called single-premium life insurance. This policy is one in which only a single premium is paid for a whole life insurance policy. The premium creates an immediate cash value. This cash value and the resulting investment income earned are sufficient to pay the policy's benefits. The excess investment income accumulates tax free.

After the elimination of many tax shelters by the Tax Reform Act of 1986, the sale of single-premium life insurance accelerated significantly because investors found this product to be an attractive tax shelter. Large amounts could be paid as a premium, the earnings grew tax free, and the owner could borrow up to the cash value without a tax liability. Single-premium life insurance thus generated significant tax-sheltered investment income.

In 1988 (via the Technical and Miscellaneous Revenue Act of 1988, TAMRA), Congress developed a new policy to discourage the use of life insurance contracts with large premiums as an investment tax shelter. The test embodied in this policy was called the seven-pay test. Consider first the effect of not meeting the seven-pay test, and then the test itself.

If an insurance policy did not meet the seven-pay test at time of issue, it was deemed to be a modified endowment contract and the tax advantages were reduced as follows. MECs have two important tax disadvantages relative to standard life insurance policies (non-MECs). First, policy loans on a MEC are made on a LIFO (last in-first out) basis—that is, the investment earnings, not the cost basis, is borrowed first and is taxable. The remainder of the loan up to the cash value of the policy is the cost basis and not subject to tax. Second, MECs are subject to a 10% penalty on any taxable gains borrowed before age 59.5 (similar provisions exist on annuities, as discussed in a later section).

Next consider the seven-pay test for determining whether the policy is an MEC. The seven-pay test is an artificial standard developed by the IRS based on the level premium concept. First, the premium for a level premium seven-year paid policy is calculated. The test or standard for determining whether an insurance policy is an MEC is that the premium actually paid on the policy during the first seven years cannot be greater than the seven-year pay level on a year-by-year basis. For example, if the seven-year pay amount calculated is $1,000 per year for seven years, the premium paid can be no more than $1,000 during the first year; similarly no more than $2,000 during the first two years; and up to $7,000 during the first seven years. If the actual premiums paid are greater than any of these amounts, the policy is an MEC. Whether or not a policy is an MEC should be determined and be divulged to the policy owner before the policy is written.

If a policy is deemed a MEC when it is written, it remains a MEC throughout its life. However, a policy that is initially a non-MEC can be subsequently deemed to be an MEC if premium payments accelerate.

The following illustrates the difference in the taxation of an MEC and a non-MEC.

Cash Value:	100
Premium Paid:	20
Earnings:	80

1. Non-MEC
   - Loan up to 100 is nontaxable (that is, neither premium paid nor earnings is taxable)
   - Rationale: withdrawal is a loan, not a distribution (that is, not included in income)
2. MEC
   - Borrow earnings (80) first—is taxable
   - Then borrow premium paid (20)—is not taxable

The characteristics of MECs and non-MECs are summarized in Table 62.4. It is important to note that MECs have no disadvantage if the policy owner does not borrow against the policy. The MEC condition serves only to disadvantage policy loans in an insurance contract.

INVESTMENT-ORIENTED LIFE INSURANCE

The major investment-oriented insurance products can be divided into two categories—cash value life insurance and annuities. Each has several types, which are listed in Table 62.2. These products are described in the following sections.

Cash Value Life Insurance

Cash value life insurance was introduced above. There are two dimensions of cash value life insurance policies. The first is whether the cash value is guaranteed (called whole life) or variable (called variable life). The second is whether the required premium payment is fixed or flexible, that is, whether it has a universal (flexible) feature or not. They can be combined in all ways. Thus, there are four

Table 62.4 Characteristics of Non-MECs and MECs

Non-MEC	MEC
Meets seven-pay test.	Does not meet seven-pay test.
Inside buildup is tax deferred.	Inside buildup is tax deferred.
Can borrow up to cash value of the policy.	Can borrow up to cash value of the the policy.
Loans are tax free.	Loans are treated on LIFO basis (investment income is borrowed first).
	Pay income tax on investment income borrowed first (with 10% penalty on earning if before age 59.5); no tax on remainder of loan up to cash value.
	No disadvantage if do not borrow against or surrender the policy).

combinations, which we discuss next. The broad classification of cash value life insurance, called whole life insurance, in addition to providing pure life insurance (as does term insurance), builds up a cash value or investment value inside the policy.

Traditional cash value life insurance, usually called whole life insurance, has a *guaranteed* buildup of cash value based on the investment returns on the general account portfolio of the insurance company. That is, the cash value in the policy is guaranteed to increase by a specified minimum amount each year. This is called the cash value buildup. (The guaranteed cash value buildup of many U.S. CVWLI policies tend to be in the range of 3%–4%.) The cash value may grow by more than this minimum amount if a dividend is paid on the policy. Dividends, however, are not guaranteed. There are two types of dividends, participating and nonparticipating. Participating dividends depend on (that is, participate in) the investment returns of the general account of the insurance company portfolio (the insurance company M&E charges also affect the dividend).

The participating dividend may be used to increase the cash value of the policy by more than its guaranteed amount. Actually, there are two potential uses of the dividend. The first is to reduce the annual premium paid on the policy. In this case, while the premium decreases, the cash value of the policy increases by only its guaranteed amount (and the face value the death benefit remains constant).

The second use is to buy more life insurance with the premium (called "paid-up additions," PUA). In this case, the cash value of the entire policy increases by more than the guaranteed amount on the original policy (and the face value of the current policy is greater than the face value of the original policy).

In either case, the performance of the policy over time may be substantially affected by the participating dividends.

Contrary to the guaranteed or fixed cash value policies based on the general account portfolio of the insurance company, *variable life insurance* polices allow the policyowners to allocate their premium payments to and among several separate investment accounts maintained by the insurance company, and also to be able to shift the policy cash value among these separate accounts. As a result, the amount of the policy cash value depends on the investment results of the separate accounts the policyowners have selected. Thus, there is no guaranteed cash value or death benefit. Both depend on the performance of the selected investment portfolio.

The types of separate account investment options offered in their variable life insurance policies vary by insurance companies. Typically, the insurance company offers a selection of common stock and bond fund investment opportunities, often managed by the company itself and also by other investment managers. If the investment options perform well, the cash value buildup in the policy will be significant. However, if the policyholder selects investment options that perform poorly, the variable life insurance policy will perform poorly. There could be little or no cash value buildup, or, in the worst case, the policy could be terminated because there is not enough value in the contract to pay the mortality charge. This type of cash value life insurance is called *variable life insurance*.

Table 62.5 Types of Cash Value Life Insurance

Premium	Guaranteed	Variable
Fixed	Whole life	Variable life
Flexible	Universal life	Variable universal life

The key element of *universal life* is the flexibility of the premium for the policyowner. The flexible premium concept separates the pure insurance protection (term insurance) from the investment (cash value) element of the policy. The policy cash value is set up as a cash value fund (or accumulation fund) to which the investment income is credited and from which the cost of term insurance for the insured (the mortality charge) is debited. The policy expenses are also debited.

This separation of the cash value from the pure insurance is called the "unbundling" of the traditional life insurance policy. Premium payments for universal life are at the discretion of the policyholder, that is, are flexible with the exceptions that there must be a minimum initial premium to begin the coverage, and there must also be at least enough cash value in the policy each month to cover the mortality charge and other expenses. If not, the policy will lapse. Both guaranteed cash value and variable life can be written on a flexible premium or fixed premium basis.

The universal feature—flexible premiums—can be applied to either guaranteed value whole life (called simply universal life) or to variable life (called variable universal life). These types are summarized in Table 62.5. Variable universal life insurance combines the features of variable life and universal life policies—that is, the choice of separate account investment products and flexible premiums.

Over the last decade, term and variable life insurance have been growing at the expense of whole life insurance. The most common form of variable life is variable universal.

Most whole life insurance policies are designed to pay death benefits when one specified insured dies. An added dimension of whole life policies is that two people (usually a married couple) are jointly insured, and the policy pays the death benefit not when the first person dies, but when the second person (the "surviving spouse") dies. This is called survivorship insurance or second-to-die insurance. This survivorship feature can be added to standard cash value whole life, universal life, variable life, and variable universal life policies. Thus, each of the four policies discussed could also be written on a survivorship basis.

In general, the annual premium for a survivorship insurance policy is lower than for a policy on a single person because, by construction, the second of two people to die has a longer life span than the first. Survivorship insurance is typically sold for estate planning purposes.

Table 62.6 provides a summary of the various types of cash value life insurance, with (annual renewable) term insurance included for contrast.

Table 62.6 Life Insurance Comparison (By Type and Element)

Type	Description	Death Benefit	Premium	Cash Value (CV)	Advantages to Owner	Disadvantages to Owner
Annual renewable term	"Pure" life insurance with no cash value; initially, the highest death benefit for the lowest premium; premium increases exponentially	Fixed, constant	Increases exponentially	None	Low premium for coverage	Increasing premium; most term insurance is lapsed
Whole life	Known maximum cost and minimum death benefit; dividends may: reduce premiums; pay-up policy; buy paid-up additions; accumulate at interest; or be paid in cash	Fixed, constant	Fixed, constant	Fixed	Predictable; forced savings and conservative investment	High premiums given death benefit
Variable life	Whole life contract; choice of investment assets; death benefits depend on investment results	Guaranteed minimum; can increase based on investment performance	Fixed, constant	Based on investment performance; not guaranteed.	Combines life insurance and investments on excess premiums	All investment risk is to the owner
Universal life	Flexible premium, current assumption adjustable death benefit policy; policy elements unbundled	Adjustable; Two options: 1. like ordinary life; 2. like ordinary life plus term rider equal to cash value.	Flexible at option of policy owner	Varies depending on face amount and premium; minimum guaranteed interest; excess increases cash value.	Flexibility	Some investment risk to owner
Variable universal life	Features of universal and variable life	Adjustable	Flexible at option of policy owner	Varies depending on face amount, premium, and investment performance; not guaranteed.	Flexibility and choice of investments	All investment risk is to owner

Uses of Life Insurance

The standard use of life insurance is to protect the survivors of an income earner. In this case, the insured is the income earner and the survivors are the beneficiaries. This is still a major use of life insurance. For this use, life insurance protects against premature death.

There are, however, many other uses. The life insurance death benefits are used to pay the estate taxes on the deceased's assets in their estate. There are also many business uses of life insurance. Split dollar life insurance, whereby the business pays for a portion or all of the premium on a life insurance policy on the executive, is used as a fringe benefit for its executives. Life insurance policies may also be written on both participants in a partnership to fund the purchase by the surviving partner of the ownership of the deceased partner according to a buy-sell agreement. There are also other business uses of life insurance.

Annuities

By definition, an annuity is simply a series of periodic payments. Annuity contracts have been offered by insurance companies and, more recently, by other types of financial institutions such as mutual fund companies.

There are two phases to annuities according to cash flows, the accumulation period and the liquidation period. During the accumulation period, the investor is providing funds, or investing. Annuities are considered primarily accumulation products rather than insurance products. During the liquidation period, the investor is withdrawing funds, or liquidating the annuity. One type of liquidation is annuitization, or withdrawal via a series of fixed payments, as discussed below. This method of liquidation is the basis for the name of annuities.

There are several ways to classify annuities. One is the method of paying premiums. Annuities are purchased with *single premiums*, *fixed periodic premiums*, or *flexible periodic premiums* during the accumulation phase. All three are used in current practice.

A second classification is the time the income payments commence during the liquidation phase. An *immediate annuity* is one in which the first benefit payment is due one payment interval (month, year or other) from the purchasing date. Under a *deferred annuity*, there is a longer period before the benefit period begins. While an immediate annuity is purchased with a single premium, a deferred annuity may be purchased with a single, fixed periodic, or flexible periodic payments, although the flexible periodic payment is most common.

An important basis for annuities is whether they are fixed or variable annuities. Fixed annuities, as discussed in more detail below, are expressed in a fixed number of dollars, while variable annuities are expressed in a fixed number of annuity units, each unit of which may have a different and changing market value. Fixed versus variable annuities is the key distinction between annuities currently provided.

Now we will look at the various types of annuities. The most common categories are variable annuities and fixed annuities.

While cash value life insurance has the appearance of life insurance with an investment feature, annuities, in contrast, have the appearance of an investment product with an insurance feature. The major advantage of an annuity is its inside buildup, that is, its investment earnings are tax deferred. However, unlike life insurance where the death benefit is not subject to income taxes, withdrawals from annuities are taxable. There are also restrictions on withdrawals. Specifically, there are IRS requirements for the taxability of early withdrawals (before age 59.5) and required minimum withdrawals (after age 70.5). These requirements and the other tax issues of annuities are very complex and considered only briefly here.

The most common types of annuities, variable and fixed annuities, are discussed below.

Variable Annuities

Variable annuities are, in many ways, similar to mutual funds. Given the above discussion, variable annuities are often considered to be "mutual funds in an insurance wrapper." The return on a variable annuity depends on the return of the underlying portfolio. The returns on annuities are, thus, in a word, "variable." In fact, many investment managers offer similar or identical funds separately in both a mutual fund and an annuity format. Thus, variable annuity offerings are approximately as broad as mutual fund offerings. For example, consider a large capitalization, blended stock fund. The investment manager may offer this fund in both a mutual fund and annuity format. But, of course, the two portfolios are segregated. The portfolios of these two products may be identical and, thus, the portfolio returns will be identical.

Before considering the differences, however, there is one similarity. Investments in both mutual funds and annuities are made with after-tax dollars; that is, taxes are paid on the income before it is invested in either a mutual fund or an annuity.

But there are important differences to investors in these two products. First, all income (dividend and interest) and realized capital gains generated in the mutual fund are taxable, even if they are not withdrawn. However, income and realized capital gains generated in the annuity are not taxable until withdrawn. Thus, annuities benefit from the same *inside buildup* as cash value life insurance.

There is another tax advantage to annuities. If a variable annuity company has a group of annuities in its family (called a "contract"), an investor can switch from one annuity fund to another in the contract (for example from a stock fund to a bond fund) and the switch is not a taxable event. However, if the investor shifts from a stock fund in one annuity company to a bond fund in another annuity company, it is considered a withdrawal and a reinvestment, and the withdrawal is a taxable event (there are exceptions to this, however, as will be discussed). The taxation of annuity withdrawals will also be considered.

While the inside buildup is an advantage of annuities, there are offsetting disadvantages. For comparison, there are no restrictions on withdrawals from (selling shares of) a mutual fund. Of course, withdrawals from a mutual fund are a taxable event and will generate realized capital gains or losses, which will generate long-term or short-term gains or losses and, thus, tax consequences. There are, however, significant restrictions on withdrawals from annuities. First, withdrawals before age 59.5 are assessed a 10% penalty (there are, however, some "hardship" exceptions to this). Second, withdrawals must begin by age 70.5 according to the IRS required minimum distribution rules (RMD). These mandatory withdrawals are designed to eventually produce tax revenues on annuities to the IRS. Mutual funds have no disadvantages to withdrawing before 59.5 nor requirements to withdraw after 70.5.

There is an exception to the taxation resulting from a shift of funds from one variable annuity company to another. Under specific circumstances, funds can be so moved without causing a taxable event. Such a shift is called a 1035 exchange after the IRS rule that permits this transfer.

Another disadvantage of annuities is that all gains on withdrawals, when they occur, are taxed as ordinary income, not capital gains, whether their source was income or capital gains. For many investors, their income tax rate is significantly higher than the long-term capital gains tax rate and this form of taxation is therefore a disadvantage.

The final disadvantage of annuities is that the heirs of a deceased owner receive them with a cost basis equal to the purchase price (which means that the gains are taxed at the heir's ordinary income tax rate) rather than being stepped up to a current market value as with most investments.

Why has the IRS given annuities the same tax advantage of inside buildup that insurance policies have? The answer to this question is that annuities are structured to have some of the characteristics of life insurance, commonly called "features." There are many such features. The most common feature is that the minimum value of an annuity fund that will be paid at the investor's death is the initial amount invested. Thus, if an investor invests \$100 in a stock annuity, the stock market declines such that the value of the fund is \$90, and the investor dies, the investor's beneficiary will receive \$100, not \$90. This is a life insurance characteristic of an annuity.

The above feature represents a death benefit (DB), commonly called a return of premium. However, new, and often more complicated, death benefits have been introduced, including a periodic lock-in of gains (called a "stepped up" DB); a predetermined annual percentage increase (called a "rising floor" DB); or a percentage of earnings to offset estate taxes and other death expenses (called an "earnings enhancement" DB). In addition to these death benefit features, some living benefit features have also been developed, including premium enhancements and minimum accumulation guarantees.

Obviously, these features have value to the investor and, as a result, a cost to the provider. The value of a feature depends on its design and can be high or approximately worthless. And the annuity company will charge the investor for the value of these features.

The cost of the features relates to another disadvantage of annuities, specifically their expenses. The insurance company will impose a charge for the potential death benefit payment (called mortality) and other expenses, overall called M&E charges, as discussed previously for insurance policies. These M&E charges will be in addition to the normal investment management, custody, and other expenses experienced by mutual funds. Thus, annuity expenses will exceed mutual fund expenses by the annuity's M&E charges. The annuity investor does, however, receive the value of the insurance feature for the M&E charge.

Thus, the overall trade-offs between mutual funds and annuities can be summarized as follows. Annuities have the advantages of inside buildup and the particular life insurance features of the specific annuity. But annuities also have the disadvantages of higher taxes on withdrawal (ordinary income versus capital gains), restrictions on withdrawals, and higher expenses. For short holding periods, mutual funds will have a higher after-tax return. For very long holding periods, the value of the inside buildup will dominate and the annuity will have a higher after-tax return.

What is the breakeven holding period, that is, the holding period beyond which annuities have higher after-tax returns? The answer to this question depends on several factors, such as the tax rates (income and capital gains), the excess of the expenses on the annuity, and others.

Fixed Annuities

There are several types of fixed annuities but, in general, the invested premiums grow at a rate—the credited rate—specified by the insurance company in each. This growth is accrued and added to the cash value of the annuity each year (or more frequently, such as monthly) and is not taxable as long as it remains in the annuity. Upon liquidation, it is taxed as ordinary income (to the extent that is represents previously untaxed income).

The two most common types of fixed annuities are the *flexible-premium deferred annuity* (FPDA) and the *single-premium-deferred annuity* (SPDA). The FPDA permits contributions which are flexible in amount and timing. The interest rate paid on these contracts—the credited rate—varies and depends on the insurance company's current interest earnings and its desired competitive position in the market. There are, however, two types of limits on the rate. First, the rate is guaranteed to be no lower than a specified contract guaranteed rate, often in the range 3% to 4%. Second, these contracts often have bailout provisions, which stipulate that if the credited rate decreases below a specified rate, the owner may withdraw all the funds (lapse the contract) without a surrender charge. Bailout credited rates are often set at 1% to 3% below the current credited rate and are designed to limit the use of a "teaser rate" (or "bait and switch" practices), whereby an insurance company offers a high credited rate to attract new investors and then reduces the credited rate significantly, with the investor limited from withdrawing the funds by the surrender charges.

An initial credited rate, a minimum guaranteed rate, and a bailout rate are set initially on the contract. The initial credited rate, thus, may be changed by the insurance company over time. The reset (or renewal) period must also be specified—this is, the frequency with which the credited rate can be changed.

Another important characteristic of annuities is the basis for the valuation of withdrawals prior to maturity. The traditional method has been book value, that is, withdrawals are paid based on the purchase price of the bonds (bonds rather than stocks are used to fund annuities). Thus, if yields have increased, the insurance company will be paying the withdrawing investor more than the bonds are currently worth. And at this time, there is an incentive for the investor to withdraw and invest in a new higher yielding fixed annuity. Thus, book value fixed annuities provide risk to the insurance company. Surrender charges, discussed next, mitigate this risk. Another way to mitigate this risk is via market value adjusted (MVA) annuities, whereby early withdrawals are paid on the basis of the current market value of the bond portfolio rather than the book value. This practice eliminates the early withdrawal risk to the insurance company. (Obviously, all variable annuities are paid on the basis of market value rather than bonds value.)

Another characteristic of both variable and fixed annuities relates to one aspect of their sales charges. These charges are very similar for annuities and mutual funds. Mutual funds and annuities were originally provided with front-end loans, that is, sales charges imposed on the initial investment. For example, with a 5% front-end load of a $100 initial investment, $5 would be retained by the firm for itself and the agent, and $95 invested in the fund for the investor.

More recently, back-end loads have been used as an alternative to front-end loads. With a back-end load, the fixed percentage charge is imposed at the time of withdrawal. Currently, the most common form of back-end load is the contingent deferred sales charge (CDSC), also called simply a surrender charge. This approach imposes a load which is gradually declining over time. For example, a common CDSC is a "7%/6%/5%/4%/3%/2%/1%/0%" charge according to which a 7% load is imposed on withdrawals during the first year, 6% during the second year, 5% during the third year, and so forth. There is no charge for withdrawals after the seventh year.

Finally, there are level loads, which impose a constant load (1% for example) every year. Currently on annuities, a front-end load is often used along with a CDSC surrender charge.

Annuities have become very complex instruments. This section provides only an overview.

Guaranteed Investment Contracts

The first major investment-oriented product developed by life insurance companies, and a form of fixed annuity, was the *guaranteed investment contract* (GIC). GICs were used extensively for retirement plans. With a GIC, a life insurance company agrees, in return for a single premium, to pay the principal amount and a predetermined annual crediting rate over the life of the investment, all of which are paid at the maturity date of the GIC. For example, a $10 million five-year GIC with a predetermined crediting rate of 10% means that at the end of five years, the insurance company pays the guaranteed crediting rate and the principal. The return of the principal depends on the ability of the life insurance company to satisfy the obligation, just as in any corporate debt obligation. The risk that the insurer faces is that the rate earned on the portfolio of supporting assets is less than the guaranteed rate.

The maturity of a GIC can vary from 1 year to 20 years. The interest rate guaranteed depends on market conditions and the rating of the life insurance company. The interest rate will be higher than the yield on U.S. Treasury securities of the same maturity. These policies are typically purchased by pension plan sponsors as a pension investment.

A GIC is a liability of the life insurance company issuing the contract. The word guarantee does not mean that there is a guarantor other than the life insurance company. Effectively, a GIC is a zero-coupon bond issued by a life insurance company and, as such, exposes the investor to the same credit risk. This credit risk has been highlighted by the default of several major issuers of GICs. The two most publicized defaults were Mutual Benefit, a New Jersey–based insurer, and Executive Life, a California-based insurer, which were both seized by regulators in 1991.

The basis for these defaults is that fixed annuities are insurance company general account products and variable annuities are separate account products. For fixed annuities, the premiums become part of the insurance company, are invested in the insurance company's general account (which are regulated by state laws), and the payments are the obligations of the insurance company. Variable annuities are separate account products, that is, the premiums are deposited in investment vehicles separate from the insurance company, and are usually selected by the investor. Thus, fixed annuities are general account products and the insurance company bears the investment risk, while variable annuities are separate account products and the investor bears the investment risk.

SPDAs and GICs

SPDAs and GICs with the same maturity and crediting rate have much in common. For example, for each the value of a $1 initial investment with a 5-year maturity and a fixed crediting rate for the 5 years at r% would have a value at maturity of $(1 + r)^5$.

However, there are also significant differences. SPDAs have elements of an insurance product and so its inside buildup is not taxed as earned (it is taxed as income at maturity). SPDAs are not qualified products, that is, they must be paid for in after tax-dollars. GICs are not insurance products. GICs, however, are typically put into pension plans (defined benefit or defined contribution), which are qualified. In this case, thus, the GIC investments are paid for in after-tax dollars and receive the tax deferral of inside buildup. SPDAs are also put into qualified plans. Specifically, banks often sell IRAs funded with SPDAs.

Another difference between SPDAs and GICs is that since SPDAs are annuities, they usually have surrender charges, typically the 7%/6%/5%/4%/3%/2%/1%/0%, mentioned previously. Thus, if a 5-year SPDA is withdrawn after three years, there is a 4% surrender charge. GICs do not have surrender charges and can be withdrawn with no penalty (under benefit responsive provisions).

Another feature of SPDAs is the reset period, the period after which the credited rate can be changed by the writer of the product. For example, a 5-year SPDA may have a reset period after three years, at which time the credited rate can also be increased or decreased. For SPDAs, there can also be an interaction between the reset period and the surrender charge. For example, a 5-year SPDA with a 3-year reset period could be liquidated after 3 years due to a lowered crediting rate, but only with a 4% surrender charge. GICs have no reset period, that is, the credited rate is constant throughout the contract's life. Early withdrawals of GICs are at book value; they are interest rate insensitive.

SPDAs typically have a reset period of 1 year but with an initial M-year minimum guarantee ($M = 1,2,3,5,7,9$). SPDAs typically have a maturity based on the age of the annuitant (such as age 90 or 95), not a fixed number of years. Thus, while SPDAs typically have a maturity greater than the guarantee period, for GICs the maturity period equals the guarantee period. Common maturities for GICs and SPDAs are 1, 3, 5, and 7 years.

Annuitization

Strictly speaking, an annuity is a guaranteed (or fixed) amount of periodic income for life. Both variable and fixed annuities are accumulation products rather than income products. Either product can, however, be annuitized—that is, converted into a guaranteed lifetime income. Annuitization refers to the liquidation rather than the accumulation period. As a matter of fact, very few of variable annuities are annuitized. One reason that few investors annuitize is that they fear they will die early and receive very little for the initial investment. On the other hand, the risk to individuals is that they will outlive their savings. Annuitization eliminates this risk. Traditionally, defined benefit retirement plans have provided a lifetime flow of income. But with the decline in defined benefit retirement plans, annuities can fill this vacuum.

Since the fixed payments of an annuity are for life, there is mortality risk for the annuity writer. If the annuitant dies soon, the payout by the annuity writer will be small. However, if the annuitant lives a long life, the payments by the annuity writer will be large. This characteristic introduces an underwriting element to annuities by the annuity writer. Some fixed annuities also have a survivorship feature. That is, when the annuitant dies, the payments will continue and be paid to a named survivor, usually a spouse.

Many variable annuity owners who wish to annuitize elect for a variation on a strict annuity called a systematic withdrawal plan (SWP) instead. While there are many types of SWPs, the most common type is based on a specified term rather than lifetime payments in order to assure that the payments last at least a certain amount of time (called a period-certain payout option). These plans, thus, do not eliminate the risk of outliving one's savings. Under a SWP, annuity shares are liquidated to pay regular payments that are either a fixed dollar amount or a percentage of the investor's account balance. Thus, unlike annuitization, SWPs cause a continual decline in the investor's account balance. There are also variations of the standard lifetime payout option which include life with a guaranteed period, joint and survivor life, and joint and survivor life with a guaranteed period.

SUMMARY

Fundamentally, insurance and investment products are distinct. Insurance products provide risk protection against a wide variety of risks and have no cash value. Investment products, often called accumulation products, provide returns on an initial investment.

However, two types of products provide elements of both insurance and investments. These two are cash value life insurance and annuities. Cash value life insurance is a combination of pure life insurance with a buildup of cash value as a result of the higher premium paid relative to a pure life insurance policy. The types of cash value life insurance include whole life and variable life, and universal versions of both of these.

The second type is annuities. There are two types of annuities, variable and fixed. Variable annuities are essentially mutual funds in an insurance wrapper. The insurance elements may include both death benefits and living benefits. The returns on variable annuities depend on the particular type of investment portfolio selected by the investor.

Fixed annuities are a guaranteed yield over an investment term. The return over the term is specified at the time of the investment and is certain. Very few annuities, variable or fixed, are annuitized, that is, converted into a lifetime stream of fixed payments, despite the attractive characteristics of annuitization.

The investment element of these hybrid insurance/investment products benefits from their tax advantages. The major tax advantage of both of these types of investment-oriented insurance products is inside buildup, although the cash value life insurance products also have other significant tax benefits. Congress provided the tax advantages to these products due to their insurance characteristics, not their investment characteristics. There are, as a result, limits on the investment characteristics of these hybrid products to qualify them for the tax advantages.

REFERENCES

Baldwin, B. (2002). *New Life Insurance Investment Advisor: Achieving Financial Security for You and Your Family Through Today's Insurance Products*, 2nd ed., New York: McGraw-Hill.

Dellinger, J. K. (2006). *The Handbook of Variable Income Annuities*. Hoboken, NJ: John Wiley & Sons.

Graves, E. E., and Hayes, L. (eds.) (1994). *McGill's Life Insurance*. Bryn Mawr, PA: Hueber School Series, The American College.

Milevsky, M. A. (2006). *The Calculus of Retirement Income: Financial Models for Pension Annuities and Life Insurance*. New York: Cambridge University Press.

Shunt, L. S. (2003). *The Life Insurance Handbook*. New York: Marketplace Books.

Weber, R. M. (2005). *Revealing Life Insurance Secrets: How the Pros Pick, Design, and Evaluate Their Own Policies*. New York: Marketplace Books.

CHAPTER 63

Stable Value Investment Options for Defined Contribution Plans

BRIAN K. HAENDIGES, FSA, CRC, CRA
Head, Institutional Defined Contribution Plans, ING Retirement Services

Abstract: One of the most prevalent investment options used by defined contribution plans in the United States is a stable value option, one that provides protection of invested principal and accrued interest to individual plan participants. The funding vehicles that back stable value options are diverse, yet retain many common elements, including the accumulation of interest, some level of assurance of stability of principal, and the participant right to withdraw funds for plan benefits at book value. Both the purchasing of these products by defined contribution plan sponsors, and the issuance of them by providers, require careful consideration of those product features that are different from one product to another. This is especially important given an aging Baby Boomer population that is spurring a renewed interest in product capabilities that allow them to protect and manage the income on assets that they've accumulated for retirement. Stable value products and features are also attracting interest for applications outside of defined contribution plans, and the lessons learned from using these products in a defined contribution environment lend themselves to extrapolation in other venues.

Keywords: stable value option, stable value products, guaranteed interest rate, stability of principal, book value withdrawals, book value benefits, credited interest rate, market value adjustment, defined contribution plans, guaranteed investment contract or guaranteed income contract (GIC), portfolio rate, separate account GIC, synthetic GIC, GIC pools, stable value pooled fund, defined contribution funding, Health Reserve Account (HRA) funding, protection of principal, guaranteed fund, safe fund, defined benefit liability defeasance, credited rate formula, general account, separate account, trust account, rate reset frequency, deferred sales charge, fixed account, commercial mortgages, private placement, mortgage pass-through securities, guaranteed account

This chapter will provide a description of the market drivers of *stable value products,* an overview of the different types of stable value products available and how they are structured and used, buyers of stable value products, common features of stable value products, some of the issues faced by both users and issuers of stable value products, and some of the positives and negatives of different stable value products. It also explores some of the lessons learned from the history of stable value products and how the products and their features might be applied as principal-protected products experience market growth driven by demographic and psychographic changes in the retirement markets.

THE NEED FOR STABLE VALUE INVESTING

Virtually every participant-directed or defined contribution retirement plan, regardless of plan size or market, maintains a participant investment option that returns to its investors a stable return over time. In other words, much like a savings account at a bank, funds deposited into this *stable value option* are credited with interest, and invested principal is protected. Often called the *guaranteed account,* or the *fixed account,* this option is available in 401(k) and 401(a) plans offered by privately held corporations, 403(b) plans offered to teachers and not-for-profit workers, and 457 deferred compensation plans offered to governmental employees, and it often holds 15% to 35% of participant investments. Why are these options so popular, especially in a world where mutual fund companies and financial advisers have promoted diversification and emphasize equity investing?

Participant Demand for Safety

Behavioral finance research has shown a strong preference by investors to avoid losses, even if it means forgoing an even larger opportunity for gains. For participants concerned about safety, stable value options provide some measure of protection at all stages of the asset accumulation cycle. Stable value investments are popular with conservative investors or investors just starting out who may be nervous about loss of their investment. Mid-career investors use stable value options to complement and provide more flexibility with their equity investments. Near retirees and retirees use these options to protect accumulated assets and guarantee an income stream.

Plan Design

Most employers that offer *defined contribution plans* make available a number of investment options to participants, often across a range of investment classes and offering varying degrees of risk. Though the standard varies by type of plan, all plan sponsors have some fiduciary obligation to provide a diversified menu to their participants. Stable value helps fill one of the options at the conservative end of the spectrum for many employers.

Returns

Over a full interest rate cycle, longer-term fixed income investments will generally provide higher returns than shorter-term investments. However, in order to earn those higher returns, the investment provider needs to be able to lock up the funds on a basis consistent with the maturity schedule of the investments. Stable value options are designed to create just such a return paradigm. By investing longer, but limiting asset liquidity to participant benefit events, stable value options can capture the returns of longer-term investing and still provide participants liquidity when they need it. As a result, stable value options generally will generate a yield for participants of 2% to 3% more than money market funds over a full investment cycle. That incremental return can mean a substantial benefit to the participant in retirement as compared with other principal-protected investments.

Historical Legacy

When defined contribution plans were first made available to employees, many providers started by offering participants only a single investment option—a fixed or stable value option. This may perhaps be driven by the long legacy that life insurers have had with fixed income products backing other types of liabilities such as life insurance and defined benefit plans. Within defined contribution plans, providers gradually made other options available, at first proprietary variable investment funds, and later nonproprietary funds. However, many investors who started in stable value still maintain a substantial allocation. Interestingly, this focus on stable value as a starting point may have implications for providers in countries where individually directed defined contribution plans are just now emerging. In order to attract investors at the outset, some *protection of principal,* and a perception of "simple and safe" may be required.

TYPES OF STABLE VALUE FUNDING VEHICLES

There is a wide variety of funding vehicles from which employers may choose when funding a stable value option. One consideration is whether to purchase a single product or multiple products. Many plans, even large ones, will manage their stable value option so that it is invested in a single product. Others will buy multiple products and pool the returns on each to create a "blended rate" to be credited to participants. Either way, there is an extensive array of choices. Table 63.1 shows the distribution of assets backing stable value investment options by product type, with total assets approaching half a trillion dollars.

Table 63.1 Distribution of Stable Value Assets by Product

Stable Value Product Type	Assets as of December 31, 2006
General account portfolio rate	\$100–\$200 billion
General account GICs	\$31 billion
Separate account GICs	\$16 billion
Synthetic GICs	\$198 billion
Other	\$65 billion
Total	\$400–\$500 billion

Source: 2007 LIMRA International, Inc. and Stable Value Investment Association, and industry estimates.

Table 63.2 Typical Insurer General Account Assets

Asset Type	% of Portfolio
Mortgage loans and real estate	12%
Mortgage-backed securities	11%
Investment-grade bonds, public and private	42%
Asset-backed securities	11%
Foreign- and emerging-market debt	8%
High-yield bonds, public and private	4%
Other	12%

Source: Lehman Brothers analysis of top 54 companies, statutory assets as of December 31, 2004.

Portfolio Rate General Account

The earliest form of stable value investment provided in defined contribution (DC) plans was the insurance company *general account.* It's called the general account because the assets backing promises to DC investors are commingled with assets backing other promises by the insurer to its customers. These assets are typically almost all fixed income, and include less liquid higher-yielding securities like *commercial mortgages* and *private placement* bonds as well as publicly traded fixed income securities. For purposes of determining *credited interest rates,* the assets may be segmented into different portfolios, but from an ownership perspective, they are essentially in one big pool. Table 63.2 shows a typical distribution of assets in an insurance company general account.

Credited rates in a *portfolio rate* product are typically set periodically, again like a bank savings account, without being tied to a specific formula. Rather, the insurer's incentive to maintain rates is simple competitive pressure, comparing the rate to other offerings available to participants in other products or from competitors. The account will often have minimum interest rate guarantees associated with it, which can be lifetime guarantees, annual guarantees, or guarantees for a shorter time period. A handful of products tie their guarantees to an index or use a formula, but they are infrequent.

General account products are more frequently seen in IRS Section 403(b) and 457 defined contribution plans, where it makes it easier for the provider to meet state insurance law and securities law requirements. They are also an effective investment for smaller plans, allowing investing in a substantial diversified pool.

GICs

A guaranteed investment contract or guaranteed income contract (GIC) works notionally much like a certificate of deposit at a bank, only with the employer rather than an individual as a buyer. In essence, the issuer accepts a block of funds from the contract holder, usually a defined contribution plan sponsor, and then guarantees a rate of return on those funds over some time period. This rate of return will reflect the yields currently available on the type and quality of assets purchased with the funds. GICs reached their heyday in the early 1980s when long-term interest rates peaked at over 14%, but are still used by plans today. As rates occasionally edge up, their popularity also climbs. A variation of GICs has also been used as a funding vehicle for foreign entities.

The most basic structure of a traditional GIC is a fixed-rate, fixed-term contract where an insurer accepts deposits to its general account over a specified time period or in specified amounts. These deposits are then left with the insurer for the remaining term of the contract, and are credited with a declared rate of interest. For example, an insurer might agree to accept deposits to a GIC over the next year, and credit the GIC with an interest rate of 8% provided the funds are left with the insurer for the following four years. The 8% interest rate would be guaranteed for the entire five-year term of the contract.

An alternative approach would be to guarantee a minimum rate of interest on all deposits (e.g., 3%) and then allocate each deposit to a calendar-year cell corresponding to the year in which the deposit was made. Each cell would be credited with its own rate of interest and have its own rate guarantees reflecting the investment conditions at the time the deposits were made. This is often referred to as the "investment year method." The credited rate reported to the contract holder under this approach is usually a composite or blend of the rates in each cell. The investment year method has declined in use due to administrative complexity and insurer underwriting issues, but persists in legacy form.

At the end of the contract term, the insurer may be obligated to return the contract holder's principal and interest in a lump sum, or in a series of installments over time. The contract holder may also elect to roll maturing principal and interest into another GIC that is able to accept new deposits. In this case, the contract holder could then have several GICs in force with the same insurer at one time.

Bank Investment Contracts

Bank investment contracts (BICs) are agreements between a bank and the plan sponsor or trustee that have deposit and maturity characteristics similar to those of GICs. A plan sponsor might choose a BIC over a GIC in order to lower their credit exposure to the insurance industry. BICs have declined in use due to issues relating to FDIC deposit insurance coverage, and the availability of more popular alternative products. Again, they may still be seen occasionally in an older portfolio.

Separate Account Products

Insurers developed *separate account* products in order to offer the contract holder the same investment expertise as through a traditional GIC, but with a decreased credit exposure to the insurer's general account. In the case of insolvency, the holder of a general account GIC would be treated as any other policyholder in determining the disposition of the insurer's assets. With separate account products, the assets underlying the contract are held in a separate account and are generally not chargeable with the insurer's other liabilities, thereby reducing the credit risk faced by the plan sponsor.

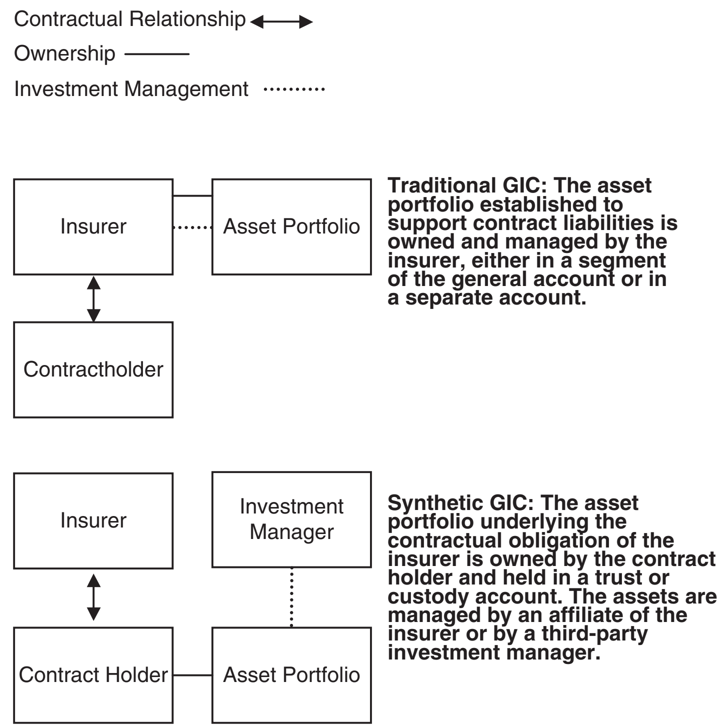


Figure 63.1 Operation of a Typical Synthetic GIC

Synthetic GICs

Synthetic GICs have become the most popular stable value funding option for large plans, as plan sponsors have sought additional diversification of credit risk. Under a synthetic GIC, the fixed income assets are owned by the plan sponsor and held in trust for plan participants. The plan sponsor or trustee selects an investment manager or managers for the fixed income portfolio, and the synthetic GIC issuer then "wraps" the portfolio in a contract that guarantees a rate of interest based on the performance of the investment portfolio and provides benefit responsiveness for plan participants. The fixed income portfolio manager may be either a third party or an affiliate of the synthetic GIC issuer. Figure 63.1 shows the operation of a typical synthetic GIC.

Insurance insolvency laws vary by state, but most states have adopted a law similar to the National Association of Insurance Commissioners' (NAIC) Insurers Rehabilitation and Liquidation Model Act. Table 63.3 shows the priority of insurance company claimants under the Model Act in the event of insolvency. Holders of traditional GICs, separate account products, and synthetics would generally be treated as Class 3 claimants (that is, on a par with other policyholders of the insurer). However, holders of separate account products have an advantage over general account contract holders in that, if the assets in the separate account are "insulated" from the general account, the assets in the separate account cannot be used to satisfy any claims of the insurer other than those of investors in the separate account. In this case, the credit exposure to the general account is equal to only that portion of the insurer's obligation, if any, in excess of the value of the

Table 63.3 Priority of Insurance Company Claimants in Insolvency

Priority Class	Claimants Included
Class 1	Administrative expenses approved by the receiver, including filing fees, the costs of recovering assets of the insurer, and compensation for services rendered in the rehabilitation or liquidation.
Class 2	Administrative expenses of guaranty associations, not including payments and expenses incurred as direct policy benefits.
Class 3	Policyholder claims for insured losses and unearned premiums, including those of federal, state, and local governments, as well as covered claims incurred by guaranty associations.
Class 4	Claims of the federal government other than those included in Class 3.
Class 5	Debts due to employees (other than principal officers and directors) for services and benefits.
Class 6	Claims of general creditors and persons not elsewhere classified.
Class 7	Claims of any state or local government for a penalty or forfeiture.
Class 8	Surplus or contribution notes, or similar obligations.
Class 9	Claims of shareholders or other owners arising from ownership.

Source: NAIC Insurers Rehabilitation and Liquidation Model Act, Copyright © NAIC 1995.

separate account assets. Holders of synthetic GICs face a similar credit position, but have an added layer of protection on the underlying assets since they are legally owned by the contract holder rather than the issuer.

Global Wraps

One particular subset of the synthetic GIC structure worth mentioning is the global wrap. Under this type of arrangement, a single plan sponsor will purchase several synthetic GICs, that is, several different fixed income asset managers and several different wraps from wrap providers. The plan will then create cross-coverage connections so that the wrap providers support each other in the event that one fails or shows signs of credit weakness. Also, either a wrapper or manager may be replaced for underperformance or credit concerns. This structure would be mostly used with very large plans with several billion dollars of retirement funds.

Pooled Funds

Pooled funds, still sometimes referred to informally as *GIC pools* even though GIC usage to fund them has declined, require less decision making on the part of the plan sponsor. Rather than the plan selecting a single particular stable value product issuer, the manager of the pool will bundle together stable value products from a number of different issuers, and then credit a blended interest rate to the funds invested in the pool. Spreading the assets among multiple issuers in this way provides diversification of credit risk. Also, by investing in such pooled accounts, small plans are often able to obtain higher yields than would otherwise be possible. One disadvantage of this approach is that there is no credited rate declared in advance. Rather, interest is calculated and applied retrospectively. Also, most of these products do not comply with securities law and state insurance law requirements that must be met to offer them to employers operating plans under IRC Section 403(b).

Individual CD-Type Products

There are a small number of providers that continue to offer a product to defined contribution plans that at the individual investor level is similar to a series of bank certificates of deposit (CDs). Investors allocate all deposits for a given time period, or window, to a particular CD tranche. All of the deposits made in the window receive a specified rate through maturity. The products have a perception of fairness, but that is significantly offset by complexity. Participants see many different "buckets" for their investments, and retirement plan administrators struggle to record keep all of the various generations of CDs.

Bond Mutual Funds and Other Alternatives

Nonstable value investments are still available as the fixed option in some plans. They may include bond mutual funds, money market funds, bank savings accounts, or credit union accounts. They often suffer from one of two disadvantages. In the case of a bond mutual fund, there is no *stability of principal.* Assets can lose value. This flies in the face of participant demand for stable value investing. In the case of money market funds, credit union accounts, or savings accounts, returns are usually lower over a full market cycle.

BUYERS OF STABLE VALUE PRODUCTS

Book-Value Accounting

The customers who purchase stable value products have been predominantly employer-sponsored retirement plans. Today, stable value products are used mostly to fund defined contribution plans. Stable value products lost their appeal to defined benefit plans in 1992, when the American Institute of Certified Public Accountants (AICPA) issued an accounting opinion that requires that investments used to fund such plans typically be held at something called "fair value," a shorthand in most cases for market value. This removed the ability of defined benefit plans to capitalize on the stable value aspect of these products.

Stable value accounting was confirmed for defined contribution plans with the AICPA's Statement of Position 94–4, which essentially states that defined contribution plans may hold benefit-responsive contracts at book value if they meet certain requirements. More recently, this position was confirmed by the Financial Accounting Standards Board (FASB), in Statement of Position AAG-INV-1.

In essence, a defined contribution plan can hold assets at book value, and thereby avoid the fluctuations and uncertainties of changes in market value, if participants can be confident of receiving plan benefits at book value. In this case, book value is based not on the fluctuation of value based on market conditions, but on the interest rate credited to participant deposits.

Types of Buyers

Buyers include:

For-profit employers. Private corporations usually offer plans with an Internal Revenue Code (IRC) Section 401(k) feature. Many similar plan types exist. Most would be eligible for stable value products.

Not-for-profit employers. Teachers, health care workers, and other not-for-profit workers often participate in IRC Section 403(b) plans. These are eligible for some stable value products, but not all. A provider must offer certain features to comply with state insurance law and securities requirements.

Governmental employers. Governmental employees (other then federal workers) participating in an IRC Section 457(b) deferred compensation plan may use stable value products much like participants in a 401(k) plan.

Mixed-plan structures. It is not unusual to see an employer offer several plan types to employees when they are eligible for multiple plans. This may occur, for example, when a not-for-profit also has status as a governmental entity. A plan might have a 401(k), 457, and 403(b) plan. Stable value may be used by each, but there may be restrictions on commingling the invested assets.

Taft-Hartley plans. As a rule of thumb, a Taft-Hartley plan may find stable value useful if it is more like a defined contribution plan in structure than a defined benefit plan.

Other potential buyers. Some of the buyers that have approached stable value providers looking for protection of principal on products have included foundations and endowments, managers of general treasury funds, municipal bond issuers, nuclear decommissioning trusts, and others. In general, these situations are limited by the availability of book value accounting to benefit-responsive arrangements, and by provider willingness to offer the requested features. However, there has been some success providing products with features similar to stable value products in other venues, such as short-term funding pools and for foreign entities.

COMMON FEATURES OF STABLE VALUE PRODUCTS

As previously noted in this chapter, there are many variations of features from one stable value product to another. However, there are a few areas that are virtually universal.

Interest Crediting

Though almost all products have some commonality in the way that they credit interest to participant accounts, such as daily interest crediting and a 365-day year, there are differences by product type. The most significant differences for each of the major product types still in use are described in Table 63.4.

Some products announce a rate to participants in advance of when it is credited. This is often an attractive feature for plans and their safety-conscious participants. Pooled funds do not use this approach. Rather, they credit interest after the fact when the blended return has been calculated.

Rates may be periodically reset. While pooled funds change rates daily, and GIC rates don't change at all (absent some unique contract forms), most other options use a quarterly, semiannual, or annual rate adjustment. For 403(b) plans, rates generally may decrease no more frequently than annually.

Products may have minimum guaranteed rates. These are usually specified percentages, but may be tied to an external reference benchmark, or to a formula. Guarantees may be lifetime or for a specified time period.

Table 63.4 Typical Interest Crediting Approaches by Stable Value Product Type

	General Account Portfolio	Traditional GIC	Separate Account GIC	Synthetic GIC	Pooled Fund
Rate declaration in advance?	Generally, yes	Yes	Yes	Yes	No
Rate change frequency	Periodic at issuer's discretion based on competitive landscape	None, except in unique constructs	Quarterly, semiannually, or annually	Quarterly, semiannually, or annually	Daily, in arrears
Minimum guarantees	Lifetime; annual; periodic	Lifetime at a specified rate	Lifetime, often of principal; periodic sometimes higher	Lifetime, often of principal; periodic sometimes higher	No, but underlying contracts usually have guarantees
Experience rating	Not explicitly to a single plan	Generally, no	Yes, over asset duration	Yes, over asset duration	Not explicitly to a single plan
Termination date	None	On a specific date or dates	None	None	None

When the interest rate is periodically reset, it may be at the discretion of the provider, or by formula. In separate account and synthetic GICs in particular, there is a relatively standardized industry approach:

$$MV(1+i)^d = BV(1+CR)^d$$

where

MV = The current market value of assets held in the separate account or trust under the synthetic.
BV = The book value of assets. Note that this should parallel the sum of participant balances, and reflects net deposits and withdrawals carried forward at the credited rate. *It is not the book value of the underlying fixed income instruments.*
i = The yield currently available on invested assets. Often an index yield is used.
d = The duration of the invested assets.
CR = The new credited rate to be used for participants.

The intent of the formula is notionally to spread the difference between the current market value and book value of assets over the duration of assets (similar to defeasance of a known liability). It is a notional approach only, in that market value and book value will never be equal, except by coincidence, but serves as an effective and fair way to smooth returns over time.

Figure 63.2 illustrates this approach graphically.

The act of periodically adjusting rates based on plan cash flows and investment performance is often referred to as "experience rating," though sticklers go a bit further and define contracts where rates change based on cash flow experience as "experience rated," and those that change based on investment performance as "participating."

There are a few additional considerations with regard to interest crediting. The first is that when there are multiple products backing a stable value option, there needs to be a methodology for blending the rates on the different products. A simple averaging approach is usually enough, and some way to incorporate into subsequent rates any current difference that has resulted from poor prior guesses.

The second is ensuring that the stable value providers can connect to the trading and administrative systems of

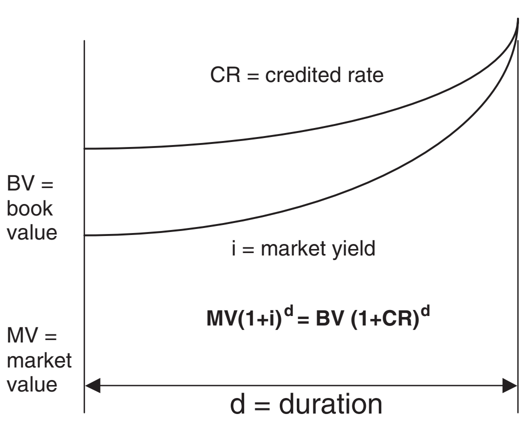


Figure 63.2 Graphic Explanation of a Stable Value Experience-Rating Formula

the record keeper. Most record keepers today work using a daily valuation system. The provider needs to produce a daily asset value. Alternatively, the provider can calculate a daily interest rate factor in advance, and use that to process a series of asset valuation factors.

Deposit and Withdrawal Limitations

A stable value issuer assumes a certain level of risk by making a credited rate guarantee to the contract holder. To help manage this risk, the issuer may include contract provisions within the stable value option specifying how frequently, and in what amounts, funds may be deposited to or withdrawn from the contract.

Deposits

When the issuer and plan sponsor enter into a contract, they will negotiate limits as to the aggregate amount and timing of new deposits. For example, the issuer may agree to a maximum deposit amount, often referred to as a "door" or "cap" provision, and the deposit period is generally known as the deposit "window." There may be a specified minimum deposit amount as well, known as a "floor."

The purpose of these deposit limitations is to help the issuer manage its interest rate risk. If market yields decline and the plan sponsor or participants are able to obtain an above market rate by depositing funds into the stable value option, this could result in a marked increase in deposits. Since the investments purchased with these funds will offer the issuer a lower yield, there will be downward pressure on the credited rate in the case of experience-rated products, and downward pressure on investment margins in cases where the insurer has guaranteed a fixed credited rate. Deposit limitations are much more prevalent in products with stronger guarantees, such as traditional GICs.

Withdrawals

In contrast to deposit restrictions, withdrawal restrictions help the issuer control its risk in the event of volatile or rising interest rates. When interest rates rise, the issuer of a stable value product offering a guaranteed fixed credited rate could face capital losses if the contract holder or participants elect to transfer funds out of the stable value option and into other investment options offering a higher yield.

Participant withdrawals for plan benefits—for example, retirement, death, disability, termination of employment, interfund transfers, loans, or emergencies—are generally not limited. Participants don't generally withdraw funds for these purposes simply for arbitrage. However, participant withdrawals for nonbenefit purposes, or sponsor withdrawals, are usually limited.

A *market value adjustment* provision limits the plan's ability to transfer funds out of the stable value option at book value in times of volatile interest rates. A provision of this type requires an adjustment to amounts paid out to approximate the changes in market value of the securities underlying the provider's asset portfolio backing the option. A book value "corridor" specifies the maximum percentage of a contract's book value (e.g., 20%) that the contract holder may transfer in a given year without a market value adjustment. The "corridor" approach is often also used as a way of limiting withdrawals paid for employer events, such as layoffs.

Transfers

Competing fund restrictions also help manage disintermediation risk by requiring that the plan sponsor not offer another fixed income investment option (such as a bond mutual fund or money market fund) alongside the stable value option. An alternative or supplement to the competing fund restriction would be an "equity wash" provision. Under an equity-wash provision, alternative fixed income investment options could be available, but participants would be unable to transfer money directly from the stable value option into these options. Instead, the money would have to be transferred to another fund involving some form of market risk (e.g., an equity mutual fund) for some period of time (most commonly 90 days) before it could be transferred to another fixed income option. This element of market risk helps prevent participants from arbitraging across the various available credited rates and allows the provider to credit a higher interest rate.

Withdrawal Hierarchy

If a plan sponsor's stable value investment option is supported by a number of stable value products, there will be a withdrawal hierarchy specifying the order in which the various contracts may have funds withdrawn from them. The withdrawal hierarchy is included in the contract, and is agreed upon at the time the contract is underwritten. Examples of withdrawal hierarchies include last in, first out (LIFO), first in, first out (FIFO), net pro rata, and gross pro rata. In a LIFO contract, the amount of any withdrawals in excess of current deposits and cash flows from maturing contracts will be taken from the contract with the most recent effective date. FIFO contracts make these withdrawals from the contract with the earliest effective date. In net pro rata contracts, withdrawals in excess of deposits are taken from each contract in proportion to the percentage of stable value funds held in that contract, while a gross pro rata hierarchy places all current deposits in the contract with the most recent effective date, and then makes all withdrawals on a pro rata basis.

Need for Exit Provisions

Stable value products may have a specified maturity date, the way that a GIC matures after a certain number of years like a bank CD. However, most stable value products work as evergreen structures. That is, they don't mature on a specific date. The provider must then allow the contract holder one or more ways to exit the contract.

In this event, there will often be an adjustment to reflect any change in the market value of the underlying assets. For example, if interest rates have risen significantly since the deposits were made, the contract holder will likely receive less than book value (that is, the amount of any deposits plus the interest credited under the contract) in the event of early withdrawal.

Market Value Adjustment

The difference between the book value and the value received, or "market value," is usually called the market value adjustment (MVA). Two frequently used methods for an insurer to administer the application of a market value adjustment are (1) to maintain two accounting records (one for book value and another for market value) or (2) to apply a market value adjustment formula to the amount of any withdrawals. Other methods might include the actual sale of underlying assets or the assignment of those assets to the plan sponsor.

Market value adjustment formulas may be either one-way or two-way. One-way formulas pay the lesser of book value or the market value-adjusted amount (that is, only negative adjustments are allowed), while two-way adjustments pay the market value-adjusted amount regardless of whether the adjustment is positive or negative.

Book Value Installments

Under certain circumstances, the contract holder may be able to receive book value at contract termination. Book value settlement options allow the contract holder to elect to receive the principal and interest guaranteed by the contract provided that the balance is paid out over a specified period of time (e.g., five years). An alternative approach is to allow a certain percentage of the contract to be withdrawn at book value every year. A specific subset of this category is used for pooled funds. Under a "one-year put," all funds may be withdrawn at book value within 12 months.

Transfer-in-Kind

A third option may be available in certain circumstances, usually in a product that invests in publicly traded securities, especially for larger plans. This might be the case, for example, with a separate account or synthetic GIC. This type of transfer arrangement is called a "transfer-in-kind." With a transfer-in-kind, the successor provider after termination receives the assets directly from the prior provider in kind. The successor provider may then decide to sell assets in the portfolio to adjust to a desired investment mix. If the securities held by the prior provider prior to transfer are not held solely for that sponsor, a methodology must be developed to prorate the holdings, maintaining the underlying asset characteristics, and allowing for transaction efficiency (e.g., to maintain round lots for trading).

Annuitization

Finally, many stable value products contain annuity purchase provisions, allowing the plan sponsor to use funds from the contract to purchase retirement annuities for individual plan participants. If there are annuity purchase provisions in the contract, guaranteed purchase rates will generally be specified, and purchases made at book value.

CONTRACT ISSUANCE

Annuity

Stable value products sold by insurers are traditionally issued in the form of an unallocated group annuity contract. The contract is most often issued to the plan trustee, but may also be issued directly to the employer or plan sponsor. As an insurance product, the contract form will usually contain guaranteed annuity purchase rates, and must be approved by state insurance regulators. In an unallocated contract, the issuer maintains a record of the aggregate value of the funds invested under the contract, but no record of individual participants' account balances. Instead, the plan itself or a third-party administrator (TPA) typically keeps track of participant balances using the information provided by the issuer (e.g., credited rates), mails participant statements, and performs other administrative functions.

Trust

The contract could also be issued as a trust agreement if the arrangement is with a bank, trust company, or pooled fund manager. In a trust agreement, plan contributions are provided to an investment trustee, which is distinct from the plan's master trustee. The investment trustee then has responsibility for investing the funds, accumulating earnings, and providing funds to the plan's master trustee to pay participant benefits. The trust agreement is a flexible structure (money may be invested in commingled trust funds or in separate investment contracts, and in equities as well as bonds) and is not subject to the approval of state insurance regulators.

Funding Agreement

The funding agreement and other "unbundled" approaches were developed by insurers in an effort to match the flexibility of the trust and to avoid some of the negative associations that buyers have with annuities. These arrangements generally do not include annuity purchase rates, but allow the plan sponsor to take advantage of the insurer's investment expertise.

There are also "split-funded" structures where the plan sponsor utilizes both a group annuity contract and a bank arrangement as part of the stable value option.

Because the buyers of stable value products are adjudged to be sophisticated investors, these contract forms usually do not need to be registered with the Securities and Exchange Commission (SEC)—a distinct advantage from the issuers' perspective, since filing a product with the SEC is a lengthy process requiring a great deal of effort and legal expense. However, the nature of certain plans, such as those established under IRC Section 403(b), do require that their investment options maintain certain minimum characteristics.

PLAN SPONSOR MANAGEMENT ISSUES

Buyers of stable value products face a number of considerations if they are to maintain a plan that keeps their participants satisfied and reflects the nature of their plan design, industry characteristics, and workforce demographics.

Diversification and Credit Risk

Much like a similar concern in the fixed income markets, sponsors want to avoid overexposure to the credit risk of a single issuer. Buyers will often diversify their purchases of stable value products to mitigate this risk. If the product is one with a strong guarantee from the provider, then it will be more important to purchase from a top-quality provider, or to diversify well across providers. In experience-rated products where the performance of the underlying instruments is passed through to the plan, diversification of invested assets is more critical than diversification by issuer.

In traditional GICs, then, the diversification concern ought to be at the issuer level. In separate accounts and synthetics, by contrast, look to the underlying fixed income portfolio. In general account portfolio rate products, both are important. In a pooled fund, the level of focus on each element will depend on the mix and characteristics of the underlying contracts.

In fixed-term contracts (traditional GICs), it is also important to diversify across time (that is, by purchase date and maturity date), so that not all purchases or sales are made in the same rate environment.

Economy of Purchase

In order to obtain a competitive rate from an issuer, a buyer often has to have a large enough placement (e.g., $5 to $10 million) to interest a group of issuers. This is a concern that buyers often must manage against their other objectives of diversifying by issuer, or in the case of fixed-term products, by frequency of purchase or maturity date.

Withdrawal Provisions—Participant Benefits

Most products should pretty routinely contain provisions to pay standard participant-directed benefit withdrawals at book value, as allowed by the plan. These benefits would typically include:

- Retirement
- Death
- Disability
- Termination of employment

- Emergency or hardship
- Loans
- Interfund transfers
- Annuity purchase

If there are multiple contracts, a plan should be sure to establish a withdrawal hierarchy and to coordinate withdrawal provisions across different products.

If the plan has the potential for volatile cash flows, either because deposits are irregular or withdrawals are unpredictable, holding a cash "buffer" fund often makes sense. A cash buffer is a shorter-term fixed income option or money market fund used as a shock absorber against volatile deposit and withdrawal patterns. The buffer is targeted at a certain level. When that level is significantly exceeded, the excess is deposited to the core stable value contract(s). When there is a significant shortfall, a withdrawal is made from the core stable value contract(s) to replenish the buffer. Typical buffer targets are in the 2% to 5% range. Returns on buffers are blended with core product returns to produce a single blended credited rate for use with participants.

Withdrawal Provisions—Plan Sponsor

This is an area that requires particular focus, both to assure that the provisions are fair and to be sure that they reflect the likely needs of the employer.

As noted, sponsor-directed withdrawals, such as might occur at the termination of the contract, are usually subject to contractual limitations or adjustments so that the provider may be assured of a longer-term investment horizon and generate a higher corresponding interest rate. The construction of these limitations or adjustments ought to be designed to reflect the economics of the underlying assets, plan cash flows, and interest rate movements, rather than set arbitrarily or with undue discretion by the provider. Table 63.5 denotes some of the typical limitations and adjustments frequently encountered, by product type.

MVAs vary widely by contract type and provider. Where they do come into play, the contractual provisions describing them need close scrutiny. In a general account portfolio, assets are invested in a large pool, and no buyer owns an individual interest in any investment in the insurer's general account. Rather, they own a promise from the insurer to pay a particular rate, which may change. MVA formulas are intended to replicate the buyer's theoretical share in the underlying pool by proxy.

Many contracts, especially those that are evergreen in nature, may have old formulas or somewhat dated provisions reflecting the market environment at time of purchase. Better or newer formulas will have these elements:

- A defined formula rather than an offer to provide a value upon request.
- Direct reference to nonmanipulable variables, such as an index or base rate.
- A proxy reflective of the characteristics (e.g., duration and quality) of the underlying assets.
- Specification of any fees or "haircuts" used by the provider in the formula to defray liquidation or unrecovered sales expenses, or the lack of liquidity of some of the assets.

Book value settlements typically contain similar issues:

- The duration of payout should correspond to that on the underlying assets.
- The interest crediting calculation and benefit liquidity provisions during settlement should be specified. Specifically, issuers may assess a fee or rate reduction during settlement. This fee or reduction needs to be disclosed and reasonable.

A transfer-in-kind provision needs to provide some information about how the pro rata share of transferred assets is calculated, and needs to parallel the quality, duration, and other characteristics of the underlying portfolio in establishing a subset of that portfolio for transfer. Illiquid assets are difficult to transfer, so provision needs to be made for alternatives or substitution.

Table 63.5 Typical Plan Sponsor Exit Limitations and Adjustments to Stable Value Products

Product Type	Market Value Immediately	Book Value Settlement	Other
Traditional GIC	Usually no market value available. If so, based on a formula.	NA—these contracts already have a specified maturity date	Not usually
General account portfolio rate product	Usually yes, by formula	Yes, usually over a set time period meant to reflect general account asset duration	Deferred sales charges may apply on older contracts
Separate account	Since invested in publicly traded securities, market value directly determinable using prorated asset share	Yes, over asset duration	Transfer-in-kind may be available
Synthetic GIC	Since the assets are held on the buyer's behalf in a trust, assets may be sold and the proceeds paid to successor provider	Often yes, over asset duration	Transfer-in-kind typical
Pooled fund	Not usually available	"One-year put"	One-year put pays assets within 12 months at book value

A payout option that is almost entirely unique to pooled funds is something often referred to as a "one-year put." Under this arrangement, investors in a pooled fund may withdraw assets at book value, but may need to wait up to one year to receive the proceeds. Some providers qualify this feature by allowing immediate withdrawal if the remaining pool investors aren't damaged, or depending on the size of the queue waiting for withdrawal. A few contain a "force majeure" provision that would further limit withdrawals in extreme market conditions.

Further Considerations at Time of Transfer

A frequent purchase concern of stable value buyers is how to effect a transfer from one record keeper or issuer to another if the stable value option changes hands. Some considerations are:

Transfer at Full Value

Sponsors want comfort that, if there is a market value adjustment or *deferred sales charge* with the prior provider, participants won't lose value in their stable value option. One protection that sponsors have used is not to describe the stable value option as "fully guaranteed," and to note in participant disclosure that there may be times when participants won't receive 100% of underlying asset value. Another, more practical approach is that there is a ready market from successor providers in so-called market value make-up contracts. These are contracts under which participant book value is maintained, and under which the successor provider bears a portion of any loss, or participants receive a reduced rate in the future to defray the costs. This approach is relatively common.

Delays in Payment

When a book value settlement or one-year put provision is engaged, sponsors want to smooth the transition from one provider to another. The usual way that this is accommodated is to maintain the prior account and new account side by side, to limit deposits and withdrawals to one of the two accounts, and to blend the rates to produce a single rate for participants.

Reinvestment Risk

Another concern is that any adjustment or delay not adversely affect participant credited rates. If a market value formula is fair, two-way, and a good reflection of underlying economics, it should not present an issue. With fixed income instruments, there is an inverse relationship between interest rates and market value. When market values decline, the proceeds may be reinvested at higher rates, and participant credited rates should be unaffected.

In the situation of a one-year put, this feature may work to either the benefit or detriment of participants. If rates rise, the sponsor may withdraw assets at book value in a high-rate environment and participants receive an increase in rate. If rates decline at the time a sponsor exits a contract, a book value transfer will decrease participant credited rates.

Table 63.6 outlines some of these considerations.

Table 63.6 Pros and Cons of Various Sponsor Transfer Provisions

Provision	Pro	Con
Actual market value	Convenient and efficient	Need to find market value make-up provider
MVA	If formula is economic, maintains rate for participants	Poor proxies can throw off rate; transaction costs of trading can reduce rate
Book value settlement	Maintains book value	Need to coordinate multiple contracts
Transfer-in-kind	No approximation required	Complex transaction best suited to large clients
One-year put	Easy to understand, and good protection in a rising rate environment	Rate decrease in declining rate environment

The Trade-off among Rate, Liquidity, and Quality

Plan Structure and Potential Plan Changes

One of the most important objectives of a plan sponsor managing a stable value option is to match the funding vehicles and the portfolio structure to the character of the particular plan and its participants. Here are some examples:

- A plan with high turnover of its employees might expect significant negative cash flows from its employees—it might want to maintain a liquid portfolio or keep a substantial portion of assets in cash.
- A small plan with assets concentrated with one or two highly paid individuals who could conceivably retire or leave the firm may invest differently than one with more diversified account balances.
- A plan with many investment options and liberal transfer restrictions can expect higher volatility of participant cash flow than one without these features.
- A plan for a company undergoing merger and acquisition activity or expected layoffs may have different needs than an established stable plan.

Duration

A stable value portfolio, just like other fixed income portfolios, has a duration, and it can be calculated in a similar manner. Fundamentally, duration measures the

Table 63.7 Typical Stable Value Option Investment Characteristics

Characteristic	Range	Typically Observed	Considerations
Duration	2–5 years	3 years	Shorter duration = lower yield, but better rate tracking
Quality	A to AAA	AA	Safety-oriented investment generally leads to higher quality
Liquidity	2%–5% cash buffer; benefit withdrawals at book value	3% buffer	Affirm that contract liquidity provisions are fair and balanced

responsiveness of the stable value option's credited investment rate to changes in external market interest rates.

Generally, in a normal, positively sloped yield curve environment, the credited rates on longer-term fixed income instruments are higher than those on shorter instruments. Absent any other factors, this might mean at first blush that a buyer of stable value instruments would prefer to have a long portfolio. Curiously, however, the duration of most stable value funds tends to be around three years, whereas the duration of the fixed income market, as approximated by the Lehman Brothers Aggregate Index, is between four and five years.

Are GIC and stable value fund managers giving up incremental yield unnecessarily, or is there some other issue driving this apparent duration decision? It could result from a number of factors:

- **Rate tracking.** The contract holder may want the overall return to move very quickly in the direction of prevailing interest rates; a portfolio with a three-year duration would be more responsive than one with a four- or five-year duration.
- **Conservatism.** Stable value funds are managed to be low risk; maybe most stable value fund sponsors are conservative.
- **Liquidity bias.** Some sponsors may wish to avoid the perception of being "locked in" to a contract or option.
- **Utility.** Sponsors may receive more benefit from flexibility than from a high rate. As one sponsor of a very large plan has noted "My return is asymmetric. If I get participants a few extra basis points, no one cares. If participants think something has gone wrong with the plan, I get yelled at."
- **Liquidity and cash buffers.** To manage plan changes and their effect on stable value option cash flow, as well as to maintain liquidity, many plans will maintain a cash buffer in their stable value option. The cash buffer works by receiving all deposits from participants and maturities of any existing maturing contracts, and by paying all withdrawals. The net amount available for investment is then deposited to a current stable value product. This simplifies plan management in that a plan does not need to approach several vendors to pay withdrawals, or can get better interest rate quotes from buyers due to reduced expected volatility.
- **Asset quality.** Generally, there is an inverse relationship between the quality of fixed income assets backing a stable value option and their current expected yield. A plan will need to decide where on the spectrum of quality and yield it would like to be. Although there are outliers, most plans operate in the AA range, most likely reflective of the stable value option being a perceived safety-oriented investment.

There is no one instrument that provides the highest possible interest rate, the highest quality, and immediate liquidity. A plan will need to decide which combination is best, based on its own characteristics. Table 63.7 summarizes some of the considerations in these trade-offs.

Book Value Accounting

Book value accounting and the benefit responsiveness that makes it possible are probably the primary reasons that sponsors buy stable value products today, mainly because most buyers are now defined contribution plans. A plan qualifies to use book value accounting if it meets AICPA guideline 94–4 and FASB Statement of Position AAG-INV-1. The benefit of book value accounting is that even if the market value of a stable value portfolio changes over time, the plan is allowed to credit participant accounts based on the book value (that is, based on the rate credited on its investment options). This allows the participants to avoid the day-to-day fluctuations of the market and see a steady stable return over time. Because of its fundamental impact on plan reporting to participants, a plan should be very careful to maintain its eligibility for book value accounting.

Reporting Needs

A plan needs to keep track of its investments and manage the various aspects of its portfolio, such as distribution by investment provider, duration, upcoming maturity amounts, liquidity, and so on. It should make whatever arrangements are necessary with its provider to produce periodic reports with the information it needs.

Participant Disclosure

A sponsor should adequately disclose situations to participants under which a participant may access their account at book value, and those under which access is limited.

ISSUER CONSIDERATIONS

Just as plan sponsors have many issues with managing stable value products, so do issuers.

Stable value products generally break down into three broad categories:

1. Fully guaranteed contracts or contracts with significant elements of a guarantee (e.g., a GIC or general account portfolio rate product).
2. Actively managed products such as a separate account or synthetic GIC, under which guarantees tend to be low and most investment performance is passed on to the contract holder over time.
3. Pooled funds, where the issuer is assuming neither guarantee risk nor active fixed income management responsibilities, but is responsible for the construction and operation of the pool.

(*Note:* It is not uncommon that a pooled fund operator will also have a subsidiary investment adviser manage plan assets, subject to restrictions on fees to avoid a conflict of interest.)

Each of these types exhibits different characteristics and creates different corresponding considerations.

An issuer needs to decide what type of investments to purchase to back its stable value products, and how those assets will be managed. Considerations include:

- Asset type
- Asset quality
- Asset duration
- Liquidity elements and cash position
- Investment optionality and derivatives usage
- Active or passive management

Fully Guaranteed Contracts

Fully guaranteed contracts are typically issued by an insurance company, and backed by a segment of the insurer's general account. Assets often include the full spectrum of fixed income, including publicly traded securities, private issuers, commercial mortgages, mortgage- and asset-backed securities, and even a few other instruments. Assets are often held for a longer duration, such as five years. The portfolio usually contains adequate cash liquidity to pay net withdrawals, since it is such a large pool. Liquid assets (publicly traded securities) are often actively managed, while illiquid investments are held to maturity.

Actively Managed Assets

Assets are usually held in publicly traded securities in a well-diversified portfolio, and, since returns are passed on to the contract holder over time, construction reflects sponsor preferences. Duration is often in the three- to five-year range, and asset quality AA. Assets are actively managed against an external benchmark. Derivatives use is limited to hedging and replication. Liquidity is not typically an issue given that securities can be sold on the public market.

Pooled Funds

A pooled fund operator usually doesn't hold the assets backing the pool, but rather diversifies by issuer of the GICs, synthetics, separate account, and other instruments that comprise the pool. A pooled fund will also look to the investment guidelines of the underlying assets to appropriately diversify.

ASSET/LIABILITY MANAGEMENT

An important subset of investment management is the management of assets against liabilities. This involves managing the cash flow of the assets to meet short-term needs for payments or maturities, managing the duration of the assets so that the value of the underlying portfolio moves in tandem with that of the liabilities, and managing convexity (a portfolio's tendency for its duration to shift over time as interest rates change).

Disastrous results can occur if a portfolio of investments is not properly matched to its corresponding liabilities. If rates rise dramatically, purchased assets may have a market value well below that granted to the buyer, and untimely liquidation of those assets could result in a significant loss.

Fully Guaranteed Products

Asset liability matching is most important in fully guaranteed products or products with substantial guarantees. Issuers of these products should match:

- The maturity schedule of assets and liabilities.
- The duration (average maturity weighted by present value of maturities).
- Nonlinear changes in value (e.g., on the asset side, the tendency of *mortgage pass-through securities* to shift in value based on prepayment activity; on the liability side, any options to the sponsor to redeem assets early at book value).

Actively Managed Products

Almost all experience under these products is passed through to the plan over time through changes to future credited rates, so asset-liability matching, except in extreme circumstances under which guarantees are breached, is in essence, automatic from the issuer's perspective. However, matching the investment characteristics to the plan's expected cash flow dynamics is critical to buyer satisfaction, even if a mismatch would not violate any explicit contractual guarantees.

Pooled Funds

Pooled funds operate on a fee basis, so asset-liability matching isn't really an issue, except again from a customer satisfaction standpoint.

UNDERWRITING

Regardless of the type of product issued, a provider needs to underwrite the risks of changes in plan cash flow activity, primarily the benefit-responsive risks it is taking on.

The benefit-responsiveness risk is the risk that the stable value product used has to pay out more in benefits than expected at an inopportune time, mainly after interest rates have risen so that the asset portfolio backing the product is worth less than the value of the benefits it must pay. In essence, the issuer has issued the equivalent of a financial option to the buyer, though one contingent on certain preagreed and presumably nonselectable events. Its risk is that the option is exercised in an adverse environment.

The risk is driven by participants' rights to withdraw funds at book value, regardless of the prevailing interest rate environment, for certain events such as retirement, death, disability, termination of employment, or transfer out of the stable value option to other available participant investment options. A similar cash flow–related risk is that a participant may deposit money to the issuer at the wrong time (e.g., after rates have dropped).

Some of the tools an issuer uses to underwrite benefit responsive risks are:

- An analysis of the plan's structure and provisions.
- An analysis of the plan's participant base.
- Historical deposit and withdrawal activity.
- Use of a cash buffer fund.
- Examination of the withdrawal hierarchy.
- Competitiveness of the plan's interest rate.
- Historical plan allocation of fixed/variable assets.
- A look at what other investment vehicles fund the stable value option.

In a fully guaranteed product, underwriting is critical to protecting the insurer's assets. In actively managed products or pooled funds, underwriting is also important. Since performance ultimately accrues to the benefit or detriment of plan participants, good underwriting can prevent one group of participants from adversely affecting another. In other words, it protects generational equity. Even in a pooled fund, underwriting is important so that one plan doesn't impact others negatively. In fact, pooled funds often contain an additional underwriting feature, a maximum plan ownership in the pool to ensure some level of diversification.

Table 63.8 illustrates typical risks under a defined contribution plan, and controls that are often used as protections against these risks.

LEGAL AND REGULATORY ISSUES

Issuers of stable value products must comply with a wide variety of laws, regulations, and accounting requirements, depending on the typical product(s) issued. Requirements include:

State law. Insurance company products are issued in most states as group annuity contracts, and sometimes as funding agreements. State insurance law applies with regard to the requirements for these particular forms, and for the issuance of contracts in general. There are also state laws that apply to the marketing of insurance products.

Securities law. Most stable value products are exempt from state or federal securities registration and sales requirements by virtue of the fact that they are sold to qualified retirement plans, which are adjudged to be sophisticated investors. However, state and federal antifraud laws do apply to the marketing of these products. Stable value products issued to 403(b) plans have unique securities law requirements to meet.

ERISA. The Employee Retirement Income Security Act of 1974 (ERISA) governs some of the requirements relating to how a stable value product can be sold and how the insurer may manage the underlying assets. If

Table 63.8 Typical Stable Value Option Underwriting Risks and Controls

Risk	Typical Control
Participants are laid off after rates have risen, and request massive withdrawals at book value under "termination."	Contractual provision to limit payments at book value in the event of "employer-initiated" events such as layoffs, early retirement incentive programs, spin-offs, divestitures, etc.
Participants transfer assets out of the stable value option to another investment option.	Issuers usually accept this risk with regard to equity options, but for "competing" options, those that are fixed income or short term in nature or have a principal guarantee, do not allow transfers or allow them only with a requirement that transferring participants maintain transferred assets in an equity account for a minimum time period (an "equity wash" provision).
Hot money issues: the plan has many retired participants still in the plan that can take their money at any time at book value.	Issuers generally underwrite for the risk rather than controlling it, and set fees accordingly.
The plan sponsor coaches participants to leave the stable value fund.	Anticoaching provisions, relieving issuer of obligations if coaching is used.
Plan deposits are higher than expected after rates drop, or lower than expected after rates rise, requiring issuer to credit an inappropriate rate compared to market rates.	Minimum and maximum deposit requirements.
Plan withdrawal hierarchy allocates all withdrawals to current provider (LIFO) and current provider gets high withdrawals.	Issuers often underwrite this risk, as it is short term in nature. Deposit floors also provide some protection.

ERISA issues do arise, they are usually with respect to potential conflicts of interest.

Trust law. Bank and trust products are subject to banking and trust laws, and some to requirements of the Office of the Comptroller of the Currency (OCC).

SOME HISTORICAL LESSONS LEARNED

Issuers of stable value products of all types have made mistakes. These mistakes are useful to remember so as not to repeat them.

Fully Guaranteed Products

Lesson 1: Forward Commitments

Up until the late 1970s, interest rates had remained fairly level for decades. In order to obtain higher rates, it had become common practice to commit forward on investments, that is, to buy investments to back a liability even before corresponding liabilities (at the time, GICs) were sold. When rates spiked in the late 1970s, this left insurers holding old investments with below-market rates.

Lesson 2: Open Windows

In the 1980s, the competition to issue GICs to large plans was fierce. On more than one occasion, insurers agreed to accept an unspecified amount of deposits, dramatically underestimated the deposit amounts as rates fell, and were forced to credit an above-market rate on the excess dollars deposited.

Lesson 3: Overinvestment in the Wrong Assets

Another result of competition has been occasional overconcentration of insurer assets in certain asset classes to achieve a rate advantage. When those assets, such as commercial mortgages, mortgage pass-through securities, or Asian debt later fell into cyclical downturns, insurers with overconcentration in these areas suffered.

All three of these errors can be looked at as different versions of the same issue; slowly becoming inured to the potential risks of overinvesting in assets or mismatching assets and liabilities.

Actively Managed Products

Actively managed products present fewer guaranteed risks to issuers, as experience is passed through to the plan. However, even with these products, issuers have made mistakes.

The biggest of these was providing options to the sponsor to exit the contract early at book value. When this happened, losses were so substantial that one issuer of products of this type became insolvent.

Pooled Funds

One issue for pooled funds has been one of trying to keep up with rising interest rates. Not so much an error as a structural issue, many pooled funds set up a second pool when rates rose to keep rates competitive. This meant the customers in the "old" pool ended up with a stale rate because they had limited new investment.

Another issue for pooled funds is that by investing in products issued by multiple providers, the risk of at least one problem with an issuer is increased. Sometimes, when these issuers had problems, pools have had to freeze or otherwise encumber access to assets. It didn't cause losses to customers, but did cause some inconvenience.

PROS AND CONS OF DIFFERENT STABLE VALUE OPTION FUNDING VEHICLES

In a vacuum, there is no right or wrong vehicle that a plan should use to fund its stable value option. However, a plan can make an assessment as to its best fit based on its characteristics.

Table 63.9 summarizes relative advantages and disadvantages of different vehicles, and profiles some typically good customer matches.

NEW DEVELOPMENTS IN PRINCIPAL-PROTECTED PRODUCTS

There are a number of factors reinvigorating interest in principal-protected products:

Higher interest rates. Although interest rates are still well below their historical acme, every time rates begin to rise, an interest in stable value products reappears.

Fee transparency. The fee structure on some newer stable value products is easier to explain than on older products. In an environment where buyers more than ever want to understand fees, these products are attractive.

Demographics. The Baby Boom is a fundamental demographic affecting a broad swath of U.S. consumption. One aspect of how it affects retirement investing is that defined contribution plan participants, in an environment where they are ever more responsible for their own future retirement needs, have substantial accumulated assets. Principal-protected products become attractive as they allow investors to protect these accumulated assets. Similarly, managing income is an objective of increasing importance to aging defined contribution plan participants. Principal-protected products can guarantee an income stream.

Some of the emerging uses for stable value include:

Managing postretirement income. Aging Baby Boomers are shifting from an accumulation mind-set to a lifestyle mind-set: "What can I do with my money, and how do I protect it?" Stable value concepts help with both aims.

Health Reserve Account (HRA) funding. Postretirement health care benefits are becoming an increasingly

Table 63.9 Comparison of Stable Value Funding Vehicles

	Fully Guaranteed	Actively Managed, Experience-Rated	Pooled Funds
Benefits	Higher minimum guarantees	Control over investment strategy	Quicker tracking of external rates
	Higher rate over a full market cycle	Can produce good returns over long term	Broad diversification
	Economy of scale in purchase of investments	Easy transfer provisions	Easy to understand
		Protection against credit risk	One-year put good in rising rate environment
Drawbacks	Credit exposure to single issuer	Individual account—no pooling	No rate declared in advance
	Approximations needed to estimate ownership shares	Subject to individual plan cash flow experience	Broader exposure to (smaller) credit events
	Older contracts often contain unattractive provisions		One-year put adverse in declining rate environment
Typical best-fit customer	Plans in need of substantive guarantee	Large plans	Small plans
	Some element of unpredictability in plan cash flows	Some expertise in or desire for setting investment parameters	Little participant interest in credited rates
		Substantial employer control	Customer sensitivity to ease of understanding rights of transfer
		Stable plan cash flows	

growing concern both for working employees and their employers. Vehicles are developing to assist in the prefunding of these obligations. Principal protection may be a valuable element of prefunding approaches.

Foreign entities. The use of the defined contribution model is spreading in various formats across the globe. Stable value can be an important concept in other countries just as it has been in the United States.

Defined benefit liability defeasance. Although defined benefit plans have decreased in usage, large liabilities still remain, and they are associated with an aging employee and retiree base. Stable value may have a role in defeasing those liabilities.

Retail uses. Although the SEC's position on book value accounting for stable value mutual funds has stymied growth in that area, latent underlying demand may still drive some type of a pooled retail stable value capability.

The attractiveness of stable value products changes with interest rates and regulations, but the underlying driving forces behind stable value—the desire for protection and stability—are fundamental to human nature. Stable value is likely to endure in some form for a long time to come, though the particular manifestation may vary from what stable value looks like today.

SUMMARY

Stable value investment options provide a flexible, adaptable funding vehicle for defined contribution participants saving for retirement. Because they tap into fundamental human concerns about protection and control, they are effective with savers across the retirement savings spectrum, regardless of age or income. Because of this same versatility, they are also emerging as viable alternatives for investors approaching or in retirement, and have potential broader application to any venue in which a group of individuals with common characteristics are saving for a future need or are managing an income stream, and wish to do so with some stability.

REFERENCES

Caswell, J. R., and Tourville, K. (1998). Managing a synthetic GIC portfolio. In F. J. Fabozzi (ed.), *The Handbook of Stable Value Investments* (pp. 65–86). Hoboken, NJ: John Wiley & Sons.

Caswell, J. R., and Tourville, K. (2005). Stable value investments. In F. J. Fabozzi (ed.), *Handbook of Fixed Income Securities*, 7th edition (pp. 471–485). New York: McGraw-Hill.

Gallo, V. A. (1998). Underwriting stable value risks. In F. J. Fabozzi (ed.), *The Handbook of Stable Value Investments* (pp. 277–300). Hoboken, NJ: John Wiley & Sons.

Haendiges, H. K., and Keener, E. A. (1998). Traditional GICs. In F. J. Fabozzi (ed.), *The Handbook of Stable Value Investments* (pp. 19–44). Hoboken, NJ: John Wiley & Sons.

LeLaurin, S. F., and Guenther, J. P. (1998). Stable value management. In F. J. Fabozzi (ed.), *The Handbook of Stable Value Investments* (pp. 327–373). Hoboken, NJ: John Wiley & Sons.

Mercier, J. L., Turco, A. A., Smith, K. J., and Smith W. M. (1998). Legal, regulatory, and accounting issues. In F. J. Fabozzi (ed.), *The Handbook of Stable Value Investments* (pp. 19–44). Hoboken, NJ: John Wiley & Sons.

Pearce, T. (1998). Buy and hold synthetics. In F. J. Fabozzi (ed.), *The Handbook of Stable Value Investments* (pp. 87–112). Hoboken, NJ: John Wiley & Sons.

Rudolph-Shabinsky, I., and Psome, C. J. (1998). Mangaged synthetics. In F. J. Fabozzi (ed.), *The Handbook of Stable Value Investments* (pp. 113–150). Hoboken, NJ: John Wiley & Sons.

Smith, K. M., and Koeppel, S. E. (1998). In F. J. Fabozzi (ed.), *The Handbook of Stable Value Investments* (pp. 45–59). Hoboken, NJ: John Wiley & Sons.

PART 7

Foreign Exchange

CHAPTER 64

An Introduction to Spot Foreign Exchange

SHANI SHAMAH
Consultant, E J Consultants

Abstract: The foreign exchange market is by far the largest market in the world, and although the world's currency markets are generally thought of as the exclusive domain of the largest banks and multinational corporations, nothing could be further from the truth. Even though major currencies are traded like commodities, it is distinguished from both the commodity or equity markets by having no fixed base. In other words, the foreign exchange market exists through communications and information systems consisting of telephones, the Internet or other means of instant communications, for example, Reuters and Bloomberg. The foreign exchange market is not located in a building, nor is it limited by fixed trading hours, but is truly a 24-hour global trading system. It knows no barriers and trading activity in general moves with the sun from one major financial center to the next—so that around the clock a foreign exchange market is active somewhere in the world. Because of this decentralization, the total size of the foreign exchange market can only be guessed at. The foreign exchange market is an over-the-counter market where buyers and sellers conduct business. Many of the traders in the markets have all started with this simplest of products: just buy low and sell high, or sell high and buy low. Thus, the foreign exchange market is a global network of buyers and sellers of currencies with a foreign exchange transaction being a contract to exchange one currency for another currency at an agreed rate on an agreed date. Today, what began as a way of facilitating trade across country borders has grown into one of the most liquid, hectic, and volatile financial markets in the world—where banks (and many hedge funds) are the major players and have the potential of generating huge profits or losses.

Keywords: foreign exchange, spot rate, spot value, reciprocal rate, indirect terms, European terms, fixed currency, variable currency, American terms, bid, offer, spread, market maker, price taker, market user, big figure, direct dealing, brokered dealing, cross rates, price determinants, market/price risk, country risk, credit risk, dealing room, dealers, back office, two-way price, quote, pips, risk, hit, foreign exchange exposure

The foreign exchange (FX) market includes the cash market and the FX derivatives market. The focus in this chapter is on the cash market, which is more commonly referred to as the spot foreign exchange market.

A foreign exchange, or currency rate is simply the price of one country's money in terms of another's. Although exchange rates are affected by many factors, in the end, currency prices are a result of supply-and-demand forces. The world's currency markets can be viewed as a huge melting pot: In a large and ever-changing mix of current events, supply-and-demand factors are constantly shifting, and the price of one currency in relation to another shifts accordingly. No other market encompasses as much of what is going on in the world at any given time as foreign exchange.

Approximately 80% of foreign exchange transactions have a dollar leg. The dollar plays such a large role in the markets because:

- It is used as an investment currency throughout the world.
- It is a reserve currency held by many central banks.
- It is a transaction currency in many international commodity markets.
- Monetary bodies use it as an intervention currency for operations in their own currencies.

The most widely traded currency pairs are:

- The American dollar against the Japanese yen (USD/JPY)
- The European euro against the American dollar (EUR/USD)
- The British pound against the American dollar (GBP/USD)
- The American dollar against the Swiss franc (USD/CHF)

In general, EUR/USD is by far the rnost traded currency pair and has captured approximately 30% of the global turnover. It is followed by USD/JPY with 20% and GBP/USD with 11%. Of course, most national currencies are represented in the foreign exchange market, in one form or another. Most currencies operate under floating exchange rate mechanisms against one another. The rates can rise or fall depending largely on economic, political, and military situations in given country.

The basic information and common definitions of foreign exchange and the foreign exchange market follow:

- *Foreign exchange market* is a global network of buyers and sellers of currencies.
- *Foreign exchange* or *FX* is the exchange of one currency for another.
- *Foreign exchange rate* is the price of one currency expressed in terms of another currency.
- *Foreign exchange transaction* is a contract to exchange one currency for another currency at an agreed rate on an agreed date.
- *Spot exchange rate* is the ratio at which one currency is exchanged for another for settlement in two business days (value date)

BRIEF HISTORY

So, what is the history of the foreign exchange market? Rather than start way back in history with the barter system and discuss when coins were first introduced, let's start with when the original method for exchange and payment of international debits and credits was to use gold. To do this, countries agreed not to restrict the cross-border flow of gold and to allow their gold coins to be melted down and recast by other countries. During this period, rate fluctuations and associated risks were small. The gold standard lasted until World War I. To finance the war, most countries printed large amounts of money—far in excess of their gold reserves—leading to the demise of the gold standard. Efforts to return to the gold standard failed after the war because most currencies were either over- or undervalued and were not easily matched to a gold standard. In addition, worldwide inflation at this time caused disparities among currencies, which led to currency devaluations and further inequalities among currencies. Needless to say, this was a very difficult time to exchange foreign currencies.

At the end of World War II, in an effort to avoid the difficulties encountered after World War I, the Americans proposed a system based on fixed exchange rates and the creation of the International Monetary Fund (IMF) at the Bretton Woods conference in 1944. The new system had three goals:

1. To create a system with stable exchange rates
2. To eliminate exchange controls
3. To allow convertibility of all currencies

To do this, the Americans guaranteed that it would buy and sell gold at $35 per ounce, thereby establishing the dollar as a parity reference for all currencies and gold. In other words, the dollar replaced gold as the dominant reserve currency of the international monetary system. As a result, in addition to reserves of gold, countries held reserves of dollars, which earned interest and were easily converted into gold. The dollar became the major currency for settlement of international transactions.

This system worked well until the late 1960s when, in spite of attempts made to stabilize the markets, the widely different growth rates in individual countries caused difficulties in the fixed rate system. Some countries revalued their currencies, while others let them "float." Eventually, this led to the demise of the Bretton Woods system. In 1973, due to continued loss of confidence in the American dollar and massive American balance of payments deficits, the system of fixed exchange rates finally collapsed. Because fixed rates were not practical, most countries let their currencies float against other currencies. This was seen as a short-term solution, until a return to stable, fixed rates was possible. However, the IMF conference in 1976 formally adopted the flexible exchange rates with a "gentlemen's agreement." Member nations are required to abstain from rate manipulation and from creating unfair advantages over other member nations.

The foreign exchange market is not entirely a free market because some countries have rules regarding repatriation of funds and some currencies are fixed or semifixed to other currencies. An example of the latter is the old European Monetary System (EMS), which was an exchange rate mechanism including most European Community (EC) countries where currencies fluctuated within relatively narrow, mutually agreed upon bands.

Since currencies have been able to float, their values have fluctuated dramatically. Continuous adjustments in currencies' values have brought volatility to the foreign exchange market. As a result, any company or institution doing business which involves currencies other than its own is faced with exposure to changes in the values of those other currencies.

FOREIGN EXCHANGE EXPOSURE

By way of explanation, foreign exchange exposure is the risk of financial impact due to changes in foreign exchange rates and, in general, there are three types of foreign exchange exposures:

1. Transactions exposures principally impact a company's profit and loss and cash flow and result from transacting business in a currency or currencies different from the company's home currency.
2. Translation exposures principally impact a company's balance sheet and result from the translation of foreign assets and liabilities into the company's home currency for accounting purposes.
3. Economic exposures relate to a company's exposure to foreign markets and suppliers. More difficult to identify, economic exposure is sometimes also referred to as competitive, strategic, or operational exposure.

There are actually five basic foreign exchange products:

1. Spot transactions
2. Forward contracts
3. Foreign exchange futures contracts
4. Foreign exchange swaps
5. Currency options

The last four are referred to as foreign exchange derivative contracts.

BASIC USES

The basic uses of foreign exchange products include the following:

- For settlement and funding in order to convert cash from one currency into another for commercial transactions (e.g., import or export payables or receivables) or to convert capital flows (e.g. dividends, inter-company loans and investments)
- To hedge/manage foreign exchange exposures caused by the passage of time and exchange rate fluctuations
- For arbitrage to take advantage of short-term discrepancies between prices in different currencies or marketplaces
- For investment to take advantage of changing exchange rates and interest rates and to optimize all components of a global investment strategy
- To speculate so as to take advantage of anticipated exchange rate changes

In actual terms for spot transactions, client groups, such as corporations, investors, funds, and institutions, will use spot transactions as part of their foreign exchange management programs. Speculators will also use this market because it is an extremely active and liquid market with roughly two-thirds of all foreign exchange activity being traded. There can be plenty of movement (volatility) in any one day, which will enable a speculator to possibly benefit from such gyrations.

CHARACTERISTICS

The foreign exchange market has some unique characteristics. As has already been mentioned, it is active 24 hours a day. Generally speaking, there is no centralized market place as there is with the stock market. Rather, deals are done in an "over-the-counter" style, with individual buyers and sellers dealing verbally, or acting through brokers; over the telephone or electronically over the Internet via various foreign exchange trading platforms.

This means that rates change from dealer to dealer rather than being controlled by a central market. For example, investors do not call around to get the best price on a specific stock because the price is quoted on the stock exchange, but they do call around to different dealers to get the best exchange rate on a specific currency. They may also refer to various widely available bank/broker screens (e.g., Bloomberg and Reuters) for indicative pricing only.

There are two exceptions to the lack of a physical marketplace. First, foreign currency futures are traded on a few regulated markets, of which the better known are the International Monetary Market (IMM) in Chicago, the SIMEX (Singapore International Monetary Exchange), and the LIFFE (London International Financial Futures Exchange) in London.

Second, in some countries there are daily "fixings" where major currency dealers meet to "fix" the exchange rate of their local currency against currencies of their major trading partners at a predetermined moment in the day. Immediately after the fixing, the rates continue to fluctuate and trade freely. The fixings happen less in this day and age and are only symbolic meetings and represent less than 0.5% of all worldwide daily trading.

Although there is no central marketplace, there are major dealing centers in three regions of the world where much of the world's FX transactions take place. There are also many smaller centers in different parts of the world.

The western European market is serviced by major dealing centers in London, Frankfurt, Paris, and Zurich. The North American market is serviced by the major dealing center in New York City, but there are also active dealing rooms in Chicago, San Francisco, Los Angeles, and a few other major cities in the States. The far eastern market is serviced by major dealing centers in Tokyo, Hong Kong, and Singapore. Around 84% of the world's foreign exchange business is executed in these major dealing centers.

MAJOR PARTICIPANTS AND THEIR ROLES

To make a market means to be willing and ready to buy and sell currencies. Market makers are those market participants that both buy and sell currencies. According to market practice, a market maker, dealer, or trader will generally quote a two-way price to another market maker. (The terms *dealer* and *trader* are used interchangeably when referring to market makers.) For market makers, reciprocity is standard practice. They constantly make prices to one another, and market makers are primarily banks.

Price takers are those market participants seeking to either buy or sell currencies and are usually corporations, fund managers, or speculators. For price takers, there is no reciprocity inasmuch as they won't quote a price to other market participants.

The major participants in the market play a number of roles depending on their need for foreign exchange and the purpose of their activities:

- International money center banks are market makers and deal with other market participants.
- Regional banks deal with market makers to meet their own foreign exchange needs and those of their clients.
- Central banks are in the market to handle foreign exchange transactions for their governments, for certain state-owned entities, and for other central banks. They also pay or receive currencies not usually held in reserves and stabilize markets through intervention.
- Investment banks, like money center banks, can be market makers and deal with other market participants.
- Corporations are generally price takers and usually enter into foreign exchange transactions for a specific purpose, such as to convert trade or capital flows or to hedge currency positions.
- Brokers are the intermediaries or middlemen in the market, and as such do not take positions on their own behalf. They act as a mechanism for matching deals between market makers. Brokers provide market makers with a bid and/or offer quote left with them by other market makers. Brokers are bound by confidentiality not to reveal the name of one client to another until after the deal is done.
- Investors are usually managers of large investment funds and are a major force in moving exchange rates today. They may engage in the market for hedging, investment, and/or speculation.
- Regulatory authorities, while not actually participants in the market, impact the market from time to time. This sector includes government and international bodies. Most of the market is self-regulated, with guidelines of conduct being established by groups such as the Bank for International Settlement (BIS) and the IMF. National governments can and do impose controls on foreign exchange by legislation or market intervention through the central banks.

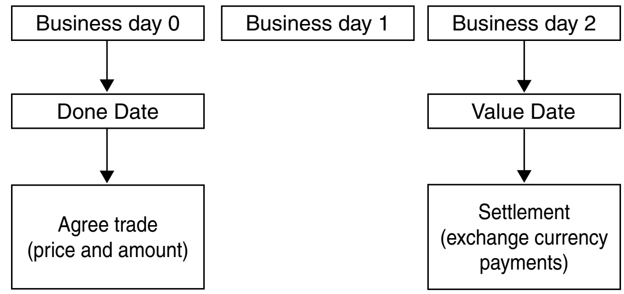


Figure 64.1 Example of a Foreign Exchange Transaction

Speculators are growing in numbers by the day as access to the Internet becomes more freely available and with ease of access to various online trading platforms. It is said that 95% of the daily foreign exchange volume is made up of trading or speculation, while the remaining 5% of daily volume consists of governments and commercial activities.

SPOT FOREIGN EXCHANGE

Foreign exchange rates are a means of expressing the value and worth of one economy as expressed by its currency as compared to that of another. Normal market usage is to quote the exchange rate for spot value, that is, for delivery two business days from the trade date (except Canadian transactions against the dollar, where the spot date is only one day). The two business days are normally required in order to get the trade information between the counterparties involved agreed and to process the funds through the local clearing systems. The two payments are made on the same date, regardless of the time zone difference.

Figure 64.1 provides an example of a foreign exchange transaction.

Spot and Reciprocal Rates

The rate used in a spot deal is the spot rate and is the price at which one currency can be bought or sold, expressed in terms of the other currency, for delivery on the spot value date. The spot exchange rate can be expressed in either currency; thus, this price has two parts, the base currency and the equivalent number of units of the other currency. For example, a rate for the U.S. dollar ($ or USD) against the Swiss franc (CHF) would be quoted as 1.2507 on July 25, 2006. (We will use exchange rates on this date throughout the chapter.) This means there are 1.2507 francs to $1. When one rate is known, the spot exchange rate expressed in the other currency (the reciprocal) is easily calculated. The price of $1, expressed in Swiss francs, is 1/$0.7996 or CHF 1.2507.

Although some newspapers calculate and publish both exchange rates, it has become standard market practice among traders to quote foreign exchange for most currencies as the amount of foreign currency that will be exchanged for $1. For example, if a bank trader were asked to quote a rate for Swiss francs against the dollar, the response would most likely be CHF 1.2507 rather than $0.7996. In this case, $1 is the traded commodity and the trader is quoting the price in Swiss francs. This type of quotation is known as European terms.

It should be noted that the U.S. dollar is the most popularly traded currency because it is the primary currency for international trade and official reserves (although the Euro is rapidly catching the U.S. dollar as a reserve currency). Therefore, the most frequently quoted exchange rates and the most liquid markets are those between dollars and various foreign currencies.

European and American Terms

Normally, the accepted practice is to quote in indirect or European terms. The price in dollar terms signifies how many dollars a single unit of the foreign currency is worth. In this case, the foreign currency is the fixed currency and the dollar is the variable currency.

However, there are exceptions to this rule. In contrast, the European euro, the British pound, Australian and New Zealand dollar, and some other old British area currencies, such as the Maltese pound, are quoted as the number of American dollars to the currency. This is an overhang from the days when these currencies were primarily quoted against sterling and therefore adopted the same quoting convention as sterling against the dollar. In this case, the currencies are always expressed in American or direct terms. For example, if a Swiss franc against the dollar price is expressed, it would read CHF 1.25/$1 or $1 equals CHF 1.25. This means that it costs CHF 1.25 to buy $1. The American dollar is the fixed currency and the Swiss franc is the variable currency. Alternatively, if a dollar against sterling price is expressed, it would read $1.84/£1 or £1 equals $1.84. This means that it costs $1.84 to buy £1. The pound is the fixed currency and the dollar is the variable currency.

Thus, the usage of either European terms or American terms is based on market practices. A market trader would quote pounds sterling as $1.8424 per pound and euro as $1.2598 per euro. This distinction is critical to understanding foreign exchange quotes and dealing screens. For example, when a corporate treasurer telephones a bank asking for foreign exchange quotes, the trader will assume the treasurer understands market conventions and will quickly rattle off prices for different currencies in the customary European or American terms.

SPOT TRANSACTIONS

Bid-Offer Spreads

Foreign currency traders are considered to be dealers when they make a two-way market price, that is, not just quoting one rate but two—that is, they provide both a bid price at which a trader is willing to buy a currency and an offer price at which a trader is willing to sell a currency. Examples of two-way quotes are:

EUR/USD	1.2598—1.2601
USD/JPY	117.06—117.09
USD/CHF	1.2507—1.2510

It should be noted that, generally, most currencies run to four figures after the point (significant figures after the decimal point) but there are exceptions like the U.S. dollar against the Japanese yen, which, as can be seen, runs to only two.

Like other financial markets, the spread favors the dealer who buys currency at one price and sells it at a slightly higher price. To determine whether a trade will take place at the dealer's bid or offer rate, a client must first know which currency the dealer is bidding or offering, that is, whether the terms are quoted in European terms such that the traded currency is $1 or in American terms, such that the traded currency is one unit of foreign currency.

Reading Foreign Exchange Rates

As with commodities and equities, foreign exchange has very specific ways of quotation and it is necessary to become familiar with these. Domestically, most countries use the direct quotation, and internationally, it is convention in the foreign exchange markets to quote most currencies against the dollar, with the dollar as the base currency. Using the dollar base simplifies currency trading and it will allow a trader to compare rates more easily.

If, for example, the dollar against the Swiss franc is quoted as 1.2507/10, what does this quote actually indicate? As has been mentioned, a market maker will normally quote a two-way price; in other words, they are obliged to make a bid and an offer for dollars against, in this case, the Swiss franc. Market makers will always quote to their advantage and to the other person's disadvantage. The left-hand side of the quote (1.2507) is the quoting person's bid for dollars, obviously surrendering as few Swiss francs as possible. Conversely, the right-hand side of the quote (1.2510) is the market maker's offer for dollars at which they will ask for as many Swiss francs as possible. The difference between the rate at which someone will buy a currency and the rate at which they will sell is called the profit (spread).

For example, when a market maker quotes spot Swiss franc (against the dollar), the trader will say: "dollar/Swiss franc is 1.2507—1.2510" where their bid is at CHF 1.2507/$1 and their offer is at CHF 1.2510/$1. That is, the market maker will buy $1. for CHF 1.2507, which means the client will sell $1 for CHF 1.2507—the client sells at the market maker's bid. Conversely, the market maker will sell $1 for CHF 1.2510, which means the client will buy $1 for CHF 1.2510—the client buys at the market maker's offer.

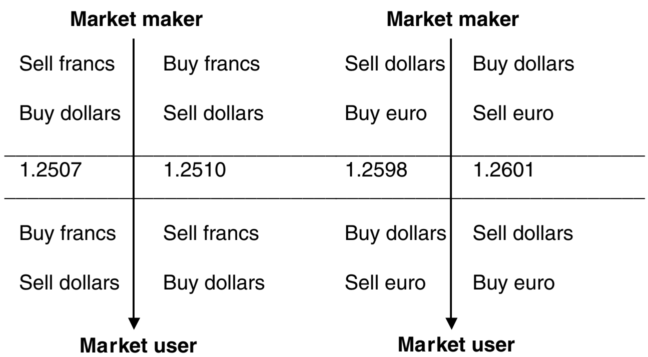


Figure 64.2 Example of a Currency Transaction

Figure 64.2 provides an example of a currency transaction.

It is important to remember that in any foreign exchange transaction, each party is both buying and selling, since it is buying one currency while selling another. One way of determining which is the buying rate and which is the selling rate is to remember that a market maker will buy dollars for another currency at a low rate (its bid rate) and sell dollars for another currency at a high rate (its offered rate). For currencies like sterling and euro, the market maker will buy sterling for dollars at the bid rate and sell sterling for dollars at the offered rate.

Big Figures

Usually, market makers will quote only the last two numbers in the price, for example 06/09, thus assuming the other party knows the rest of the price, which is known as the big figure, in this case 117 (dollar/Japanese yen). In an example of, say EUR/USD, being quoted 1.2598—1.2601, the big figure is both 1.25 and 1.26. However, in this case, the trader would say: "98—01, around 1.26."

Spread

As mentioned earlier, by quoting a higher offer than bid, the market maker ensures that if both sides of the quote are dealt on simultaneously, the market maker will profit from the difference between the bid and offer. This difference is the spread, and the size of the spread is affected by various factors. The main factors are the assessment of risk, the volatility of the market, the liquidity of the currency, and the time of day in each time zone.

Direct versus Brokered Dealing

When a dealer calls another dealer for a price, it is called *direct dealing*. When a dealer puts a bid or offer in at a foreign exchange broker or via some Internet trading platforms, it is called *brokered dealing*. Brokered dealing is somewhat like a silent auction, as the buyers and sellers are unaware of each other's identity until the deal is done, and the bid and offer process may not be accepted.

CROSS RATES

A *cross rate* is the rate of exchange between two currencies that do not involve the domestic currency. In other words, within the international marketplace, cross rates have come to mean rates that do not involve the American dollar.

Most transactions are dealt as the American dollar against another currency. However, currencies are also dealt against each other, for example the Swiss franc against the Japanese yen (CHF/JPY). In these instances, it is necessary to calculate the cross rate. In order to calculate this cross rate, start with the two rates against the dollar. The objective is to obtain the number of Japanese yen per Swiss franc Consider:

One dollar	=	francs 1.2507/10
One dollar	=	yen 117.06/09

Each quotation represents a bid and an offer for the currency against the dollar. The cross rate is achieved by taking opposite sides of the two prices. The rate for selling yen and buying francs is achieved by using the left-hand side of the dollar/yen (bid) rate and the right-hand side of the dollar/franc (offer) rate. The same logic is applied for buying yen and selling francs. Thus, the preceding quotations can be broken down as follows:

$$117.06 \div 1.2510 = 93.57$$
$$117.09 \div 1.2507 = 93.62$$

Hence, the spot cross rate for Swiss francs against the Japanese yen is 93.57/62. The number of places after the decimal point is determined by the convention of the quoted currency (the variable currency). In this example, this is the yen, since we are looking for the number of yen per franc. It is usual to quote cross currency exchange rates using the "heavier" currency as the base, for example, the number of yen per franc.

There is, of course, a variation to the rule. For currencies like the euro and sterling, it is market practice to multiply the respective currencies against each other. For example, consider the following;

One dollar	=	yen 117.06/09
One pound	=	dollar 1.8424/29

Then the sterling against yen spot cross rate calculation will be:

$$117.06 \times 1.8424 \text{ and } 117.09 \times 1.8429 = 215.67/79$$

Thus, it follows that one pound is equal to 215.67 or 215.79 yen. This is the only way it is expressed.

PRICE DETERMINANTS

Exchange rates (or prices) in the foreign exchange market are driven by the laws of supply and demand. The supply and demand for specific currencies change given the amount of trade and investment being done in that currency. If there is a high demand for a currency, its value increases. If there is a low demand, then its value decreases.

However, exchange rates are affected not only by supply and demand. The exchange rate will also be influenced by the economic, political, monetary, and social factors of the country involved and also by outside developments. Exchange rates can change quickly and significantly, reflecting the volatility in the market; and rates can also be moved by rumors and anticipated factors. Typically, currency rates can fluctuate from day to day due to small imbalances in supply and demand and to economic and political factors that affect the sentiment of market makers and investors.

Economic factors include economic policy, disseminated by government agencies and central banks, and economic conditions, generally revealed through economic reports. Economic policy comprises government fiscal policy (budget/spending practices) and monetary policy, that is, the means by which a government's central bank influences the supply and "cost" of money, which is reflected by the level of interest rates. Economic conditions include:

- *Government budget deficits or surpluses*. The market usually reacts negatively to widening government budget deficits, and positively to narrowing budget deficits.
- *Balance of trade levels and trends*. The trade flow among countries illustrates the demand for goods and services, which in turn indicates demand for a country's currency to conduct trade. Surpluses and deficits in the trade of goods and services reflect the competitiveness of a nation's economy. For example, trade deficits may have a negative impact on the currency.
- *Inflation levels and trends*. Typically, a currency will lose value if there is a high level of inflation in the country or if inflation levels are perceived to be rising. This is because inflation erodes purchasing power, thus demand, for that particular currency.
- *Economic growth and health*. Reports such as gross domestic product (GDP), employment levels, retail sales, inflation figures, and others, detail the levels of a country's economic growth and health. Generally, the more healthy and robust a country's economy, the better its currency will perform, and the more demand for it there will be.

Internal, regional, and international political conditions and events can have a profound effect on the currency markets. For instance, political upheaval and instability can have a negative impact on a nation's economy. The rise of a political faction that is perceived to be fiscally responsible can have the opposite effect. Also, events in one country in a region may spur positive or negative interest in a neighboring country and, in the process, affect its currency.

Market psychology is perhaps the most difficult to define but it does influence the foreign exchange market in a variety of ways:

- *Flight to quality*. Unsettling international events can lead to a "flight to quality," with investors seeking a "safe haven." There will be a greater demand, thus a higher price, for currencies perceived as stronger over their relatively weaker counterparts.
- *Long-term trends*. Very often, currency markets move in long, pronounced trends. Although currencies do not have an annual growing season like physical commodities, business cycles do make a difference. Cycle analysis looks at longer-term price trends that may arise from economic or political trends.
- "*Buy the rumor, sell the fact*." This market truism can apply to many currency situations. It is the tendency for the price of a currency to reflect the impact of a particular action before it occurs and when the anticipated event comes to pass, react in exactly the opposite direction. This may also be referred to as a market being "oversold" or "overbought."
- *While economic numbers can certainly reflect economic policy, some reports and numbers take on a talisman-like effect*. The number itself becomes important to market psychology and may have an immediate impact on short-term market moves. "What to watch" can change over time. In recent years, for example, money supply, employment data, trade balance figures, and inflation numbers have all taken turns in the spotlight.

RISK CONSIDERATIONS

It must be remembered that there are risks with spot transactions. First, there is credit risk. Like the risk a bank incurs when making a loan, a foreign exchange contract poses the risk that the client will not perform according to the terms of the contract (that is, will not deliver the appropriate currency on time). In a foreign exchange transaction, the market maker and the client agree that each will deliver to the other a specified amount of a currency on a specific date, at an agreed rate. Trading the currencies of countries that are in different time zones compounds this risk.

Second, there is market/price risk. Trading in any currency has a degree of risk. Exchange rate risk is inevitable because currency values rise and fall constantly in response to market pressures. When engaging in a foreign exchange trade, the client's position is open until it is closed or covered. While that position is open, the client is exposed to the risk of changes in exchange rates. A few moments can transform a potentially profitable transaction into a loss.

Third, there is country risk. Some countries (and their currencies) are more risky than others. Country risk may be due to anything from governmental regulations and restrictions to political situations, or the amount of foreign currency reserves the country has. However, this risk is usually of less significance.

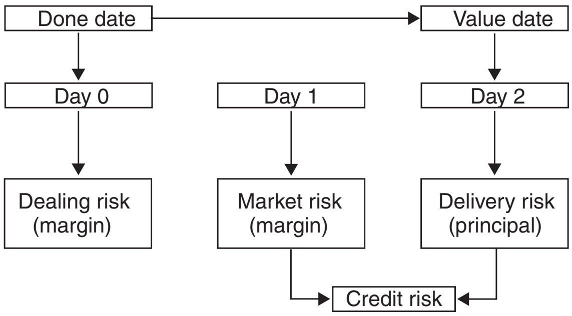


Figure 64.3 Spot Exchange Risk

The spot exchange risk is shown graphically in Figure 64.3.

ASKING FOR A QUOTE

Today, dealings in the foreign exchange market are usually done over the telephone or via the Internet. The Internet is a very quick and easy way to transact a foreign exchange transaction, with "click and deal" or "request for price" systems dominating. However, the majority of all trading is still transacted over the telephone. Using the telephone allows for prompt and timely execution but does leave room for errors in communication.

When asking for a quote, the following basic information needs to be conveyed to the market maker directly or to a corporate foreign exchange dealer: type of transaction, currency pair, and quantity. For example, a client would say: "I want to buy CHF 5 million value spot—what is your quote, please?"

Today, a market maker normally quotes a two-way price, where the trader stands ready to bid for, or offer up to, some standard amount. The difference between the two prices, as already mentioned, is the spread. Market convention, where trading is between market "professionals," is not to quote the big figures. Instead, a trader tends to quote only the last two figures of the price, the pips. For example, if the rate of dollars against Swiss francs were 1.2507/10, then the trader would quote only 07/10. That is, the trader bids for dollars at 1.2507 and offers dollars at 1.2510. If a client is checking prices with other market makers, then the client should inform the trader by saying that it is at his risk, that is, the quote can be changed by the trader. The client should then ask again when a new price is needed. If the client wishes to deal, then the trader's price would be hit, that is, where one side of the price or the other is accepted. Written confirmation, whether the deal is oral or electronic, will be exchanged and instructions taken as to whether this trade is to be settled or not and the currency amounts are transferred into the designated accounts on the value date.

As an example, suppose XYZ Corporation, based in Japan, needs to raise dollars to pay for a delivery of machine parts from America. The treasurer gets in contact with his bank dealer to arrange to buy the dollars in the spot market. The treasurer's screens display indicative prices contributed by certain major banks. This gives the treasurer a good idea of the current exchange rate. However, this is only an indicative rate and it is not a dealable price (that is, not a price to be dealt on). Therefore, the following steps are taken:

Treasurer: Please quote me dollar/yen in 10 million dollars (*at this stage there is no mention of whether the client wants to buy or sell dollars*).
Dealer: 06/09.
Treasurer: I buy 10 million dollars.
Dealer: To confirm, you buy 10 million dollars against yen at 117.09 value spot.

At this stage, the dealer fills in a deal ticket with the details of the trade including currencies, amount, which currency is bought and which one is sold, value date, exchange rate, counterparty, and settlement details if known.

The treasurer could also have said "at 09" or "mine." All three ways would be correct and within market practice. Also, the dealer knew that the treasurer was used to market conventions and, hence, did not quote the big figure of 117.

As another example, suppose that Mr. Jones is a high-net-worth individual who has a margin account with ABC International. He calls up and speaks to his favorite salesperson. After the customary pleasantries, Mr. Jones asks for a dealing price for sterling in "half a pound." The salesperson obtains the price from the trader and communicates it to Mr. Jones for consideration. It is likely that Mr. Jones has a general idea of where the price is from his computer screen. If Mr. Jones accepts the price, the salesperson immediately informs the trader, possibly via a hand signal. It is then up to the trader what happens with that position. The dealing conversation could be:

Mr. Jones: A dealing price for half a pound, please.
Corporate dealer: Sure, price for half a pound coming—Charlie (trader), cable in half?
Charlie: Who for?
Corporate dealer: Old Mr. Jones.
Charlie: What's he doing?
Corporate dealer: How do I know—just give me the price.
Charlie: 20/30.
Corporate dealer: Cable in half a pound is 1.8420/30.
Mr. Jones: Hmm, I was hoping for a better spread.
Corporate dealer: Your risk, let me try for you, you know it is only in half a pound but let me ask.
Charlie: What's he doing—trading or not?
Corporate dealer: He is looking for a better spread.
Charlie: What, in half a pound? You decide—he is your client.
Corporate dealer: I am 1.84.23/28 now.
Mr. Jones: Hmmm. . . well. . .
Corporate dealer: Mr. Jones, your risk again.
Mr. Jones: Okay, how now, please?
Corporate dealer: Charlie—how are you left on that half pound for Mr. Jones?
Charlie: Has he still not dealt—do what you like within 20/30 and let me know this side of Christmas.
Corporate dealer: 1.8424/29.
Mr. Jones: Okay, I sell.
Corporate dealer: Okay to confirm you sell half a million pounds and buy dollars at 1.8424 for value spot.
Mr. Jones: Agreed and thanks.

In the two preceding examples, the process happened in only one or two seconds, as the market can move very rapidly. The salesperson and the trader can always change the price as long as the client has not firmly accepted the last quote made. Before making a decision to actually trade, it is not unusual for a client to shop around for quotes from various market makers in order to obtain the best deal. Sometimes, clients will ask for quotes knowing that they are not ready to actually trade but just want to check where the market is.

It should be noted that the following information is needed when asking for a quote:

- The two currencies being traded (e.g., $/JPY)
- The value date of the trade (e.g., spot)
- The amount (e.g., $1 million)

If possible, the trader will try to know what side of the price you are or, at best, guess. However, wherever possible, always ask for a two-way quote. But, in all cases, be extremely clear in the details of a trade or the instructions given, in order to avoid costly errors at a later stage. It is all too easy to mishear a quote "for a half" and think it is a quote in "four and a half."

SUMMARY

In days gone by, banking institutions were the sole purveyor of the information vital to the transaction of business in the market. With no central organized market, bank dealers executed trades solely by telephone or telex, writing trade details on pieces of paper, keeping positions on blotters, and maintaining charts by hand. The resulting scarcity of information meant that price discovery was inefficient, bid-offer spreads were wide, margins were large, and major institutions played the largest role simply because they knew where activity and prices in the marketplace were occurring. As a result, foreign exchange trading was a profitable activity for these institutions. The risks of trading were somewhat controlled and isolated at the bank level, with a degree of volatility sufficient enough to warrant active participation by only the most sophisticated of market participants.

Today, most clients are pretty sophisticated when it comes to knowing where the market is and what bid-offer spreads to expect for the foreign exchange deals. Banks and brokers that are uncompetitive in their pricing don't even leave the starting blocks in the race to win foreign exchange business. What distinguishes the best from the rest is the provision of high-quality information, in the way of charting and flows of relevant market information. Also, clients are looking for systems that are Internet enabled, scalable across regions, reliable, and safeguarded against crashes. In addition, clients are looking for Internet platforms that offer real-time risk management systems, among other demands.

Today, access to price information is widely available. Both institutional investors and retail investors can now gain live access to multicontributor price feeds, which can be downloaded directly into spreadsheets, if necessary. In addition, up-to-the-minute political and economic developments are widely available through news sources such as MMS and CNBC. As a result of these developments, however, bid-offer spreads have collapsed, as have profit margins, and this in turn has hampered the growth of direct investor participation.

Also, any currency trader needs to develop a global perspective and a feel for intermarket relationships. Interest rate trends are the most important external information source. If the U.S. Federal Reserve cuts rates, the dollar should weaken. However, if the European Central Bank is expected to follow suit, interest rate trends will converge and the value of the currency may not change at all. A trader will need to follow not only the chairman of the Federal Reserve, but also his counterparts in the European Central Bank and the Bank of Japan as well.

Currencies tend to be trendier than either stocks or commodities, and it is important to understand the trend from both a fundamental and a technical point of view. Currencies cannot reverse trends very easily because economies do not reverse quickly relative to one another. However, a successful trader will recognize how fundamentals and technicals combine to indicate a trend reversal. Of course, finding the trend reversal before it happens is the "Holy Grail" of trading. No one can be expected to be successful 100 percent of the time.

Activity in the foreign exchange market still remains predominantly the domain of the large professional players. Foreign exchange is both a science and an art. Risk can be quantified and alternatives identified to reduce or eliminate it. But judgment and personal attitudes toward risk, as well as other personal and corporate orientations, are required for consistent position management. However, with liquidity and the advent of Internet trading, plus the availability of margin trading, this 24-hour market is accessible to any person with the relevant knowledge and experience. Nevertheless, a strong, disciplined approach to trading must be followed, as profit opportunities and potential loss are equal and opposite.

REFERENCES

De Grauwe, P. and Grimaldi, M. (2006). *The Exchange Rate in a Behavioral Finance Framework*. Princeton, NJ: Princeton University Press.

Lyons, R. K. (2006). *The Microstructure Approach to Exchange Rates*. Cambridge, MA: MIT Press.

Rosenberg, M. (2003). *Exchange Rate Determination*. New York: McGraw-Hill.

Samo, L. and Taylor, J.P. (2002). *The Economics of Exchange Rates*. United Kingdom: Cambridge University Press.

Shamah, S. (2003). *A Foreign Exchange Primer*. London: Wiley Finance.

Taylor, F. (1997). *Mastering Foreign Exchange and Currency Options*. Upper Saddle River, NJ. Prentice Hall.

Weisweiller, R. (1990). *How the Foreign Exchange Market Works*. New York: New York Institute of Finance.

CHAPTER 65

An Introduction to Foreign Exchange Derivatives

SHANI SHAMAH
Consultant, E J Consultants

Abstract: Not surprisingly, over the past few years the financial markets have responded to increasing price volatility and there are now a range of financial instruments and strategies that can be used to manage the resulting exposures to financial price risk. At one level, there are now financial instruments that permit the direct transfer of financial price risk to a third party, who is more willing to accept that risk. At another level, the financial markets have evolved to the point whereby financial instruments can be combined with other instruments to unbundled financial price risk from the other risks inherent in the process, for example, raising capital.

Keywords: forward contracts, interest rate differentials, short-dated contracts, long-dated contracts, broken-dated contracts, nondeliverable forward (NDF), foreign exchange swaps, currency swaps, foreign exchange futures, exchange for physical (EFP), currency options, forward points, forward pips or swap points, premium, discount, bids, offers

In today's business world, unpredictable movements in exchange rates, interest rates and commodity prices cannot only affect a company's performance but may even determine whether a firm survives. Over the past couple of decades, companies have been increasingly challenged by financial price risks. It is no longer enough to be the company with the most advanced production technology, the cheapest labor supply, or the best marketing team. Price volatility can put even well-run companies out of business and changes in exchange rates can create strong new competitors. Similarly, fluctuations in commodity prices can drive input prices to the point that substitute products (or products made from different inputs) become more affordable to end users. Changes in interest rates can put pressure on the company's costs, as higher interest rates may hurt sales and thus the company could find themselves in financial distress as sales plummet and borrowing costs skyrocket. Hence, it is not surprising that the financial markets have responded to increasing price volatility with a range of financial instruments and strategies that can be used to manage the resulting exposures to financial price risk. Such financial instruments are covered in this chapter and includes FX forward contracts, nondeliverable forwards (NDFs), FX swaps, currency swaps, FX futures, and exchange for physical (EFP).

FOREIGN EXCHANGE FORWARD CONTRACTS

Forward contracts are a common hedging product and are used by importers, exporters, investors, and borrowers. They are valuable to those with existing assets or liabilities in foreign currencies and to those wanting to lock in a specific foreign exchange rate in the future. For example, corporations that must receive or pay foreign currencies in the future because of their normal business activities usually prefer to transfer the risk that the values of these currencies will change during the intervening period. They can use the bank forward market to establish today, the exchange rate between two currencies for a value date in the future. Generally, when corporations contract to pay to or receive from a bank foreign currency in the future, no money is exchanged until the settlement on the value date.

While forwards may be used to hedge payables and receivables, corporations will also hedge other assets and liabilities on a company's balance sheet. The value dates of forward contracts are often constructed to match up with the expected dates of receipts for a foreign payment, or payment of a foreign currency obligation. A forward contract can be tailored to meet a client's specific needs in terms of delivery dates and amount. In addition to transacting with clients, banks actively trade forward currency commitments among themselves, as well.

In essence, forwards provide certainty in the uncertain world of currency movements by locking in a specific rate, and as the forward markets are quite liquid, the bid/offer spreads are relatively low for the major currencies.

Definitions

By way of definition, a forward contract (or forward outright) is a transaction executed today in which one currency is bought or sold against another for delivery on a specified date that is not the spot date, for example, three months from now. In addition, forward points are relative *interest rate differentials* expressed as units of currency, or fractions of the spot value of that currency.

Interest Rate Differentials

Forwards work much like spots, but the value date is different from the spot date and usually extends further into the future, for example, six months from the commencement date. At first sight there would seem to be no reason why the spot and the forward rate are not the same. However, one of the factors influencing a currency's forward exchange rate is the level of interest rates for that currency relative to interest rates in the other currency. There are many theories on how a forward exchange rate can be calculated, but market participants adopt the interest rate differential between two currencies and the current market spot rate, as the basis of their calculations. The forward price is often referred to as *forward points*, *forward pips* or *swap points* (pips).

For example, assume the spot and forward rates between dollars and sterling are the same, but the interest rates in sterling are 4% per annum for a three-month deposit, while in dollars they are 2%. Investors would sell their dollars and buy sterling spot for the higher yield. They would simultaneously sell sterling and buy dollars forward for delivery at the end of the investment period. In this way, the investor would end up with more dollars than if the investment had been kept in dollars.

Periods

Market makers regularly trade forward contracts for periods of 1, 2, 3, 6, and 12 months from the spot value date value date. A *broken date* or odd date forward deal is a contract with maturity other than a normal market quote of complete months. An example would be to ask for the forward pips for 24 days.

Premium or Discount

Forward contract prices are determined by two main factors: the current spot price between the two currencies and the interest rate prevailing in each of the two currencies. The forward price is calculated as the spot rate plus or minus the forward pips. To decide whether to add or subtract the forward pips, firstly determine whether the currency to be bought or sold is trading at a premium or is trading at a discount. As all exchange rates have a fixed and a variable component, if the interest rates in the variable currency are greater than those of the fixed currency, the variable currency is trading at a discount relative to the fixed currency and forward pips are added to the spot rate to obtain the forward rate. If the interest rates in the variable currency are less than those of the fixed currency, the variable currency is trading at a premium and forward pips are subtracted from the spot rate to obtain the forward rate.

There are two simple rules of thumb to decide if a currency is at a premium or a discount, and what to do with the forward points. Remember, exchange rates are quoted as units of that currency, which equal \$1 (except for sterling, euro, and a few other currencies like the Australian and New Zealand dollar).

If the forward points are ascending, for instance, if the offer is numerically higher than the bid (20/25), that is, if the forward points rise from left to right, the currency is at a discount to the dollar and, hence, the forward points are added to the spot rate. (The major exception is sterling and the euro, where they are at a premium to the dollar.) If the bid is numerically higher than the offer, that is, the points are descending (25/20), that is, the forward points decline from left to right, the currency is at a premium to the dollar and the forward points are deducted from the spot rate. (The major exception is the quotation for the euro and sterling against the dollar, where if the forward points decline from left to right, the points are deducted, but the dollar is at a premium to the euro and sterling).

Calculations

Forward rates are not determined by where the market expects the currency to be in the future, but rather by the interest rate differential. Also, the forward exchange rate is fixed at the time of the transaction, but no accounts are credited or debited until the maturity date.

The forward pips are calculated in the following way. If we assume that the spot and forward rates between dollars and sterling are the same, say 1.4400/10, but the interest rates in sterling are 4% per annum for a three-month deposit, while in dollars they are 2% per annum for the same deposit, investors would sell their dollars and buy sterling spot for the higher yield. They would simultaneously sell sterling and buy dollars forward for delivery at the end of the investment period. In this way, the investor would end up with more dollars than if the money had been kept in dollars. For example, if Mr. Jones has \$5 million to invest for three months, at 2%, the interest earned at the end of the period will be:

$$\frac{5{,}000{,}000 \times 92 \times 2}{360 \times 100} = \$25{,}555{,}56$$

Thus, the total principal and interest earned at the end of the period will be \$5,025,555.56 (\$5,000,000 + \$25,555.56).

However, if Mr. Jones buys sterling at 1.4410 and sells his dollars, he will receive £3,469,812.63, which can be invested at 4% for the same period. The interest earned at the end of the period will be:

$$\frac{3{,}469{,}812.63 \times 92 \times 4}{365 \times 100} = £34{,}983.32$$

Thus, the total principal and interest earned at the end of the period will be £3,504,795.95 (£3,469,812.63 + £34,983.32).

This can then be converted back into dollars at 1.4400, which would give an amount of \$5,046,906.17 (£3,504,795.95 × 1.4400). The total gain, at the end of the period will be \$21,350.61 (\$5,046,906.17 − \$5,025,555.56).

In a free market, however, the advantage of the higher sterling interest rate is usually neutralized by the lower value of sterling in the forward foreign exchange market and any yield pickup will be small, or nonexistent.

In calculating the forward points, users adopt a simple arithmetic formula which takes the interest rate differential per annum, converts it into a differential for the required period, and then expresses the spot rate as a percentage of the differential for the period. However, it cannot be used entirely in isolation, for it assumes knowledge of relative interest rate levels by the interested party. It is, in essence, a variation on the old banking formula:

$$\text{Principal} \times \text{rate} \times \text{time} = \text{interest}$$

where the principal is the spot rate, the rate is the interest rate differential and time is the maturity in days. Thus:

$$\frac{\text{Spot rate} \times \text{Interest rate differential} \times \text{Days}/360}{1 + (\text{Currency interest rate} \times \text{Days}/360)}$$
$$= \text{Pips/Points}$$

In other words, the formula for dollars against currency forwards is:

$$\frac{A \times D \times (B - C)}{(100 \times E) + (C - D)}$$

which equals the number of forward points of spot currency, with 360 day basis, where

A = spot exchange rate
B = currency interest rate
C = dollar interest rate
D = maturity in days
E = day basis

It has to be noted that, in the money market, all calculations are based on the actual number of days elapsed divided by 360, except for calculations involving sterling and some other currencies when 365 days are used. The formula is adjusted when the two currencies involved have a different day base. Also, when the value date is the last business day of a month, the corresponding date in any future month is also the last business day. For example, if the spot value date were February 28, the value date in a one-month forward transaction would be March 31. If the spot value date were May 31, the six-month forward transaction date would be November 30. If the last day of the month is not a business day, then the value date is the next preceding business day.

Bids and Offers

Just as there is a *bid* and *offer* in the spot market, there is a bid and offer rate in the forward market as well. This means that the forward points for both sides of the exchange rate must be quoted. A typical example of how forward rates are quoted is:

Currency	1 month	3 month
USD/JPY	19.55/19.30	62.7/61.7
USD/CHF	0.1/1.1	1.7/1.9
EUR/USD	9.07/8.99	29.5/27.8

To Add or Subtract

As already has been mentioned, the simple rule to arrive at the forward rate is if the forward pips decline from left to right, the currency is at a premium to the dollar and the forward pips are deducted from the spot rate. For example, if spot \$/JPY is 117.06/117.09, the one-month forward price is:

$$117.06 - 0.1955 \text{ and } 117.09 - 0.1930 = 116.8645/116.897$$

If the forward pips rise from left to right, the currency is at a discount to the dollar and the forward points are added to the spot. For example, if spot \$/CHF is 1.2507/10, the three-month forward price is:

$$1.2507 + 0.00017 \text{ and } 1.2510 + 0.00019 = 1.25087/1.25119$$

The major exception is the quotation for sterling against the dollar, where if the forward pips decline from left to right, the pips are deducted but the dollar is at a premium to sterling. Occasionally, it is possible to have forward pips that have a negative number for one side of the quote and a positive number for the other. An example would be –0.7/+1.3. The rules for adding or subtracting are still the same. This type of forward pips behavior occurs when the interest rates of the two currencies are so close, that the offer side of one crosses the bid side of the other.

Forward Quotes

When a market maker quotes a forward price, the trader is likely to say:

"Three-month dollar/yen 62.7 at 61.7"

Where the market maker will buy and sell JPY at −62.7 pips (sell and buy dollars) and will sell and buy JPY at −61.7 pips (buy and sell dollars). Of course, for currencies quoted in American terms, that is, euro, the market maker will quote a one-month forward price as:

"One-month euro/dollar 9.07 at 8.99"

where the market maker will buy and sell dollars at −9.07 pips (sell and buy euro) and will sell and buy dollars at −8.99 pips (buy and sell euro). Some examples of forward quotations are:

Market maker

Buy + Sell yen Sell + Buy dollars	Sell + Buy yen Buy + Sell dollars
–62.7	–61.7
Sell + Buy yen Buy + Sell dollars	Buy + Sell yen Sell + Buy dollars

Market user

Market maker

Buy + Sell dollars Sell + Buy euro	Sell + Buy dollars Buy + Sell euro
–9.07	–8.99
Sell + Buy dollars Buy + Sell euro	Buy + Sell dollars Sell + Buy euro

Market user

Forward Cross Rates

Forward cross rates are worked out in the same manner as for spot rates. First, work out the forward rate from the spot rate and the *forward points*. Then, decide what currency is being bought and which one sold. Finally, decide if the rates should be divided or multiplied by one another, as appropriate.

For example, by using the forward points in the table below,

	USD/JPY	USD/CHF	CHF/JPY
Spot:	115.90/95	1.4409/14	80.41/80.47
3-month pips:	53.9 – 53.6	27 – 26	23 – 22

CHF/JPY three-month forward can be worked out as:

3 months USD/JPY	3 months USD/CHF	3 months CHF/JPY
115.90 – 115.95	1.4409 – 1.4414	80.41 – 80.47
−53.9 −53.6	−27 −26	−23 −22
115.361 to 115.414	1.4382 to 1.4388	80.18 to 80.25

As described in Chapter 64 of Volume I where spot pricing and calculations are covered, the bid for spot Swiss francs against Japanese yen is derived by taking 115.90 (bid) and dividing it by 1.4414 (offer), in order to reach the spot price of 80.41. Likewise, in order to obtain the offer spot price of Swiss francs against Japanese yen, take the offer of dollar yen, 150.95 and divide it by the bid of dollar Swiss franc 1.4409, which gives 80.47. In other words:

Market maker

Sell yen Buy dollars	Buy yen Sell dollars
115.90	150.95
Buy yen Sell dollars	Sell yen Buy dollars

Market user

Market maker

Sell francs Buy dollars	Buy francs Sell dollars
1.4409	1.4414
Buy francs Sell dollars	Sell francs Buy dollars

Market user

Market maker

Sell yen Buy francs	Buy yen Sell francs
80.40	80.47
Buy yen Sell francs	Sell yen Buy francs

Market user

Risks Involved

Because of the time span involved in forward contracts, there can be significant risks, just like with a spot deal. Credit risk, market/price risk and country risk are all potential problems. In fact, country risk is more significant than for spot trades as unexpected events in a foreign country are more likely given the longer period of exposure.

Short-Dated Contracts

As has been stated, most foreign exchange deals are executed for value two business days forward, or longer. However, some participants could have a need for currency the same day, the next day or the day after spot. For some currencies, like sterling and euro, it is possible to trade for the same day value, but for the majority of currencies, the earliest execution would be tomorrow.

The terminology for these differing time periods is:

Value same day	Overnight	o/n
Value tomorrow	Tom next	t/n
SPOT		
Value day after	Spot next	s/n

The rates quoted for value dates occurring before spot are treated in a different way to those occurring after spot. The rules applying to a value date after spot is that if the forward pips go from high to low (20–18), they are subtracted and if the forward pips go from low to high (18–20), they are added to the spot rate. But, if the value date is before spot, the points are switched and then the normal rule is followed. The pips for "overnight" and "tom next" represent only one day each. If, therefore, a rate is calculated for value today, the pips for "overnight" and "tom next" have to be added together.

Long-Dated Contracts

With any forward transaction, the quotation is based on the relationship between the prevailing interest rates of the two currencies concerned. However, when considering forwards beyond one year, it is necessary to account for the annual interest compounding effect.

The short method for calculating a forward rate beyond one year is to use the normal formula for calculating forward points from interest rate differentials. However, this formula has to be modified to take in to account the effect of the compounding effect of the interest.

Note: The principle for short- or *long-dated contracts* is the same as with forward rates and is made on the basis of interest rate gain or loss. The exception to the rule is that prices normally added on are deducted and prices normally deducted are added. This, actually, is not as odd as it sounds. If prices are normally quoted for spot delivery and a value tomorrow quote means the market-maker will have to surrender that currency earlier than normal, there has to be some compensation.

Broken-Dated Contracts

A *broken-dated contract* is a forward contract with maturity other than the normal market quote of complete months and in order to price a broken dated contract, it is necessary to interpolate between the two standard date quotations on either side of the desired maturity. For example, to work out the forward pips for USD/JPY for one and half months (45 days), assume the following rates:

$/JPY spot	117.06/117.09
1 month forward pips	21/18
2 month forward pips	44/41

The pips for buying JPY and selling USD would be calculated according to the following:

- Work out the number of days in the period between the one and two month forward quotes, because the delivery date falls within this period. The answer is 45 days.
- Subtract the bid one-month forward pips from the bid two-month forward pips, which will then show what the two-month pips are worth over the one-month pips. The answer is 23 forward pips.
- Divide the difference in forward pips (23) by the number of days in the period between the two standard quotes (45) and multiply the answer (0.5111) by the difference in the number of days between the required date and the last day of the two month quote (15). Hence, the total of those days is worth 7.7 forward pips.
- Subtract this answer (7.7) from the two-month forward pips, giving us the forward pips for the broken date of 36.3.

The interpolated rate is the basis for the market maker's quote, but the actual rate quoted will also probably reflect the market maker's position.

Outright Forwards

An outright forward contract is the purchase or sale of a currency for delivery on any date other than spot and not forming part of a swap operation. For example, an importer might want to fix the rate today, for the delivery of a shipment in two months' time. The process for the rate on an outright is to use the spot rate and the two-month forward pips. If spot dollar yen is 123.32/37 and the two-month forward pips are 44/41, the outright price for two months time is 122.88/96 (123.32 – 44 and 123.37 – 41). An outright forward transaction is shown in Figure 65.1.

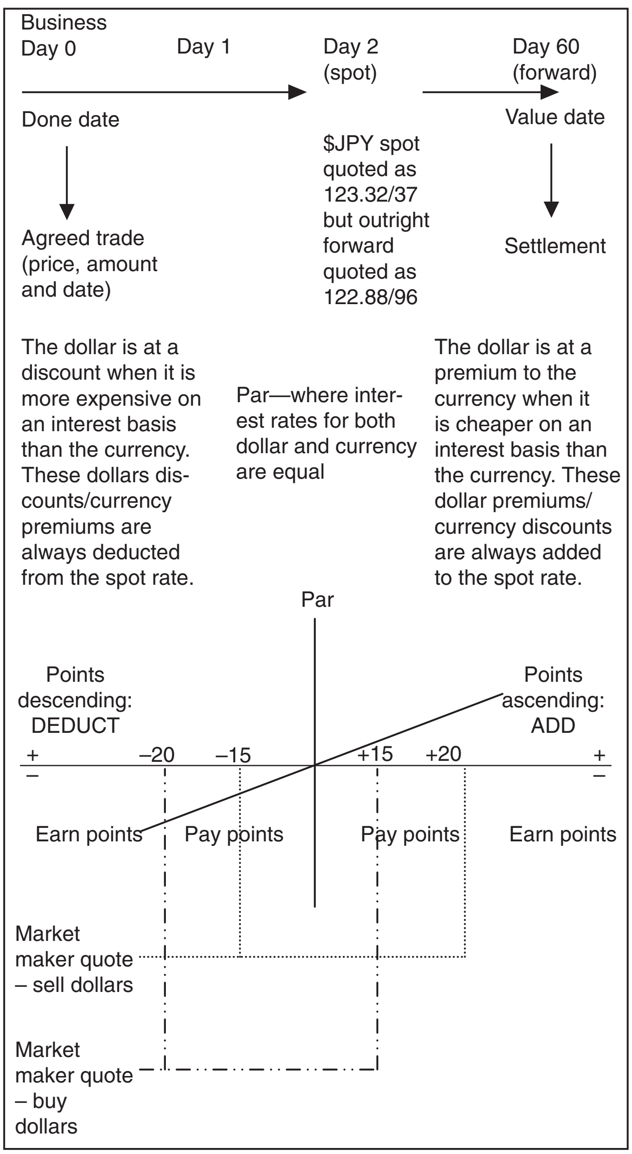


Figure 65.1 Illustration of an Outright Forward Transaction

NONDELIVERABLE FORWARDS

A *nondeliverable forward* (*NDF*) is a short-term committed forward "cash settlement" currency derivative instrument. It is essentially an outright (forward) foreign exchange contract whereby on the contracted settlement date, profit or loss is adjusted between the two counterparties based on the difference between the contracted NDF rate and the prevailing spot foreign exchange rates on an agreed notional amount.

Fixing Methodology

When an NDF deal is contracted, a fixing methodology is agreed. It specifies how a fixing spot rate is determined on the fixing date, which is normally two working days before settlement, to reflect the spot value. Generally, the fixing spot rate is based on a reference page on either Reuters or Telerate with a backup of calling between three and five market banks. Settlement is made in the major currency, paid to or by the client and reflects the differential between the agreed-upon NDF rate and the fixing spot rate.

How Quoted

The NDF is quoted using foreign exchange forward market convention, with two way prices quoted as bid/offer pips, at a premium or discount to the prevailing spot market. The spreads are more than likely wider than would be expected in the normal forward market. As with a normal forward transaction, the market user either buys or sells the NDF, depending on the position to be hedged or according to the view of the underlying currency or interest rates.

Risk Management Tool

NDFs are a risk management tool used to hedge the risk of forward currency inconvertibility, which can result from a number of factors, including credit risk, sovereign risk, regulatory restrictions, or lack of settlement procedures. NDFs are typically utilized by banks, multinational corporates, investment managers and proprietary traders to hedge currency risk. In addition, NDFs can be used for currency arbitrage, to trade currencies where formal transaction documentation does not exist (as an off-balance sheet product, documentation is not required) or as a tool to facilitate locking in the enhance yields of emerging market currencies. Volatile currencies can bring greater yields when compared to current short-term interest rates in America and Europe.

Availability

NDFs are available in several "exotic" currencies, and for most NDF products, prices are quoted for up to one year. It is not unusual to have the spot price being fully convertible, but forwards past spot being quoted only on a NDF basis. Today, most South American countries and some Far East countries operate NDFs.

For example, to hedge against currency depreciation, if the fixing rate is greater than the outright price at maturity, the purchaser of the NDF would receive from the seller the difference between the fixing rate and the outright rate in cash terms. This amount can be calculated by using the following formula:

$$\frac{(F - O) \times N}{F}$$

where

F = fixing rate
O = outright price
N = notional amount

Obviously, if the fixing rate is less than the outright price at maturity, the opposite will apply.

To hedge against currency appreciation, if the fixing rate is greater than the outright price at maturity, the seller of the NDF pays the buyer the difference between the fixing rate and the outright rate in cash terms, calculated as above. As with a purchase, if the fixing rate is less than the outright, the opposite will apply. An example of the above would be:

Notional amount:	\$10,000,000
Maturity:	90 days
Spot:	2.0000 fx/\$
90 day NDF:	0.0100
Outright:	2.0100 fx/\$
Fixing rate:	2.0200 fx/\$

At maturity, the purchaser of the NDF will receive from the seller:

$$\frac{(2.0200 - 2.0100 \times 10{,}000{,}000)}{2.0200} = \$49{,}504.95$$

Another example is where, say, an investor has invested \$2,000,000 in stock on the Korean stock market for one year. The investor expects the stock market to rise, but is worried about potential Korean won (KRW) depreciation. The investor wishes to hedge the foreign exchange exposure using an NDF. A nondeliverable forward rate of KRW 1310 per dollar is agreed between the bank and the client. The principal amount is \$2,000,000. There are three possible outcomes in one year's time:

- The KRW has reached the forward rate.
- It has depreciated further.
- It has appreciated relative to the forward rate.

Examples of the three scenarios are shown below:

	Outcome A	Outcome B	Outcome C
USD/KRW	Depreciated	—	Appreciated
Fixing spot rate	1330	1310	1290
Equivalent amount	\$1,969,925	\$2,000,000	\$2,031,008
Settlement	Bank pays client \$30,075	No net payment	Client pays bank \$31,008

In all outcomes, the client has achieved the objective of hedging the KRW exposure at 1310.

In outcome A, the exchange rate loss that the client would suffer if the investor sells the investment and exchange the KRW proceeds in the spot market, is compensated by the proceeds of the NDF. In outcome C, the client's exchange gain on realization of the investment is countered by the payment the investor makes on the NDF.

For a corporate, an example would be where the corporate is due to receive Philippine pesos (PHP) 102,000,000 in three month's time. They are concerned about potential depreciation and wish to hedge this exposure using an NDF. Assume the agreed NDF rate is PHP 51 per dollar. The principal amount of PHP 102 million is equivalent to \$2 million. Again, there are three possible outcomes in three months' time and the consequences provided in the following table:

	Outcome A	Outcome B	Outcome C
PHP/USD	Depreciated	—	Appreciated
Fixing spot rate	51.5	51.0	50.5
Equivalent amount	\$1,980,583	\$2,000,000	\$2,019,802
Settlement	Bank pays client \$19,417	No net payment	Client pays bank \$19,802

Typical Risks Encountered

- Contingent risk—exists when the dollar value of anticipated but not yet committed cash flows is subject to changes in exchange rates.
- Sovereign risk—the risk that the government of a country may interfere with the repayment of a debt. For example, a borrower in a foreign country may be economically sound and capable of replaying a loan in local currency. However, his country's government may not permit him to repay a loan to a foreign bank because of a lack of foreign exchange or for political reasons. The bank making the loan in the first place must take this sovereign risk into account and reflect it in the interest rate.
- Transaction exposure—arises whenever any company unit commits to pay or receive funds in a currency other than its national currency.
- Translation exposure—the risk that financial statements of overseas subsidiaries of a company will gain or lose value because of exchange rate movements when translated into the currency of the parent company upon consolidation.

Characteristics of Emerging Markets

- Limited currency convertibility
- Central bank regulations
- Illiquid markets
- Limited hedging vehicles
- Event/sovereign risk
- Greater volatility
- Cross-border risk
- Withholding taxes

Index-Linked Deposits

An index-linked deposit is basically a restructured NDF. It is a deposit held in a major currency with its return linked to the exchange rate of an NDF and earning an enhanced coupon. The coupon reflects the implied local interest rates derived from the NDF market, which may be significantly higher than the major currency interest rates. The index-linked deposit is particular suitable for asset managers who need to hold a physical asset, but at the same time, wish to gain access and exposure to higher yielding markets.

Index-linked deposits are available in two types, namely, those linked to principal and interest and those purely linked to principal. The former offers a higher coupon but exposes both principal and interest to exchange rate fluctuations; where as the latter exposes only the principal. Both types of deposit are not principal protected. These deposits not only have many of the same advantages as NDFs, but they also often allow depositors to assume a lower credit risk or to earn more interest than depositing onshore. Moreover, they can be used as a form of collateral for NDFs.

Summary of Characteristics

An NDF is a short-term committed forward cash settlement currency derivative instrument. It is essentially an outright (forward) foreign exchange contract whereby on the contracted settlement date, profit or loss is adjusted between the two counterparties basing on the difference between the contracted NDF rate and the prevailing spot foreign exchange rates on an agreed notional amount.

The NDF rate is the rate agreed between the two counterparties on the transaction date. This is essentially the outright (or forward) rate of the currencies dealt. The notional amount is the "face value" of the NDF, which is agreed between the two counterparties. It should again be noted that there is never any intention to exchange the two currencies principal sums—the only movement is the difference between the NDF rate and the prevailing spot market rate and this amount is settled on the settlement date.

Every NDF has a fixing date and a settlement (delivery) date. The fixing date is the day and time whereby the comparison between the NDF rate and the prevailing spot rate is made. The settlement date is the day whereby the difference is paid or received.

As it is a "cash settlement" instrument, there is no movement of the principal amounts of the two currencies contracted. The only movement is the settlement amount representing the difference between the contracted NDF rates and prevailing spot rate. Hence, NDFs are "non-cash" products, which are off the balance sheet and as the principal sums do not move, possess very much lower counterparty risks.

NDFs are committed short-term instruments. Both the counterparties are committed and are obliged to honor the deal. Of course, the user can cancel an existing contract by entering into another offsetting deal at the prevailing market rate.

The more active banks will quote NDFs from between one month to one year, although some will quote up to two years upon request. Odd-dated NDFs can also be requested. It should also be noted, that NDFs are quoted with the dollar as the reference currency, that is they are quoted in terms of dollars against other third currencies and the settlement is also in dollars.

Without an NDF, an investor who wanted to take advantage of the type of enhanced yields available in the emerging markets would have to do the following:

- Buy the spot currency and sell dollars.
- Invest in a local risk-free asset (that is, government bond).
- Fund the dollars at the London Interbank Offered Rate (LIBOR).
- At maturity, the investor receives the capital plus interest.
- Sell the currency on the spot market and purchase dollars.

FOREIGN EXCHANGE SWAPS

A *foreign exchange swap* is the simultaneous purchase and sale of one currency against another for two different value dates. Usually, one of the value dates is the spot date and the other is a date in the future. In a typical swap transaction, one currency amount is held constant for both dates of the transaction. Most foreign exchange swaps have a maturity less than one year. In addition, a forward/forward is a swap where both the near date and the end date are forward dates.

Combinations

In fact, a swap may be most easily understood as simply the combination of a spot and a forward, or the combination of two forwards. It can be the combination of a purchase with a simultaneous forward sale or a sale with a simultaneous forward purchase. Like forward contracts, swaps are regularly for periods of 1, 2, 3, 6, and 12 months from the spot value date. Frequently, however, the date is customized to meet a client's needs.

Forward contract prices are determined, as before, by the current spot price between the two currencies and the interest rates prevailing in each of the two countries. For example, a company could sell dollars and buy Swiss francs spot, and buy dollars and sell Swiss francs 3 months forward. The cash flows in such an exercise are similar to borrowing one currency (Swiss francs) and investing in another (dollars). The exposure to the company is one of interest rate risk rather than currency risk. Consequently, banks will only charge, or pay, the interest differential.

Uses

Swaps are used primarily by investors and borrowers, and for cash management purposes. They are valuable to those who have liquidity in one currency but need liquidity in another currency. Typically, a client will buy spot and sell forward to generate liquidity in the currency purchased at spot. That is, if a client exchanges dollars for francs at spot and simultaneously exchanges francs forward for dollars, the client has created liquidity in francs (that is, has them to spend) until the forward date. A foreign exchange swap is an alternative to straight borrowing in a foreign currency.

A swap allows the two parties involved to use a currency for a period in exchange for another currency not needed at that time. For example, companies can access foreign currency to finance foreign currency denominated assets, such as those of a foreign subsidiary. Hence, foreign exchange swaps can help clients to diversify their investments, to fund intracompany loans, to fund a position rather than use the money markets, to potentially improve the yield with no exchange risk in conjunction with a foreign currency investment, and to minimize borrowing costs in certain cases by using a swap rather than straight borrowing in a foreign currency. In such a contract, the exposure is therefore one of interest rate risk rather than currency risk. Consequently, market makers will only charge, or pay, the interest differential. In the swap market, this interest differential is expressed, again, in points or pips.

Formula

The formula for determining the interest rate differential underlying the swap pips is:

$$\left(C - \left(\frac{36{,}000 + C}{T}\right)\frac{S}{B}\right) = 360/360\,\text{arbitrage}$$

where:

C = currency interest rate
T = period in number of days
S = swap pips as a decimal added or deducted
B = outright forward rate

Consider the following quotation:

Spot USD/CHF	1.4791	1.4796
3-month forward points	25.5	24.5
3-month dollar deposit	1.72%	1.82%
3-month Swiss franc deposit	1.09%	1.17%

In order to choose the number of points to be applied in the swap, analyse the cash at maturity. Assume, the company is buying dollars and selling Swiss francs, which are the right hand side of a foreign exchange quote, so the number of points is 24.5. The number of points represents the interest differential when borrowing Swiss francs and investing dollars. Similarly, 25.5 points represents the interest differential when borrowing dollars and investing Swiss francs.

There are two points to note about swaps. First, as the swap pips determine the price of the swap, the spot rate used is less important. In practice, market makers tend to use the middle rate when actually processing the swap transaction. The key point to note is that whichever spot rate is chosen, the forward rate is determined by adjusting that spot rate by the swap pips. Second, the amount of one currency in a swap is kept constant. Typically, this is the dollar, thus the same amount of dollars is sold and bought in the transaction.

Uses of Swaps

As mentioned before, swaps are undertaken together with a money market operation to take advantage of imperfect exchange rate and interest rate differentials. This is particularly of use to companies, which have a borrowing advantage in one currency or type of facility over another (that is, acceptance facility).

Swaps are also used where the domestic money market may not offer the necessary investment possibilities. For example, the smaller Swiss companies and wealthy private clients place short-term Swiss franc deposits with the Swiss banks domestically. Since there is a shortage of domestic money market instruments in which to invest these deposits, the Swiss banks may place them abroad, mainly in dollars, through swaps.

Finally, a swap can be used to hedge exposure. For example, a client wishes to buy Japanese yen against dollars three months forward. The bank can cover the obligation to provide the JPY by purchasing spot and undertaking a three-month dollar against JPY swap, giving up the use of JPY, but getting the use of dollars for the period. At maturity, the bank uses the Japanese yen received under the swap to meet the obligation to the client, and the dollars received from the client to meet the dollar obligation under the swap. Alternatively, a client can use a swap to roll a hedge forward. For example, the client may have entered into a contract to buy Swiss francs against dollars forward. If, in three-months time, the dollars do not materialize, the hedge would be extended. This can be achieved using a swap, whereby the original forward is closed out by the spot transaction and the exposure is covered by the forward transaction.

Risks

Swap risks are almost identical to those for forwards. A swap effectively becomes a forward once the near date has settled. The difference between a forward and a swap is that to do a swap, there must be two transactions in opposite directions at different times.

CURRENCY SWAPS

A swap is an agreement between two counterparties to exchange future cash flows. There are two fundamental types of swap; the cross-currency, which involves the exchange of cash flows in one currency for those in another with an agreement to reverse that transaction at a future date and the interest rate (single-currency) swap, which changes the basis on which income streams or liabilities are received or paid on a specified principal amount.

From a foreign exchange point of view, the cross-currency swap is much more relevant, as they allow companies to borrow in the most efficient market, usually in one in which the company have not borrowed too heavily in the past. The major difference between cross-currency swaps and currency forwards is that there is only one contract in the case of swaps, whereas forwards require separate contracts for each payment of interest and principal.

Technique Involved

An interest rate swap is exclusively concerned with the exchange of cash flows relating to the interest payments on the designated notional amount. However, there is no exchange of notional at the inception of the contract. The notional amount is the same for both sides of the currency and it is delineated in the same currency, that is, principal exchange is redundant.

In the case of a *currency swap*, however, principal exchange is not redundant. The exchange of principal on the notional amounts is done at market rates, often using the same rate for the transfer at inception as is employed at maturity.

For example, consider an American-based company that has raised money by issuing a Swiss franc denominated Eurobond with fixed semiannual coupon payments of 6% on CHF 100 million. Up front the company receives CHF 100 million from the proceeds of the Eurobond issue. In essence, they are using the Swiss francs to fund their American operations. Because this issue is funding American based operations, the company is going to have to convert the CHF 100 million into dollars. This can be done by entering into a currency swap whereby the Swiss franc debt can be converted into a dollar like debt.

The American company can agree to exchange the CHF 100 million at inception into dollars, receive the Swiss franc coupon payments on the same dates as the coupon payments are due to the company's Eurobond investors, pay dollar coupon payments tied to a preset index and reexchange the dollar notional into Swiss francs at maturity.

Thus, lays the fundamental difference between a currency swap and the classic foreign exchange swap. During the life of the transaction, each currency bears an agreed rate of interest, which is usually paid or received at intervals.

No Interest Payable

Under a foreign exchange swap, no interest is payable on either currency. Rather, the price at which the currencies will be exchanged at maturity takes account of the interest differential between the two. Thus, if sterling rates for one year are at 5% and the dollar rates are at 2%, the theoretical forward exchange rate between the two currencies is 3% less than the spot rate prevailing. Under a one-year currency swap between the two, the rate for the reexchange at the end of one year will be the same as that used at the start, but interest will be payable or receivable on each currency. In the simple case of a one-year swap between two currencies at a fixed rate of interest the two techniques are little different. Consider, however, a five-year traditional foreign exchange swap between dollars and Swiss francs. The forward foreign exchange rate will represent the compounded interest rate differential between the two currencies and only two cash flows will occur, namely the spot transaction and the forward leg in five years, at a radically different exchange rate.

Under a currency swap between dollars and Swiss francs for five years, an amount of each currency would be exchanged at the start (determined by the spot rate prevailing), the party receiving the francs would pay an agreed interest rate periodically, as would the party receiving the dollars. At the end of five years, the same amount of each currency would be reexchanged. There is no need for the rate of interest applicable to each currency to be on a fixed basis, it can be a floating rate tied to LIBOR, for example. Indeed, the vast majority of currency swaps currently transacted are between dollars at six months LIBOR and another currency at a fixed rate of interest, payable either biannually or annually as agreed between the two parties.

Flexibility

Currency swaps give companies extra flexibility to exploit their comparative advantage in their respective borrowing markets. Also, currency swaps allow companies to exploit advantages across a matrix of currencies and maturities.

Liquid and Cost Effective

The currency swap market has become a liquid and cost-effective market for corporate treasurers to achieve long-term currency hedges for their liabilities. Today, one of the most common transactions in the currency swap market is that related to a capital market debt issue, which is then swapped in its entirety to another currency that the borrower requires. Also, an interesting application of the currency swap has been to generate foreign exchange prices by combining two or more zero-coupon swaps against floating-rate dollars, which cancel out the floating-rate flows and leave one with an exchange of a given amount of one currency against another at a future date, which is precisely similar to a long-term foreign exchange transaction. Under a zero-coupon swap, the fixed rate interest payable/receivable is not paid until maturity and is compounded at the same time as it is paid.

Exposure

Because of the exchange and reexchange of notional principal amounts, the currency swap generates a larger credit exposure than an interest rate swap.

Graphic Example

Graphically, a currency swap can be shown by the three stages in Figure 65.2.

FOREIGN EXCHANGE FUTURES

Although futures contracts on commodities have been traded on organized exchanges since the 1860's, financial futures are relatively new, dating from the introduction of foreign currency futures in 1972. The basic form of the futures contract is identical to that of the forward contract, whereby a futures contract obligates its owner to purchase a specified asset at a specified exercise price on the contract maturity date. Likewise, currency futures are defined as a standardized contract/agreement to sell or buy a specific amount of a currency at a particular price on a stipulated future date.

In fact, futures developed in response to the substantial volatility for currency trading that occurred following the 1971 shift from fixed to flexible currency exchange rates.

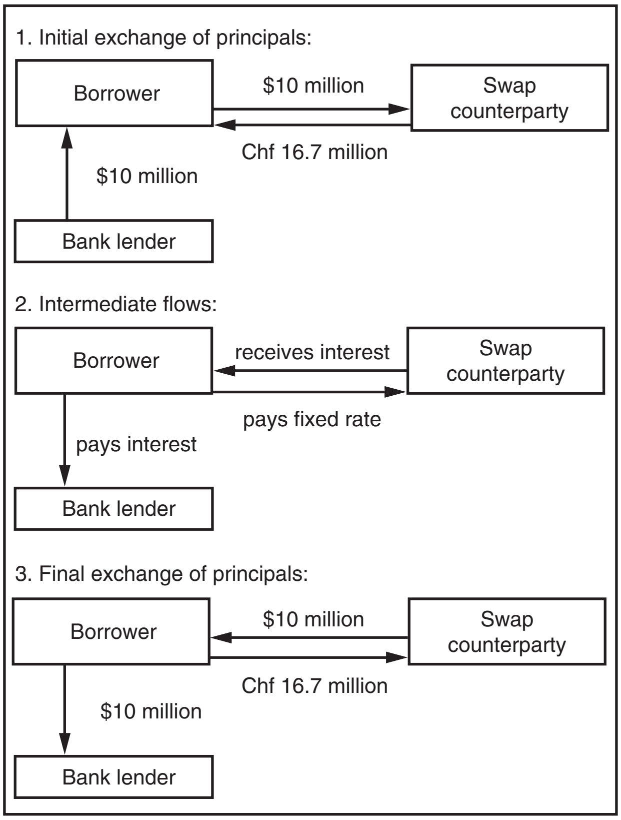


Figure 65.2 Graphical Presentation of the Three Stages of a Currency Swap

In brief, buyers and sellers of *foreign exchange futures* are required to post initial margin or security deposits for each contract. Participants also have to pay brokerage commissions that can be fixed or negotiated depending on the size of the trade. Foreign exchange futures are only traded on regulated exchanges. In general, futures are used by banks, commodity trading advisors and arbitrage houses, that is, by "professional" traders rather than by corporations.

Two-Sided Risk

Like a forward contract, the futures contract has two-sided risk. However, in marked contrast to forwards, credit or default risk can be virtually eliminated in a futures market. Firstly, instead of conveying the value of a contract through a single payment at maturity, any change in the value of a futures contract is conveyed at the end of the day in which it is realized. For example, suppose that, on the day after origination, the financial price rises and, consequently, the financial instrument has a positive value. In the case of a forward foreign exchange contract, this value change would not be received until maturity. With a futures contract, this change in value is received at the end of the day. In the language of the futures markets, the futures contract is cash-settled or marked-to-market daily.

Since the value of the futures contract is paid or received at the end of each day, a futures contract can be likened to a series of forward contracts. That is, a futures contract is like a sequence of forwards in which the forward contract written on day 0 is settled on day 1 and is replaced, in effect, with a new "forward" contract reflecting the new day 1 expectations. This new contract is itself settled on day 2 and replaced, and so on until the day the contract ends. In other words, a futures contract can be thought of as "rolling over" a forward contract on a daily basis. Strictly speaking, the futures price and the forward price are not quite the same but as a practical matter they are so close that little accuracy is lost in viewing them as identical. Therefore, analogous to forward contracts, the futures price is that contract price which results in the futures contract having zero value to both the buyer and the seller each day the contract is settled and refix.

Margin

All market participants, sellers and buyers alike, post a performance bond (that is, margin). If a futures contract increases in value during the trading day, this gain is added to the margin account at the end of the day. Conversely, if the contract loses value, this loss is deducted from the margin account. If the margin account balance falls below some agreed-upon minimum, the holder will be required to post additional bond. Hence, the margin account must be replenished or the holder's position will be closed out.

Exchange Members

There are two types of exchange members who can trade any futures contract. First, there are commission brokers, or floor brokers, who execute orders for nonmembers. These orders from nonmembers will originate through futures commission merchants, which are organizations, for example, brokers and commercial banks. These types of organizations will solicit orders for futures trading. Futures commission merchants also hold their clients margin monies and handle all margin accounting. The floor broker executing the order may or may not be affiliated with the futures commission merchant, which originated the order.

The other type of exchange member who will be trading the futures contract is called a "local" who is simply an individual trading for his or her own account. Essentially, locals are willing to hold positions, inter- or intraday, acting much like a market maker who hopes to profit from the bid/offer spread or market moves.

Clearing Corporation

An important feature of an organized futures exchange is the Clearing Corporation. Essentially, the Clearing Corporation interposes itself as the seller to every buyer and the buyer to every seller. In other words, the Clearing Corporation becomes the counterparty to every trade, guaranteeing the opposite side of every transaction. This has several attractive features:

- The buyer of a futures contract need not be concerned with the creditworthiness of the seller. If the buyer's position is doing well, that is, the futures price is rising, then the buyer is guaranteed the daily receipt of the variation margin by the Clearing Corporation, independent of whether the original seller was able to pay that same variation margin. In this example, the Clearing Corporation looks to the member firm, who originated the futures sale, for the timely payment of the daily variation margin independent of whether the original seller has paid in sufficient margin into the account.

- The other major advantage of the Clearing Corporation from the viewpoint of the member firms is that the margin accounting problem is significantly simplified. There is now only one entity with which the member firm must deal in settling margin calls, as opposed to having to exchange monies with all other member firms.
- In addition, the Clearing Corporation will net out all margin calls and receipts for a single member firm across all of their positions, such that only one net amount of funds must be transferred at the end of each trading day.

Major Exchanges

The major exchanges for financial futures include the Chicago Board of Trade (CBT), the Chicago Mercantile Exchange (CME), the International Monetary Market (IMM), the London International Financial Futures Exchange (LIFFE), the New York Futures Exchange (NYFE), and the Kansas City Board of Trade (KC).

Quoting Currency Futures

Generally, in the foreign exchange market, currencies are quoted against the American dollar. For example, a rate of 1.67 Swiss francs per dollar means that it takes 1.67 francs to buy/sell 1 dollar. Of course, there are the exceptions to this rule, for example sterling. However, currency futures are priced in American terms, in that it quotes how many dollars it takes to buy one unit of foreign currency. They are the reciprocal of those used in the cash market. Thus, a rate of 1.67 francs per dollar would be quoted in the futures market as 0.5988 dollars per franc (1 divided by 1.67), which means it costs 60 cents to buy one franc. For each contract, there is a specific contract size, for example, one Swiss franc contract is worth 125,000 francs, the Japanese yen is worth 12,500,000 yen, and sterling is worth 62,500 pounds, while the euro is worth 125,000 euros.

Ticks and Delivery Months

The minimum price movement of a currency futures contract is called a tick. The value of a tick is determined by multiplying the minimum tick size by the size of the contract. For example, using the Swiss franc against the dollar, one point is $.0001 per Swiss franc, which equals $12.50 per contract, while one-point sterling is worth $.0001 per pound, which equals $6.25 per contract. The contract trading months are on the same quarterly cycle as other financial instruments: March, June, September, and December. They are also known as the delivery months, because the seller of a contract must be prepared to deliver the specified amount of foreign currency to the buyer if the seller has not cancelled the obligation with an offsetting purchase. It must be said that the vast majority of market participants close out their positions before delivery.

Table 65.1 Summary of Contract Specifications for Currencies against the Dollar Futures

Product	Trading Unit	Point Description
Australian dollar	100,000 dollars—physically delivered	1 point = $0.0001 per dollar = $10.00 per contract
Brazilian real	100,000 real—cash settled	1/2 point = $0.0005 per real = $5.00 per contract
British pounds	62,500 pounds—physically delivered	1 point = $0.0001 per pound = $6.25 per contract
Canadian dollars	100,000 dollars—physically delivered	1 point = $0.0001 per dollar = $10.00 per contract
Euro	125,000 euro—physically delivered	1 point = $0.0001 per euro = $12.50 per contract
Japanese yen	12,500,000 yen—physically delivered	1 point = $0.000001 per yen = $12.50 per contract
Mexican peso	500,000 peso—physically delivered	1 point = $0.00001 per peso = $5.00 per contract
New Zealand dollars	100,000 dollars—physically delivered	1 point = $0.0001 per dollar = $10.00 per contract
"New" Russian rouble	2,500,000 rouble—cash settled	1 point = $0.00001 per rouble = $25.00 per contract
South African rand	500,000 rand—physically settled	1 point = $0.00001 per rand = $5.00 per contract
Swiss franc	125,000 franc—physically delivered	1 point = $0.0001 per franc = $12.50 per contract
Swedish krona	2,000,000 krona—physically delivered	1 point = 0.00001 dollar/krona = $20.00 per contract
Norwegian krone	2,000,000 krone—physically delivered	1 point = 0.00001 dollar/krone = $20.00 per contract

Contract Specifications

The contract specifications for currencies against the dollar futures are shown in Table 65.1. There are, of course, contracts for crosses as well.

EXCHANGE FOR PHYSICAL

An *exchange for physical* (EFP) refers to exchanging a physical (cash) position for a futures position. This is where a spot interbank transaction can be converted into a futures position via an exchange for physical. Consequently, when a cash position is exchanged for a future position, the EFP is simply a mechanism by which the cash position is converted to its IMM or Finex equivalent. The EFP represents the current spot (cash) price plus or minus the interest rate differential (cost of carry) between the two currencies, expressed in the futures price and it is essentially an ex-pit transaction.

Example 1

The mechanics of an EFP transaction would be where on, say December 13, the interbank spot price for francs for value December 15 is 1.6700 bid–offer 1.6705. The market user sells 25 million francs at 1.6705 for value December 15 and then decides to convert the short cash position of 25 million francs to the IMM equivalent for short contracts for March. Assume the forward pips for three months $/CHF is 1.7/1.9. The resulting transaction can be viewed as:

Buys 25 million CHF at 1.6705 value December 15
Sells 200 March IMM contracts at 0.5986
(1.6705 plus 1.9 = 1.67069 and then 1 divided by 1.67069 giving 0.5986)

The result for the cash position is:

Short 25 million CHF at 1.6705 value December 15
Long 25 million CHF at 1.6705 value December 15

so that the net cash position is flat and the market user is left with an open IMM futures position of being short 200 March IMM futures contracts at 0.5986.

Example 2

On December 13, the interbank spot price for British pounds for value December 15 was 1.5000 bid–offer 1.5005. The trader sells 10 million pounds (equivalent of 160 IMM contracts) at 1.5000 value December 15. The trader decides to convert the short cash position of 10 million pounds to the IMM equivalent of short 160 contracts for March. The resulting transaction can be viewed as:

Dealer interbank swap price for March 16 is 70-67, thus the trader executes a simultaneous swap transaction.
Buys 10 million pounds at 1.5000 value December 15.
Sells 160 March IMM contracts at 1.4930 (1.5000 – 0.0070).

Hence, the result is:

Short 10 million pounds at 1.5000 value December 15
Long 10 million pounds at 1.5000 value December 15

This results in a net cash position which is flat (square) and the trader is short 160 March IMM futures contracts at 1.4930.

Point of the Exercise

In brief, the main points are that with the EFP execution, the client's spot cash position is flat. No profit or loss will be generated. The client will have a futures position at only one average price for the full amount traded. With the execution of the EFP, the futures price will be posted (reported) to the exchange and the client will receive a confirmation as they would an IMM or Finex transaction. Additionally, fees and commissions will be recorded in exactly the same manner as if the transaction was executed on any of the exchanges. Also, there are no commissions or fees charged on the cash side of the transaction.

Table 65.2 Differences Between Interbank Spot and Futures

Interbank Spot	IMM Futures
Single counterparty risk	Counterparty risk with exchange
Unregulated market	Regulated market
Tailored maturity dates	Limited delivery months
Tailored currency amounts	Specific contract specifications
Greater liquidity	Lower average volume
Single average price	Multiple price fills
No exchange fees	
No reporting levels	
Unrealized gains can only be withdrawn upon maturity date	

Interbank versus Futures

A summary of the differences between interbank spot and futures is provided in Table 65.2.

SUMMARY

There are now financial instruments that permit the direct transfer of financial price risk to a third party, who is more willing to accept that risk. At another level, the financial markets have evolved to the point whereby financial instruments can be combined with other instruments to unbundled financial price risk from the other risks inherent in the process, for example, raising capital.

Forwards provide certainty in the uncertain world of currency movements by locking in a specific rate, and as the forward markets are quite liquid, the bid/offer spreads are low. The interest rate differential between the dollar and another currency is expressed in points, which are fractions of that currency's exchange value against one dollar. In terms of short- and long-dated contracts, the principal is the same as with regular forward rates and is again made on the basis of interest rate gain or loss. The exception to the rule, in the case of short-dated contracts, is that prices normally added on are deducted and prices normally deducted are added. This, actually, is not as odd as it sounds. If prices are normally quoted for spot delivery and a value tomorrow quote means the market maker will have to surrender dollars earlier than normal, there has to be some compensation for it.

An NDF is a short-term committed forward "cash settlement" currency derivative instrument. It is essentially an outright (forward) foreign exchange contract whereby on the contracted settlement date, profit or loss is adjusted between the two counterparties basing on the difference between the contracted NDF rate and the prevailing spot foreign exchange rates on an agreed notional amount. The NDF rate is the rate agreed between the two counterparties on the transaction date. This is essentially the outright (or forward) rate of the currencies dealt. The notional amount is the face value of the NDF, which is agreed between the two counterparties. It should again be noted that there is never any intention to exchange the two currencies principal sums—the only movement is the difference between the NDF rate and the prevailing spot market rate and this amount is settled on the settlement date.

Every NDF has a fixing date and a settlement (delivery) date. The fixing date is the day and time whereby the comparison between the NDF rate and the prevailing spot rate is made. The settlement date is the day whereby the difference is paid or received. As it is a cash settlement instrument, there is no movement of the principal amounts of the two currencies contracted. The only movement is the settlement amount representing the difference between the contracted NDF rates and prevailing spot rate. Hence, NDFs are noncash products, which are off the balance sheet and as the principal sums do not move, possess very much lower counterparty risks.

NDFs are committed short-term instruments. Both the counterparties are committed and are obliged to honor the deal. Of course, the user can cancel an existing contract by entering into another offsetting deal at the prevailing market rate.

The more active banks will quote NDFs from between one month to one year, although some will quote up to two years upon request. Odd-dated NDFs can also be requested. NDFs are quoted with the dollar as the reference currency, that is they are quoted in terms of dollar against other third currencies and the settlement is also in dollars.

Without an NDF, an investor who wanted to take advantage of the type of enhanced yields available in the emerging markets would have to buy the spot currency and sell dollars; invest in a local risk free asset (that is, government bond); fund the dollars at LIBOR; at maturity, receive the capital plus interest; and sell the currency on the spot market and purchase dollars.

Swap risks are almost identical to those for forwards. A swap effectively becomes a forward once the near date has settled. The difference between a forward and a swap is that to do a swap there must be two transactions in opposite directions at different times. The currency swap market has become a liquid and cost-effective market for corporate treasurers to achieve long-term currency hedged for their liabilities. Today, one of the most common transactions in the currency swap market is that related to a capital market debt issue, which is then swapped in its entirety to another currency that the borrower requires. Also, an interesting application of the currency swap has been to generate foreign exchange process by combining two or more zero-coupon swaps against floating-rate dollars, which cancel out the floating-rate flows and leave one with an exchange of a given amount of one currency against another at a future date, which is precisely similar to a long-term foreign exchange transaction. Under a zero-coupon swap, the fixed rate interest payable/receivable is not paid until maturity and is compounded at the same time as it is paid. Because of the exchange and reexchange of notional principal amounts, the currency swap generates a larger credit exposure than an interest rate swap.

A foreign exchange futures contract is a forward contract for standardized currency amounts and for standard value dates. Buyers and sellers of foreign exchange futures are required to post initial margin or security deposits for each contract. Participants also have to pay brokerage commissions that can be fixed or negotiated depending on the size of the trade. Foreign exchange futures are only traded on regulated exchanges. In general, futures are used by banks, commodity trading advisors and arbitrage houses, that is, by professional traders rather than by corporations.

The main point to remember with the EFP execution is that the client's spot cash position is flat. No profit or loss will be generated. The client, however, will have a futures position at only one average price for the full amount traded. The EFP transaction is posted on the required exchange. Also, there are no commissions or fees charged on the cash side of the transaction.

REFERENCES

Henderson, H. (2002). *Currency Strategy: The Practitioner's Guide to Currency Investing, Hedging and Forecasting*. West Essex, UK: John Wiley & Sons.

Horner, R. (2006). *30 Days of Forex Trading—Trades, Tactics and Technique*. Hoboken, NJ: John Wiley & Sons.

Martinez, J. (2007). *The 10 Essentials of Forex Trading*. New York: McGraw Hill.

Oberlechner, T. (2004). *The Psychology of the Foreign Exchange Market*. West Essex, England: John Wiley & Sons.

Rosenstreich, P. (2005). *Forex Revolution: An Insider's Guide to the Real World of Foreign Exchange Trading*. Upper Saddle River, NJ: Financial Times Prentice Hall.

Shamah, S. (2003). *A Foreign Exchange Primer*. London: Wiley Finance.

Steiner, R. (2002). *The Foreign Exchange and Money Markets: Theory, Practice and Risk Management*. Oxford: Butterworth-Heinemann.

Taylor, F. (2004). *Mastering Foreign Exchange and Currency Options: A Practical Guide to the New Marketplace-Second Edition*. Upper Saddle River, NJ: Financial Times Prentice Hall.

Walmsley, J. (2000). *The Foreign Exchange and Money Markets Guide*. Hoboken, NJ: John Wiley & Sons.

Weithers, T. (2006). *Foreign Exchange: A Practical Guide to the FX Markets*. Hoboken, NJ: John Wiley & Sons.

PART 8

Inflation-Hedging Products

CHAPTER 67

Inflation-Linked Bonds

P. BRETT HAMMOND, PhD
Managing Director and Chief Investment Strategist, TIAA-CREF Asset Management

Abstract: In many regimes, inflation is arguably the largest systematic bond risk factor. As such, inflation-linked bonds are as close as the market has gotten to the riskless asset, the ultimate real return investment. The issuer market is dominated by sovereign entities willing to take on inflation risk to reduce interest payments, manage the economy, and match certain payments to cash flows. They are used by long and short-term investors interested in inflation hedging, managing risk, consumption smoothing, asset allocation, and as a basis for inflation-linked derivatives. The basic structure of inflation-linked bonds is unique in that it pays a return equal to actual accrued inflation plus a real interest rate. Consequently, an inflation bond responds uniquely to market and other forces such as economic growth, expected inflation, interest rates and taxes. Investors should therefore consider the effect of such forces on the behavior of an inflation bond's return, current yield, volatility, duration, beta, term structure, and other factors. These factors suggest uses for inflation-linked bonds in liability matching, diversified portfolios, and inflation "trading."

Keywords: inflation, inflation expectations, breakeven inflation, inflation-linked bonds, nominal bonds, inflation risk premium, real interest rate, real return, triple duration, volatility, riskless asset, inflation indexation, consumer price index (CPI-U), Treasury inflation-protected securities (TIPS), linkers, consumption

U.S. inflation-linked securities were first issued in 1997. Originally called Treasury inflation-indexed securities (TIIS), they are now commonly referred to as *Treasury inflationprotected securities (TIPS)*. Other nicknames for *inflation-linked bonds* include *linkers* (United Kingdom), *real return* bonds (Canada), inflation-protected bonds, inflation-indexed bonds, and *inflation* bonds.

When they were first introduced in early 1997, it was not clear that TIPS would become a permanent fixture. Despite the existence of inflation-linked bonds in the United Kingdom and other countries, the new U.S. TIPS were thought by many observers to be experimental. Most primary U.S. Treasury dealers and many others expressed indifference or even antipathy toward TIPS. And it wasn't until 2004 that the U.S. Treasury clearly indicated that inflation-linked bonds were on its agenda for continued issuance. Perhaps it was not apparent to many investors exactly what inflation-linked bonds were, how they would trade, and for what purposes they would be used.

The modern inflation-linked bond market didn't start with TIPS. These bonds began to appear in the latter half the twentieth century, when Finland, followed later by Israel, Iceland, and other countries, issued inflation-linked bond–like securities in response to post–World War II inflation. (It is puzzling that inflation securities were not introduced or discussed more in Germany and other countries experiencing post–World War I hyperinflation or that Lord John Maynard Keynes did not discuss the idea in light of his work on interwar economic developments.) Then, beginning in the late 1950s, economists such as Milton Friedman, Paul Samuelson, and many others endorsed the idea of creating inflation-linked government securities. In 1981, the United Kingdom first issued "linkers," their version of inflation-linked sovereign bonds, followed by Australia, New Zealand, Canada, and other developed nations.

Since 1997, U.S. and global inflation-linked bond issuance, ownership, and understanding have deepened and broadened. As of 2007, 22 countries issued sovereign inflation-linked bonds. (Germany remained the only member of the G7 not to do so.) TIPS represent roughly 8% of the total tradable U.S. government securities market and in the United Kingdom, "linkers" are roughly 20% of the total sovereign bond market. Figure 67.1 shows that the size of the sovereign global inflation-linked bond market at the beginning of 2007 was over \$1 trillion, with the U.S. representing about \$400 billion of that total (Barclays Capital, 2007). In the United Kingdom, over 50 agencies and private firms have also issued inflation-linked bonds, while in the Eurozone there are about 10 such issues, with a combined total value of approximately \$50 billion. Between 2002 and 2007, the global inflation-linked bond market grew by an average of 30% per year. Also, as inflation-linked strips and derivatives markets emerged, conceptual and practical experience with the behavior of inflation-linked bonds increased as well.

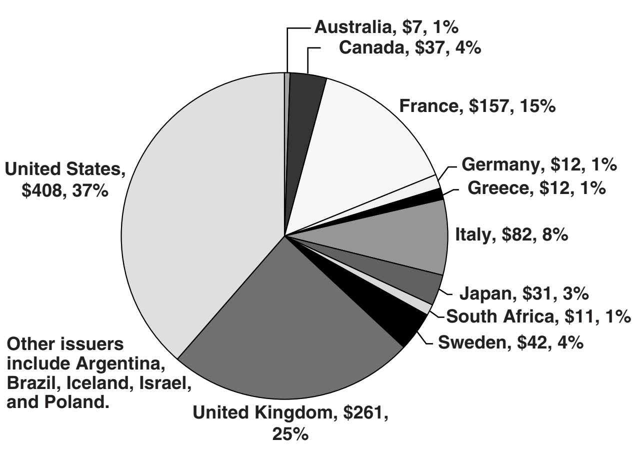


Figure 67.1 Global Sovereign Inflation-Indexed Bond Issuance Total = US\$1.06 Trillion. *Source:* Barclays Global.

This chapter develops a basic understanding of inflation-linked bonds, including their structure, pricing, and common uses. It focuses on some of the key distinctions that set inflation-linked bonds apart from other asset classes, particularly *nominal bonds.* Its basic argument is that, in concept, inflation-linked bonds are not only a separate asset class, they are as close as the market has ever come to the risk-free asset, with the practical limitation that they are not as readily available as nominal "risk-free" alternatives. We will also discuss some other not-so-obvious practical characteristics of inflation bonds as the risk-free asset, such as the *inflation risk premium* and the *triple duration.*

INFLATION, INVESTING, AND CONSUMPTION

Inflation is arguably the most fundamental systematic investment risk. Inflation poses a significant challenge in the effort to smooth *consumption* over time or preserve the value of an asset as it moves into the future.

In economic theory, consumers do or should smooth real consumption over time. As such, inflation introduces a discount factor as well as more uncertainty about the relationship between income and consumption. In other words, inflation is a major component of the discount rate and net present value. Interestingly, this issue has been around at least since the Roman Empire, where at one point soldiers' pay was regularly adjusted to the cost of food staples. And in the early nineteenth century some states paid Revolutionary War pensions according to the changing price of corn and hogs (see Shiller, 2003).

Figure 67.2 shows the long-term trends and short-term spikes in U.S. inflation since wartime price controls were removed at the end of 1945. At one level, inflation directly determines the half-life of asset value and purchasing power. For example, at the long-term U.S. inflation rate of 3% (dotted line), the purchasing power of money (or real value of an asset with zero return) drops by over 50% in 28 years. Even at the 2% inflation rate target used by many central banks, real asset value drops by more than half in 39 years. As such, it is easy to see how inflation is intimately connected to the vast literature on the time value of money and fixed income investing (see Homer and Leibowitz, 2004; and Fabozzi, 2005).

In addition, inflation can also be volatile, and a short-term inflationary shock can reduce real purchasing power and asset values quickly. As such, it is easy to see that inflation's long- and short-term behavior reduces the ability of an investor to anticipate or achieve an expected return through time. In the absence of inflation, an investor interested in evaluating an investment opportunity will still require a basic, or real, return to compensate for the opportunity cost associated with other foregone investments. The presence of inflation both increases the required compensation in order to preserve the real return *and* it

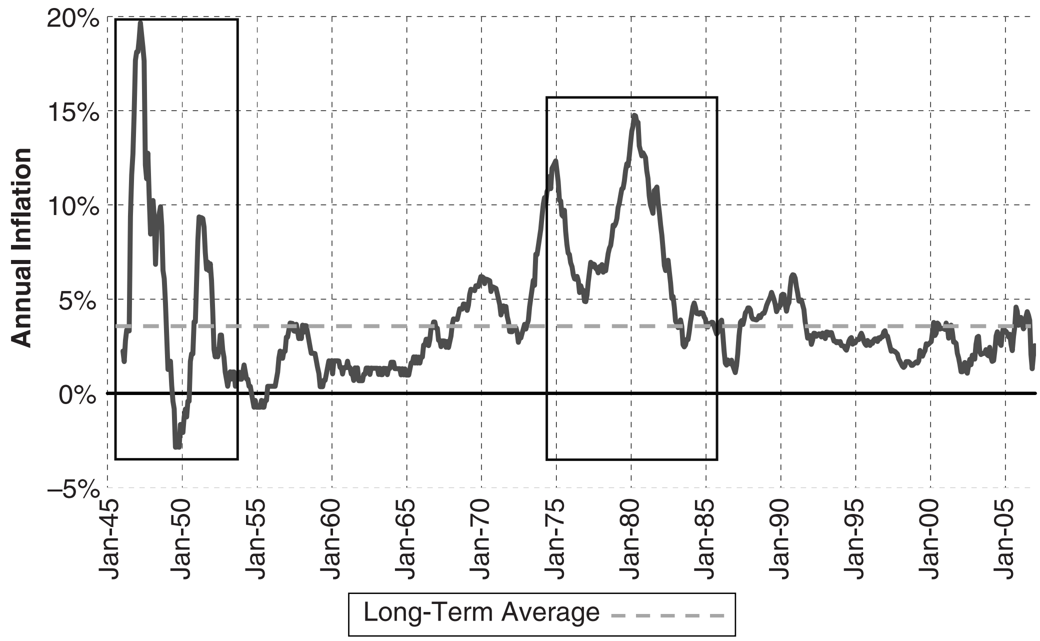


Figure 67.2 Inflation Is Fairly Benign, But Hasn't Always Been (Rolling Annual U.S. CPI-U, January 1946–December 2006). *Source:* U.S. Bureau of Labor Statistics.

introduces more uncertainty about what that total return (basic or real return plus expected inflation) will be.

INFLATION BOND BASICS: THEORY AND STRUCTURE

Following this logic, a systematic approach to investment, consumption, and inflation appeared in the 1920s with the well-known Fisher equation, which decomposed a nominal bond's return into components:

$$r_n = [1 + E(r_r)][1 + E(\pi)] \tag{67.1}$$

where current nominal return or yield (r_n) is composed of two components: the expected (or required) real return or yield (r_r) and expected inflation (π). The implication of the Fisher equation is that *ex ante,* a nominal bond's return is not associated with actual inflation, but rather what investors believe inflation will be over the life of the investment. In other words, investors must assess whether or not realized inflation will equal, exceed, or trail actual inflation. If it turns out that

$$E(\pi) < \pi \tag{67.2}$$

then the investor will be disappointed because realized higher-than-expected inflation will reduce the real, inflation-adjusted return (that is, the consumption value of the dividend stream and return of principal). If, however,

$$E(\pi) \geq \pi \tag{67.3}$$

then the bond's returns will either meet expectations or exceed them, because actual inflation will be either the same or less than the *ex ante* estimate and the consumption value of the dividend stream and return of principal will rise. Of course, in this buy-and-hold example, the bond's nominal return will be the same in both of these inflation regimes.

Instead of this static, buy-and-hold view, when we look at this bond more dynamically with daily pricing and yields, we see the familiar relationship between prices, interest rates and inflation. If interest rates (real interest rate and/or inflation rate) rise, then current bondholders are likely to experience a drop in the bond's market price. However, if rates fall, then current bondholders could experience a welcome price rise.

The static and dynamic descriptions of fixed income assets, real rates, and inflation are familiar, but incomplete. What they leave out is the uncertainty associated with inflation. While a bond's return is fixed in nominal terms over the life of the asset and thus reflects what investors think inflation will be, as we have seen an inflation surprise, up or down, will change its marked-to-market value. And while bondholders might be happy with a positive surprise (lower-than-expected inflation), they are surely averse to a negative surprise. We need to modify the classic Fisher equation to account for inflation uncertainty:

$$r_n = [1 + E(r_r)][1 + E(\pi)][1 + p] \tag{67.4}$$

where p is the inflation risk premium an investor requires in order to compensate for the extra uncertainty associated with future inflation rates. So a nominal bond's return is really composed of three terms, a *real interest rate,* expected inflation, and an inflation uncertainty (risk) premium.

So what does an inflation-linked bond do that a nominal bond doesn't? We can begin to see this by looking

Table 67.1 Effect on Existing Bond Prices of Changes in . . .

	GDP Growth Rate		Real Interest Rates		Inflation Expectations		Cash Flows	
	Rise	Fall	Rise	Fall	Rise	Fall	Rise	Fall
Inflation Bond	−	+	−	+	+	−	+	−
Nominal Bond	−	+	−	+	−	+	+	−

Source: Author's estimates.

at the return or yield components of an inflation-linked bond:

$$r_i = [1 + E(r_r)][1 + \pi] \tag{67.5}$$

One key difference between r_i in equation (67.5) and r_n in equation (67.4) is that in (67.5) the right-hand term reflects actual or realized inflation instead of expected inflation, as in (67.4). An inflation-linked bond guarantees to return an amount equivalent to actual inflation as well as a real interest rate. A second difference is that an inflation-linked bond's total return does not include the inflation risk premium component for inflation uncertainty since there is no uncertainty about whether the return will match inflation, at least in a static, buy-and-hold setting.

With daily market pricing, the effects are also different. In Table 67.1 we see that changes in the GDP, real rates, and cash flows affect nominal and inflation-linked sovereign bond prices similarly. However, unanticipated changes in inflation will have opposing effects on nominal and inflation-linked bond prices. With an inflationary surprise, nominal bond prices will fall as investors realize that their future cash flows will be worth less than they had earlier expected. In contrast, inflation-linked bond prices should rise as investors demand an asset whose cash flows preserve purchasing power and asset value in response to an inflationary surprise.

The practical structure of an inflation bond follows from this simple theory. Most (nonzero) nominal bonds are structured to pay a dividend twice a year by applying an interest rate set by the market at auction, with a return of original principal at maturity. For the Canadian-style inflation-linked bond structure, which has been adopted by almost all countries, a real coupon rate is set at auction. (At the end of 2004, the United Kingdom became the last country to switch to the Canadian structure.) Then, the principal is readjusted to reflect changes in the country's selected retail or *consumer price index (CPI-U)* lagged by three months. Twice a year the real coupon rate is applied to the adjusted principal in order to produce the current coupon payment. The principal and coupon payment can adjust upward as well as downward in response to consumer price index changes during the life of the bond, except that at maturity a principal amount that is "underwater" must be reset to the original par amount.

To calculate an inflation-linked bond's current price and yield, most countries use a reference index. In the United States the index for the first of the month is based on the value of the CPI-U index three months previous (for June 1 price, use the CPI-U index value for March). Since the CPI-U is a monthly index, for prices on other days, the index calculation is linearly interpolated. (There are slight variations in reference formulas used by some countries, but all follow this basic approach.)

Inflation-linked bond carry is another structural feature. For a nominal bond, carry is defined as the difference between the bond's current yield and the cost of funding the same bond in the repo market. Forward yield is the level at which the bond's value and funding costs are the same. For an inflation-linked bond, the carry is calculated using the bond's real yield and the inflation accrual. One implication of this structure is that while the inflation accrual within any one month is smooth and linear, the use of a monthly price index can make the current yield volatile from month to month. A second implication is that the lag in inflation accrual means that some inflation information is not contained within the yield (between two and six weeks, depending on the day) so that a forward real yield to the last day of inflation accretion may be superior estimate of the bond's actual value. The difference between spot and forward yields can sometimes be quite large, especially when incremental price index changes are extreme and with shorter maturities.

WHY INFLATION-LINKED BONDS?

Inflation-linked bonds' structure and purpose can make them attractive for borrowers as well as lenders in terms of gauging *inflation expectations* as well as managing inflation, liabilities, risk diversification, and active trading opportunities.

Gauging Inflation Expectations and Preventing Policy Errors

The yield difference between similar nominal and inflation-linked bonds, known as *break-even inflation,* can be used as a gauge of inflation expectations. Formally, break-even inflation is known as

$$(1 + \pi_{be}) = \frac{\left(1 + \frac{R_n}{2}\right)^2}{\left(1 + \frac{R_r}{2}\right)^2} \tag{67.6}$$

where π_{be} is breakeven inflation rate between a nominal and inflation-linked bond with semiannual coupon

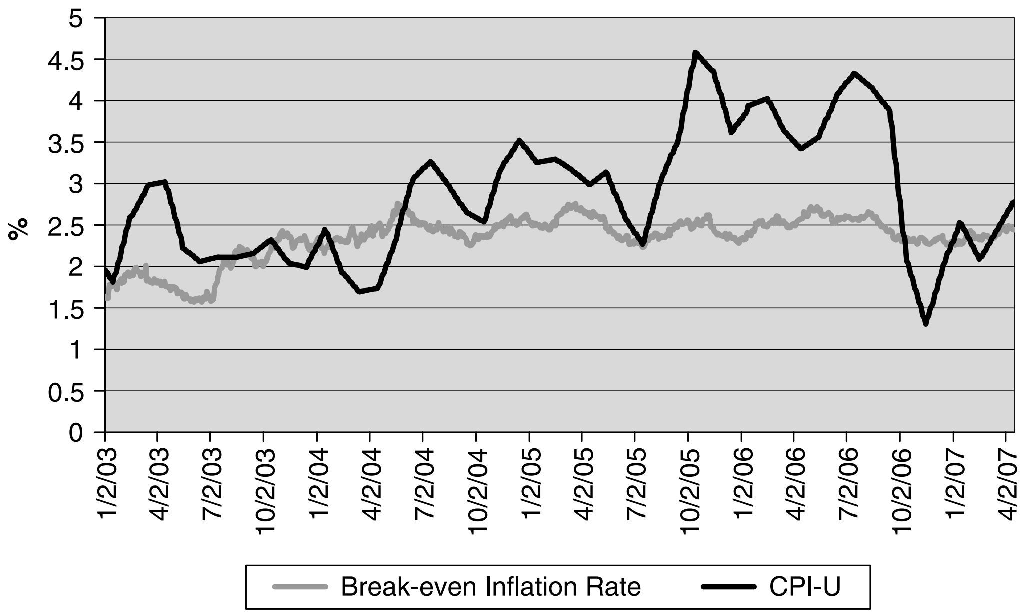


Figure 67.3 U.S. Inflation and Break-Even Inflation Rate for Constant-Maturity 10-Year U.S. TIPS.
Source: Author's calculations based on TIPS yield, T-bond yield, and CPI-U.

payments. In practice, this formal approach is usually truncated as simply

$$\pi_{bei} = R_n - R_r \tag{67.7}$$

or the current nominal yield minus the real yield. Of course, in theory, the market should anticipate inflation over the full life of the bond, since the current yield to maturity should take into account all future cash flows.

Figure 67.3 shows the break-even inflation rate and actual inflation beginning in 2003 for a constant 10-year maturity TIPS series versus a similar nominal Treasury bond series, as well as realized inflation (CPI-U). In this figure, the CPI-U is lagged by 15 days to correspond to the Bureau of Labor Statistics' inflation announcement date each month. Although the break-even rate is indicative of the market's view of longer-term inflation, it is still instructive to compare it with contemporary inflation. Central banks, such as the U.S. Federal Reserve, other government agencies, and investors use the break-even rate to gauge inflation sentiment, as a check on contemplated monetary policy changes, and to sense whether inflation-linked bonds seem expensive or cheap relative to nominal bonds.

Managing Inflation and Inflation Expectations

As the proportion of an issuer's total debt issued in inflation-linked securities rises, incentives increase for the issuer to control inflation so as to limit dividend increases associated with inflation adjustments. The effects may be both real (set low inflation targets) and perceptual (increased credibility for the issuer). Increased inflation-linked issuance during an inflationary spike can also reflect a bet by the issuer that it can reduce inflation faster than inflation expectations can fall. For example, the United Kingdom made this bet at the height of stagflation in 1981 and successfully reduced its nominal interest payments over the next few years as inflation moderated.

Liability Management

Governments and some other entities can experience revenue and expenditure streams that are highly inflation dependent. Income tax revenues, for example, are correlated with inflation, as are transportation tolls. Inflation-linked bonds enable changes in revenues to be more directly linked to debt service. On the buy side, these bonds are the ultimate inflation immunizer. As such, inflation-linked bonds are particularly useful in pension plans where liabilities are determined, in part, by future salaries and benefit cost-of-living increases. As a result of regulatory pressure to improve asset-liability matching, many U.K. pension managers, who face partial or full inflation-adjusted benefit promises, have become large long-term holders of U.K. linkers. In the United States, pressure to improve pension liability matching is more recent and benefits are less frequently indexed to cost of living. Consequently, U.S. pensions are less interested in TIPS as part of an immunization strategy.

Risk Diversification

For both borrowers and lenders, inflation-linked bonds are portfolio diversifiers. Since the basic structure of these bonds differs from that of nominal bonds, their return behavior also differs from that of other assets (see the section on the behavior of inflation-linked bonds later in this chapter). So investors and issuers alike can benefit from the diversification benefits. For issuers, the presence of inflation-linked bonds may also be a source of a different kind of diversity by attracting new types of purchasers who are less interested in nominal bonds and more interested in longer-term, inflation-sensitive holdings. On the lender side, foreign governments and banks, especially those in Asia that are heavily exposed to U.S. fixed income markets, have been active buyers of TIPS as they manage their international capital flows.

Active Management

Institutional fixed income managers have had success with active management of portfolios containing inflation-linked bonds with or without other securities. Opportunities for active management include the use of seasonality, spots and forwards, and relative value.

Seasonality

In countries where inflation accruals are based on an index that is not seasonally adjusted, break-even inflation rates often display cyclical patterns. In the United States, break-even rates tend to rise in the first half of the year and fall in the latter half of the year (see Figure 67.3). In addition, various inflation-linked bond maturities can be affected differently by seasonality. These effects represent buying and selling opportunities, both among different inflation-linked bonds along the yield curve and break-even inflation trades between inflation-linked bonds and other asset classes. At any time of the year, a breakeven rate that is below expected inflation can signal the time to sell inflation-linked bonds; when the rate is above expected inflation, it can be a good time to buy.

Spots and Forwards

The structural features (previously described) of the spot and forward inflation-linked bond values also provide trading opportunities. Forwards on inflation-linked bonds are highly sensitive to new inflation rate announcements, because the information arrives intermittently rather than smoothly. Further, the carry shouldn't be correlated with current monthly return, but when it is, a trader can go long or short depending on the movement. A simple way of thinking about this is that a trader can compare the forward yield and break-even inflation rate with his or her own estimate of expected inflation and whether it is priced into current and forward yields.

Relative Value

The growth of the global inflation-linked bond market provides opportunities for trading among various countries' sovereign bonds. Differences in maturities, structures, price indices, seasonality, liquidity, and macroeconomic conditions, can all affect the relative value of sovereign inflation-linked bonds. And the trend in declining bid-ask spreads allows these factors to be used for active management.

Arguments against the use of inflation-linked bonds often focus on the incentives for sovereign issuers and the economy. The principal criticism is that they encourage more general *inflation indexation,* such as wider use of cost-of-living wage, benefit, and price adjustments. As indexation becomes ubiquitous, it could encourage a continuing wage and price spiral. Such an outcome is more likely in countries where there is already a tendency toward very high inflation and central bank control is relatively weak. But even there (Brazil, Israel), governments have successfully used new issuance of inflation-linked bonds to signal a strong intention to control inflation.

The other major objection is that inflation-linked bonds trade less frequently than nominal bonds. One analysis showed that U.K. linker and TIPS turnover was about one-half (linkers) to two-thirds (TIPS) that of similar maturity nominal bonds (Hammond, 2003), and bid-ask spreads are often somewhat higher than for nominal bonds. So traders who are familiar with nominal bond markets find that the inflation-linked market can offer fewer traditional active opportunities. As we mentioned, linkers and TIPS have attracted many purchasers because of their liability matching and foreign exchange reserve management. Because of these properties, it is not surprising that inflation-linked bond purchasers are more likely to be long-term holders than nominal bond purchasers.

BEHAVIOR OF INFLATION-LINKED BONDS

Although some of the uses of inflation-linked bonds are now familiar and even obvious, there are several aspects of their behavior that are becoming much better understood as we have gained experience with these assets. These aspects include return attribution, *volatility,* correlation, and duration.

Return Decomposition

Figure 67.4 shows the annual return composite for a weighted index of all 12 U.S. TIPS returns. It shows that the index's total return averaged almost 6.7% from 1997 through the end of 2006 (individual years are shown from 1999 to 2006). In one year, 2001, TIPS were the best performing major asset class. Overall, long-dated TIPS provided better returns than short-dated TIPS. And TIPS in general outperformed nominal U.S. bonds over the same period (with a larger outperformance at the long end of the yield curve). This is surprising given that TIPS, at least in theory, should provide a lower return due to the greater certainty of real cash flows (the inflation risk premium).

When we look at the composition of returns in Figure 67.4, we can see the contribution of the real coupon, inflation, and price changes. Overall, the real coupon contributed about 3.25% per year to total annual return, with little variation from year to year. This figure masks a declining trend in real yields over the period from an impressive 3.4% to 3.5% for the first issues in 1997 to 1.5% to

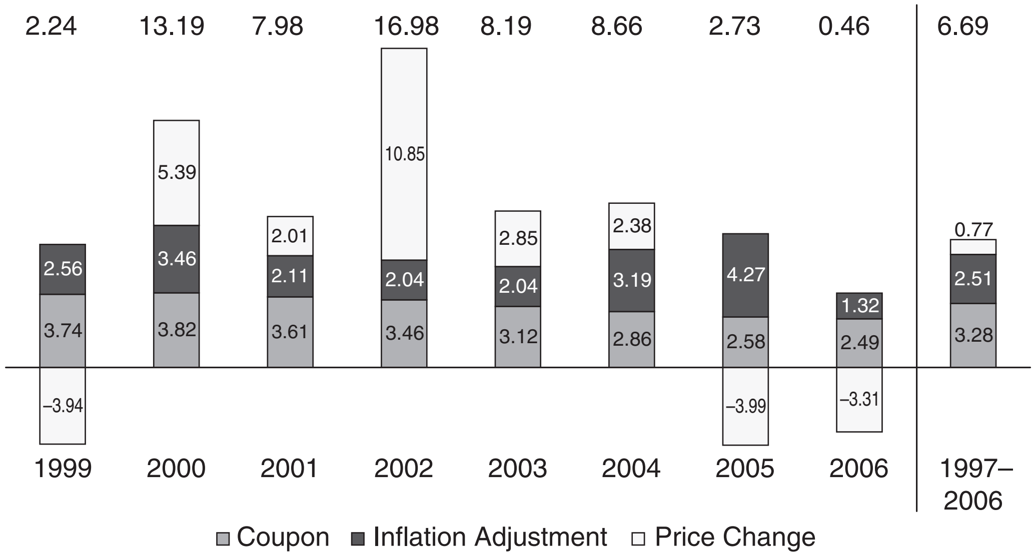


Figure 67.4 Composition of U.S. TIPS Return (Percentages Based on Monthly Returns from 3/97 through 12/06). *Source:* Barclays Capital.

2.5% for new issues at the end of the period. The overall inflation return for the full period was about 2.5% per year, reflecting the relatively benign inflationary environment.

More interesting, price returns, which as we have seen should be positively correlated with inflation surprises, in fact seem uncorrelated. As we might expect, overall price return was small, averaging less than 1% per year. But the overall figure masks considerable annual swings. When inflation rose from 1999 to 2000 and then fell slightly in 2001, price returns followed a similar but more exaggerated path. However, in 2002, when inflation remained fairly flat, price return rose dramatically to nearly 11%. And, in 2005, when inflation reached its zenith during this period, price return was decidedly negative, as it was again in 2006 when inflation hit its nadir. The explanation is that these returns weren't primarily inflation driven, since inflation remained fairly low and changes in inflation expectations were muted. Rather, a large component of price changes was changing expectations in the equity markets and the relatively limited supply of inflation-linked bonds relative to other fixed income securities. As equity markets plunged in the early years of the decade and inflation expectations rose slightly, there was a rapid flight to quality emphasizing inflation-protected instruments. As equity markets entered better years, price returns fell to negative.

Saving the Inflation Risk Premium?

Theory tells us that there should be an inflation risk premium and that it should be positive (see equation 67.4). There have been several efforts to uncover the size of the premium (Hammond, Fairbanks, and Durham, 1999b; Roll, 2004; and Campbell and Shiller, 1996). These range from a calculation that uses the simplest definition of the break-even inflation rate to produce an implied risk premium to more sophisticated analyses (Roll, 2004). Most, however, fail to show any persistent positive premium; some show that in the case of many nations' inflation-linked bonds, it has been negative. A negative risk premium implies that, instead of being willing to pay more (that is, receive a lower yield or return) for inflation protection, investors may demand a higher yield (lower price) for a new asset whose properties and continued existence are uncertain.

A somewhat more conceptual approach to understanding the inflation-risk premium starts by assuming that investors don't react rationally to inflation (Hammond, 1999b). In fact, investors are likely to be much more adverse to inflation increases than decreases, they are likely to give excess prominence to the possibility of extreme inflation, and their memory of inflation "events" fades over time. Combined with the presence of serial correlation in inflation measures, these behavioral tendencies, if true, suggest that modeling the inflation premium requires modifications to the usual expected utility models. Applying these assumptions to a model that uses a series of bootstrapped 10-year inflation-linked bond returns produces a predicted average inflation-risk premium in the range of a half of a percent, but one that varies considerably depending on the sequence and proximity of inflation patterns. In the end, even as the market gains full liquidity and depth, it is possible that the inflation risk premium's volatility will prevent issuers from fully realizing its benefits. In the interim, it is possible that investors have been receiving a somewhat greater return than they might have anticipated from the theory.

Low Volatility

TIPS volatility is also lower than for nominal bonds. For each TIPS issue, volatility has been about one-half to two-thirds that of similar nominal bonds. For example, the annualized standard deviation of daily yields for the 2007 TIPS note from 1997 through 2005 was about 8%, while the nominal 10-year 2007 U.S. note's standard deviation over the same period was about 13% (author's calculations). Similarly, the 30-year 2028 and 2030 TIPS have had annualized standard deviations in the 7% range, while the 30-year nominal Treasury bond had an 11% standard deviation (Roll, 2004). These differences are understandable since actual inflation-related volatility ought to be factored out of the TIPS real yield.

Correlation

In their 10-year history, the cross-sectional correlations among all TIPS bond returns have been high (at least 0.80 and in most cases above 0.95, based on annualized quarterly data), with the exception of a near-maturity TIPS, which behaved more like a very short-term bond. Unlike most nominal bonds, TIPS returns exhibit some serial correlation, perhaps because inflation is itself autocorrelated.

In theory, we would expect to see low or even negative correlations between inflation-linked bonds and other asset classes because of their unique response to changes in inflation and inflation expectations. As shown in Table 67.2, TIPS' correlation with nominal bonds since 1997 was highly positive (nearly 0.80) and their correlations with real estate and equities was –0.25 and –0.50, respectively (similar to nominal bond correlations). Note that the 10-year time frame may not be long enough to really gauge long-term correlations as we move through additional interest rate, inflation, and equity market cycles (e.g., TIPS and nominal bonds showed contrasting correlations with other asset classes in the 1997–2002 period). Moreover, structural supply-and-demand forces are being replaced by economic forces as the supply of TIPS increases, so we may see additional changes in relationships between TIPS and other asset classes. We should also expect to see temporal volatility in TIPS correlations. For example, because they represent the ultimate safe harbor, future market flights to quality may provide a stimulus to inflation-linked bond returns relative to nominal government securities.

If we switch from nominal to real (*ex post*) return measurement, the correlation between inflation-linked bonds and other assets is positive for all major asset classes, except commodities. It is also estimated that inflation-linked bonds have the shortest positive-return "holding" periods (holding period required to achieve a positive real return through the economic cycle) of any major asset class (Barclays Capital, 2006).

Triple Duration

For nominal bonds, duration is an analytical relationship that has several useful variations (Macaulay, modified, effective). For purpose of this discussion, we will focus on the percentage change in price of a bond as a function of a unit change in nominal and real interest rates or inflation. For an inflation-linked bond, determining its duration with respect to the real interest rate is straightforward, much like calculating the nominal duration of a nominal bond. And, since both the coupon and real yield of an inflation-linked bond are lower than those for a similar-maturity nominal bond, the inflation-linked bond's real duration should be longer than the nominal bond's nominal duration.

Because inflation-linked bonds are evaluated, not just in real terms, but also in relation to nominal markets, we would also like to understand their sensitivity to changes in nominal factors, such as interest rates and inflation. And the existence of inflation-linked bonds suggests the importance of evaluating nominal bond duration with respect to real factors and inflation. We could think of this as the "triple duration" (real rate, nominal rate, and inflation duration). Leibowitz et al. (1989) first identified a "double duration" for nominal and real rates and Siegel and Warning (2004) applied that notion to inflation bonds in their analysis of pension plans. For a nominal bond, the triple duration is given as

$$D_n \approx D_r \approx D_\pi \tag{67.8}$$

Table 67.2 TIPS Historical Experience through 2006 (%)

	Starting Date	Geometric Return	Std Dev	Correlations Inflation Bonds	U.S. Stocks	Int'l Stocks	Nom Bonds	Real Estate	T Bills
U.S. inflation bonds	1997	6.69	4.40	1.00	−0.51	−0.48	0.78	−0.26	−0.08
U.S. stocks	1979	13.35	18.00		1.00	0.87	−0.53	0.18	−0.09
International stocks	1970	11.61	20.52			1.00	−0.53	0.20	−0.19
Nominal bonds	1976	8.58	7.40				1.00	−0.06	0.19
Real estate	1978	10.11	3.65					1.00	0.35
T bills	1926	3.73	1.55						1.00

Source: Author's calculations based on annualized quarterly data from Ibbotson Associates.

Notes:

Returns are since starting date; correlations are from the period 12/97–12/07.

U.S. inflation bonds = Lehman Bros. Global Real U.S. TIPS Index

Domestic stocks = Russell 3000 Index

International stocks = Morgan Stanley EAFE Index

Regular bonds = Lehman Bros. Aggregate Bond Index

Real estate = National Council of Real Estate Investment Fiduciaries (NCREIF) Property Index

Table 67.3 Inflation Bonds Have Triple Duration

		Price Sensitivity to Changes in . . .		
		Inflation	Real Interest Rates	Nominal Interest Rates
5-Year				
TIPS	TII $3^3/_8$, 1/15/07	–0.001	4.10	1.10
T bond	T $6^3/_8$, 2/15/07	3.932	3.91	3.90
10-Year				
TIPS	TII $3^3/_8$, 1/15/12	–0.001	8.02	3.40
T bond	T $4^7/_8$, 2/15/12	7.440	7.44	7.40
30-Year				
TIPS	TII $3^7/_8$, 4/15/29	–0.018	17.10	5.40
T bond	T $6^1/_8$, 8/15/30	13.306	13.45	13.30

Source: Author's calculations.

where D_n is the duration of the nominal bond with respect to the nominal interest rate, D_r is the duration of the bond with respect to the real interest rate, and D_π is duration with respect to inflation.

However, an inflation-linked bond's price will not respond similarly to changes in nominal rates, real rates, and inflation rates. In addition, it isn't possible to obtain a closed-form solution for all of the elements of the triple duration. Instead, we can get a statistical picture of these relationships by looking at the actual behavior of inflation-linked bonds with respect to the critical factors. Another approach is to use the inflation-linked bond return beta with respect to nominal rates as described in Barclays Global (2006). Some readers may find this more accessible, although the results are statistically equivalent.

Table 67.3 shows the results of regressions of similar-maturity inflation-linked and nominal bonds. Of course, duration is dynamic and declines over time, so these figures are snapshots. Moreover, although they are statistically significant, they don't explain all of the variance (that is, $R^2 \neq 1$). Not surprisingly, nominal bonds of various maturities show nearly identical empirical duration with respect to expected inflation, implied real rates, and nominal rates. And the analytically calculated durations with respect to changes in nominal rates are close to the comparable empirically calculated durations. More important, the empirical triple durations for each inflation-linked bond can assist us in better understanding their distinctive behavior.

Looking at the first column, the empirical duration of nominal bonds with respect to inflation is, as we might expect, essentially zero ($D_\pi = 0$). As we have seen, since inflation-linked bonds incorporate actual inflation, they should reflect current inflation (with a slight lag). Turning to the second column, with respect to real rates, inflation-linked bond durations are consistently longer than nominal bond durations, because of the inflation-linked bond's lower real coupons and yields (in that respect they behave a little more like a zero-coupon bond than a nominal bond). Finally, in the third column we see that, with respect to nominal rates, inflation-linked bond duration is considerably and consistently shorter. This reflects the fundamental difference between the two types of bonds in that nominal bonds reflect inflation expectations, for which changes flow through to affect their prices, while inflation-linked bonds reflect actual inflation, for which changes don't flow through.

Finally, by tracing the triple durations through time, we know that inflation-linked bond duration with respect to real rates descends steadily (low volatility) as maturity approaches. In contrast, inflation-linked duration with respect to real rates displays significant volatility around zero over time. Note that these are relative trends; exact calculations will vary with economic and market conditions.

Taxes

One challenge to the foregoing conclusions about the triple duration is that the tax treatment of inflation-linked bonds will indeed affect their sensitivity with respected to expected inflation, as described in Roll (2004). The U.S. Internal Revenue Service treats inflation accruals to principal for TIPS as current income taxable at regular rates. So changes in inflation expectations can affect expected taxes, which in turn can affect the demand for TIPS. In contrast, for U.K. linkers and other sovereign inflation-linked bonds, inflation-related adjustments are not taxed as ordinary income and so we would not expect to see the tax effect there. Tempering the tax effect on duration in the United States is that the bulk of TIPS ownership is in tax-deferred or institutional accounts, which are insensitive to the inflation-accrual tax treatment.

Asset Allocation and Portfolio Construction

With the theoretical characteristics and behavioral features of inflation-linked bonds, it is easy to assert that they are a truly new asset class that can add value to portfolios of other assets. Figure 67.5 shows efficient frontiers with and without TIPS, based on expected returns, volatility, and covariances obtained from Ibbotson Associates. Adding TIPS to a portfolio of traditional bonds and equities extends the efficient frontier and improves the risk-return characteristics conservative portfolios by largely replacing other U.S. Treasuries with TIPS.

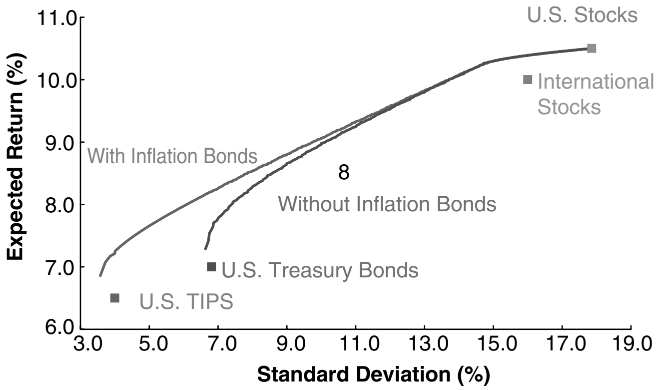


Figure 67.5 Inflation Bonds Improve Diversified Portfolios.
Source: Author's calculations based on Ibbotson data.

Perhaps the most important question about asset allocation and inflation-linked bonds is the search for the risk-free asset "holy grail." In modern portfolio theory, the risk-free rate is approximated in practice with short-term (e.g., 30- or 90-day) government bills on the theory that they will quickly reflect changes in real rates and inflation and not expectations. However, in theory and—as we have seen from their behavior—in practice, inflation-linked bonds come close to being the risk-free holy grail since they remove one of the largest systematic risks affecting fixed income assets. If so, then in Figure 67.6 we would expect the efficient frontier's risk-free extension tangent line to be composed of TIPS instead of Treasury bills. The principal effect of this substitution would be to increase the nominal (and perhaps real) risk-free rate compared to Treasury bills, thereby changing the tangent point where the risk-free rate line grazes the efficient frontier. This should affect the choice of efficient portfolios of risky assets and the risk control provided by allocations to the risk-free rate.

One implication of viewing inflation-linked bonds this way is the argument that they should be used by individuals and institutions to match assets and liabilities. They can be used to "immunize" against future liabilities, since liabilities are most affected by future inflation.

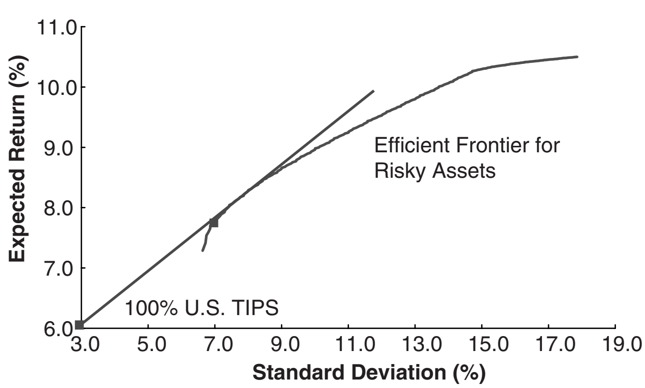


Figure 67.6 Are Inflation Bonds the "Risk-Free" Asset?
Source: Author's calculations based on Ibbotson data.

Table 67.4 Defined Contribution Retirement Spending

	Probability that Payments Last . . .		
	Until Death	**Until Age 100**	**Until Age 115**
Regular bonds	0.84	0.61	0.50
Inflation bonds	0.93	0.76	0.67

Source: Author's calculations based on Monte Carlo simulations of asset returns (and in the case of "until death" simulations of mortality).
Age = 65; annual spending = 4% of initial assets; asset allocation = 60% equities 40% bonds.

On the institutional side, for example, a defined benefit pension could use the triple-duration feature of inflation-linked bonds to provide a guarantee "floor" for benefit promises so that it takes into account both the nominal duration and the real duration of the fund. So a laddered portfolio of inflation bonds could guarantee a personal or institutional pension in real terms. Then, any assets that are held in excess of that guarantee could be invested in risky assets to provide an additional return (Bodie, 2005). The theoretical attraction of using inflation-linked bonds in this manner, however, have not been enthusiastically endorsed in practice. Although there are some signs of the use of linkers for immunization in the United Kingdom, few pension portfolios—defined benefit or defined contribution—in the United States have followed suit. Perhaps the perception of low initial yields has dampened interest.

Spending Policy

For portfolios with significant inflation-driven liabilities, inflation-linked bonds can also be used in implementing a spending policy. As the use of defined contribution retirement plans increases, the issue of the individual spending policy is becoming increasingly visible, particularly the spend-down rate. The usual advice is for the individual to create a conservative portfolio and then draw down each year approximately 4% of initial assets. But there is a significant probability that an individual will live longer than his or her assets under this arrangement. Even with a relatively high allocation to equities, a 60% equities/40% bond portfolio (as shown in Table 67.4) has a 0.16 probability of running out before the recipient dies. Replacing the nominal bond portfolio with inflation-linked bonds increases the chances of finishing in-the-money from 0.84 to 0.93. The probability differential increases with age. However, as with portfolio immunization, the use of inflation-linked bonds for pension payout purposes is also limited at this time, perhaps again because of low initial yields.

CONTINUING ISSUES

The attributes and advantages of inflation-linked bonds, some of which were predicted and others of which have emerged since their wider introduction in the last couple of decades of the twentieth century, seem clear. However, there are a number of continuing issues that remain either unresolved or in need of clarification.

Why so Little Nonsovereign Issuance?

While the global sovereign inflation-linked bond market has grown, the agency and private markets have seen little issuance. With the existence of national bonds as an anchor, it might be expected that other organizations could use that anchor to issue their own bonds. This might be especially attractive for organizations whose cash flows are inflation related, including utilities, state governments, toll authorities, and retail establishments, among others. Some agencies and private organizations have indeed issued inflation-linked bonds, principally in Australia, Canada, and the United Kingdom. But the private market in the United States and elsewhere is minimal. Perhaps one explanation is that, compared to sovereign nations, no other organization or institution relies so heavily on revenue streams, such as income taxes, that are so tightly tied to inflation. Another reason may be that inflation has declined significantly over the past two decades levels. Organizations that could, in the past, raise prices in response to inflation, are faced with competitive pressures or other factors that discourage price increases. Finally, around the world, many central banks have set inflation targets and have managed monetary policy to those targets. Even in the United States, there appears to be an implicit inflation "guideline" of about 2%. So inflation-bond issuance may seem to many organizations to be irrelevant or unattractive at present. One major factor that could influence future issuance will be actual inflation. As inflation rises beyond current expectations, there may be renewed interest among a wider variety of organizations.

What Is the Future of the Inflation-Linked Bond Market?

The agency and private organization issuance question is part of the more general question regarding the future of the inflation-linked bond market. While sovereign bond issuance continues to grow, it is apparent from yield and price movements as well as turnover that these markets, compared to nominal bond markets, are far from as deep and liquid. Large institutional trades can significantly move markets where issuance is only a fraction of the nominal bond market and a percentage of that fraction is infrequently traded. Further, the U.S. Treasury eliminated the 30-year TIPS when it ended issuance of the 30-year nominal Treasury. On the positive side, the U.S. Treasury has expressed continuing commitment to the TIPS program and has begun issuing a 20-year TIPS. Institutional investors now see inflation-linked bonds as an attractive asset class for long-term portfolios as well as short-term trading. In addition, inflation-linked swaps and other derivatives markets have emerged, along with ideas for new products. Finally, as world populations age, larger numbers of individuals will be dependent on public and private pension income that may or may not be explicitly indexed to inflation. In any event, retirees cannot rely on salary and wage increases (which are correlated with inflation) and they or their political representatives may seek additional inflation-protected solutions for retirement income. One such idea could be based on an equity-linked real bond product (Bhansali, 1998). All of these developments suggest that the market, while not exploding, is growing and should continue to grow wider and deeper over time.

SUMMARY

Whether we count the emergence of the modern inflation-linked bond market from 1981, when the U.K. began issuing linkers; 1997, when the U.S. began issuing TIPS, or the beginning of 2007, when the global market exceeded $1 trillion, inflation-linked bond theory and practice have developed and broadened considerably. Special or unique concepts, such as the inflation risk premium, the triple duration, and the break-even inflation carry; special trading opportunities and practices; and special products, such as real-return funds and inflation-linked annuities, all represent significant advances. They also presage more to come as our understanding of these bonds and the derivatives markets and solutions based on them continue to develop.

REFERENCES

Arnott, R. D. (2003). Editor's corner: The mystery of TIPS. *Financial Analysts Journal* 59, 5: 4–7.

Barclays Capital. (2006). *Global Inflation-Linked Products: A User's Guide*. New York: Barclays.

Barclays Capital. (2007). *Linkers Monthly*, January (appendix).

Bhansali, J. D. (1998a). Equity linked real bonds. *Derivatives Quarterly*, Fall: 1–10.

Bhansali, J. D. (1998b). "Inflation-indexed U.S. Treasury bonds: An analysis. *Journal of Investing*, Fall: 1–7.

Bodie, Z. (1990). "Inflation, indexed-linked bonds and asset allocation. *Journal of Portfolio Management* 16, 2: 48–53.

Bodie, Z. (2003). *Worry-Free Investing*. New York: FT Prentice Hall.

Bootle, R. (1991). *Index-Linked Gilts: A Practical Investment Guide*, 2nd Edition. Cambridge, UK: Woodhead-Faulkner.

Brynjolfsson, J., and Fabozzi, F. J. (eds.). (1999). *Handbook of Inflation Indexed Bonds.* Hoboken, NJ: John Wiley & Sons.

Campbell, J. Y., and Shiller, R. (1996). A scorecard for indexed government debt. In B. Bernanke and J. Rotemberg (eds.), *NBER Macroeconomics Annual* (pp. 155–196). Cambridge, MA: National Bureau of Economic Research.

Fabozzi, F. J. (2005). *Handbook of Fixed Income Securities*. New York: McGraw-Hill Professional.

Hammond, P. B. (2003). Turnover of inflation-linked bonds. TIAA-CREF working paper.

Hammond, P. B. (1999a). Using inflation-indexed securities for retirement saving and income: The TIAA-CREF experience. In J. Brynjolfsson and F. J. Fabozzi (eds.), *Handbook of Inflation Indexed Bonds* (pp. 21–41). Hoboken, NJ: John Wiley & Sons.

Hammond, P. B., Fairbanks, A., and Durham, B. J. (1999b). Understanding the inflation risk premium. In J. Brynjolfsson and F. J. Fabozzi (eds), *Handbook of Inflation Indexed Bonds* (pp. 157–182). Hoboken, NJ: John Wiley & Sons.

Homer, S., and Leibowitz, M. L. (2004). *Inside the Yield Book*, 2nd Edition. Princeton, NJ: Bloomberg Press.

Jarrow, R., and Yildirim, Y. (2003). " Pricing Treasury inflation protected securities and related derivatives using an HJM model." *Journal of Financial Quantitative Analysis* 38, 2: 337–358.

Leibowitz, M. L., Sorensen, E. H., Arnott, R. D., and. Nicholas, N. H. (1989). A total differential approach to equity duration. *Financial Analysts Journal* 45, 5: 30–37.

Roll, R. R. (2004). Empirical TIPS. *Financial Analysts Journal* 60, 1: 31–53.

Shiller, R. J. (2003). The invention of inflation-indexed bonds in early America. *NBER Working Paper Series*, 10183.

Siegel, L. B., and Waring, M. B. (2004). TIPS, the dual duration, and the pension plan. *Financial Analysts Journal* 60, 5: 52–64.

Tobin, J. (1971). An essay on the principles of public debt management. *Macroeconomics,* Vol. 1 *of Essays in Economics* (pp. 378–455) New York: Markham Publishing Co.

CHAPTER 68

Introduction to Inflation Derivatives

JEROEN KERKHOF PhD
Vice President, Morgan Stanley

Abstract: Numerous financial institutions face inflation risk in their activities. Using inflation derivatives allows them to transfer their inflation risk. The payoff of inflation derivatives depends on the value of inflation indices. Inflation indices are constructed by statistical agencies using the value of representative baskets of goods. By linking the payoff of inflation derivatives to representative baskets of goods inflation derivatives allow investors to focus on real rather than nominal returns. A variety of inflation derivatives exist in market and the most common types are the zero-coupon inflation swap and the year-on-year inflation swap.

Keywords: inflation risk, inflation indices, seasonality, nominal returns, real returns, zero-coupon inflation swap, year-on-year inflation swap, real swaptions

Inflation derivatives allow the transfer of inflation risk. For example, pension funds may want to cover their natural liabilities to inflation risk, while utility companies may want to shed some of their natural exposure to inflation risk. In short, inflation derivatives provide an efficient way to transfer inflation risk. Their flexibility allows users to replicate in derivative form the inflation risks embedded in other instruments such as standard cash instruments (that is, inflation-linked bonds). For example, as is explained later, an inflation swap can be theoretically replicated using a portfolio of a zero-coupon inflation-linked bond and a zero-coupon nominal bond.

As in other markets inflation derivatives have the advantage over the more standard inflation-linked bonds that they can be tailored to fit particular investor needs. Although in essence very similar the variety of inflation derivatives is much greater than that of inflation-linked bonds. For instance, in the cash market most of the inflation-linked bonds have the same maturity date and thereby exposure to the same inflation index fixings. Using inflation swaps one can get exposure to the January, February, . . . , or December index fixing. Besides having standard swap structures the inflation derivatives market also allows for exotic coupon structures linked to inflation. For instance, a real estate company might be interested in hedging the inflation adjustment in his rental income. In order to do so the company enters into a swap that replicates the cash flows from its rental contracts.

Contrary to their cash inflation-linked products, inflation derivatives are unfunded. Separating the issue of funding from inflation risk has made the inflation markets more accessible to parties with high funding costs and has made it cheaper to leverage inflation risk. For instance, hedge funds are increasingly involved in inflation markets and often use the derivative format rather than cash.

INFLATION DERIVATIVES MARKET

Since 2002 the inflation derivatives market has grown from a fairly exotic branch of the interest rate market to a well-established market of its own. Low returns on traditional fixed income assets in 2003 led to a demand for structured inflation products in 2003 on the back of which the zero-coupon swap market (due to hedging demand) increased substantially. Although the market has gathered critical mass and the growth has been swift, it still represents only a small fraction of the total interest rates market. In the longer run, one could expect the ratio of outstanding notional in nominal derivatives versus inflation derivatives to move towards the ratio of outstanding nominal bonds versus inflation-linked bonds.

The actively traded indices in the inflation derivatives market are the same as in the cash market. The main markets are the European market which uses the Euro Consumer Price Index (HICPxT) published by Eurostat, the French market using the FRCPI published by the Institut National de la Statistique et des Études Économiques (INSEE), and the U.K. market using the Retail Prices Index (RPI) published by National Statistics. As of early 2007, the U.S. market is relatively small, despite having the largest cash market (the U.S. Treasury Inflation Protection Securities, TIPS); as the TIPS market the derivatives market also uses the CPURNSA index published by the Bureau of Labor Statistics (BLS). The compositions of the various indices can vary considerably as can be seen in Figure 68.1.

The payers of inflation in the market are typically entities that receive inflation in the natural line of their business. Prime examples are sovereigns and utility companies that receive inflation via taxes and fees, respectively. For them having inflation-linked debt fits very well in their asset-liability management. On the receiving side, we typically have investors that need to pay inflation in the natural line of business. Pension funds are the prime example. As their liabilities are (either explicitly or implicitly) linked to inflation, inflation-linked securities fit very well in their asset-liability management. Besides for the above mentioned natural players inflation markets are also attractive for inflation neutral investors as they allow for diversification benefits for their portfolios.

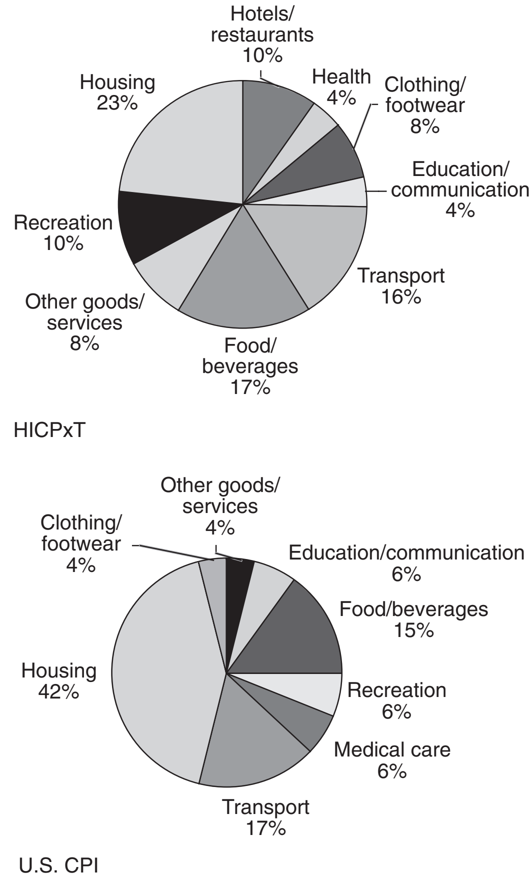


Figure 68.1 Constituents of HICP Ex-Tobacco and U.S. CPI
Source: Eurostat, Bureau of Labor Statistics.

INFLATION BASICS

Inflation, Nominal Value, and Real Value

A simple, but important and fundamental, economic axiom states that economic agents are concerned about the relative value of money rather than its absolute value. This well-known economic axiom underlies the existence of the inflation-linked market. In order to represent the concept of real value a basket of goods and services is constructed that tries to represent the basket of goods and services used by a representative customer. The nominal value of the basket of goods and services is computed at regular intervals (typically monthly). An inflation index is defined as the relative value of the basket. A base date (period) is chosen at which the nominal value of the index is set equal to (typically) 100. If the nominal value of the basket equals €1,000 at the base date it means that the index will rise 1 point if the value of the basket of goods increases by €10.

For example, let us consider an investor with assets equal to €100,000. The investor can currently buy 100 baskets of goods with his assets. The value of his basket of goods is represented by an inflation index. A year later the index has risen from 100 to 104 (the cost of the basket has increased to €1,040). This means that inflation was equal to 4% (= 104/100 – 1). Besides the increase in the index, the nominal value of the assets of the investor has increased to €102,000. The nominal increase in income for the investor equalled:

$$\text{Nominal change} = \frac{102{,}000 - 100{,}000}{100{,}000} = +2.00\%$$

However, due to the inflation of 4% the investor can now only buy 98.08 (= 102,000/1,040) baskets of goods and services. The real income change for the investor is therefore equal to:

$$\text{Real change} = \frac{102{,}000/1{,}040 - 100}{100} = -1.92\%$$

Of course, there is no need to compute the real change via the value of the reference basket. We find the same result using the nominal values and the inflation index:

$$\text{Real change} = \frac{102{,}000/104}{100{,}000/100} - 1 = -1.92\%$$

Even though the value of the investor's assets grew in nominal terms (+2%), in real terms his assets have actually decreased (–1.92%).

Real Bonds and Inflation-Linked Cash Flows: Ideal World

As investors care about real income rather than nominal income, they prefer to invest in securities guaranteeing them a real return rather than a nominal one. In this section, how investors can get guaranteed real returns instead of nominal returns by using inflation-linked bonds in a world with inflation is demonstrated.

An inflation-linked zero-coupon bond is a bond that has a single payment at time T, its maturity date. Its value today (denoting today with time 0) is denoted as $D_{IL}(0,T)$. The nominal payment at maturity equals

$$D_{IL}(T, T) = I(T)$$

the value of the inflation index at the maturity date. As investors are interested in real returns, they value all cash flows relative to the index, I. Therefore, the inflation-linked bond has a real value equal to

$$I(T)/I(T) = 1$$

real unit at maturity. In order to get the real value of this inflation-linked bond today, $D_r(0,T)$, we need to divide the value of the inflation-linked bond by the current value of the inflation index, $I(0)$. The real value of an inflation-linked zero-coupon bond is given by

$$D_r(0, T) = \frac{D_{IL}(0, T)}{I(0)}$$

The current nominal value of €1 at time T is denoted by $D_n(0,T)$, a nominal zero-coupon bond. Similarly, the current real value of 1 real unit at time T is denoted by $D_r(0,T)$, a real zero-coupon bond. The cash flows and values of a nominal and an inflation-linked zero-coupon bond in both nominal and real terms are illustrated in Table 68.1.

The real return (that is, return in real units) on an inflation-linked zero-coupon bond is given by

$$y_r(0, T) = D_r(0, T)^{-1/T} - 1$$

where $y_r(0,T)$ denotes the annualized real zero-yield and $D_r(0,T)$ denotes real value of an inflation-linked bond maturing at time T. Thus, using inflation-linked bonds one can get a guaranteed real return in the same way as nominal bonds allow one to get a guaranteed nominal return.

Table 68.1 Inflation-Linked Payments in Nominal and Real Terms

	Today ($t = 0$)	**Maturity (T)**
Inflation-linked zero-coupon bond		
Nominal units	$I(0)D_r(0,T) = D_{IL}(0,T)$	$I(T)$
Real units	$D_r(0,T)$	1
Nominal zero-coupon bond		
Nominal units	$D_n(0,T)$	1
Real units	$D_n(0,T)/I(0)$	$1/I(T)$

The nominal return on an inflation-linked bond is uncertain and given by:

$$\left(\frac{I(T)}{I(0)D_r(0, T)}\right)^{1/T} - 1 = \left(\frac{I(T)}{I(0)}\right)^{1/T} D_r(0, T)^{-1/T} - 1,$$

For example, without loss of generality, assume that the current index equals 100, $I(0) = 100$. The market trades an inflation-linked zero-coupon bond at 97.09% and a nominal bond at 95.24%, both with a time to maturity equal to 1 year ($T = 1$). From these values we can get the annualised nominal yields and real yields in the following manner.

Nominal yield on nominal zero-coupon bond:

$$y_n(0, T) = D_n(0, T)^{-1/T} - 1 = \frac{1}{0.9524} - 1 = 5.00\%$$

Real yield on inflation-linked zero-coupon bond:

$$y_r(0, T) = D_r(0, T)^{-1/T} - 1 = \frac{1}{0.9709} - 1 = 3.00\%$$

An investor can thus lock-in a guaranteed nominal return of 5% or a guaranteed real return of 3%. Given the growth of the inflation index, we can calculate the real yield on a nominal bond and the nominal yield on an inflation-linked bond. Assuming the inflation index grows to 102, $I(T) = 102$, these can be computed in the following manner.

Real yield on nominal zero-coupon bond:

$$\frac{1 \times I(0)/I(T)}{0.952} - 1 = \frac{100/102}{0.952} - 1 = 2.94\%$$

Nominal yield on inflation-linked zero-coupon bond:

$$\left(\frac{I(T)}{I(0)D_r(0, T)}\right)^{1/T} - 1 = \left(\frac{102}{100}\right)\frac{1}{0.9709} - 1 = 5.06\%$$

We see in Figure 68.2 that the nominal bond has a certain nominal return but uncertain real return, whereas the inflation-linked bond has a certain real return and an uncertain nominal return. Figure 68.2 also shows that for 1.94% inflation the nominal bond and the inflation-linked bond have the same nominal and real returns. This inflation level is denoted the breakeven inflation and I will discuss this later on in more depth.

As is the case in the nominal market, issuers do not issue zero-coupon bonds, but typically issue inflation-linked coupon bonds. In the same manner as for nominal bonds we can show that an inflation-linked coupon bond is nothing more than a portfolio of inflation-linked

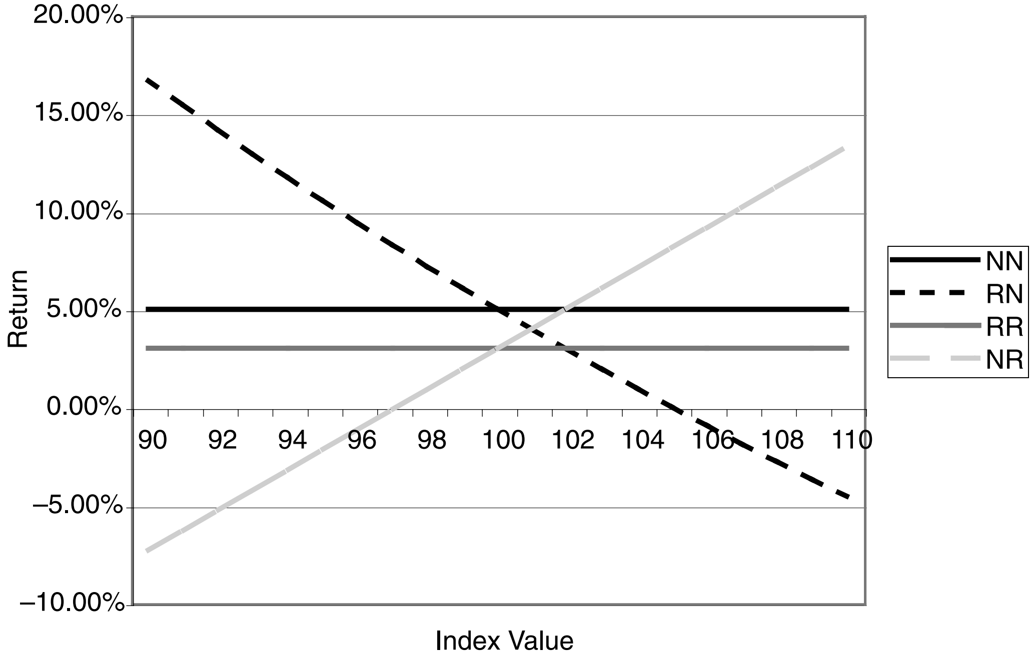


Figure 68.2 Nominal versus Real Returns of Nominal and Inflation-Linked Bonds

NN = nominal return on the nominal bond, RN = real return on the nominal bond, NR = nominal return on the real bond, and RR = real return on the real bond.

zero-coupon bonds with different maturities. Consider an inflation-linked coupon bond that pays a coupon equal to c at times $T_1, \ldots, T_N$. We can write the nominal value of this bond at time $0 \leq T_1$ as follows

$$\begin{aligned} B_{IL}(0, T_N) &= \sum_{i=1}^{N} c D_{IL}(0, T_i) + D_{IL}(0, T_N) \\ &= I(0) \left(\sum_{i=1}^{N} c D_r(0, T_i) + D_r(0, T_N) \right) \\ &= I(0) B_r(0, T_N). \end{aligned}$$

A detailed account of inflation-linked bonds is given in Deacon et al. (2004).

Real Bonds and Inflation-Linked Cash Flows: Practice

Basket

As discussed before, the main purpose of inflation-linked securities is to provide real value certainty. In the previous section, an ideal world was described. In practice, there are certain constraints due to which it is not possible to exactly guarantee a real return using inflation-linked bonds. First, it should be noted that only a limited number of inflation indices exist in the market. Because consumers are a heterogeneous group, one cannot hope to find a basket of goods which represents the different preferences of all consumers. At best, one can find a basket that represents the average consumer's preferences accurately. In measuring a real return an inflation index can therefore only be seen as an approximation.

Lags

In order to achieve a high degree of real value certainty the inflation-linked cash flows should be linked as closely as possible to contemporaneous inflation. However, in practice, the value of the index is not yet known for the cash flow date and a lagged index value is taken. As a result, investors have no inflation protection over the last period (typically, three months) of their inflation-protected security. They are compensated for this by receiving the inflation of the period preceding the purchase of the security. This is illustrated in Figure 68.3. In general, the inflation over the perfect indexation period is not equal to the inflation over the lagged inflation period leading to a lower degree of real value certainty.

Differences are likely to be bigger for longer indexation lags and more volatile inflation environments. Furthermore, the influence of the lag increases with decreasing time to maturity. Therefore, a small indexation lag is preferred for a high degree of real value certainty.

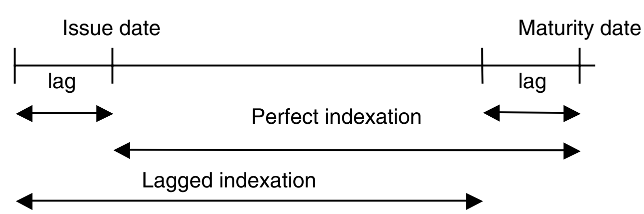


Figure 68.3 Indexation Lag

This figure graphically illustrates the influence of lagged inflation on real value certainty.

The indexation lags stems from two reasons. First, it takes time to process consumer price data and compute inflation numbers. Due to the processing time, typically inflation is announced about two weeks after the month under consideration (for example, January inflation is announced on about February 15). Second, a lag is needed due to trading and settling of bonds between coupon payment dates. As for nominal bonds, inflation-linked bonds usually pay coupons; if the bond trades between coupon dates sellers should be compensated for having held the bond for part of the coupon period even though they will not receive the coupon. As for nominal bonds, this compensation is effected via the payment of accrued interest. Two main methods of accrued interest payment are seen in practice. The oldest is the one employed by inflation-linked gilts in the U.K. market issued before 2005, where the next coupon is known at all times. This is achieved by using an eight-month lag consisting of a two-month period allowing for publication of the inflation index and six months for the accrued interest calculation (the inflation-linked gilts pay semiannual coupons).

A more common and preferred method these days is to base the accrued interest on the cumulative movements in the associated inflation index. This calculation method was initiated by Canada for their inflation-linked bonds and has been adopted in continental Europe, the United States, and in the United Kingdom for bonds issued from 2005 onwards. The method computes (daily) reference numbers for dates using a linear interpolation of the index values of, typically, two and three months ago. The reference number for the first of any calendar month equals the index value of the calendar month three months earlier. I(01-Apr-07) = CPI(Jan-07), I(01-May-07) = CPI(Feb-07), and so on. The reference numbers for other dates can then be computed using linear interpolation of the reference numbers of the first days of the calendar months. For example, in Figure 68.4 we compute the reference number, I(19-Sep-06) at 19 September 2006 for the French Consumer Price Index (CPI). In general, the daily reference number can be computed as follows:

$$I(dd/mm/yy) = I(01/mm/yy) + \frac{dd-1}{TDM}[I(01/mm+1/yy) - I(01/mm/yy)]$$

where TDM denotes the number of the total days in the month for all days between the first of January and the first of December. For the days in December we have:

$$I(dd/12/yy) = I(01/12/yy) + \frac{dd-1}{TDM}[I(01/01/yy+1) - I(01/12/yy)]$$

Using the (daily) reference numbers, inflation-linked bonds can be quoted in the standard manner, that is, as real bonds. However, in order to get the value of the inflation-linked bond, this price in real terms should be multiplied by the index ratio which is the current daily reference number computed in the manner suggested by the Canadian Treasury divided by the daily reference number at the start of the bond.

Breakeven Inflation

To explain the concept of breakeven inflation, we consider two products available in the market today. The first is a nominal zero-coupon bond with maturity date T, whose nominal value today is indicated by $D_n(0,T)$ and pays off 1 at maturity. The second is a zero-coupon inflation-linked bond with maturity date T, whose nominal value today we indicate by $D_{IL}(0,T) = I(0)D_r(0,T)$, where $I(0)$ denotes the current reference number and $D_r(0,T)$ denotes the real value of a real bond with maturity date T. The final payoff of this inflation-linked bond at maturity will equal $I(T)$, the reference number at maturity.

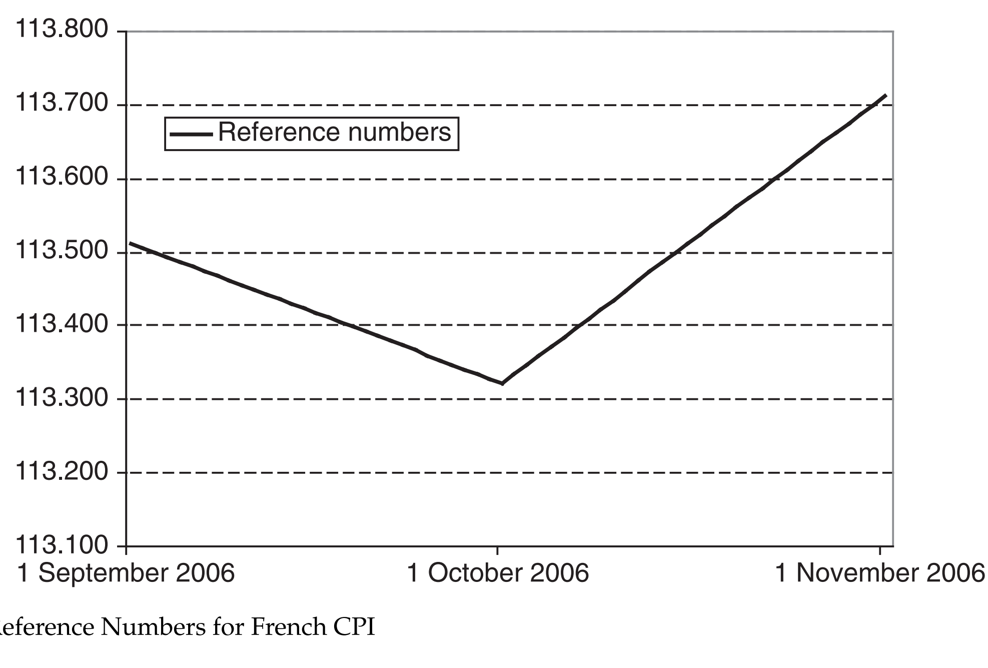


Figure 68.4 Reference Numbers for French CPI

We assume an investor has €100 to invest and needs to choose between the following two investments. Investment 1 is in nominal zero-coupon bonds, while investment 2 is in inflation-linked zero-coupon bonds.

Invest €100 in zero-coupon nominal bonds, that is, $100/D_n(0,T)$ units. The nominal payoff of this investment at maturity is given by

$$\frac{100}{D_n(0,T)} \times 1 = 100(1 + y_n(0,T))^T$$

where $y_n(0,T)$ is the annualised nominal yield on the nominal zero-coupon bond. Assuming $D_n(0,T) = 0.9524$ with $T = 1$, we have a final payoff of

$$\frac{100}{0.9524} \times 1 = 105.00$$

Invest €100 in zero-coupon inflation-linked bonds, that is, $100/(I(0)D_r(0,T))$ units. The nominal payoff of this investment at maturity is given by:

$$\frac{100}{I(0)D_r(0,T)} \times I(T) = 100(1 + i(0,T))^T(1 + y_r(0,T))^T$$

where $i(0,T)$ denotes the annual realised inflation and $y_r(0,T)$ is the annualized real yield on the real bond. Assuming $I(0) = 100$ and $D_r(0,T) = 0.9709$ for $T = 1$ we have as a final payoff:

$$\frac{100}{97.09} \times I(T) = 103.00 \times (1 + i(0,T))$$

which will depend on the inflation realized in the next year. The returns are illustrated in Table 68.2. The payoff from the nominal investment can be contracted today, while the payoff from the inflation-linked bond depends on realized inflation.

For the nominal investment, the nominal payoff at maturity is known today as $D_n(0,T)$, and thereby $y_n(0,T)$ are known today. For the inflation-linked investment, the nominal payoff at maturity depends on the realized inflation from today to maturity, $i(0,T)$. If realized inflation, $i(0,T)$, turns out to equal:

$$\frac{1 + y_n(0,T)}{1 + y_r(0,T)} - 1 = \frac{1.05}{1.03} - 1 = 1.94\%$$

the investor would, *ex post*, be indifferent between investment 1 and 2. We define this quantity as the breakeven inflation rate, $b(0,T)$:

$$b(0,T) = \frac{1 + y_n(0,T)}{1 + y_r(0,T)} - 1$$

It is easy to check that if inflation equalled 1.94% investors would have been indifferent between investing in the inflation-linked and the nominal bond. In the case of the nominal bond, they would have invested €100 in $100/0.9524 = 105.00$ nominal bonds, which returned €105.00 at maturity. In the case of the inflation-linked bond, they would have invested €100 in $100/0.9709 = 103.00$ inflation-linked bonds resulting in $103.00 \times 101.94 =$ €105.00 at maturity. The payoffs in both nominal and real terms thus coincide for both the nominal and inflation-linked bond if realised inflation equals the breakeven inflation. If the realised inflation turns out to be higher (lower) than $b(0,T)$, investors would have been better off investing in the inflation-linked (nominal) bond.

The breakeven rate gives us the indifference point of the realized inflation rate between the inflation-linked and the nominal investment. Another quantity of interest is the reference level for which the investor would be indifferent. This reference level is called the breakeven reference number and is denoted by $I(0,T)$. If the reference level at maturity, $I(T)$, equals:

$$\begin{aligned} I(0,T) &= \frac{I(0)D_r(0,T)}{D_n(0,T)} \\ &= I(0)(1 + b(0,T))^T \\ &= 100 \times (1.0194) \\ &= 101.94 \end{aligned}$$

the investor would, *ex post*, be indifferent between investment 1 and 2. If the reference index at maturity turns out to be higher (lower) than $I(0,T)$, investors would have been better off investing in the inflation-linked (nominal) bond.

With the introduction of the breakeven reference level, we can write the current nominal value of an inflation-linked payment at time T as $D_n(0,T) \times I(0,T)$, the discounted nominal value of the breakeven reference number. This follows from the fact that, by definition, we have:

$$I(0,T)D_n(0,T) = I(0)D_r(0,T)$$

Table 68.2 Nominal versus Inflation-Linked Investment

	Today	⟶	Maturity
Nominal	100	⟶	$100(1 + y_n\ (0,T))T$
Inflation-linked	100		$100(1 + y_n\ (0,T))^T\ (1 + i(0,T))T$

This table presents the payoffs of a nominal and an inflation-linked zero-coupon bond.

Components of Breakeven Inflation Rate

It is tempting to think that the breakeven rate should equal expected inflation. However, although expected inflation typically comprises the largest component of the breakeven swap rate, there are several reasons why they will not usually be the same. First, there is the *compounding effect*, which is a mathematical point. If the annualized inflation for the period from 0 to T, that is $i(0,T)$, is random, the expected payoff of an inflation-linked security would be higher than if it were to grow at the expected annualized inflation rate as a consequence of Jensen's inequality. In formulas, the compounding effect can be presented as

$$E[(1 + i(0,T))^T] \geq (1 + E[i(0,T)])^T$$

where E denotes expectation. The equality applies only if $i(0,T)$ is deterministic. Thus, the compounding effect has upward pressure on breakeven inflation rates.

Second, the inflation convexity, meaning the second-order price effect in case of inflation changes, increases with the maturity of the bond. High convexity is attractive for investors: it means that prices rise more than inflation duration predicts if breakeven inflation rates increase, and decrease less than inflation duration predicts if breakeven inflation rates decrease. As convexity is attractive for investors, it pushes down the breakeven rates.

Finally, as inflation-linked bonds provide a high degree of real value certainty, investors are willing to pay an *inflation risk* premium to receive inflation. The inflation risk premium pushes breakeven inflation higher than expected inflation. More specifically, consider risk-averse investors who are interested in real income which is perfectly matched by the daily reference numbers, I. At time t these investors can invest in either an inflation-linked bond with maturity T offering them a real return of $y_r(0,T)$ or a nominal bond with maturity T offering them a nominal return of $y_n(0,T)$. This gives the following real returns of the nominal and the inflation-linked bond, respectively.

$$\frac{I(0)(1+y_n(0,T))^T}{I(T)} \quad \text{versus} \quad (1+y_r(0,T))^T$$

Because the real return on the nominal bond is uncertain and the real return on the inflation-linked bond is certain, risk-averse investors will only consider investing in the nominal bond if they are compensated for bearing the inflation risk (or get diversification benefits). This will be the case if the expected real return on the nominal bond is higher than the real return on the inflation-linked bond, or if the nominal return on the nominal bond is higher than the expected nominal return on the inflation-linked bond, that is,

$$(1+y_n(0,T))^T \geq (1+y_r(0,T))^T E\left[\frac{I(T)}{I(0)}\right]$$

The additional return that sovereigns (or other issuers) need to pay on nominal issues compared with inflation-linked issues is called the inflation risk premium, which we denote by $p(0,T)$. We can now write the nominal rate as a Fisher equation:

$$\begin{aligned} 1+y_n(0,T) &= (1+y_r(0,T)) \times (1+E[i(0,T)]) \\ &\quad \times (1+c(0,T)) \times (1+p(0,T)) \end{aligned}$$

Thus, the nominal return equals the real return times the expected index increase times the risk premium. The size of the inflation risk premium depends on the volatility of inflation (higher volatility leads to higher premium) and the risk-averseness of investors (the more risk-averse the higher the premium).

INFLATION PRODUCTS

Zero-Coupon Inflation Swap

The basic building block of the inflation derivatives market is the zero-coupon inflation swap. Its appeal is its simplicity and the fact that it offers investors and hedgers a wide range of possibilities that did not previously exist in the cash market.

A fixed zero-coupon inflation swap is a bilateral contract that enables an investor or hedger to secure an inflation-protected return with respect to an inflation index. The inflation buyer (also called the inflation receiver) pays a predetermined fixed rate, and in return receives from the inflation seller (also called the inflation payer) an inflation-linked payment.

Figure 68.5 Cash Flows of Zero-Coupon Inflation Swap

The mechanics are fairly simple; today an inflation payer and an inflation receiver agree to exchange the change in the inflation index value from a base month (say, November 2006) to an end month (say, November 2012) versus a compounded fixed rate (see Figure 68.5).

If the value of the index in the base month is known at the time of the inception of the contract, we call the inflation swap spot starting. If the value of the index in the base month is not yet known, we speak of a forward starting inflation swap (it is shown later that forward starting inflation swaps are the building blocks of period-on-period inflation swaps).

The inflation market trades inflation swaps using two different conventions. The first convention, the CPI convention, uses the value of the CPI in the payoff while the second convention, the interpolated convention, uses the daily reference numbers to compute the payoff. Table 68.3 gives an overview of the conventions used in the main inflation markets. Table 68.4 provides an example term sheet of a spot starting inflation swap for the European HICPxT market.

Table 68.3 Market Conventions for Zero-Coupon Swaps

Market	Method
European (HICPxT, HICP – all items)	Monthly index level (3-month lag)
French (FR CPI)	Interpolated values
United Kingdom (UK RPI)	Monthly index level (2-month lag)
United States (CPI-NSA)	Interpolated values

Table 68.4 Example of a Term Sheet for an HICPxT Zero-Coupon Inflation Swap

Notional:	€100,000,000
Index:	HICPxT (non revised)
Source:	First publication by Eurostat as shown on Bloomberg CPTFEMU
Trade date:	21-Sep-2006
Start date:	25-Sep-2006 (Trade date + 2 business days)
End date:	25-Sep-2011 (Start date + tenor (5 years))
First fixing:	102.51 (June 2006)
Fixed leg:	$(1+2.12\%)^5 - 1$
Inflation leg:	$\frac{HICPxT(Jun/11)}{HICPxT(Jun/06)} - 1 = \frac{1(01-Sep-11)}{I(01-\mathrm{Sep}-06)} - 1$

Table 68.5 Example of a Term Sheet for French CPI Zero-Coupon Inflation Swap

Notional:	€100,000,000
Index:	FRCPI (nonrevised)
Source:	First publication by INSEE as shown on Reuters OATINFLATION01
Trade date:	21-Sep-2006
Start date:	25-Sep-2006 (Trade date + 2 business days)
End date:	25-Sep-2011 (Start date + tenor (5 years))
First fixing:	113.358 (reference number for 25-Sep-2006)
Fixed leg:	$(1 + 2.00\%)^5 - 1$
Inflation leg:	$\dfrac{I(25 - \text{Sep} - 11)}{I(25 - Sep - 06)} - 1 = \dfrac{\frac{24}{30}CPI(Jun/11) + \frac{6}{30}CPI(Jul/11)}{\frac{24}{30}CPI(Jun/06) + \frac{6}{30}CPI(Jul/06)} - 1$

The inflation swap starts on 25 September 2006 and ends on 25 September 2011 with the exchange of cash flows. The only unknown quantity in the payoff is the value of the HICPxT index for June 2011.

As the value of the HICPxT index in the contract month of 2011 needs to be known at the payment, the contract month is lagged to the current month (usually by two to three months). In the above example, the contract month is June and the HICPxT index values for June are normally published mid-July, which is well before the payment/ end date. As the value of the HICPxT index in June 2006 (it equalled 102.51) is known by September 21, 2006, the inflation swap in our example is spot starting. It is market standard to quote fixed inflation swaps whose initial lifetime equals a multiple of whole years (5 years in our example). Note that as the contract month is June, the inflation leg payoff in terms of reference numbers is based on September 1, not September 25. This has the advantage that all contracts trading with the same contract month and maturity have the same final payoff. This simplifies closing out of the position.

Not all markets use the convention to pay in terms of index levels. The market standard for the French FR-CPI and US-CPI is to define the payout on the inflation leg in terms of reference numbers. A term sheet would look like the one shown in Table 68.5.

The less liquid markets, such as the Dutch CBS index, are typically quoted as a spread to the HICPxT. For example, if the five-year breakeven rate for HICPxT equals 2% and the Dutch CBS index is quoted at 50 basis points, this means that the five-year breakeven rate for the Dutch CBS index equals 2.50%.

Valuing Zero-Coupon Swaps

An inflation swap has an inflation period starting at time S and ending at T over which the inflation is computed and a single payment date, which typically equals T, when the inflation payment (on the inflation leg) is exchanged with a fixed amount (on the fixed leg). The inflation leg thus pays the net increase in reference numbers from S to T, where $I(S)$ is known. The fixed leg pays a fixed amount which is conveniently written as an accumulated rate, b. The rate b is quoted in the market and called the breakeven swap rate. The rate b will differ depending on the current time and the inflation period, and therefore we use the notation $b = b(0;S,T)$ for the breakeven swap rate today for an inflation period S to T. In general, S can be different from today. In the term sheet given in Table 68.3 we take S = 01-Sep-06 and T = 01-Sep-11 and assume today is given by the September 21, 2006. The breakeven inflation swap rate quoted in the market equals b(0;01-Sep-06, 01-Sep-11) = 2.12% and the discount factor for September 25, 2011 equals 0.83. Assuming a notional equal to €100,000,000 the value of the fixed leg can then be computed as:

$$\begin{aligned}&\text{Current value of fixed leg}\\&= D_n(0, T)[(1 + b(0; S, T))^{T-S} - 1]\\&= 0.83 \times [(1 + 2.12\%)^5 - 1] \times 100{,}000{,}000\\&= 9.179{,}027.73\end{aligned}$$

The cash flow at maturity remains constant and therefore the fixed leg only varies with the discount factor. At inception the breakeven swap rate is set at such a level that the market considers the value of the fixed leg to equal the value of the uncertain inflation leg:

$$\begin{aligned}\text{Current value of inflating leg} &= \text{Current value of fixed leg}\\&= 9.179{,}027.73\end{aligned}$$

The only unknown on the inflation leg is the reference number at T. Using the concept of the inflation-linked zero-coupon bond introduced before we know that the current value of a payoff of $I(T)$ at T equals $I(0)D_r(0,T)$. This allows us to write the current value of the inflation leg as follows:

$$\begin{aligned}&\text{Current value of inflation leg}\\&= \left[\frac{I(0)D_r(0, T)}{I(T)} - D_n(0, T)\right] \times 1{,}000{,}000\\&= 9.179{,}027.73\end{aligned}$$

where the real discount bond, $D_r(0,T)$, is the remaining unknown. Using the fact that the value of the fixed leg and inflation leg are equal at inception, we find that the value of $D_r(0,T)$ consistent with the quoted breakeven swap rate equals:

$$\begin{aligned}D_r(0, T) &= \frac{I(S)}{I(0)} D_n(0, T_e)(1 + b(0; S, T))^{T-S}\\&= \frac{102.51}{102.41} \times 0.83 \times (1 + 2.12\%)^5 = 0.923\end{aligned}$$

where the value of the HICPxT index for June 2006 equals 102.51, the reference number today (September 21, 2006) equals 102.41, and as before the discount factor for September 25, 2011 equals 0.83.

Besides a breakeven swap rate for the swap, we can also compute a breakeven reference number for the zero-coupon inflation swap which we denote by $I(0;S,T)$. It is given by:

$$\begin{aligned}I(0; S, T) &= I(S)(1 + b(0; S, T))^{T-S}\\&= 102.51 \times (1 + 2.12\%)^5\\&= 113.85\end{aligned}$$

It is also easy to show that the start date of the period does not matter. Plugging in the bootstrapped value for $D_r(0,T)$ gives:

$$I(0,T) = \frac{I(0)D_r(0,T)}{D_n(0,T)} = I(T)(1 + b_e(0;S,T)^{T-S} = I(0;S,T)$$

One can easily check that $I(0)D_r(0,T) = I(0,T)D_n(0,T)$. We have $102.41 \times 0.923 = 113.85 \times 0.83$. As a special case of our extended definition of the breakeven swap rate, we have:

$$b(0,T) = b(0;0,T)$$

for a zero-coupon inflation swap with inflation period from today ($t = 0$) to T.

Period-on-Period Inflation Swaps

In the previous section I introduced and described the (fixed) zero-coupon inflation swap. Besides zero-coupon inflation swaps a number of other inflation swaps are traded, which are typically portfolios of zero-coupon swaps in one way or the other. Most are straightforward portfolios of spot starting zero-coupon inflation swaps. Here, a period-on-period (p-o-p) inflation swap which is a portfolio of forward starting zero-coupon swaps is described.

The inflation buyer (also called the inflation receiver) pays a predetermined fixed or floating rate (usually minus a spread). In return, the inflation buyer receives from the inflation seller (also called the inflation payer) inflation-linked payment(s). Two main types of (zero-coupon) inflation swaps exist: fixed inflation swaps (inflation versus fixed rate) and floating inflation swaps (inflation versus floating rate, usually Libor).

We call an inflation swap a payer inflation swap if you pay inflation and a receiver inflation swap if you receive inflation. Using an interest rate swap (IRS), we can find a no-arbitrage relationship between fixed and floating inflation swaps which we call the inflation swap parity:

Floating payer inflation swap
= Fixed payer inflation swap
+ Payer fixed-for-floating swap

Floating receiver inflation swap
= Fixed receiver inflation swap
+ Receiver fixed-for-floating swap

A period-on-period swap has multiple payments during its life. It pays the inflation over a number of accrual periods. The most common structure is the year-on-year (y-o-y) inflation swap, which pays annual inflation at the end of each year. An example term sheet is given in Table 68.6. The y-o-y inflation swap in Table 68.6 is initiated on September 21, 2006 and the inflation payer pays five times annual inflation from June to June every September 25 in the years 2006, . . . , 2011.

The cash flows on the inflation leg can be replicated using a series of forward starting zero-coupon inflation swaps. In the above example, we enter into forward starting zero-coupon swaps paying June 2006–2007 inflation, . . . , June 2010–2011 inflation. Therefore, the valuation can be done in terms of forward starting zero-coupon swaps. The valuation of forward starting zero-coupon swaps is not straightforward and involves modeling assumptions. Although the *year-on-year inflation swap* is the most popular instrument, other period-on-period swaps trade as well. We make a distinction between what we call pure period-on-period inflation swaps and annualized period-on-period inflation swaps.

Table 68.6 Example of a term sheet for HICPxT year-on-year inflation swap

Notional:	€100,000,000
Index:	HICPxT (non revised)
Source:	First publication by Eurostat as shown on Bloomberg CPTFEMU
Trade date:	21-Sep-2006
Start date:	25-Sep-2006 (Trade date + 2 business days)
End date:	25-Sep-2011 (Start date + tenor (5 years))
Rolls:	25th
Payment:	Annual, modified following
Day count:	30/360 unadjusted
First fixing	102.51 (June 2006)
First rate:	2.10%
Fixed leg:	day count fraction × fixed rate
Inflation leg:	$\left(\frac{HICPxT(Jun/yy+1)}{HICPxT(Jun/yy)} - 1\right) = \left(\frac{I(01-Sep-yy+1)}{I(01-Sep-yy)} - 1\right)$ for yy = 06, . . . , 10

A pure p-o-p inflation swap pays the inflation over the period on the inflation leg. For example, a semi-annual pure p-o-p inflation swap with contract months June and December pays the net increase in the index from June to December and the net increase from June to December. As the inflation payments are not on an annual basis, seasonality is an important issue when valuing these swaps.

Just like a y-o-y inflation swap, an annual p-o-p inflation swap pays annual inflation, but it pays it at a higher frequency and weighted with the appropriate day count fraction. For example, a semiannual annual p-o-p inflation swap with contract months June and December pays half the net index increase from last June to June and half the net index increase from last December to December. (See Figure 68.6.)

As the period for a year-on-year swap equals a year, a y-o-y inflation swap falls in both categories.

As seasonality plays a role in valuing the pure period-on-period swaps, they would trade at a premium as the inflation seller needs to be rewarded for taking on the seasonal risk. Using high frequency (e.g., monthly) annualised period-on-period inflation swaps seems quite attractive as it spreads inflation payments over the year instead of one lump sum payment each year without a seasonality premium as seasonality does not affect the valuation.

Inflation Futures

The Chicago Mercantile Exchange (CME) started trading futures on the U.S. CPI inflation index in February 2004.

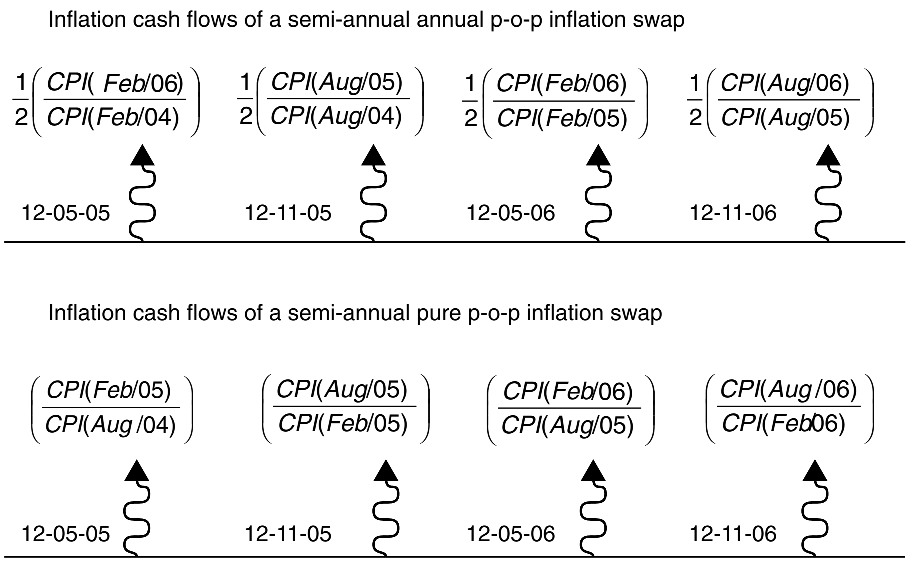


Figure 68.6 Inflation Leg Cash Flows of Two-Year Semiannual Period-on-Period Swaps

The figure plots the inflation leg cash flows from a semiannual p-o-p inflation swap with annual inflation periods and pure semi-annual inflation swaps. The fixed/floating leg payments are omitted.

The main advantage of CPI futures over zero-coupon inflation swaps is mitigated counterparty risk. The CPI futures traded on the CME are designed to resemble the Eurodollar futures contract. Likely due to the ill design of the CPI future (the contract traded annualised quarterly inflation), the market so far never really took off.

The CME started trading inflation futures on the Euro Consumer Price Index (HICPxT) in September 2005. One of the main advantages of the euro contract over the U.S. is that the inflation is annual. The main characteristics of the HICPxT contract are summarized in Table 68.7.

Considering their short maturity inflation futures complement the inflation-linked bond markets and allow investors to hedge short-term inflation exposures. As the futures trade on 12 consecutive months, investors can also take a view on inflation seasonality or hedge seasonality risk.

Limited Price Index Swaps

The Limited Price Index (LPI) swap is a typical U.K. instrument as U.K. pension funds that have limited indexation schemes. LPI swaps can come in a variety of flavors,

Table 68.7 Contract specification for the Euro Consumer Price Index HICPxT) inflation futures

Reference index	100 – annual inflation rate in the 12 month period preceding the contract month based on the Eurozone Harmonised Index of Consumer Prices excluding tobacco published by Eurostat. The same index is used for the French, Italian and Greek Euro inflation-linked bonds.
Settlement price	Final settlement amounts to 100 less the annual %-change in the HICPxT over the past 12 months and is rounded to four decimal places. Thus: $100 - \left[100 - \left(\frac{HICPxT(T_i)}{HICPxT_{(i-12)}} - 1\right)\right]$ where T_i denotes the contract month and T_{i-12} the base month. For example, for the July 2004 contract the relevant HICPxT index values are June 2004 (115.10, released July 16, 2004) and July 2003 (112.70, released July 18, 2003). The final settlement price would have been: $97.8705 = 100 - \left[100 - \left(\frac{115.10}{112.70} - 1\right)\right]$ A price of over 100 indicates deflation during the past 12-month period.
Contract months	12 consecutive calendar months.
Contract size	€10,000 times reference index.
Minimum tick size	0.01 index points, which amounts to €100.00.
Expiry date	Trading finishes 4:00 P.M. Greenwich Mean Time (GMT) on the business day preceding the scheduled day the HICPxT announcement is made in the contract month. In case the announcement is postponed beyond the contract month, trading ceases at 4:00 P.M. GMT on the last business day of the contract month.

Source: CME

but what they have in common is that, in one way or another, the inflation payment is capped and/or floored. If the inflation payment is both capped and floored, we call it collared. The most common traded LPI swap is the zero-coupon LPI swap. The zero-coupon LPI swap is particularly useful for pension funds that have liabilities related to the limited indexation of pensions in deferment as introduced in the 1995 Pension Act. They have liabilities of the following form:

$$L(0) = RPI(0),$$
$$L(T_i) = L(T_{i-1}) \times \max\left\{\min\left\{\frac{RPI(T_i)}{RPI(T_{i-1})}, 1.05^T\right\}, 1.0\right\}$$

where $L(0)$ denotes the liability when the retiree retires. The difference between T_{i-1} and T_i is usually a year. (See Table 68.8 for an example term sheet.)

The zero-coupon LPI swap is a highly path-dependent product. Other values for the cap and floor can be used (e.g. 0% and 3% is not uncommon). The 0% and 5% used are a result of the 1995 Pension Act legislation.

Real Swaptions

In the nominal interest rate market swaptions are (with caps and floors) the primary interest rate volatility instruments. Although nominal swaptions trade as exchanges of fixed versus floating coupons they can be seen as an option to exchange a fixed coupon bond versus a floating rate bond by including an offsetting notional exchange at the final payment. For the real swaption we discuss here, we use a similar concept. We apply the idea of having an option to exchange an inflation-linked bond versus a floating rate bond. The difference with the nominal case is the inflation uplift on the notional for an inflation-linked bond. Where a nominal bond pays the notional at maturity an inflation-linked bond pays the notional times the inflation uplift, that is,

$$\frac{I(T)}{I(0)} vs\ 1$$

In order to correct for this we add a payment of $I(T)/I(0) - 1$ to the inflation leg at maturity of the swap.

Note that we have chosen the unknown index for 2007 as the base in the inflation leg rather than the index for 2006.

Table 68.8 Example of a Term Sheet for LPI (0%–5%) Swap

Notional:	£ 100,000,000
Index:	UK RPI (nonrevised)
Source:	First publication by National Statistics as shown on Bloomberg UKRPI
Trade date:	21-Sep-2006
Start date:	25-Sep-2006 (Trade date + 2 business days)
End date:	25-Sep-2036 (Start date + tenor (30 years))
First fixing:	198.50 (June 2006)
Fixed leg:	(1 + 3.02%)30
Inflation leg:	$\prod_{y=2007}^{2036} \min\left(\max\left(\frac{\text{RPI(Jul/y)}}{\text{RPI(Jul/y}-1)}, 1\right), 1.05\right)$

Table 68.9 Example of a Term Sheet for Real Swaption on HICPxT

Notional:	100,000,000
Index:	HICPxT (non revised)
Source:	First publication by Eurostat as shown on Bloomberg CPTFEMU
Trade date:	21-Sep-2006
Maturity date:	25-Sep-2007
Swap Start date:	25-Sep-2007
Swap end date:	25-Sep-2027 (Start date + tenor (20 years))
Rolls:	25th September, modified following.
First fixing:	102.51 (June 2006)
Floating leg:	12m EURIBOR
Inflation leg:	$2.00\% \times \frac{\text{HICPxT(Jun/yy)}}{\text{HICPxT(Jun/2007)}}$ for yy = 2008, ..., 2027 and $\frac{\text{HICPxT(Jun/2007)}}{\text{HICPxT(Jun/2007)}} - 1$ for 2027.

This has advantages for the valuation of the real swaption. Using the current construction the real swaption can be valued using similar techniques as nominal swaptions. In case we would have used 2006 as the base the valuation becomes substantially harder as certain convexity adjustments need to be made.

ISDA INFLATION DERIVATIVES DOCUMENTATION

In 2005 the International Swap and Derivatives Association (ISDA) published documentation on inflation definitions supplementing the *ISDA Master Agreements*. The main issues relate to delay and disruption in the publication of the inflation index. Furthermore, it defines the most relevant indices. Here I will discuss the most important problems that can occur while settling an inflation derivatives contract.

Delay of Publication

If an inflation index is not published on time a substitute index is used. If the inflation index has not been published five business days prior to the next payment date for the transaction related to that index, the calculation agent, which is specified in the term sheet, shall use a substitute index level using the following methodology:

1. If applicable, the calculation agent takes the same action to determine the substitute index level as that specified in the terms and conditions of the related bond. The related bond, if any, is specified in the confirmation of the trade. A related bond is typically specified for asset swaps, but not for inflation swaps.
2. If (1) does not result in a substitute index level for the affected payment date the calculation agent determines

the substitute index level as follows:

Substitute index level
= Base level × (Latest level/Reference level)

where

Base level means the level of the index 12 calendar months prior to the month for which the substitute index level (definitive or provisional) is being determined, for example, December 2006.

Latest level means the latest available level (definitive or provisional) for the index, for example, November 2006.

Reference level means the level (definitive or provisional) of the index 12 calendar months prior to the month to which the latest level is referring, for example, November 2005.

If a relevant level is published after five business days prior to the next payment date, no adjustments will be made to the transaction. The determined substitute reference level will be the definitive level for that reference month.

Successor Index

If the inflation index no longer gets published, a replacement index will be used. If, during the term of the transaction, the index sponsor announces that an index will no longer be published or announced but will be superseded by a replacement index specified by the index sponsor, and the calculation agent determines that the replacement index is calculated using the same or similar methodology as the original index, this index is deemed the successor index.

Cessation of Publication

If an index has not been published for two consecutive months or if the index sponsor (publisher) has announced that it will no longer publish the index, the calculation agent shall determine a successor index for the purpose of the transaction using the following methodology:

1. If a successor index has been designated by the calculation agent pursuant to the terms and conditions of the related bond, such successor index shall be designated a successor index hereunder.
2. If no related bond exists, the calculation agent shall ask five leading independent dealers to state what the replacement index shall be. If three or more dealers out of at least four responses state the same index, this index will be deemed the successor index. If, out of three responses, two or more dealers state the same index, this index will be deemed the successor index. If no successor index has been decided following responses from dealers by the third business day prior to the next payment date or by the date that is five business days after the last payment date (if no further payment dates are scheduled), the calculation agent determines an appropriate alternative index. This alternative index will be deemed the successor index. If the calculation agent determines that there is no appropriate alternative index, a termination event occurs and both parties are affected parties as defined in the 2002 ISDA Master Agreement.

Rebasing the Index

If an index is rebased the rebased index will be used going forward. If the calculation agent determines that the index has been or will be rebased at any time, the rebased index will be used from then on. However, the calculation agent shall make adjustments pursuant to the terms and conditions of the related bond, if any. If there is no related bond, the calculation agent shall make adjustments to the past levels of the rebased index so that rebased index levels prior to the rebase date reflect the same inflation rate as before it was rebased.

Material Modification Prior to Payment Date

If prior to five business days before a payment date an index sponsor announces that it will make a material change to an index, then the calculation agent shall make any adjustments to the index consistent with adjustments made to the related bond. If there is no related bond, only those adjustments necessary for the modified index to continue as the index will be made.

Manifest Error in Publication

If, within 30 days of publication, the calculation agent is notified that the index level has to be corrected to remedy a material error in its original publication, the calculation agent will notify the parties of the correction and the amount payable as a result of that correction.

SUMMARY

In this chapter I have discussed a number of underlying key concepts for understanding the need for inflation derivatives. It was shown that nominal cash flows can be transformed into real cash flows using inflation cash products and more flexible using inflation derivatives. The key instrument in the inflation market is the zero-coupon inflation swap which naturally performs the role of transforming nominal cash flows into real cash flows.

REFERENCES

Belgrade, N., Benhamou, E., and Koehler, E. (2004). A market model for inflation. Working Paper, CDC Ixis Capital Markets.

Deacon, M., Derry, A., and Mirfendereski, D. (2004). *Inflation-Indexed Securities*, 2nd edition. London: John Wiley & Sons.

Hughston, L. (1998). Inflation derivatives. Working Paper, Merrill Lynch.

Mercurio, F. (2005). Pricing inflation-indexed derivatives. *Quantitative Finance* 5: 289–302.

Mercurio, F., and Moreni, N. (2006). Inflation with a smile. *Risk* 68: 70–75.

PART 9

Securities Finance

CHAPTER 69

An Introduction to Securities Lending

MARK C. FAULKNER
Managing Director, Spitalfields Advisors

Abstract: Securities lending—the temporary transfer of securities on a collateralized basis—is a major and growing activity providing significant benefits for issuers, investors, and traders alike. These are likely to include improved market liquidity, more efficient settlement, tighter dealer prices and, perhaps, a reduction in the cost of capital.

Keywords: securities lending, triparty agent, buy/sellbacks, custodian banks, prime brokers, beneficial owners, repos

This chapter describes securities lending, the motivation for lenders and borrowers to participate, the role of intermediaries, market mechanics, and the risks faced by the lenders of securities.

WHAT IS SECURITIES LENDING?

Securities lending is an important and significant business that describes the market practice whereby securities are temporarily transferred by one party (the lender) to another (the borrower). The borrower is obliged to return the securities to the lender, either on demand, or at the end of any agreed term. For the period of the loan the lender is secured by acceptable assets delivered by the borrower to the lender as collateral.

Securities lending today plays a major part in the efficient functioning of the securities markets worldwide. Yet it remains poorly understood by many of those outside the market.

In some ways, the term "securities lending" is misleading and factually incorrect. Under English law and in many other jurisdictions, the transaction commonly referred to as "securities lending" is, in fact . . .

> a disposal (or sale) of securities linked to the subsequent reacquisition of equivalent securities by means of an agreement.

Such transactions are collateralized and the "rental fee" charged, along with all other aspects of the transaction, is dealt with under the terms agreed between the parties. It is entirely possible and very commonplace that securities are borrowed and then sold or on-lent.

There are some consequences arising from this clarification:

1. Absolute title over both the securities on loan and the collateral received passes between the parties.
2. The economic benefits associated with ownership—e.g., dividends, coupons, etc.—are "manufactured"

back to the lender, meaning that the borrower is entitled to these benefits as owner of the securities but is under a contractual obligation to make equivalent payments to the lender.
3. A lender of equities surrenders its rights of ownership, e.g., voting. Should the lender wish to vote on securities on loan, it has the contractual right to recall equivalent securities from the borrower.

Appropriately documented securities lending transactions avoid taxes associated with the sale of a transaction or transference fees.

Different Types of Securities Loan Transaction

Most securities loans in today's markets are made against collateral in order to protect the lender against the possible default of the borrower. This collateral can be cash or other securities or other assets.

Transactions Collateralized with Other Securities or Assets

Noncash collateral would typically be drawn from the following collateral types:

- Government bonds
  - Issued by G7, G10 or non-G7 governments
- Corporate bonds
  - Various credit ratings
- Convertible bonds
  - Matched or unmatched to the securities being lent
- Equities
  - Of specified indices
- Letters of credit
  - From banks of a specified credit quality
- Certificates of deposit
  - Drawn on institutions of a specified credit quality
- Delivery by value (DBV)
  - Concentrated or unconcentrated
  - Of a certain asset class
- Warrants
  - Matched or unmatched to the securities being lent
- Other money market instruments

Note that *delivery by value* is a mechanism in some settlement systems whereby a member may borrow or lend cash overnight against *collateral*. The system automatically selects and delivers collateral securities, meeting predetermined criteria to the value of the cash (plus a margin) from the account of the cash borrower to the account of the cash lender and reverses the transaction the following morning.

The eligible collateral will be agreed upon between the parties, as will other key factors including:

- Notional limits
  - The absolute value of any asset to be accepted as collateral
- Initial margin
  - The margin required at the outset of a transaction

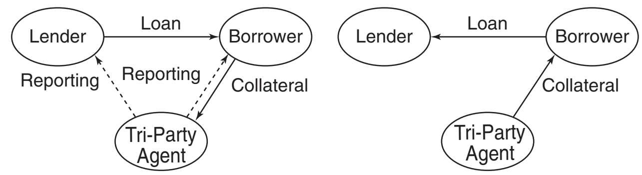


Figure 69.1 Noncash Collateral Held by a Third-Party Agent

- Maintenance margin
  - The minimum margin level to be maintained throughout the transaction
- Concentration limits
  - The maximum percentage of any issue to be acceptable, for example, less than 5% of daily traded volume
  - The maximum percentage of collateral pool that can be taken against the same issuer, that is, the cumulative effect where collateral in the form of letters of credit, CD, equity, bond and convertible may be issued by the same firm

Figure 69.1 shows collateral being held by a *triparty agent*. This specialist agent (typically a large custodian bank or international central securities depository) receives only eligible collateral from the borrower and hold it in a segregated account to the order of the lender. The triparty agent marks this collateral to market, with information distributed to both lender and borrower (in the diagram, dotted "Reporting" lines). Typically, the borrower pays a fee to the triparty agent.

Table 69.1 provides an illustration of cash flows on a securities against collateral other than cash for a transaction in the United Kingdom.

There is debate within the industry as to whether lenders, which are flexible in the range of noncash collateral that they are willing to receive, are rewarded with correspondingly higher fees. Some argue that they are; others claim that the fees remain largely static, but that borrowers are more prepared to deal with a flexible lender and, therefore, balances and overall revenue rise.

The agreement on a fee is reached between the parties and would typically take into account the following factors:

- Demand and supply
  - The less of a security available, other things being equal, the higher the fee a lender can obtain
- Collateral flexibility
  - The cost to a borrower of giving different types of collateral varies significantly, so that they might be more willing to pay a higher fee if the lender is more flexible
- The size of the manufactured dividend required to compensate the lender for the posttax dividend payment that it would have received had it not lent the security
  - Different lenders have varying tax liabilities on income from securities; the lower the manufactured dividend required by the lender, the higher the fee it can negotiate. (An explanation of how securities lending can be motivated by the different tax status of borrowers and lenders is discussed later in this chapter.)

Table 69.1 Cash Flows on a Securities Loan Against Collateral Other than Cash. The return to a lender of securities against collateral other than cash derives from the fee charged to the borrower. A cash flow of this transaction reads as follows:

Transaction date	June 13, 2007
Settlement date	June 16, 2007
Term	Open
Security	XYZ Limited
Security price	£10.00 per share
Quantity	100,000 shares
Loan value	£1,000,000.00
Lending fee	50 basis points (100ths of 1%)
Collateral	UK FTSE 100 Concentrated DBVs
Margin required	5%
Collateral required	£1,050,000.00 in DBVs
Daily lending income	£1,000,000.00 × 0.005 × (1/365) = £13.70

Should the above transaction remain outstanding for one month and be returned on July 16, 2005 there will be two flows of revenue from the borrower to the lender.
On June 30 fees of £191.80 (£13.70 × 14 days)
On July 31 fees of £219.20 (£13.70 × 16 days)
Thus, total revenue is £411.00 against which the cost of settling the transaction (loan and collateral) must be offset.
Note: For purposes of clarity, the example assumes that the value of the security on loan has remained constant, when in reality the price would change daily resulting in a mark to market event, different fees chargeable per day and changes in the value of the collateral required. Open loan transactions can also be re-rated or have their fee changed if market circumstances alter. It is assumed that this did not happen either.

- The term of a transaction
  - Securities lending transactions can be open to recalls or fixed for a specified term; there is much debate about whether there should be a premium paid or a discount for certainty. If a lender can guarantee a recall-free loan then a premium will be forthcoming. One of the attractions of repo and swaps is the transactional certainty on offer from a counterpart
- Certainty
  - As explained later in this chapter, there are trading and arbitrage opportunities, the profitability of which revolves around the making of specific decisions. If a lender can guarantee a certain course of action, this may mean it can negotiate a higher fee

Transactions Collateralized with Cash

Cash collateral is, and has been for many years, an integral part of the securities lending business, particularly in the United States. The lines between two distinct activities, securities lending and cash reinvestment, have become blurred and to many U.S. investment institutions securities lending is virtually synonymous with cash reinvestment. This is much less the case outside the United States, but consolidation of the custody business and the important role of U.S. custodian banks in the market means that this practice is becoming more prevalent. The importance of this point lies in the very different risk profiles of these increasingly intertwined activities.

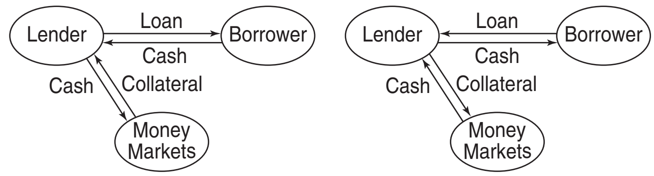


Figure 69.2 Cash-Collateral Securities Transaction

The revenue generated from cash-collateralized securities lending transactions is derived in a different manner from that in a noncash transaction (see Figure 69.2). It is made from the difference or "spread" between interest rates that are paid and received by the lender (see Table 69.2).

Reinvestment guidelines are typically communicated in words by the beneficial owner to their lending agent, and some typical guidelines might be as follows:

Conservative
- Overnight G7 government bond repo fund
- Maximum effective duration of 1 day
- Floating-rate notes and derivatives are not permissible
- Restricted to overnight repo agreements

Quite Conservative
- AAA-rated government bond repo fund
- Maximum average maturity of 90 days
- Maximum remaining maturity of any instrument is 13 months

Quite Flexible
- Maximum effective duration of 120 days
- Maximum remaining effective maturity of 2 years
- Floating-rate notes and eligible derivatives are permissible
- Credit quality: Short-term ratings: A1/P1, long-term ratings: A-/A3 or better

Flexible
- Maximum effective duration of 120 days
- Maximum remaining effective maturity of 5 years
- Floating-rate notes and eligible derivatives are permissible
- Credit quality: Short-term ratings: A1/P1, long-term ratings: A-/A3 or better

Some securities lending agents offer customized reinvestment guidelines while others offer reinvestment pools.

Other Transaction Types

Securities lending is part of a larger set of interlinked securities financing markets. These transactions are often used as alternative ways of achieving similar economic outcomes, although the legal form and accounting and tax treatments can differ. The other transactions are described in the following subsections.

Sale and Repurchase Agreements

Sale and repurchase agreements or repos involve one party agreeing to sell securities to another against a transfer of cash, with a simultaneous agreement to repurchase

Table 69.2 Cash Flows on a Securities Loan Collateralized with Cash

Transaction date	June 13, 2007
Settlement date	June 16, 2007
Term	Open
Security	XYZ Limited
Security price	£10.00 per share
Quantity	100,000 shares
Loan value	£1,000,000.00
Rebate rate	80 basis point
Collateral	USD cash
Margin required	5%
Collateral required	£1,718,850.00 (£1,050,000.00 × 1.67)
Reinvestment rate	130 basis points
Daily Lending Income	£23.87 or £14.58 (£1,718,850.00 × 0.005 × (1/360))

FX Rate assumed of £1.00 = $1.637

If the above transaction remains outstanding for one month and is returned on July 16, 2007, there will be two flows of cash from the lender to the borrower. These are based upon the cash collateral, and the profitability of the lender comes from the 50 basis points spread between the reinvestment rate and the rebate rate.

$1,718,850 × 0.008 × (1/360)) = $38.20

Payments to the borrower:

On June 30, $534.80 ($38.20 × 14 days)
On July 31, $611.20 ($38.20 × 16 days)

The lender's profit will typically be taken as follows:

On June 30, £204.12 (£14.58 × 14 days)
On July 31, £233.28 (£14.58 × 16 days)

Thus, total revenue is £437.40 against which the cost of settling the transactions (loan and collateral) must be offset.

Note: For purposes of clarity, this example assumes that the value of the security on loan has remained constant for the duration of the above transaction. This is most unlikely; typically the price would change daily resulting in a mark to market and changes to the value of the collateral required. Open loan transactions can also be re-rated or have their rebate changed if market circumstances alter. It is assumed that this did not happen either.

The marginal increase in daily profitability associated with the cash transaction at a 50-bps spread compared with the noncash transaction of 50-bps is due to the fact that the cash spread is earned on the collateral which has a 5% margin as well as the fact that the USD interest rate convention is 360 days and not 365 days as in the United Kingdom.

the same securities (or equivalent securities) at a specific price on an agreed date in the future. It is common for the terms "seller" and "buyer" to replace the securities lending terms "lender" and "borrower." Most *repos* are governed by a master agreement called the TBMA/ISMA Global Master Repurchase Agreement (GMRA) created by the Securities Industry and Financial Markets Association (SIFMA), a U.S. trade association.

Repos occur for two principal reasons: either to transfer ownership of a particular security between the parties or to facilitate collateralized cash loans or funding transactions.

The bulk of bond lending and bond financing is conducted by repurchase agreements (repos) and there is a growing equity repo market. An annex can be added to the GMRA to facilitate the conduct of equity repo transactions.

Repos are much like securities loans collateralized against cash, in that income is factored into an interest rate that is implicit in the pricing of the two legs of the transaction.

At the beginning of a transaction, securities are valued and sold at the prevailing "dirty" market price (that is, including any coupon that has accrued). At termination, the securities are resold at a predetermined price equal to the original sale price together with interest at a previously agreed rate known as the repo rate.

In securities-driven transactions (that is, where the motivation is not simply financing) the repo rate is typically set at a lower rate than prevailing money market rates to reward the "lender" who invests the funds in the money markets and, thereby, seek a return. The "lender" often receives a margin by pricing the securities above their market level.

In cash-driven transactions, the repurchase price typically is agreed at a level close to current money market yields, as this is a financing rather than a security-specific transaction. The right to substitute repoed securities as collateral is agreed by the parties at the outset. A margin is often provided to the cash "lender" by reducing the value of the transferred securities by an agreed "haircut" or discount.

Buy/Sellbacks

Buy/sellbacks are similar in economic terms to repos but are structured as a sale and simultaneous purchase of securities, with the purchase agreed for a future settlement date. The price of the forward purchase is typically calculated and agreed by reference to market repo rates.

The purchaser of the securities receives absolute title to them and retains any accrued interest and coupon payments during the life of the transaction. However, the price of the forward contract takes account of any coupons received by the purchaser.

Buy/sellback transactions are normally conducted for financing purposes and involve fixed income securities. In general a cash borrower does not have the right to substitute collateral. Until 1996, the bulk of buy/sellback transactions took place outside of a formal legal framework with contract notes being the only form of record. In 1995, the GMRA was amended to incorporate an annex that dealt explicitly with buy/sellbacks. Most buy/sellbacks are now governed by this agreement.

Table 69.3 compares the three main forms of collateralized securities loan transaction.

LENDERS AND INTERMEDIARIES

The securities lending market involves various types of specialist intermediary which take principal and/or agency roles. These intermediaries separate the underlying owners of securities—typically large pension or other funds, and insurance companies—from the eventual borrowers of securities, whose usual motivations are described later in this chapter.

Table 69.3 Summary of Collateralized Loan Transactions

	Securities Lending		Repo		
Characteristic	**Cash Collateral**	**Securities/Other Noncash Collateral**	**Specific Securities (securities driven)**	**General Collateral (cash driven)**	**Buy/sellback**
Formal method of exchange	Sale with agreement to make subsequent reacquisition of equivalent securities	Sale with agreement to make subsequent reacquisition of equivalent securities	Sale and repurchase under terms of master agreement	Sale and repurchase under terms of master agreement	Sale and repurchase
Form of exchange	Securities vs. cash	Securities vs. collateral (Note: Often free of payment but sometimes delivery versus delivery)	Securities vs. cash (*Note:* Often delivery versus payment)	Cash vs. securities (*Note*: Often delivery versus payment)	Cash vs. securities (*Note:* Often delivery versus payment)
Collateral type	Cash	Securities (bonds and equities), letters of Credit, DBVs, CDs	Cash	General collateral (bonds) or acceptable collateral as defined by buyer	Typically bonds
Return is paid to the supplier of	Cash collateral	Loan securities (not collateral securities)	Cash	Cash	Cash
Return payable as	Rebate interest (that is, return paid on cash lower than comparable cash market interest rates)	Fee, e.g., standard fees for FTSE 100 stocks are about 6–8 basis points (that is, 0.06–0.08% p.a.)	Quoted as repo rate, paid as interest on the cash collateral (lower than general collateral repo rate)	Quoted as repo rate, paid as interest on the cash	Quoted as repo rate, paid through the price differential between sale price and repurchase price
Initial margin	Yes	Yes	Yes	Yes	Possible
Variation margin	Yes	Yes	Yes	Yes	No (possible only) through close out and repricing)
Overcollateralization	Yes (in favor of the securities lender)	Yes (in favor of the securities lender)	No	Possible (if any, in favor of the cash provider)	Possible (if any, in favor of the cash provider)
Collateral substitution	Yes (determined by borrower)	Yes (determined by borrower)	No	Yes (determined by the original seller)	No (possible only through close out and repricing)
Dividends and coupons	Manufactured to the lender	Manufactured to the lender	Paid to the original seller	Paid to the original seller	No formal obligation to return income normally factored into the buyback price
Legal set off in event of default	Yes	Yes	Yes	Yes	No
Maturity	Open or term	Open or term	Open or term	Open or term	Term only
Typical asset type	Bonds and equities	Bonds and equities	Mainly bonds, equities possible	Mainly bonds, equities possible	Almost entirely bonds
Motivation	Security specific dominant	Security specific	Security specific	Financing	Financing dominant
Payment	Monthly in arrears	Monthly in arrears	At maturity	At maturity	At maturity

Intermediaries

Agent Intermediaries

Securities lending is increasingly becoming a volume business and the economies of scale offered by agents that pool together the securities of different clients enable smaller owners of assets to participate in the market. The costs associated with running an efficient securities lending operation are beyond many smaller funds for which this is a peripheral activity. Asset managers and custodian banks have added securities lending to the other services they offer to owners of securities portfolios, while third-party lenders specialize in providing securities lending services.

Owners and agents "split" revenues from securities lending at commercial rates. The split will be determined by many factors including the service level and provision by the agent of any risk mitigation, such as an indemnity.

Securities lending is often part of a much bigger relationship and therefore the split negotiation can become part of a bundled approach to the pricing of a wide range of services.

Asset Managers

It can be argued that securities lending is an asset-management activity—a point that is easily understood in considering the reinvestment of cash collateral. Particularly in Europe, where custodian banks were, perhaps, slower to take up the opportunity to lend than in the United States, many asset managers run significant securities lending operations.

What was once a back-office low profile activity is now a front office growth area for many asset managers. The relationship that the asset managers have with their underlying clients puts them in a strong position to participate.

Custodian Banks

The history of securities lending is inextricably linked with the custodian banks. Once they recognized the potential to act as agent intermediaries and began marketing the service to their customers, they were able to mobilize large pools of securities that were available for lending. This in turn spurred the growth of the market.

Most large custodians have added securities lending to their core custody businesses. Their advantages include: the existing banking relationship with their customers; their investment in technology and global coverage of markets, arising from their custody businesses; the ability to pool assets from many smaller underlying funds, insulating borrowers from the administrative inconvenience of dealing with many small funds and providing borrowers with protection from recalls; and experience in developing as well as developed markets.

Being banks, they also have the capability to provide indemnities and manage cash collateral efficiently—two critical factors for many underlying clients.

Custody is so competitive a business that for many providers it is a loss-making activity. However, it enables the custodians to provide a range of additional services to their client base. These may include foreign exchange, trade execution, securities lending, and fund accounting.

Third-Party Agents

Advances in technology and operational efficiency have made it possible to separate the administration of securities lending from the provision of basic custody services, and a number of specialist third-party agency lenders have established themselves as an alternative to the custodian banks. Their market share is currently growing from a relatively small base. Their focus on securities lending and their ability to deploy new technology without reference to legacy systems can give them flexibility.

Principal Intermediaries

There are three broad categories of principal intermediary: broker dealers, specialist intermediaries, and prime brokers. In contrast to the agent intermediaries, principal intermediaries can assume principal risk, offer credit intermediation, and take positions in the securities that they borrow. Distinctions between the three categories are blurred. Many firms would be in all three.

In recent years securities lending markets have been liberalized to a significant extent so that there is little general restriction on who can borrow and who can lend securities. Lending can, in principle, take place directly between beneficial owners and the eventual borrowers. But typically a number of layers of intermediary are involved. What value do the intermediaries add?

A beneficial owner may well be an insurance company or a pension scheme while the ultimate borrower could be a hedge fund. Institutions are often reluctant to take on credit exposures to borrowers that are not well recognized, regulated, or who do not have a good credit rating, which would exclude most hedge funds. In these circumstances, the principal intermediary (often acting as prime broker) performs a credit intermediation service in taking a principal position between the lending institution and the hedge fund.

A further role of the intermediaries is to take on liquidity risk. Typically they will borrow from institutions on an open basis—giving them the option to recall the underlying securities if they want to sell them or for other reasons—while lending to clients on a term basis, giving them certainty that they will be able to cover their short positions.

In many cases, as well as serving the needs of their own propriety traders, principal intermediaries provide a service to the market in matching the supply of beneficial owners that have large stable portfolios with those that have a high borrowing requirement. They also distribute securities to a wider range of borrowers than underlying lenders, which may not have the resources to deal with a large number of counterparts.

These activities leave principal intermediaries exposed to liquidity risk if lenders recall securities that have been on lent to borrowers on a term basis. One way to mitigate this risk is to use in-house inventory where available. For example, proprietary trading positions can be a stable source of lending supply if the long position is associated with a long-term derivatives transaction. Efficient inventory management is seen as critical and many securities lending desks act as central clearers of inventory within their organizations, only borrowing externally when netting of in-house positions is complete. This can require a significant technological investment. Other ways of mitigating "recall risk" include arrangements to borrow securities from affiliated investment management firms, where regulations permit, and bidding for exclusive (and certain) access to securities from other lenders.

On the demand side, intermediaries have historically been dependent upon hedge funds or proprietary traders that make trading decisions. But a growing number of securities lending businesses within investment banks have either developed "trading" capabilities within their lending or financing departments, or entered into joint ventures with other departments or even in some cases their hedge fund customers. The rationale behind this trend is

Table 69.4 Services Provided by Prime Brokers

Profitable Activities	Part of the Cost of Being in Business
Securities lending	Clearance
Leverage of financing provision	Custody
Trade execution	Reporting

that the financing component of certain trading strategies is so significant that without the loan there is no trade.

Broker-Dealers

Broker-dealers borrow securities for a wide range of reasons:

- Market making
- To support proprietary trading
- On behalf of clients

Many broker-dealers combine their securities lending activities with their prime brokerage operation (the business of servicing the broad requirements of hedge funds and other alternative investment managers). This can bring significant efficiency and cost benefits. Typically, within broker-dealers the fixed income and equity divisions duplicate their lending and financing activities.

Prime Brokers

Prime brokers serve the needs of hedge funds and other "alternative" investment managers. The business was once viewed, simply, as the provision of six distinct services, although many others such as capital introduction, risk management, fund accounting, and startup assistance have now been added (see Table 69.4).

Securities lending is one of the central components of a successful prime brokerage operation, with its scale depending on the strategies of the hedge funds for which the prime broker acts. Two strategies that are heavily reliant on securities borrowing are long/short equity and convertible bond arbitrage.

The cost associated with the establishment of a full-service prime broker is steep, and recognized providers have a significant advantage. Some of the newer entrants have been using total return swaps, contracts for difference, and other derivative transaction types to offer what has become known as "synthetic prime brokerage." Again, securities lending remains a key component of the service as the prime broker will still need to borrow securities in order to hedge the derivatives positions it has entered into with the hedge funds, for example, to cover short positions. But it is internalized within the prime broker and less obvious to the client.

Beneficial Owners

Those beneficial owners with securities portfolios of sufficient size to make securities lending worthwhile include pension funds, insurance and assurance companies, mutual funds/unit trusts, and endowments.

Beneficial Owner Considerations

When considering whether and how to lend securities, beneficial owners first need to consider organization characteristics and portfolio characteristics.

Organization characteristics include management motivation, technology investment, and credit risk appetitive. With regards to management motivation, some owners lend securities solely to offset custody and administrative costs, while others are seeking more significant revenue. Lenders vary in their willingness to invest in technological infrastructure to support securities lending. The securities lending market consists of organizations with a wide range of credit quality and collateral capabilities. A cautious approach to counterpart selection (AAA only) and restrictive collateral guidelines (G7 bonds) will limit lending volumes.

Portfolio characteristic include size, holdings size, investment strategy, investment strategy, tax jurisdiction and position, and inventory attractiveness. With respect to size, other things being equal, borrowers prefer large portfolios. Loan transactions generally exceed $250,000. Lesser holdings are of limited appeal to direct borrowers. Holdings of under $250,000 are probably best deployed through an agency programme, where they can be pooled with other inventories. Active investment strategies increase the likelihood of recalls, making them less attractive than passive portfolios. Borrowers want portfolios where they need liquidity. A global portfolio offers the greatest chance of generating a fit. That said, there are markets that are particularly in demand from time to time and there are certain borrowers that have a geographic or asset class focus.

With respect to tax jurisdiction and position, borrowers are responsible for "making good" any benefits of share ownership (excluding voting rights) as if the securities had not been lent. They must "manufacture" (that is, pay) the economic value of dividends to the lender. An institution's tax position compared to that of other possible lenders is therefore an important consideration. If the cost of manufacturing dividends or coupons to a lender is low then its assets will be in greater demand. Finally, regarding inventory attractiveness, "hot" securities are those in high demand while *general* collateral or general collateral securities are those that are commonly available. Needless to say, the "hotter" the portfolio, the higher the returns to lending.

Routes to the Market

Having examined the organization and portfolio characteristics of the beneficial owner, we must now consider the various possible routes to market. The possible routes to the securities lending market are briefly discussed below. For a more detailed discussion see Faulkner (2003).

Using an Asset Manager as Agent A beneficial owner may find that the asset manager they have chosen, already operates a securities lending programme. This route poses few barriers to getting started quickly.

Using a Custodian as Agent This is the least demanding option for a beneficial owner, especially a new one. They will already have made a major decision in selecting an appropriate custodian. This route also poses few barriers to getting started quickly.

Appointing a Third-Party Specialist as Agent A beneficial owner, who has decided to outsource, may decide it does not want to use the supplier's asset manager(s) or custodian(s), and instead appoint a third-party specialist. This route may mean getting to know and understand a new provider prior to getting started. The opportunity cost of any delay needs to be factored into the decision.

Auctioning a Portfolio to Borrowers Borrowers demand portfolios for which they bid guaranteed returns in exchange for gaining exclusive access to them. There are several different permutations of this auctioning route:

- Do-it-yourself auctions
- Assisted auctions
- Agent assistance
- Consultancy assistance
- Specialist "auctioneer" assistance

This is not a new phenomenon but one that has gained a higher profile in recent years and discussed in more detail in other chapters in this book. A key issue for the beneficial owner considering this option is the level of operational support that the auctioned portfolio will require and who will provide it. The key issue here is finding the best auctioneer.

Selecting One Principal Borrower Many borrowers effectively act as wholesale intermediaries and have developed global franchises using their expertise and capital to generate spreads between two principals that remain unknown to one another. These principal intermediaries are sometimes separately incorporated organizations, but more frequently, parts of larger banks, broker-dealers or investment banking groups. Acting as principal allows these intermediaries to deal with organizations that the typical beneficial owner may choose to avoid for credit reasons such as, hedge funds.

Lending Directly to Proprietary Principals Normally, after a period of activity in the lending market using one of the above options, a beneficial owner that is large enough in its own right, may wish to explore the possibility of establishing a business "in house," lending directly to a selection of principal borrowers that are the end users of their securities. The proprietary borrowers include broker-dealers, market makers and hedge funds. Some have global borrowing needs while others are more regionally focused.

Choosing Some Combination of the Above Just as there is no single or correct lending method, so the options outlined above are not mutually exclusive. Deciding not to lend one portfolio does not preclude lending to another; similarly, lending in one country does not necessitate lending in all. Choosing a wholesale intermediary that happens to be a custodian in the United States and Canada does not mean that a lender cannot lend Asian assets through a third-party specialist, and European assets directly to a panel of proprietary borrowers.

THE BORROWING MOTIVATION

One of the central questions commonly asked by issuers and investors alike is "Why does the borrower borrow my securities?" Before considering this point, let us examine why issuers might care.

If securities were not issued, they could not be lent. Behind this simple tautology lies an important point. When initial public offerings are frequent and corporate merger and acquisition activity is high, the securities lending business benefits. In the early 2000s, the fall in the level of such activity depressed the demand to borrow securities leading to a depressed equity securities lending market (that is, fewer trading opportunities, less demand, and fewer "specials") and issuer concern about the role of securities lending, such as whether it is linked in any way to the decline in the value of a company's shares or whether securities lending should be discouraged.

How many times does an issuer discussing a specific corporate event stop to consider the impact that the issuance of a convertible bond, or the adoption of a dividend reinvestment plan might have upon lending of their shares? There is a significant amount of information available on the "long" side of the market and correspondingly little on the short side. Securities lending activity is not synonymous with short selling. But it is often, although not always, used to finance short sales (discussed next) and might be a reasonable and practical proxy for the scale of short-selling activity in the absence of full short sale disclosure. It is, therefore, natural that issuers would want to understand how and why their securities are traded.

Borrowers, when acting as principals, have no obligation to tell lenders or their agents why they are borrowing securities. In fact, they may well not know themselves as they may be on-lending the securities to proprietary traders or hedge funds that do not share their trading strategies openly. Some prime brokers are deliberately vague when borrowing securities as they wish to protect their underlying hedge fund customer's trading strategy and motivation.

This section explains some of the more common reasons behind the borrowing of securities. In general, these can be grouped into: (1) borrowing to cover a short position (settlement coverage, naked shorting, market making, arbitrage trading); (2) borrowing as part of a financing transaction motivated by the desire to lend cash; and (3) borrowing to transfer ownership temporarily to the advantage of both lender and borrower (tax arbitrage, dividend reinvestment plan arbitrage).

Borrowing to Cover Short Positions

Settlement Coverage

Historically, settlement coverage has played a significant part in the development of the securities lending market. Going back a decade or so, most securities lending businesses were located in the back offices of their organizations and were not properly recognized as businesses in their own right. Particularly for less liquid securities—such as corporate bonds and equities with a limit free float—settlement coverage remains a large part of the demand to borrow.

The ability to borrow to avoid settlement failure is vital to ensure efficient settlement and has encouraged many securities depositories into the automated lending business. This means that they remunerate customers for making their securities available to be lent by the depository automatically in order to avert any settlement failures.

Naked Shorting

Naked shorting can be defined as borrowing securities in order to sell them in the expectation that they can be bought back at a lower price in order to return them to the lender. Naked shorting is a directional strategy, speculating that prices will fall, rather than a part of a wider trading strategy, usually involving a corresponding long position in a related security.

Naked shorting is a high-risk strategy. Although some funds specialize in taking short positions in the shares of companies they judge to be overvalued, the number of funds relying on naked shorting is relatively small and probably declining.

Market Making

Market makers play a central role in the provision of two-way price liquidity in many securities markets around the world. They need to be able to borrow securities in order to settle "buy orders" from customers and to make tight, two-way prices.

The ability to make markets in illiquid small capitalization securities is sometimes hampered by a lack of access to borrowing, and some of the specialists in these less liquid securities have put in place special arrangements to enable them to gain access to securities. These include guaranteed exclusive bids with securities lenders.

The character of borrowing is typically short term for an unknown period of time. The need to know that a loan is available tends to mean that the level of communication between market makers and the securities lending business has to be highly automated. A market maker that goes short and then finds that there is no loan available would have to buy that security back to flatten its book.

Arbitrage Trading

Securities are often borrowed to cover a short position in one security that has been taken to hedge a long position in another as part of an "arbitrage" strategy. Some of the more common arbitrage transactions that involve securities lending are described in the following subsections.

Convertible Bond Arbitrage

Convertible bond arbitrage involves buying a convertible bond and simultaneously selling the underlying equity short and borrowing the shares to cover the short position. Leverage can be deployed to increase the return in this type of transaction. Prime brokers are particularly keen on hedge funds that engage in convertible bond arbitrage as they offer scope for several revenue sources:

- Securities lending revenues
- Provision of leverage
- Execution of the convertible bond
- Execution of the equity

Pairs Trading or Relative Value "Arbitrage" This in an investment strategy that seeks to identify two companies with similar characteristics whose equity securities are currently trading at a price relationship that is out of line with their historical trading range. The strategy entails buying the apparently undervalued security while selling the apparently overvalued security short, borrowing the latter security to cover the short position. Focusing on securities in the same sector or industry should normally reduce the risks in this strategy.

Index Arbitrage In this context, arbitrage refers to the simultaneous purchase and sale of the same commodity or stock in two different markets in order to profit from price discrepancies between the markets.

In the stock market, an arbitrage opportunity arises when the same security trades at different prices in different markets. In such a situation, investors buy the security in one market at a lower price and sell it in another for more, capitalizing on the difference. However, such an opportunity vanishes quickly as investors rush in to take advantage of the price difference.

The same principle can be applied to index futures. Being a derivative product, index futures derive their value from the securities that constitute the index. At the same time, the value of index futures is linked to the stock index value through the opportunity cost of funds (borrowing/lending cost) required to play the market.

Stock index arbitrage involves buying or selling a basket of stocks and, conversely, selling or buying futures when mispricing appears to be taking place.

Financing

As broker dealers build derivative prime brokerage and customer margin business, they hold an increasing inventory of securities that requires financing.

This type of activity is high volume and takes place between two counterparts that have the following coincidence of wants: One has cash that they would like to invest on a secured basis and pickup yield. The other has inventory that needs to be financed.

In the case of bonds, the typical financing transaction is a repo or buy/sellback. But for equities, securities lending and equity, repo transactions are used.

Triparty agents are often involved in this type of financing transaction as they can reduce operational costs for the cash lender and they have the settlement capabilities the cash borrower needs to substitute securities collateral as their inventory changes.

Temporary Transfers of Ownership
Temporary transfers of ownership are driven by tax arbitrage and dividend plan reinvestment arbitrage opportunities.

Tax Arbitrage Tax driven trading is an example of securities lending as a means of exchange. Markets that have historically provided the largest opportunities for tax arbitrage include those with significant tax credits that are not available to all investors—examples include Italy, Germany, and France.

The different tax positions of investors around the world have opened up opportunities for borrowers to use securities lending transactions, in effect, to exchange assets temporarily for the mutual benefit of purchaser, borrower, and lender. The lender's reward comes in one of two ways: either a higher fee for lending if they require a lower manufactured dividend, or a higher manufactured dividend than the posttax dividend they would normally receive (quoted as an "all-in rate").

For example, an offshore lender that would normally receive 75% of a German dividend and incur 25% withholding tax (with no possibility to reclaim) could lend the security to a borrower that, in turn, could sell it to a German investor who was able to obtain a tax credit rather than incur withholding tax. If the offshore lender claimed the 95% of the dividend that it would otherwise have received, it would be making a significant pick-up (20% of the dividend yield), while the borrower might make a spread of between 95% and whatever the German investor was bidding. The terms of these trades vary widely and rates are calculated accordingly.

Dividend Reinvestment Plan Arbitrage Many issuers of securities create an arbitrage opportunity when they offer shareholders the choice of taking a dividend or reinvesting in additional securities at a discounted level.

Income or index tracking funds that cannot deviate from recognized securities weightings may have to choose to take the cash option and forgo the opportunity to take the discounted reinvestment opportunity.

One way that they can share in the potential profitability of this opportunity is to lend securities to borrowers that then take the following action:

- Borrow as many guaranteed cash shares as possible, as cheaply as possible.
- Tender the borrowed securities to receive the new discounted shares.
- Sell the new shares to realize the "profit" between the discounted share price and the market price.
- Return the shares and manufacture the cash dividend to the lender.

MARKET MECHANICS

This section outlines the processes in the life of a securities. Specifically, the following are discussed:

- Negotiation of loan deals
- Confirmations
- Term of loan
- Term trades
- Putting securities "on hold"
- Settlements, including how loans are settled and settlement concerns
- Termination of loans
- Redelivery, failed trades, and legal remedies
- Corporate actions and voting

There are other issues that are unique to specific countries. These include any tax arrangements and reporting of transactions to an exchange or other authority/regulator.

Loan Negotiation

Traditionally securities loans have been negotiated between counterparts (whose credit departments have approved one another) on the phone, and followed up with written or electronic confirmations. Normally the borrower initiates the call to the lender with a borrowing requirement. However, proactive lenders may also offer out in-demand securities to their approved counterparts. This would happen particularly where one borrower returns a security and the lender is still lending it to others in the market, they will contact them to see if they wish to borrow additional securities.

Today, there is an increasing amount of bilateral and multilateral automated lending whereby securities are broadcast as available at particular rates by email or other electronic means. Where lending terms are agreeable, automatic matching can take place.

An example of an electronic platform for negotiating equity securities loan transactions is EquiLend, which began operations in 2002 and is backed by a consortium of financial institutions. EquiLend's stated objective is to:

> Provide the securities lending industry with the technology to streamline and automate transactions between borrowing and lending institutions and . . . introduce a set of common protocols. EquiLend will connect borrowers and lenders through a common, standards-based global equity lending platform enabling them to transact with increased efficiency and speed, and reduced cost and risk.

EquiLend is not alone in this market; for example, SecFinex offers similar services in Europe.

Confirmations

Written or electronic confirmations are issued, whenever possible, on the day of the trade so that any queries by the other party can be raised as quickly as possible. Material changes during the life of the transaction are agreed between the parties as they occur and may also be confirmed if either party wishes it. Examples of material changes are collateral adjustments or collateral substitutions. The parties agree who will take responsibility for issuing loan confirmations.

Confirmations would normally include the following information:

- Contract and settlement dates
- Details of loaned securities

- Identities of lender and borrower (and any underlying principal)
- Acceptable collateral and margin percentages
- Term and rates
- Bank and settlement account details of the lender and borrower

Term of Loan, and Selling Securities While on Loan

Loans may be either for a specified term or open. Open loans are trades with no fixed maturity date. It is more usual for securities loans to be open or "at call," especially for equities, because lenders typically wish to preserve the flexibility for fund managers to be able to sell at any time. Lenders are able to sell securities despite their being on open loan because they can usually be recalled from the borrower within the settlement period of the market concerned. Nevertheless, open loans can remain on loan for a long period.

Term Trades—Fixed or Indicative?

The general description "term trade" is used to describe differing arrangements in the securities lending market. The parties have to agree whether the term of a loan is "fixed" for a definite period or whether the duration is merely "indicative" and, therefore, the securities are callable. If fixed, the lender is not obliged to accept the earlier return of the securities; nor does the borrower need to return the securities early if the lender requests it. Accordingly, securities subject to a fixed loan should not be sold while on loan.

Where the term discussed is intended to be "indicative," it usually means that the borrower has a long-term need for the securities but the lender is unable to fix for term and retains the right to recall the securities if necessary.

Putting Securities "On Hold"

Putting securities "on hold" (referred to in the market as "icing" securities) is the practice whereby the lender will reserve securities at the request of a borrower on the borrower's expected need to borrow those securities at a future date. This occurs where the borrower must be sure that the securities will be available before committing to a trade that will require them.

While some details can be agreed between the parties, it is normal for any price quoted to be purely indicative and for securities to be held to the following business day. The borrower can "roll over" the arrangement (that is, continue to "ice" the securities) by contacting the holder before 9 A.M.; otherwise, it terminates.

Key aspects of icing are that the lender does not receive a fee for reserving the securities and they are generally open to challenge by another borrower making a firm bid. In this case the first borrower would have 30 minutes to decide whether to take the securities at that time or to release them.

"Pay-to-Hold" Arrangements

A variation of icing is "pay-to-hold," where the lender does receive a fee for putting the securities on hold. As such, they constitute a contractual agreement and are not open to challenge by other borrowers.

Settlements

Securities lenders need to settle transactions on a shorter timeframe than the customary settlement period for that market. Settlement will normally be through the lender's custodian bank and this is likely to apply regardless of whether the lender is conducting the operation or delegating to an agent. The lender will usually have agreed a schedule of guaranteed settlement times for its securities lending activity with its custodians. Prompt settlement information is crucial to the efficient monitoring and control of a lending program, with reports needed for both loans and collateral.

In most settlement systems securities loans are settled as "free-of-payment" deliveries and the collateral is taken quite separately, possibly in a different payment or settlement system and maybe a different country and time zone. For example, U.K. equities might be lent against collateral provided in a European International Central Securities Depository or U.S. dollar cash collateral paid in New York. This can give rise to what is known in the market as "daylight exposure," a period during which the loan is not covered as the lent securities have been delivered but the collateral securities have not yet been received. To avoid this exposure some lenders insist on precollateralization, thereby transferring the exposure to the borrower.

The CREST system for settling U.K. and Irish securities is an exception to the normal practice as collateral is available within the system. This enables loans to be settled against cash intraday and for the cash to be exchanged, if desired, at the end of the settlement day for a package of DBV securities overnight. The process can be reversed and repeated the next day.

Termination of the Loan

Open loans may be terminated by the borrower returning securities or by the lender recalling them. The borrower will normally return borrowed securities when it has filled its short position. A borrower will sometimes refinance its loan positions by borrowing more cheaply elsewhere and returning securities to the original lender. The borrower may, however, give the original lender the opportunity to reduce the rate being charged on the loan before borrowing elsewhere.

Redelivery, Failed Trades, and Legal Remedies

When deciding which markets and what size to lend in, securities lenders consider how certain they can be of having their securities returned in a timely manner when called, and what remedies are available under the legal agreement (discussed later) in the event of a failed return.

Procedures to be followed in the event of a failed redelivery are usually covered in legal agreements or otherwise agreed between the parties at the outset of the relationship. Financial redress may be available to the lender if the borrower fails to redeliver loaned securities or collateral on the intended settlement date. Costs that would typically be covered include:

- Direct interest and/or overdraft incurred.
- Costs reasonably and properly incurred as a result of the borrower's failure to meet its sale or delivery obligations.
- Total costs and expenses reasonably incurred by the lender as a result of a "buy-in" (that is, where the lender is forced to purchase securities in the open market following the borrower's failure to return them)

Costs that would usually be excluded are those arising from the transferee's negligence or willful default and any indirect or consequential losses. An example of that would be when the nonreturn of loaned securities causes an onward trade for a larger amount to fail. The norm is for only that proportion of the total costs which relates to the unreturned securities or collateral to be claimed. It is good practice, where possible, to consider "shaping" or "partialing" larger transactions (that is, breaking them down into a number of smaller amounts for settlement purposes) so as to avoid the possibility of the whole transaction failing if the transferor cannot redeliver the loaned securities or collateral on the intended settlement date.

Corporate Actions and Votes

The basic premise underlying securities lending is to make the lender "whole" for any corporate action event—such as a dividend, rights, or bonus issue—by putting the borrower under a contractual obligation to make equivalent payments to the lender, for instance by "manufacturing" dividends. However a shareholder's right to vote as part owner of a company cannot be manufactured. When securities are lent, legal ownership and the right to vote in shareholder meetings passes to the borrower, who will often sell the securities on. Where lenders have the right to recall securities, they can use this option to restore their holdings and voting rights. The onus is on the borrower to find the securities, by borrowing or purchasing them in the market if necessary. This can damage market liquidity, which is a risk that intermediaries manage.

It is important that beneficial owners are aware that when shares are lent the right to vote is also transferred. For example, in the United Kingdom, the Securities Lending and Repo Committee's (SLRC's) code of guidance states in Section 2.5.4 that lenders should make it clear to clients that voting rights are transferred. A balance needs to be struck between the importance of voting and the benefits derived from lending the securities. Beneficial owners need to ensure that any agents they have made responsible for their voting and stock lending act in a coordinated way.

Borrowing securities in order to build up a holding in a company with the deliberate purpose of influencing a shareholder vote is not necessarily illegal in the United Kingdom. However, institutional lenders have recently become more aware of the possibility, and tend not to see it as a legitimate use of securities borrowing.

A number of market bodies throughout the world have been addressing the relationship between securities lending and voting. Internationally, a working group of the International Corporate Governance Network is currently examining best practices for longterm investors in relation to securities lending and voting. The SLRC is also considering additions to its code in this area.

FINANCIAL RISKS AND RISK MANAGEMENT

This section reviews the main financial risks in securities lending and how lenders usually manage them. More detailed discussion of these risks is provided in other later chapters.

Financial risks in securities lending are primarily managed through the use of collateral and netting. As described earlier in this chapter, collateral can be in the form of securities or cash. The market value of the collateral is typically greater than that of the lent portfolio. This margin is intended to protect the lender from loss and reflect the practical costs of collateral liquidation and repurchase of the lent portfolio in the event of default. Any profits made in the repurchase of the lent portfolio are normally returned to the borrower's liquidator. Losses incurred are borne by the lender with recourse to the borrower's liquidator along with other creditors.

When Taking Cash as Collateral

Because of its wide acceptability and ease of management, cash can be highly appropriate collateral. However, the lender needs to decide how best to utilize this form of collateral. As described earlier in this chapter, a lender taking cash as collateral pays rebate interest to the securities borrower, so the cash must be reinvested at a higher rate to make any net return on the collateral. This means the lender needs to decide on an appropriate risk-return trade-off. In simple terms, reinvesting in assets that carry one of the following risks can increase expected returns: a higher credit risk: a risk of loss in the event of defaults or a longer maturity in relation to the likely term of the loan. Many of the large securities lending losses over the years have been associated with reinvestment of cash collateral.

Typically, lenders delegate reinvestment to their agents (e.g., custodian banks). They specify reinvestment guidelines, such as those set out earlier in this chapter. There is a move toward more quantitative, risk-based approaches, often specifying the "value at risk" in relation to the different expected returns earned from alternative reinvestment profiles. Agents do not usually offer an indemnity against losses on reinvestment activity so that the lender retains all of the risk while their agent is paid part of the return.

When Taking Other Securities as Collateral

Compared with cash collateral, taking other securities as collateral is a way of avoiding reinvestment risk. In addition to the risks of error, systems failure and fraud always present in any market, problems then arise on the default of a borrower. In such cases the lender will seek to sell the collateral securities in order to raise the funds to replace the lent securities. Transactions collateralized with securities are exposed to a number of different risks that are described below.

Reaction and Legal Risk

If a lender experiences delays in either selling the collateral securities or repurchasing the lent securities, it runs a greater risk that the value of the collateral will fall below that of the loan in the interim. Typically, the longer the delay, the larger the risk.

Mispricing Risk

The lender will be exposed if either collateral securities have been overvalued or lent securities undervalued because the prices used to mark-to-market differ from prices that can actually be traded in the secondary market. One example of mispricing is using mid rather than bid prices for collateral. For illiquid securities, obtaining a reliable price source is particularly difficult because of the lack of trading activity.

Liquidity Risk

Illiquid securities are more likely to be released at a lower price than the valuation used. Valuation "haircuts" are used to mitigate this risk (that is, collateral is valued at, for example, 98% or 95% of the current market value). The haircuts might depend upon:

- The proportion of the total security issue held in the portfolio—the larger the position, the greater the haircut
- The average daily traded volume of the security: the lower the volume, the greater the haircut
- The volatility of the security; the higher the volatility, the greater the haircut

Congruency of Collateral and Lent Portfolios (Mismatch Risk)

If the lent and collateral portfolios were identical, then there would be no market risk. In practice, of course, the lent and collateral portfolios are often very different. The lender's risk is that the market value of the lent securities increases but that of the collateral securities falls before rebalancing can be effected. Provided the counterpart has not defaulted, the lender will be able to call for additional collateral on any adverse collateral/loan price movements. However, following default, it will be exposed until it has been able sell the collateral and replace the lent securities.

The size of mismatch risk depends on the expected covariance of the value of the collateral and lent securities. The risk will be greater if the value of the collateral is more volatile, the value of the lent securities is more volatile, or if their values do not tend to move together, so that the expected correlation between changes in their value is low.

Many agent intermediaries will offer beneficial owners protection against these risks by agreeing to return (buy-in) lent securities immediately for their clients following a fail, taking on the risk that the value of the collateral on liquidation is lower.

SUMMARY

Securities lending is a global activity that the majority of financial organizations engage in to some greater or lesser extent. Some outsource what they see as a noncore activity to others (e.g., pension funds and hedge funds); others specialize in the activity and add value as either agent or principal intermediaries (e.g., custodian banks acting as agents or broker dealers acting as prime brokers.

The importance of this business to individual organizations varies significantly as they take their share of a $10 billion dollar gross revenue pool. The importance to the global markets overall is much more important as this activity facilitates liquidity, enables pricing efficiency, and provides hedging opportunities.

ACKNOWLEDGMENTS

The author is grateful for the assistance provided by David Rule, Simon Hills, Dagmar Banton, John Serocold, Andrew Clayton, Joyce Martindale, Susan Adeane, Habib Motani, Niki Natarajan, Andrew Barrie, Jackie Davis, and Bill Cuthbert.

REFERENCES

Blount, E., and Gerdeman, A. J. (2005). Managing liquidity risks in cash-based lending programs. In Fabozzi, F. J., and Mann, S. V. (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 127–140). Hoboken, NJ: John Wiley & Sons.

Dropkin, C. E. (2005). Developing effective guidelines for managing legal risks—U.S. guidelines. In Fabozzi, F. J., and Mann, S. V. (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 167–178). Hoboken, NJ: John Wiley & Sons.

Economou, P. (2005). Risk, return, and performance measurement in securities lending. In Fabozzi, F. J., and Mann, S. V. (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 151–166). Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. (ed.) (1997). *Securities Lending and Repurchase Agreements*, Hoboken, NJ: John Wiley & Sons.

Faulkner, M. C. (2005a). Finding a route to market: An institutional guide to the securities lending labyrinth. In Fabozzi, F. J., and Mann, S. V. (eds.), *Securities Finance:*

Securities Lending and Repurchase Agreements (pp. 57–78). Hoboken, NJ: John Wiley & Sons.

Faulkner, M. C. (2005b). Quantifying risks in securities lending transactions. In Fabozzi, F. J., and Mann, S. V. (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 141–150), Hoboken, NJ: John Wiley & Sons.

Kiefer, D. E., and Mabry, J. G. (2005). The auction process and its role in the securities lending markets. In Fabozzi, F. J., and Mann, S. V. (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 87–106). Hoboken, NJ: John Wiley & Sons.

Nazzaro, A. A. (2005). Evaluating lending options. In Fabozzi, F. J., and Mann, S. V. (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 79–86). Hoboken, NJ: John Wiley & Sons.

Peters, S. C. (2005). Accounting treatment of loans of securities. In Fabozzi, F. J., and Mann, S. V. (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 205–217). Hoboken, NJ: John Wiley & Sons.

Shapiro, R. J. (2005). Tax issues associated with securities lending. In Fabozzi, F. J., and Mann, S. V. (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 179–204). Hoboken, NJ: John Wiley & Sons.

CHAPTER 70

Mechanics of the Equity Lending Market

JEFF COHEN
Securities Lending Manager, Susquehanna Intl Group, LLLP

DAVID HAUSHALTER, PhD
Corporate Research and Educational Associate, Susquehanna Intl Group, LLLP

ADAM V. REED, PhD
Assistant Professor of Finance, University of North Carolina at Chapel Hill

Abstract: Short selling involves a transaction outside the stock market; short sellers borrow stock for delivery to buyers. The equity lending market allows short sellers and other market participants to borrow shares from stock owners for a price. Supply and demand factors determine each loan's price. Since supply of shares is usually high, most loans have low prices. However, episodic events can lead to significantly higher prices. Although there are risks for borrowers and lenders in the equity loan market, cash collateral allows the risk for lenders to be minimized.

Keywords: equity loans, securities lending, short sales, delivery, settlement, rebate rate, specialness, delivery failure

Short sellers sell stock they do not own. The equity lending market exists to match these short sellers with owners of the stock willing to lend their shares for a fee. Despite its obvious importance to the operation of financial markets, the equity lending market is arcane. The market is dominated by loans negotiated over the phone between borrowers and lenders. Although there have been significant improvements in recent years, there is no widely used electronic quote or trade network in the equity lending market.

In this chapter, we discuss the mechanics of equity loans, the participants and their roles, and how rebate rates (prices) are determined in the market.

THE LENDING PROCESS

An investor who wants to sell a stock short must first find a party willing to lend the shares. One exception to this rule is for market makers. For example, the New York Stock Exchange (NYSE) requires affirmative determination (a locate) of borrowable or otherwise attainable shares for members who are not market makers, specialists or odd lot brokers in fulfilling their market-making responsibilities. Similar rules exist for the National Association of Securities Dealers (NASD) and American Exchange (AMEX) (see Evans et al., 2003).

Once a lender has been located and the shares are sold short, exchange procedures generally require that the short-seller deliver shares to the buyer on the third day after the transaction ($t + 3$) and post an initial margin requirement at its brokerage firm. Under Regulation T, the initial margin requirement is 50%. Self-regulatory organizations (e.g., NYSE and NASD) require the short seller to maintain a margin of at least 30% of the market value of the short position as the market price fluctuates.

As described in Figure 70.1, the proceeds from the short sale are deposited with the lender of the stock. For U.S.

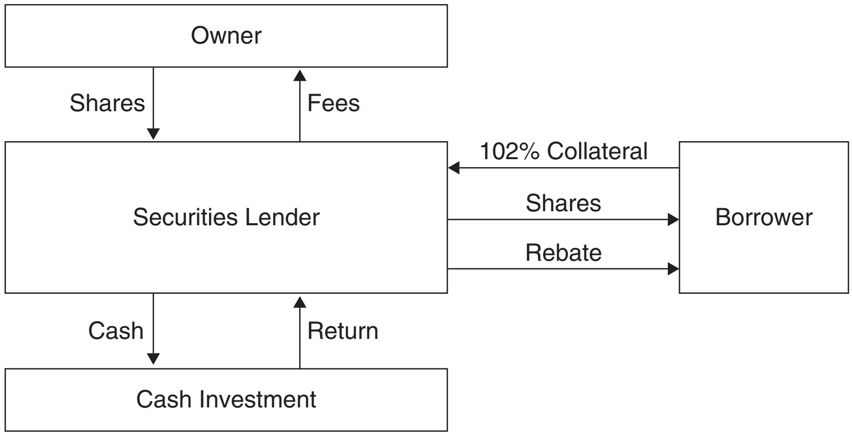


Figure 70.1 Equity Loan Structure

stocks, the lender requires 102% of the value of the loan in collateral. The value of the loan is marked to market daily; an increase in the stock price will result in the lender requiring additional collateral for the loan, and a decrease in the stock price will result in the lender returning some of the collateral to the borrower. When the borrower returns the shares to the lender, the collateral will be returned.

While a stock is on loan, the lender invests the collateral and receives interest on this investment. Generally, the lender returns part of the interest to the borrower in the form of a negotiated *rebate rate*. Therefore, rather than fees, the primary cost to the borrower is the difference between the current market interest rate and the rebate rate the lender pays the borrower on the collateral. A lender's benefit from participating in this market is the ability to earn the spread between these rates. Although the earnings from this interest spread are often split between several parties participating in the lending process, the interest can add low risk return to a lender's portfolio.

LENDERS

Traditionally, custodian banks that clear and hold positions for large institutional investors have been the largest equity lenders. With the beneficial owner's permission, custodian banks can act as lending agents for the beneficial owners by lending shares to borrowers. The custodian bank and the beneficial owners share in any revenue generated by securities lending with a prearranged fee sharing agreement. A typical arrangement would have 75% of the revenue going to the beneficial owner and 25% going to the agent bank (see Bargerhuff & Associates, 2000). Depending on the type of assets being lent and the borrowing demand, lending revenue earned by the owner of the security may completely offset custodial and clearance fees for institutional investors.

In addition to traditional custodian bank lenders, a number of specialty third-party agent lenders have entered the equity lending market over the past several years. Under this structure, the assets are lent by an agent firm who represents the beneficial owner but is not the custodian of the assets. Once a loan is negotiated between the agent lender and the borrower, the agent facilitates settlement by working with a traditional custodian bank in arranging delivery of the shares to the borrower. In comparison with custodian banks, these noncustodial lenders often offer advantages to the beneficial owner such as more specialized reporting, flexibility, and more lending revenue.

As an alternative to agency-lending arrangements, the beneficial owner may decide to lend assets directly to borrowers. Increasingly, owners choose to lend their assets via an exclusive arrangement, where the owner commits his assets to one particular borrower for a specific period of time. For example, in recent years, the California Public Employees Retirement System (CalPERS) has lent its portfolios through an auction system with the winning bidder gaining access to the portfolio for a predetermined period of time. This arrangement guarantees a return to the beneficial owner for loaning out the assets. Another avenue that some institutions have explored is managing their own internal lending department, therefore having total control over the lending process and keeping all of the revenues generated. Due to the large costs involved in setting up a lending department and the infrastructure needed, this option is only available to the largest institutional investors.

Lender's Rights

The owner of a stock retains beneficial ownership of the shares it lends. This status gives the owner the right to receive the value of any dividends or distributions paid by the issuing company while the stock is on loan. However, rather than being paid by the company, the dividend and distributions are paid by the borrower. This is referred to as a "substitute payment." The beneficial owner is also entitled to participate in any corporate actions that occur while the security is on loan. For example, in the case of a tender offer, if the beneficial owner wishes to participate in the offer and the borrower is unable to return the security prior to the completion of the offer, the borrower is required to pay the beneficial owner the tender price. The only right the lender gives up when lending their assets is the right to vote on a security. (For a discussion of lending and voting, see Christoffersen et al. [2004].) However, the lender generally has the right to recall the loaned security from the borrower for any reason, including to exercise voting rights.

In the event of a recall, the borrower is responsible for returning the shares to the lender within the normal settlement cycle. For example, if the beneficial owner sells a security that is on loan, the agent lender will send a recall notice to the borrower on the first business day after the trade date ($T + 1$) instructing borrower that the shares need to be returned to the agent within two business days ($T + 3$). If the shares are returned within this period, the custodian can settle the pending sell trade. If the borrower fails to return the shares by ($T + 3$), the agent may buy shares to cover the position, therefore closing out the loan.

Lender's Risks

There are three types of risk the beneficial owner faces when lending stock: investment risk, counterparty risk, and operational risk. Investment risk involves the choices that the beneficial owner or their agent makes in investing collateral. Some lenders are reluctant to take risk in their reinvestment of collateral, and they invest primarily in overnight repurchase agreements or other very low risk investments. Other lenders look to achieve extra income by investing in higher risk assets. For example, lenders can earn more return by investing in longer term investments and short-term corporate debt with lower credit ratings. It is the beneficial owner's responsibility to monitor the investment of the collateral to manage these risks. Even if there is a loss from investing the borrower's collateral, the beneficial owner is still responsible for returning the borrower's full collateral when the security is returned. An example of this risk is provided by Citibank which, acting as an agent lender, is estimated to have lost approximately $80 million in collateral on an investment in asset-backed security issued by National Century Financial Enterprises. After this event occurred, it was unclear whether Citibank would cover the beneficial owners for this loss of collateral.

Counterparty risk is the risk that the borrower fails to provide additional collateral or fails to return the security. The beneficial owner can manage this risk by approving only the most creditworthy borrowers and by imposing credit limits on these borrowers. Furthermore, the fact that collateral is marked to market daily allows lenders to buy shares to cover the loan if the borrower will not return the shares.

The last major risk to the beneficial owner is operational risk. This is the risk that various responsibilities of the agent lender or borrower are not met. This could be the failure to collect dividend payments, the failure to instruct clients on corporate actions resulting in missed profit opportunities, the failure to mark a loan to market, and the failure to return a security in the event of a recall. These risks can be minimized by maintaining a good lending system which tracks dividends, corporate actions, and the collateralization of loans.

BORROWERS

The largest borrowers of stocks are prime brokerage firms facilitating the short demand for their own proprietary trading desks, for their hedge fund clients, and for other leveraged investors. Trading desks often borrow stock to enable long-short trading strategies. Furthermore, tremendous growth in the hedge fund industry during the past decade has resulted in an increase in the use of other sophisticated strategies that require borrowing stock according to a September 2003 staff report of the Securities and Exchange Commission (SEC) on hedge funds. Because lending firms are reluctant to approve hedge funds as creditworthy borrowers, hedge funds have traditionally used prime brokers to gain access to the lending markets.

The two risks that a borrower faces are the risk of a loan recall and the risk of a decrease in rebate rates. A borrower's challenge is to find a lender that best balances these risks. Recall risk is the risk of the stock's, being recalled by the lender before the borrower is prepared to close out his position, which happens in approximately 2% of the loans in the sample of a study by D'Avolio (2002). Borrowers would prefer to have loans lasting the duration of the short position, but guaranteed term loans are rare. (For a discussion of loan terms, see D'Avolio (2002) and Geczy, Musto, and Reed [2002].) So, borrowers need to manage recall risk by working with a lender that is likely to be willing to loan the stock for an extended period of time. Often, the most stable sources of stock loans are portfolios with little turnover, such as index funds.

There are no rules governing which loans will be recalled if a beneficial owner recalls its stock. If the agent for the lender has loaned the stock to several prime brokerage firms and some of those shares need to be returned, the lending agent has discretion in deciding which prime brokers' loans will be recalled. Moreover, if the prime broker, whose loan has been selected, has allocated these shares to several borrowers, the broker has flexibility in selecting which of the borrowers will have their shares recalled. If the borrower's loan does get recalled by the lender, it is the borrowers' responsibility to return shares to the lender either by buying shares in the market or by borrowing the shares from another lender. If the borrower fails to return the shares, the lender can use the borrower's collateral to buy shares to cover the loan, which is known as a buy-in. In other words, recalls can force borrowers to unwind their trading strategies suboptimally or expose the borrowers to potentially poor execution in the case of a buy-in.

THE DETERMINANTS OF REBATE RATES

The rebate rate, or the rate a borrower is paid on his cash collateral, effectively determines the price of a stock loan. This rate is determined by supply and demand in the market for borrowing stock. For highly liquid stocks that are widely held by institutional lenders, the borrower can expect to earn the full rebate or general collateral rate, on the collateral. This rate is generally 5 to 25 basis points below the Fed funds rate for each day. (In a Fitch IBCA's report ("Securities Lending and Managed Funds") it is estimated that the industry average spread from the Fed funds rate to the general collateral rate on U.S. equities is 21 basis points.) When there is less available supply in the equity lending market, as with middle-capitalization stocks, the spread generally increases to around 35 basis points according to Bargerhuff & Associates (2000).

The majority of loans in the equity lending market are made in widely held stocks that are cheap to borrow. However, on less widely held securities or securities with large borrowing demand, rebate rates may be reduced, in which case, the securities are said to be "trading special" or just "special." This means that the rebate rate is negotiated on a case by case basis, and the rate earned by the borrower on the collateral is below the general collateral rate paid on easily available securities. Only a few stocks are on special each day; a one-year sample in a study by Geczy, Musto, and Reed (2002) had approximately 7% of its securities on special. And, the specials aren't necessarily limited to small stocks; 2.77% of large stocks were found to be on special in the same sample (Reed, 2003). In rare cases, when a stock is in high demand, the rebate rate can be significantly negative. For example, shares of Stratos Lightwave, Inc. had a rebate rate more than 4,000 basis points below the general collateral rate in late August 2000, just after the firm's initial public offering (IPO) (see Mitchell, Pulvino, and Stafford, 2002). In these cases, the lender is keeping the full investment rate of return on the collateral and also earning a premium for lending the securities.

Although specials are identified by their low rebate rates, the difficulty of borrowing specials goes beyond the increase in borrowing costs. Only well-placed investors (e.g., hedge funds) will be able to borrow specials and receive the reduced rebate. Generally, brokers will not borrow special shares on behalf of small investors; the order to short sell will be denied. Loans in stock specials will be expensive for well-placed investors and impossible to obtain for retail investors.

Specials tend to be driven by episodic corporate events that increase the demand for stock loans or reduce the supply of stocks available for loan. For example, initial public offerings, dividend reinvestment discount programs, and dividend payments of foreign companies often lead to an increase in borrowing demand and/or a reduction in the supply of available shares. In the case of IPOs, even though shares are available in the first settlement days, they are generally on special. At issuance, the average IPO's rebate rate is 300 basis points below the general collateral rate, but this spread from the general collateral rate falls to 150 basis points within the first 25 trading days. Similarly, the short selling of merger acquirers' stock drives *specialness*. Loans of merger acquirers' stock have average rebate rates 23 basis points below general collateral rates according to Geczy, Musto, and Reed (2002). Additionally, because brokers prohibit their clients from buying stocks with prices below $5 on margin, there can be a limited supply of stock available for loan from broker dealers for these low-price shares. (Broker dealers usually have the right to loan out any stock held in individual investors' margin accounts. However, shares that are paid in full cash rather than in margin accounts are generally not available to borrow from a broker dealer without consent of the owner.) Some factors that can improve liquidity in a stock and therefore improve its rebate rate include a secondary issue of the security, an expiration of an IPO lock-up period, and the reduction in short-selling demand as a result of the completion of a merger or corporate action.

SUMMARY

As investors continue to become more sophisticated and new arbitrage opportunities develop, the securities lending markets will continue to expand and see new entrants. Beneficial owners have been increasing their participation in the lending markets, and they view the market as a low risk way to achieve increased return on their assets. Broker-dealers eager to attract the very profitable client base of hedge funds and other leveraged investors continue to expand their securities lending infrastructures. As a result, the securities lending markets have seen tremendous growth over the last decade. New entrants on both the lending and borrowing side combined with new technologies improving the transparency in the lending markets continue to increase the importance of this market.

REFERENCES

D'Avolio, G. (2002). The market for borrowing stock. *Journal of Financial Economics* 66 (November): 271–306.

Duffie, D., Garleanu, N., and Pedersen, L. (2002). Securities lending, shorting, and pricing. *Journal of Financial Economics* 66 (November): 307–339.

Bargerhuff & Associates (2000). *Securities Lending Analytics: 2nd Quarter*.

Christoffersen, S., Geczy, C., Musto, D., and Reed, A. (2005). Crossborder dividend taxation and the preferences of taxable and nontaxable investors: Evidence from Canada. *Journal of Financial Economics* 78: 121–144.

Christoffersen, S., Geczy, C., Musto, D., and Reed, A. (2007). Vote trading and information aggregation. *Journal of Finance*, forthcoming.

Evans, R., Geczy, C., Musto, D., and Reed, A. (2005). Failure is an option: Impediments to short-selling and options prices. *Review of Financial Studies*, forthcoming.

Fabozzi, F. J. and Mann, S. V. (eds.). (2005). *Securities Finance: Securities Lending and Repurchase Agreements*. Hoboken, NJ: John Wiley & Sons.

Geczy, C., Musto, D., and Reed, A. (2002). Stocks are special too: An analysis of the equity lending market. *Journal of Financial Economics* 66 (November): 241–269.

Jones, C., and Lamont, O. (2002). Short sale constraints and stock returns. *Journal of Financial Economics* 66 (November): 207–239.

Lamont, O. (2004). Short sale constraints and overpricing. In F. J. Fabozzi (ed.), *Short Selling: Strategies, Risks, and Rewards* (pp. 179–204). Hoboken, NJ: John Wiley & Sons.

Mitchell, M., Pulvino, T., and Stafford, E. (2003). Limited arbitrage in equity markets. *Journal of Finance* 57, 2: 551–584.

Reed, A. (2003). Costly short selling and stock price adjustment to earnings announcements. University of North Carolina working paper (June).

CHAPTER 71

Securities Lending, Liquidity, and Capital Market-Based Finance

STATE STREET CORPORATION

Abstract: As the advantages that deep, liquid capital markets offer to national economies—most notably, enhanced capacity for economic growth—become more evident, policy makers in nations around the world are seeking ways to foster capital market growth. As capital markets evolve, they divide risk even more finely—by evolving new financial instruments such as options and futures (derivatives) and new investment vehicles and strategies such as mutual funds and hedged investments. The single most important quality that securities markets need to function successfully and to grow is liquidity—the ability to buy or sell substantial investment positions quickly, smoothly and with minimal market impact. One of the most important factors in fostering liquidity is the evolution of a broad array of securities lending functions. The ability to borrow securities is, in fact, a key element in the development of advanced capital markets. Wherever securities lending has not yet become accepted practice, the evolution of national or regional capital markets is stunted—limiting their ability to allocate capital more efficiently to economic development.

Keywords: bank, borrowing, borrowing of securities, capital market, counterparty risk, financial services, lending, lending of securities, lending securities, securities borrowing transactions, securities finance, securities lending, securities lending transactions

Capital markets play an indispensable role in economic development and *securities lending* enables these markets to work much better and to evolve. Many nations have been moving to remove legal and regulatory obstacles to securities lending and to encourage more participation in the practice as a way to spur the growth of their domestic capital markets.

The world's leading *central banks* themselves engage in the closely related practice of using repurchase transactions (repos) in government *securities markets* as an element in managing their monetary policies. These institutions have also been encouraging a wider array of private securities firms to participate in this market alongside them. The growing official consensus in favor of capital markets and the increased recognition by policymakers of securities lending's function as an important market lubricant will ensure that securities lending remains a central element in twenty-first century capital markets.

This chapter discusses the central role that securities lending, liquidity, and a strong capital-based market system play in creating a robust *economy.* Securities lending and capital markets evolve in tandem. The evolution of the securities lending market has been a significant component to increasing *market liquidity,* globally, and the ability to lend and borrow securities is an essential element in the development of advanced capital markets. Securities lending has the effect of increasing the total supply of assets, and nations are acknowledging its significance and encouraging the injection of new liquidity into their capital market systems through this practice.

THE GREAT TRANSITION: THE RISE OF *CAPITAL MARKET–BASED FINANCE*

The financial history of the world's most developed economies through the twentieth century centers on a single theme—the securitization of *finance*—as capital markets grow to supplement, even to displace, traditional banks as the prime intermediaries between borrowers and lenders of capita (see Chernow, 1997).

The last decades of the twentieth century saw capital markets in the most developed nations come to eclipse traditional, bank-dominated financial systems. Fueled by a multitrillion dollar-wave of pension and retirement savings, capital markets in the United States, the United Kingdom, and other leading economies have grown well past the scale of the total holdings of their national banking systems.

Capital markets in these nations have, in fact, replaced banks as the dominant source of corporate finance. In the United States, for example, this process of disintermediation is so far advanced that less than 30% of corporate finance now comes from traditional commercial banks. Some of the most dynamic growth areas in these nations' banking industries now center on the transformation of traditional bank products such as mortgages or credit card debt into securitized products that can be traded on the capital markets.

The forces driving the rise of capital markets remain strong. These range from the aging of the global population, the attendant multitrillion-dollar rise in retirement savings, the continuing triumph of capitalism itself, progress in the application of both raw computing power and quantitative strategies to investing, the explosive growth of *derivatives* and *hedge funds*, and the ubiquitous availability of information to guide and execute investment and trading strategies on a global basis.

A capital markets–based financial system can systematically provide seed capital to entirely new high-tech industries. It is almost inconceivable, for example, that the Internet, biotechnology, and other new economy industries now rising in the United States and elsewhere could have been financed so rapidly or on today's scale through traditional lending by an old economy banking-dominated financial system.

A growing number of economists and policy makers, backed up by day-to-day experience, now share a new consensus view robust capital markets, which offer a full array of modern financial products and practices, contribute to long-term national *economic growth* by encouraging entrepreneurship and innovation, even given periodic market corrections (see Levine and Zervos, 1999).

Capital markets can finance economic growth more efficiently than traditional bank lending systems that depend on making a spread of interest rate revenue over the banks' costs of funds. Capital markets can more easily diversify and distribute risk by dividing shares in the equity ownership or portions of the debt involved in financing enterprises into stocks and bonds, which in turn can be much more widely dispersed among investors than traditional bank loans.

The availability of active markets for shares in new enterprises then enables venture capitalists to make a range of investments in a variety of high-risk ventures—in the hope that one or more spectacularly successful initial public offerings (IPOs) will more than make up for other ventures' failures and losses. Traditional commercial banks, by contrast, cannot risk lending to an array of unproven start-ups—however promising—because banks can not earn enough additional interest on those new firms that succeed to make up for capital they are likely to lose when other, unproved borrowers fail.

In addition, as capital markets evolve further, they can split the atom of risk even more finely—by creating new financial instruments such as *options* and futures and new investment vehicles and strategies such as mutual funds, exchange-traded funds, and hedged investments. These provide investors with new ways to increase returns and manage risks, and to do so more cost effectively.

Given these dynamic, growth-fostering advantages, it is no surprise that both developed and emerging nations are actively seeking to follow the same process of financial evolution so evident in the United States, the United Kingdom, and other capital market leaders. The movement away from communist economic regimes in the 1980s and 1990s has spawned a huge expansion in the number of global *stock markets* that money managers and institutional investors have to consider—from fewer than 80 in the early 1980s to more than 160 by the turn of the twenty-first century.

Much of that growth has been concentrated among the world's most advanced securities markets—notably New York and London—where market capitalization on the leading exchanges multiplied 10-fold in the 1990s. As the strategic growth advantage that developed capital markets offer to national economies become more evident, policy makers in many nations are coming to view capital *market development* as imperative to their nations' futures, to their ability to finance new high-tech industries and to their competitiveness in a globalizing economy.

Over and above their growth advantage, the development of deep, liquid capital markets also offers nations the benefit of greater financial system stability. As former Federal Reserve Board chairman Alan Greenspan noted, the existence of strong capital markets alongside well-regulated banking systems may help insulate a nation's whole financial system from systemic risk by providing alternate sources of liquidity and financing that can be tapped when either banking systems or securities markets are in short-term crisis (Greenspan, 1999).

This is not to suggest that capital markets represent some magic elixir for economic growth, or that traditional banks are moribund. Even when accompanied by well-developed rule of law, advanced accounting standards, and free flow of information, capital markets can, at times, overinflate or depress underlying economic value, creating bubbles and panics. As the Fed chairman noted, the central banks' ability to inject liquidity into the financial system through banks was essential in containing the financial contagion that had frozen many securities markets in the wake of the 1997–1998 Asian financial crisis.

The key point is that having both advanced capital markets and strong banking systems gives nations both greater competition in the provision of capital (and the possibility of turning to complementary financing systems) and eliminates the need to simply rely on one or the other. The securitization of the U.S.-based mortgage industry, for example, helped keep housing finance flowing, which limited the depth of the 1990–1991 U.S. recession precisely because banks could repackage and sell their mortgage loans into capital markets.

FINANCIAL COMPLEXITY AND INTENSIFICATION

Besides their sheer scale, the world's most developed capital markets have become vastly more complex and transnational in scope. Companies doing business in and serving these capital markets—both buy-side institutional investors and sell-side brokerage firms—have expanded their horizons from national to global markets as they seek to manage the largest pools of long-term investment capital in history.

Investors have also changed their own investment and trading habits in a process that some analysts have dubbed financial *intensification*. This refers to both the vast proliferation of new financial instruments—mainly, options, futures, and other derivatives that investors use to manage and mitigate risk—and to the dramatic rises in trading volumes as investors engage these new instruments to conduct trading and investment strategies that often produce vastly higher turnover.

Taken together, the rise of cross-border investing and the proliferation of financial instruments that serve to arbitrage differences between national capital markets points to the emergence of a single truly global capital market which is subject to the law of one price as domestic price and regulatory differences erode (see Bryan and Farrell, 1996). Individual nations' markets, then, become nodes in this emerging global network, and their success depends on the extent that national policymakers make their markets attractive to domestic and foreign investors.

Clearly, what capital markets need above all to grow, to become liquid and to sustain increasing volumes of transactions is capital—preferably sustainable flows of long-term, patient investment. The prime source for funding the rise of late-twentieth-century capital markets has been domestic pension savings and the evolution of collective investment vehicles. It is no coincidence that the nations with the highest ratios of equity market capitalization to gross domestic product (GDP)—for example, the United States, the United Kingdom, and the Netherlands—also have the most well-developed systems of pension, collective fund, and personal retirement savings. While domestic pension savings have been a prime fuel for their growth, the most advanced capital markets also benefit from their openness to cross-border investing, which grew explosively in the 1990s.

For nations whose capital markets are less developed, one clear lesson is the removal of obstacles to foreign investment is itself a prerequisite for the development of effective capital markets (World Trade Organization, 1997). Improving regulatory *transparency* is also necessary to boost foreign investment. Transparency leads to business predictability for foreign entities that are expanding to new markets and taking the risk of dealing with many uncertainties. In turn, the ability of a given national or regional securities market to attract capital—whether from domestic savings and pension funds or from offshore investors—depends critically on the creation of efficient, well-regulated mechanisms for handling rising transaction flows, settling trades, and mitigating risk.

THE CENTRAL ROLE OF LIQUIDITY

The single most important quality that successful securities markets must foster is *liquidity*—the ability to buy or sell substantial investment positions quickly, smoothly and with minimal market impact. An analysis from the Counterparty Risk Management Policy Group (1999) states market liquidity is a precondition for the smooth pursuit of all financial activities, including the pricing of financial products, the *risk management* of financial institutions, and the conduct of monetary policy.

There is, of course, a notorious circularity in analyzing the root sources of liquidity, because it is, to a large degree, a self-fulfilling phenomenon. *Investor confidence* spurs a general willingness to trade, the participation of many transactors deepens markets and smooths trading and these qualities of a market further raise investors' confidence. Liquidity is, or can be, the function of such a virtuous circle.

Definitions of liquidity range beyond the ability to deploy capital into and out of a market in an efficient way—that is, without excessive transaction costs or impacts on securities prices. Microanalysis of a given market measures its liquidity in at least three dimensions (Bank for International Settlements, Committee on the Global Financial System, 1999):

1. **Tightness**—how far transaction prices diverge from mid-market prices—a metric generally visible in the size of bid-asked spreads.
2. **Depth**—how large a volume of trades can be processed without significantly affecting prevailing market prices or the amount of orders on market makers' books in a given time frame.
3. **Resiliency**—how quickly price fluctuations resulting from trade are dissipated and/or how quickly order flow imbalances are adjusted and price recovery occurs.

In a somewhat broader sense, liquidity includes the ability of market participants to make money by trading when a market is moving downward as well when that market is trending upwards. Liquidity also relies on efficient price information and settlement systems, low transactions pricing and spreads and low infrastructure and tax costs.

These overall features of a market's *efficiency,* all of which contribute liquidity to traders and investors in a market, are continually evolving. National laws and regulations, systems for trade settlements and record keeping, provisions for the security of investors' own data and that for greater transparency of financial information provided to the market can all enhance liquidity if they are well designed and implemented. Alternatively, regulatory restrictions on *short selling* or hedge funds and other, even more inhibiting measures—such as capital controls or transaction taxes—can discourage investors and erode liquidity.

Ultimately, liquidity is a function of investors' confidence that they have the ability to buy and sell their investments when they want in markets that may fluctuate but will not stall or fail. Clearly investor confidence—or its withdrawal—has a self-reinforcing impact on any market's liquidity. Fostering such confidence, then, has to be a central aim for national authorities intent on developing their capital markets. One way that governments and central banks foster liquidity directly is through implicit assurances that they will provide market participants with funds to keep orderly trading under way and mitigate trading freeze-ups or panicky sell-offs if market crises do occur.

Offering a specific asset class—such as long-term government bonds—with specific policy assurances that the government will keep the market liquid, can also be a useful way to ensure that even amid the evaporation of liquidity from some markets, at least some benchmark asset that the rest of the market relies on to price other risks and values will continue trading freely until confidence generally can be restored (Bank for International Settlements, Committee on the Global Financial System, 1999).

Another way to encourage liquidity is for national regulatory authorities to allow and encourage more market participants to engage in lending and *borrowing securities* already outstanding in the nations' equity and bond markets. (Such permission, even encouragement, is already common in most markets for government bonds, because most central banks are themselves major players in these markets.) Regulators can further assist by understanding and encouraging the use of *swaps,* options, and other derivatives, which encourage liquidity by enabling traders and investors to mitigate their risks.

Evolution of the U.S. Securities Lending Market

The development of a broad array of securities lending activities can provide a very significant source of liquidity to any well-developed capital market. A brief review of how securities lending has evolved in U.S. markets—the world's deepest and most liquid securities investment and trading arena—can help illustrate the critical role that securities lending practices play in providing liquidity to increasingly vast capital markets which are executing increasingly complex trading strategies.

Historically, the earliest evidence of securities lending in the United States can be traced back to the market for U.S. government war debt following the Declaration of Independence in 1776. But a considerably more robust market for private securities lending in both the American and British stock and bond markets developed throughout the 1800s.

From those centuries-old origins well into the mid-twentieth century, the lending and borrowing of securities evolved as a private, ad hoc practice usually transacted directly between investors or broker-dealers. It was not until the 1960s, in the United States, that securities lending began to develop as a substantial day-to-day market of its own served by specialized institutions and practitioners.

The most important factor driving the emergence of the modern securities lending industry was the revival of interest in stock market investing brought about in the 1960s by the U.S. economy's booming growth. Many of the leading firms on Wall Street not only notched record profits, but also drew a level of individual and institutional investment not seen since before the crash of 1929 and the subsequent Great Depression.

As rapid economic growth fueled a booming equity market on Wall Street, first individuals and, increasingly, pension funds, rushed to invest. Many corporations also took advantage of rising share prices to issue equity-related hybrid securities convertible into common stock. Other companies used their rising stock as currency to take part in a wave of corporate takeovers and restructurings.

Both of these developments opened new opportunities for professional traders to arbitrage between common stock and hybrids—or between the stocks of acquiring or target firms engaged in takeover battles. The bull market also revived interest in American Depositary Receipts (ADRs), an instrument developed in the late 1920s to represent foreign shares traded in markets in other countries. Not least, as stock prices soared, more bearish speculators sold shares short in hopes that prices would later decline.

By the early 1970s, both stock exchanges and securities firms were struggling to cope with the huge upsurge of trading brought on by these overlapping waves of change. The result was a series of major back-office snarls, some severe enough to lead to the collapse of major Wall Street trading firms and an explosion in settlement failures. These symptoms of operational dysfunction—and classic market illiquidity—were eased in the course of the 1970s by two developments. First, the trade settlement process was increasingly automated and the back-office paper jams eased. Second, a true securities lending industry began to emerge, which was able to reduce trade fails substantially by providing borrowed assets to arbitrageurs, short sellers, and other traders who needed securities that they did not own to conduct their investment strategies.

The growth of institutionalized securities lending was a timely development for U.S. markets since it paralleled a further surge in the demand side of the securities lending equation. This was brought about by the boom in option trading and other derivatives in the mid-1970s set off by the application to capital markets of the Black-Scholes option-pricing model. This analytical tool provided traders with a more reliable formula for gauging the value of put and call options on stocks. With a reliable metric for measuring values in the options markets, volume exploded. Trading strategies based on options required the borrowing and lending of shares for hedging as well as for arbitrage. As the so-called derivatives revolution rolled on, the investment strategies born of the Black-Scholes model laid the groundwork for a fresh wave of financial innovation centered on new derivatives, index arbitrage and other complex investment and trading strategies throughout the 1980s and 1990s—all of which drove demand from dealers and investors for borrowed securities to execute their trades and hedge their market risks.

On the supply side, U.S. custodian banks moved to meet demand for borrowed securities in the 1970s by devising lending services for such institutional clients as insurance companies, corporate investment portfolios and, later, college endowment funds. Legislation soon permitted pension funds to join the quest for enhanced returns by engaging in securities lending. By the mid-1980s, the majority of institutional investors in the United States were using securities lending routinely as a way to earn extra income to offset custodial fees—and securities lending in the United States had itself become a thoroughly institutionalized industry.

The key lesson of this U.S. experience is simple: Securities lending and capital markets evolve in tandem.

Securities Lending: Key to Market Liquidity

In mobilizing the securities already outstanding in a market, securities lending has the effect of increasing the total supply of assets available to support trading and settlement. This enables the outstanding stock of assets, in effect, to do double duty in the service of market liquidity by converting otherwise sterile holdings into a dynamic, internally generated source of finance that can support higher trading volumes and more sophisticated trading strategies.

By turning existing stocks and bonds into financing sources for further transactions, a well-developed securities lending business can minimize trading friction, improve efficiency, reduce settlement failures, and lower transaction costs across an entire capital market. The benefits are multiple. *Risk mitigation* is made easier by the options that securities lending provides to investors wanting to balance long positions with offsetting short positions. Indeed, all market participants benefit—not just those who engage in securities lending or borrowing.

The development of a sophisticated securities lending industry has, in fact, played a central role in enhancing the liquidity of those markets that have managed to leap to maturity. Indeed, market maturity may best be defined as the level of liquidity that can attract significant investment from large global investors.

In country after country through the 1980s and 1990s, new or revitalized capital markets began their economic take-offs by first attracting increased attention from domestic investors and from the most venturesome of foreign investors. Almost by definition, it is this first wave of inward investment that makes an emerging market actually emerge. To continue growing, a capital market needs to draw investment that is more stable and longer term—from larger investors who are typically much more risk averse than the pioneers.

This has required capital marketplaces around the world to improve and automate their settlement processes, to establish central securities depositories (where they did not yet exist) and to decertify securities ownership and unclog paper flows.

As these changes take hold and investment in a given market rises, further pressures build—for better data, for greater transparency and for the creation of derivatives or short-selling practices that, increasingly, larger investors need to hedge their investment risks. The demand for means to hedge exposures is particularly acute among global pension funds.

Bound by fiduciary standards of prudence, many institutional investors are virtually obliged to use derivatives, repos, and other instruments to manage their investment exposures. The rise of markets in derivatives instruments, in turn, depends on the ability of players in the real or underlying securities markets to engage in substantial short selling and securities lending, and so to sustain liquidity amid rising transaction volumes.

In markets where securities lending is underdeveloped—or explicitly discouraged by regulatory or cultural barriers—evolution to a world-class level is simply stunted, at least until these barriers are removed.

As this market development pattern has replayed time after time, more and more governments, multilateral agencies like the *World Bank,* and economists have come to acknowledge the catalytic role of capital markets in economic development. Institutions like the Bank for International Settlements are also now acknowledging the role of securities lending in helping securities markets to function well.

A new consensus is emerging according to the Technical Committee of the International Organization of Securities Commissions (1999): the ability to borrow securities is an indispensable element in the development of advanced, effective capital markets. Indeed, the greater the turnover in a market, the more important securities lending becomes. Securities lending, in short, is no longer an ad hoc, back-office operation that enables borrowers to trade on securities they currently do not own. Nor is securities lending merely a low-risk way for institutional investor lenders to earn a few more basis points or cut their custody fees on their holdings. Securities lending as an industry has matured to become a major source of internal financing that any capital market needs to achieve a world-class, twenty-first-century practice.

It is little wonder that a report by the International Organization of Securities Commissions and the Bank for International Settlements (1999, p. 2) concluded:

> Securities lending markets are a vital component of domestic and international finance markets, providing liquidity and greater flexibility to securities, cash and derivatives markets. . . . Securities lending activity will continue to increase and become an even more integral component of financial markets in the future.

Sophisticated regulators and policy makers in many nations now recognize securities lending provides the liquidity that lubricates their capital market engines. As a 1998 report by the Bank for International Settlements, Committee on the Global Financial System (1998) notes, investors are more willing to transact and take positions in markets where they expect liquidity to continue at a high level for the foreseeable future . . . and market liquidity tends to be enhanced when instruments can be substituted for one another, since the market for each of them will be less fragmented.

This growing recognition by governments and regulators of the value of securities lending should not be surprising. It stems, in large part, from central banks' and monetary authorities' own reliance on the closely related practice of using *repurchase agreements (repos)* in their government debt markets as a key element in monetary policy—a development we will turn to shortly. However, we will first explain how securities lending finances liquidity.

How Securities Lending Finances Liquidity

To understand how securities lending concretely contributes to market liquidity, consider the structure of a specific equity lending transaction in its simplest form. In basic equity lending, a *counterparty* borrows stocks against a collateral obligation. The borrowed shares are cycled back into the trading market and the collateral (if cash) is used to purchase additional instruments, generally short-term money market or other fixed income instruments. Both components of the transaction—the lent securities and the reinvested collateral—inject additional securities or cash into capital markets, enhancing liquidity both directly and indirectly.

The increased supply of assets that lending makes available to support transactions in a given market facilitates that market's efficiency in the pricing and settlement of transactions, which helps the market's trading flow move more smoothly and with less market impact. This is virtually a dictionary definition of what liquidity means.

However, securities lending also enhances liquidity indirectly. The smoother transaction flows that lending facilitates contribute to investors' confidence that they can trade with less risk of fails or market freeze-ups. This holds true not only for the simple example cited above, but for the whole array of complex trading strategies that have evolved over recent decades, all of which depend on a robust securities lending market for their execution.

As a market grows in value and trading volume, market participants create new instruments and trading strategies that increase demand for borrowed securities. Securities lending thus evolves from a settlement and back-office function to the supplying of securities to cover short positions to the supplying of lent securities to support global trading strategies.

By reintroducing shares, bonds, or other financial instruments into the market on a cost-efficient and timely basis, securities lending enables market participants to use these assets in ways that rebalance prices, diversify risk, minimize trade and settlement failures, and allow positions to be exchanged even when parties to a trade do not own the securities being traded.

Here is a further example from the world of arbitrage, one of the heavy generators of demand for securities lending in today's marketplace. Arbitrage trading, the object of which is to capture differences in prices for the same security or its equivalent in different markets, generates continual demand for *securities borrowing* as arbitrageurs seek to exploit often minimal and transitory price differences between securities they may not own. The arbitrageur's profit is often minuscule. But he repeats this strategy all day long, whenever the price spread gets out of line on either the high or low side. That makes him an omnipresent rebalancer of prices—and an incessant contributor to liquidity on both sides of the market. Although arbitrageurs seek profit from inefficient pricing, it is their trading, often supported by borrowed securities, that keeps bringing prices back in line and makes overall markets more efficient.

In ADR arbitrage, for instance, the arbitrageur trades back and forth between a depositary receipt traded in the United States and the actual shares traded in, say, Frankfurt, capturing price discrepancies as they arise. The arbitrageur borrows securities as needed to execute her trade—and in the process deepens trading volume and pushes the prices back in line. Similarly, index arbitrage keeps pricing in line between a basket of shares and an index futures contract.

The more complex strategy of risk arbitrage in corporate merger and acquisition deals also rebalances and adds liquidity to securities markets. When one company offers its shares to buy another company, the arbitrage strategy is often to purchase the target company's shares, borrow shares of the acquiring company, and sell them short to capture the premium (often 20% or more) the acquirer is offering as an incentive to the target company's shareholders. When, or if, the deal goes through, the arbitrageur can capture the premium by delivering her shares in the target company in exchange for the acquiring company's shares, which she then returns to the securities lender.

Target companies sometime object to risk arbitrage activity on grounds that a large proportion of its shares in arbitrageurs' hands will swing a shareholder vote in favor of the deal. Risk arbitrage, however—and the securities lending that makes it possible—benefits the market by absorbing a large portion of the acquisition risks, bringing pricing in line with those risks and adding trading liquidity that permits shareholders in the target company to sell their shares and capture a portion of the premium before the deal goes through.

Recognizing Securities Lending's Key Roles

The decade of the 1990s was bracketed by two major policy reports that resoundingly endorsed the role of securities lending in capital market development—and urged nations to do more to encourage it. The first was a 1989 report by the *Group of 30 (G30)* on clearing and settlement systems. One of the report's recommendations urged governments and regulators to facilitate securities lending in order to reduce the high rates of trading fails that were discouraging cross-border investors and rendering domestic capital markets illiquid and prone to paralysis.

The G30s call to take down regulatory and taxation barriers that inhibit securities lending has received increasingly positive response through the 1990s (International Organization of Securities Commissions, Bank for International Settlements, July, 1999). Japan, Australia, the United Kingdom, Switzerland, Italy, France, and other nations acted to remove legal and regulatory obstacles to securities lending and to encourage more participation in lending, swaps and securities sell–buy-back agreements by both domestic and foreign entities.

At the same time, in the form of repurchase agreements, securities lending has become a vital tool of modern monetary policy though the activity of central banks themselves in government securities markets. Leading central banks all now use an active repo trading strategy to add liquidity to their sovereign debt markets, to stabilize their currencies, and to attract foreign investment.

Amid the explosive growth of global capital markets—and the increasing use of securities lending and hedging techniques by central banks and governments—the decade of the 1990s closed with this conclusion from a joint study by the International Organization of Securities Commissions and Bank for International Settlements (July 1999, p. 5):

> Securities lending has become a central part of securities market activity in recent years, to the point where the daily volume of securities transactions for financing purposes considerably exceeds that of outright purchase and sale transactions.

Repo and Securities Lending

Securities lending and the market in repos have similar characteristics but with different legal structures. They both follow the same transaction structure whereby a security is transferred versus a collateral obligation. Repo transactions are outright sales of a security accompanied by an agreement to buy the security back at a specified price on a specified date—sometimes as soon as the next day. Thus, they can be used as either a securities borrowing or cash borrowing vehicle. In effect, the repo seller lends the security against cash collateral, while the repo buyer lends cash against the security as collateral.

Like a securities loan, repo may have the effect of bringing divergent prices back into line, of lowering the cost of financing and trading strategies and of splitting the atom of risk.

By the 1990s, the *repo market* was quite sophisticated (International Organization of Securities Commissions and Bank for International Settlements, July 1999, p. 10):

> In the U.S. Treasury repo market, brokers began to run matched book portfolios to provide liquidity to their customers and to use the repo market to take positions on the short end of the yield curve. For example, a broker might lend securities on repo for one month and finance them for one week, in the expectation that repo financing rates would fall. Thus repo grew beyond a straightforward financing market to become a money market instrument in its own right, as an alternative to interbank deposit and treasury bill/certification of deposit markets.

Perhaps most significantly, repo has evolved to be an important tool in managing monetary policy for a number of central banks around the world. As noted by the Bank for International Settlements, Committee on the Global Financial System in its report on Implications of the Repo Market for Central Banks (March 1999, p. 11):

> For the central banks that use them, repos have often become the most important monetary policy instrument. In a number of G-10 central banks, the proportion of repos used in the refinancing of domestic financial sector is over 70%.

Repo and securities lending are related transactions with related functions. They are linked by their similarity in providing a supply of securities, increasing trading volumes, diversifying risk and helping to keep *financial markets* running smoothly. These very similar practices are, in fact, linked across markets. As the Bank for International Settlements, Committee on the Global Financial System (March 1999, p. 8) notes:

> In some instances, the supply of securities in repo markets can be increased by stock-lending agreements . . . (such agreements) allow institutions that hold securities but do not want to (or are not allowed to) participate in the repo markets to earn a higher return. . . . Since repo markets support securities markets, securities issuers sometimes take steps to promote them.

In addition to the increased liquidity that loans of securities inject into a capital market by directly facilitating various trading strategies, the collateral that is posted against borrowed securities also benefits the markets.

When cash is pledged as collateral, the general practice is to reinvest it in short-term, money-market instruments because securities lenders have to price, purchase, sell, and settle on a daily basis, and holding any illiquid instrument in a short-term fund would be excessively risky.

The need to invest such collateral, in turn, generates substantial, continuing demand from securities lenders for reliable money-market investments—adding breadth and depth to markets for supranational, corporate, and securitized short-term debt. Where noncash collateral is accepted, lenders will generally approve only issues that can readily be priced, traded, and liquidated for a cash position in order to protect securities loans.

THE EMERGING OFFICIAL CONSENSUS: FOSTERING CAPITAL MARKETS AND SECURITIES LENDING

As governments, multinational agencies and scholars recognize both the catalytic role of capital markets in economic development and the ways that securities lending keeps markets liquid, a growing number of nations are removing legal and regulatory obstacles to securities lending. Some are actively encouraging more participation in the practice by both domestic and foreign entities.

Nations are continuing to recognize the merits of securities lending and encourage the practice through reforms. Official support is particularly notable in the closely related arena of repo transactions in government securities markets—which have become central to the operations of the largest and most powerful monetary authorities in the world.

SUMMARY

If the 1990s saw the rise of capital markets as the prime vehicles for financing the most dynamic economies in the world, the first decade of the twenty-first century will see these markets truly come of age. Growing awareness of the powerful competitive advantages that well-developed capital markets bring to national economies will spur their further development worldwide. The continued global movement towards pension and savings reforms will provide trillions of dollars in mass-based investment capital to help world securities markets grow.

Nations that want to harness these vast, stable flows of long-term funds to spur their capital markets will, in turn, need to open themselves to the full array of legal, regulatory and transaction mechanisms that make securities markets work. Derivatives, hedging, short selling, and securities lending are among the key elements that any market will need to make available to attract investors and grow.

The increased official recognition by policy makers and central banks that they need to stimulate securities lending in general and repo markets in particular promise to make securities lending a central element in the growth of twenty-first century capital markets. This implies that as astonishing as the rise of securities lending has been over the past 20 years, the industry's best days are yet to come.

REFERENCES

Bank for International Settlements, Committee on the Global Financial System (1999a). *Market Liquidity: Research Findings and Selected Policy Implications*, May.

Bank for International Settlements, Committee on the Global Financial System (1999b). *A Review of Financial Market Events in Autumn 1998* October.

Bank for International Settlements, Committee on the Global Financial System (1999c). *Implications of Repo Markets for Central Banks*, March.

Bryan, L., and Farrell, D. (1996). *Market Unbound: Unleashing Global Capitalism*. New York: John Wiley & Sons.

Chernow, R. (1997). *The Death of the Banker: The Decline and Fall of the Great Financial Dynasties and the Triumph of the Small Investor*. New York: Vintage Books, Random House.

Counterparty Risk Management Policy Group. (1999). *Improving Counterparty Risk Management Practices*, June.

Fabozzi, F. J., and Mann, S. V. (eds.) (2005). *Securities Finance: Securities Lending and Repurchase Agreements*. Hoboken, NJ: John Wiley & Sons.

Greenspan, A. (1999). Remarks before the World Bank group and the International Monetary Fund, program of seminars, Washington, D.C., September 27.

Group of 30. (1989). Clearance and settlement in the world's securities markets, March 1989.

Levine, R., and Zervos, S. (1999). *Stock Markets, Banks, and Economic Growth*. IFC, World Bank.

Technical Committee of the International Organization of Securities Commissions. (1999). *Securities Lending Transactions: Market Development and Implications*, Bank for International Settlements, Committee on Payment and Settlement Systems, July.

World Trade Organization. (1997). Opening Markets in Financial Services and the Role of the GATS, September 22.

CHAPTER 72

Repurchase Agreements and Dollar Rolls

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

STEVEN V. MANN, PhD
Professor of Finance, Moore School of Business, University of South Carolina

Abstract: Access to short-term financing and the ability to borrow securities to cover short positions are essential elements to a liquid, well-functioning bond market. Repurchase and reverse repurchase agreements are mechanisms used by dealers to accomplish these needs. Repurchase agreements occupy a central position in the money market and provide a relatively safe investment opportunity for short-term investors. Structured repurchase agreements introduce variations on the basic agreement and are designed to accommodate specialized clienteles of users. Similarly, dollar rolls developed in the mortgage-backed securities market because of the need to borrow these more complicated securities to cover short positions.

Keywords: repurchase agreement, reverse repurchase agreement, repo rate, overnight repo, repo term, repo margin, collateral, delivered out, hold-in-custody (HIC) repo, triparty repo, general collateral, special collateral, buy/sell-back agreement, structured repurchase agreements, dollar roll, pass-through, TBA (to be announced), collateralized mortgage obligation (CMO), forward drop, prepayment speed

One of the largest segments of the money markets worldwide is the market in repurchase agreements or repos. A most efficient mechanism by which to finance bond positions, repo transactions enable market makers to take long and short positions in a flexible manner, buying and selling according to customer demand on a relatively small capital base. In addition, repos are used extensively to facilitate hedging and speculation. Repo is also a flexible and relatively safe investment opportunity for short-term investors. The ability to execute repo is particularly important to firms in less developed countries who might not have access to a deposit base. Moreover, in countries where no repo market exists, funding is in the form of unsecured lines of credit from the banking system which is restrictive for some market participants. A liquid repo market is often cited as a key ingredient of a liquid bond market. In the United States, repo is a well-established money market instrument and is developing in a similar way in Europe and Asia.

A major sector of the bond market in the United States is the mortgage-backed securities (MBS) market. A specialized type of collateralized loan is used in the mortgage-backed securities (MBS) market because of the characteristics of these securities and the need of dealers to borrow these securities to cover short positions. This arrangement is called a *dollar roll* and can be thought of as a specialized form of a reverse repurchase agreement with pass-through securities serving as collateral. A dollar roll is so named because dealers are said to "roll in" securities they borrow and "roll out" securities when returning the securities to the investor.

In this chapter, we discuss repurchase agreements and dollar roll agreements.

REPURCHASE AGREEMENTS

A *repurchase agreement* or "repo" is the sale of a security with a commitment by the seller to buy the same security back from the purchaser at a specified price at a designated future date. For example, a dealer who owns a 10-year U.S. Treasury note might agree to sell this security (the "seller") to a mutual fund (the "buyer") for cash today while simultaneously agreeing to buy the same 10-year note back at a certain date in the future (or in some cases on demand) for a predetermined price. The price at which the seller must subsequently repurchase the security is called the repurchase price and the date that the security must be repurchased is called the repurchase date. Simply put, a repurchase agreement is a collateralized loan where the collateral is the security that is sold and subsequently repurchased. One party (the "seller") is borrowing money and providing collateral for the loan; the other party (the "buyer") is lending money and accepting a security as collateral for the loan. To the borrower, the advantage of a repurchase agreement is that the short-term borrowing rate is lower than the cost of bank financing, as we will see shortly. To the lender, the repo market offers an attractive yield on a short-term secured transaction that is highly liquid. This latter aspect is the focus of this section.

The Basics

Suppose that on September 28, 2006, a government securities dealer purchases a 4.875% coupon on-the-run 10-year U.S. Treasury note that matures on August 15, 2016. The face amount of the position is $1 million, and the note's full price (that is, flat price plus the accrued interest) is $1,025,672.55. Further, suppose the dealer wants to hold the position until the next business day, which is Friday, September 29, 2006. Where does the dealer obtain the funds to finance this position?

Of course, the dealer can finance the position with its own funds or by borrowing from a bank. Typically, though, the dealer uses a repurchase agreement or "repo" to obtain financing. In the repo market, the dealer can use the purchased Treasury note as collateral for a loan. The term of the loan and the interest rate a dealer agrees to pay are specified. The interest rate is called the *repo rate.* When the term of a repo is one day, it is called an *overnight repo.* Conversely, a loan for more than one day is called a term repo. The transaction is referred to as a repurchase agreement because it calls for the security's sale and its repurchase at a future date. Both the sale price and the purchase price are specified in the agreement. The difference between the purchase (repurchase) price and the sale price is the loan's dollar interest cost.

Let us return to the dealer who needs to finance a long position in the Treasury note for one day. The settlement date is the day that the collateral must be delivered and the money lent to initiate the transaction, which, in our example, is September 28, 2006. Likewise, the termination date of the repo agreement is September 29. At this point, we need to address who might serve as the dealer's counterparty (that is, the lender of funds) in this transaction. Suppose that one of the dealer's customers has excess funds in the amount of $1,025,672.55 and is the amount loaned in the repo agreement. Accordingly, on September 28, 2006, the dealer would agree to deliver ("sell") $1,025,672.55 worth of 10-year U.S. Treasury notes to the customer and buy the same 10-year notes back for an amount determined by the repo rate the next business day on September 29, 2006. (We are assuming in this example that the borrower will provide collateral that is equal in value to the money that is loaned. In practice, lenders usually require borrowers to provide collateral in excess of the value of money that is loaned. We will illustrate how this is accomplished when we discuss repo margins.)

Suppose the repo rate is this transaction is 5.15%. Then, as will be explained below, the dealer would agree to deliver the 10-year Treasury notes for $1,025,819.28 the next day. The $146.73 between the "sale" price of $1,025,672.55 and the repurchase price of $1,025,819.28 is the dollar cost of the financing.

Repo Interest

The following formula is used to calculate the dollar interest on a repo transaction:

$$\text{dollar interest} = (\text{dollar principal}) \times (\text{repo rate}) \times (\text{repo term}/360)$$

In our illustration, using a repo rate of 5.15% and a repo term of one day, the dollar interest is $146.73 as shown below:

$$\$1{,}025{,}672.55 \times 0.0515 \times (1/360) = \$146.73$$

The advantage to the dealer of using the repo market for borrowing on a short-term basis is that the rate is lower than the cost of bank financing for reasons explained shortly. From the customer's perspective (that is, the lender), the repo market offers an attractive yield on a short-term secured transaction that is highly liquid.

Reverse Repo and Market Jargon

In the illustration presented above, the dealer is using the repo market to obtain financing for a long position. The repo market can correspondingly be used to borrow securities. Securities are routinely borrowed for a number of reasons including opening a short position, the need to deliver securities against the exercise of a derivative contract, and the need to cover a failed transaction in the securities settlement system. Many arbitrage strategies involve the borrowing of securities (e.g., convertible bond arbitrage).

Suppose a government dealer established a short position in the 30-year Treasury bond one week ago and must now cover the position—namely, deliver the securities. The dealer accomplishes this task by engaging in a *reverse repurchase agreement*. In a reverse repo, the dealer agrees to buy securities at a specified price with a commitment to sell them back at a later date for another specified price. (Of course, the dealer eventually would have to buy the 30-year bonds in the market in order to cover its short position.) In this case, the dealer is making a collateralized loan to its customer. The customer is lending securities and borrowing funds obtained from the collateralized loan to create leverage.

There is a great deal of Wall Street jargon surrounding repo transactions. In order to decipher the terminology, remember that one party is lending money and accepting a security as collateral for the loan; the other party is borrowing money and providing collateral to borrow the money. By convention, whether the transaction is called a repo or a reverse repo is determined by viewing the transaction from the dealer's perspective. If the dealer is borrowing money from a customer and providing securities as collateral, the transaction is called a repo. If the dealer is borrowing securities (which serve as collateral) and lends money to a customer, the transaction is called a reverse repo.

When someone lends securities in order to receive cash (that is, borrow money), that party is said to be "reversing out" securities. Correspondingly, a party that lends money with the security as collateral for the loan is said to be "reversing in" securities.

The expressions "to repo securities" and "to do repo" are also commonly used. The former means that someone is going to finance securities using the securities as collateral; the latter means that the party is going to invest in a repo as a money market instrument.

Lastly, the expressions "selling collateral" and "buying collateral" are used to describe a party financing a security with a repo on the one hand, and lending on the basis of collateral on the other.

Types of Collateral

While in our illustration, we use a Treasury security as collateral, the collateral in a repo is not limited to government securities. Money market instruments, federal agency securities, and mortgage-backed securities are also used. In some specialized markets, even whole loans are used as collateral.

Credit Risks

Just as in any borrowing/lending agreement, both parties in a repo transaction are exposed to credit risk. This is true even though there may be high-quality collateral underlying the repo transaction. Let us examine under which circumstances each counterparty is exposed to credit risk.

Suppose the dealer (that is, the borrower) defaults such that the Treasuries are not repurchased on the repurchase date. The investor gains control over the collateral and retains any income owed to the borrower. The risk is that Treasury yields have risen subsequent to the repo transaction such that the market value of collateral is worth less than the unpaid repurchase price. Conversely, suppose the investor (that is, the lender) defaults such that the investor fails to deliver the Treasuries on the repurchase date. The risk is that Treasury yields have fallen over the agreement's life such that the dealer now holds an amount of dollars worth less then the market value of collateral. In this instance, the investor is liable for any excess of the price paid by the dealer for replacement securities over the repurchase price.

Repo Margin

While both parties are exposed to credit risk in a repo transaction, the lender of funds is usually in the more vulnerable position. Accordingly, the repo is structured to reduce the lender's credit risk. Specifically, the amount lent should be less than the market value of the security used as collateral, thereby providing the lender some cushion should the collateral's market value decline. The amount by which the market value of the security used as collateral exceeds the value of the loan is called *repo margin* or "haircut." Repo margins vary from transaction to transaction and are negotiated between the counterparties based on factors such as the following: term of the repo agreement, quality of the collateral, creditworthiness of the counterparties, and the availability of the collateral. Minimum repo margins are set differently across firms and are based on models and/or guidelines created by their credit departments. Repo margin is generally between 1% and 3%. For borrowers of lower creditworthiness and/or when less liquid securities are used as collateral, the repo margin can be 10% or more.

To illustrate the role of the haircut in a repurchase agreement, let us once again return to the government securities dealer who purchases a 4.875% coupon, 10-year Treasury note and needs financing overnight. The face amount of the position is $1 million and the note's full price is $1,025,672.55.

When a haircut is included, the amount the counterparty is willing to lend is reduced by a given percentage of the security's market value. Suppose the collateral is 102% of the amount being lent. To determine the amount being lent, we divide the Treasury note's full price of $1,025,672.55 by 1.02 to obtain $1,005,561.33, which is the amount the counterparty will lend. Suppose the repo rate is 5.15%. As a result, the transaction is structured as follows. The dealer would agree to deliver to the 10-year Treasury notes for $1,005,561.33 and repurchase the same securities for $1,005,705.18 the next day. The $143.85 difference between the "sale" price of $1,005,561.33 and the repurchase price of $1,005,705.18 is the dollar interest on the financing. Using a repo rate of 5.15 and a repo term of one day, the dollar interest is calculated as shown below:

$$\$1,005,561.33 \times 0.0515 \times (1/360) = \$143.85$$

Marking the Collateral to Market

Another practice to limit credit risk is to mark the collateral to market on a regular basis. Marking a position to market means simply recording the position's value at its market value. When the market value changes by a certain percentage, the repo position is adjusted accordingly. For complex securities that do not trade frequently, there is considerable difficulty in obtaining a price at which to mark a position to market.

Delivery of the Collateral

One concern in structuring a repurchase agreement is delivery of the collateral to the lender. The most obvious procedure is for the borrower to actually deliver the collateral to the lender or to the cash lender's clearing agent. If this procedure is followed, the collateral is said to be *delivered out*. At the end of the *repo term*, the lender returns collateral to the borrower in exchange for the repurchase price (that is, the amount borrowed plus interest).

The drawback of this procedure is that it may be too expensive, particularly for short-term repos (e.g., overnight) owing to the costs associated with delivering the collateral. Indeed, the cost of delivery is factored into the repo rate of the transaction in that if delivery is required this translates into a lower repo rate paid by the borrower. If delivery of collateral is not required, an otherwise higher repo rate is paid. The risk to the lender of not taking actual possession of the collateral is that the borrower may sell the security or use the same security as collateral for a repo with another counterparty.

As an alternative to delivering out the collateral, the lender may agree to allow the borrower to hold the security in a segregated customer account. The lender still must bear the risk that the borrower may use the collateral fraudulently by offering it as collateral for another repo transaction. If the borrower of the cash does not deliver out the collateral, but instead holds it, then the transaction is called a *hold-in-custody (HIC) repo*. Despite the credit risk associated with a HIC repo, it is used in some transactions when the collateral is difficult to deliver (e.g., whole loans) or the transaction amount is relatively small and the lender of funds is comfortable with the borrower's reputation.

Investors participating in a HIC repo must ensure: (1) they transact only with dealers of good credit quality since an HIC repo may be perceived as an unsecured transaction and (2) the investor (that is, the lender of cash) receives a higher rate in order to compensate them for the higher credit risk involved. In the U.S. market, there have been cases where dealer firms that went into bankruptcy and defaulted on loans were found to have pledged the same collateral for multiple HIC transactions.

Another method for handling the collateral is for the borrower to deliver the collateral to the lender's custodial account at the borrower's clearing bank. The custodian then has possession of the collateral that it holds on the lender's behalf. This method reduces the cost of delivery because it is merely a transfer within the borrower's clearing bank. If, for example, a dealer enters into an overnight repo with Customer A, the next day the collateral is transferred back to the dealer. The dealer can then enter into a repo with Customer B for, say, five days without having to redeliver the collateral. The clearing bank simply establishes a custodian account for Customer B and holds the collateral in that account. In this type of repo transaction, the clearing bank is an agent to both parties. This specialized type of repo arrangement is called a *triparty repo*. For some regulated financial institutions (e.g., federally chartered credit unions), this is the only type of repo arrangement permitted.

Determinants of the Repo Rate

Just as there is no single interest rate, there is not one repo rate. The repo rate varies from transaction to transaction depending on a number of factors: quality of the collateral, term of the repo, delivery requirement, availability of the collateral, and the prevailing federal funds rate.

Table 72.1 presents repo and reverse repo rates for maturities of one day, one week, two weeks, three weeks, one month, two months, and three months using U.S. Treasuries as collateral. These data are obtained from Bloomberg. Repo and reverse repo rates differ by maturity and type of collateral. Another pattern evident in these data is that repo rates are lower than the reverse repo rates when matched by collateral type and maturity. These repo (reverse repo) rates can viewed as the rates the dealer will borrow (lend) funds. Alternatively, repo (reverse repo) rates are prices at which dealers are willing to buy (sell) collateral. While a dealer firm primarily uses the repo market as a vehicle for financing its inventory and covering short positions, it will also use the repo market to run a "matched book." A dealer runs a matched book by simultaneously entering into a repo and a reverse repo for the same collateral with the same maturity. The dealer does so to capture the spread at which it enters into a repurchase agreement (that is, an agreement to borrow funds) and a reverse repurchase agreement (that is, an agreement to lend funds).

For example, suppose that a dealer enters into a term repo for one month with a money market mutual fund and a reverse repo with a corporate credit union for one month for which the collateral is identical. In this arrangement,

Table 72.1 Repo and Reverse Repo Rates

Maturity	Repo (%)	Reverse (%)
1 day	5.10	5.15
1 week	5.06	5.11
2 weeks	5.10	5.15
3 weeks	5.10	5.15
1 month	5.11	5.16
2 months	5.14	5.19
3 months	5.15	5.20

the dealer is borrowing funds from the money market mutual fund and lending funds to the corporate credit union.

From Table 72.1, we find that the repo rate for a one-month repurchase agreement is 5.11% and the repo rate for the one-month reverse repurchase agreement is 5.16%. If these two positions are established simultaneously, then the dealer is borrowing at 5.11% and lending at 5.16% thereby locking in a spread of 5 basis points.

However, in practice, traders deliberately mismatch their books to take advantage of their expectations about the shape and level of the short-dated yield curve. The term matched book is therefore something of a misnomer in that most matched books are deliberately mismatched for this reason. Traders engage in positions to take advantage of (1) short-term interest rate movements and (2) anticipated demand and supply in the underlying bond.

The delivery requirement for collateral also affects the level of the repo rate. If delivery of the collateral to the lender is required, the repo rate will be lower. Conversely, if the collateral can be deposited with the bank of the borrower, a higher repo rate will be paid.

The more difficult it is to obtain the collateral, the lower the repo rate. To understand why this is so, remember that the borrower (or equivalently the seller of the collateral) has a security that lenders of cash want for whatever reason. (Perhaps the issue is in great demand to satisfy borrowing needs.) Such collateral is said to "on special." Collateral that does not share this characteristic is referred to as *general collateral.* The party that needs collateral that is "on special" will be willing to lend funds at a lower repo rate in order to obtain the collateral.

There are several factors contributing to the demand for *special collateral.* They include:

- Government bond auctions—the bond to be issued is shorted by dealers in anticipation of new supply and due to client demand.
- Outright short selling whether a deliberate position taken based on a trader's expectations or dealers shorting bonds to satisfy client demand.
- Hedging including corporate bonds underwriters who short the relevant maturity benchmark government bond that the corporate bond is priced against.
- Derivative trading such as basis trading creating a demand for a specific bond.
- Buy-back or cancellation of debt at short notice.

Financial crises will also impact a particular security's "specialness." Specialness is defined the spread between the general collateral rate and the repo rate of a particular security. Michael Fleming found that the on-the-run 2-year note, 5-year note, and 30-year bond traded at an increased rate of specialness during the Asian financial crisis of 1998. In other words, the spread between the general collateral rate and the repo rates on these securities increased. Moreover, these spreads returned to more normal levels after the crisis ended (see Fleming (2000)).

While these factors determine the repo rate on a particular transaction, the federal funds rate determines the general level of repo rates. The repo rate generally will trade lower than the federal funds rate, because a repo involves collateralized borrowing while a federal funds transaction is unsecured borrowing.

Special Collateral and Arbitrage

As noted earlier in the chapter, there are a number of investment strategies in which an investor borrows funds to purchase securities. The investor's expectation is that the return earned by investing in the securities purchased with the borrowed funds will exceed the borrowing cost. The use of borrowed funds to obtain greater exposure to an asset than is possible by using only cash is called leveraging. In certain circumstances, a borrower of funds via a repo transaction can generate an arbitrage opportunity. This occurs when it is possible to borrow funds at a lower rate than the rate that can be earned by reinvesting those funds.

Such opportunities present themselves when a portfolio includes securities that are "on special" and the manager can reinvest at a rate higher than the repo rate. For example, suppose that a manager has securities that are "on special" in the portfolio, Bond X, that lenders of funds are willing to take as collateral for two weeks charging a repo rate of say 3%. Suppose further that the manager can invest the funds in a 2-week Treasury bill (the maturity date being the same as the term of the repo) and earn 4%. Assuming that the repo is properly structured so that there is no credit risk, then the manager has locked in a spread of 100 basis points for two weeks. This is a pure arbitrage and the manager faces no risk. Of course, the manager is exposed to the risk that Bond X may decline in value but this the manager is exposed to this risk anyway as long as the manager intends to hold the security.

The results of a study examining the relationship between cash prices and repo rates for bonds that have traded special appeared in the February 1997 and August 1997 market sections of the Bank of England's *Quarterly Bulletin*. The results of the study suggest a positive correlation between changes in a bond trading expensive to the yield curve and changes in the degree to which it trades special. This result is not surprising. Traders maintain short positions in bonds which have associated funding costs only if the anticipated fall in the bond's is large enough to engender a profit. The causality could run in either direction. For example, suppose a bond is perceived as being expensive relative to the yield curve. This circumstance creates a greater demand for short positions and hence a greater demand for the bonds in the repo market to cover the short positions. Alternatively, suppose a bond goes on special in the repo market for whatever reason. The bond would appreciate in price in the cash market as traders close out their short positions which are now too expensive to maintain. Moreover, traders and investors would try to buy the bond outright since it now would be relatively cheap to finance in the repo market.

Participants in the Market

The repo market has evolved into one of the largest sectors of the money market because it is used continuously by dealer firms (investment banks and money center banks acting as dealers) to finance positions and cover short positions. The primary borrowers of securities include major security dealers and hedge funds. Conversely, the primary lenders of securities include institutional investors with long investment horizons (e.g., insurance companies, pension funds, mutual funds). These institutional investors view securities lending as an additional source of revenue. Alternatively, viewing the repo market as a mechanism to borrow and lend cash, the primary borrowers of cash are the same institutions that also borrow securities, namely, dealer firms and hedge funds. Lenders of cash include financial institutions, nonfinancial corporations, money market mutual funds, and municipalities.

Another repo market participant is the repo broker. To understand the repo broker's role, suppose that a dealer has shorted $50 million of the current 10-year Treasury note. It will then query its regular customers to determine if it can borrow, via a reverse repo, the 10-year Treasury note it shorted. Suppose that it cannot find a customer willing to do a repo transaction (repo from the customer's perspective, reverse repo from the dealer's perspective). At that point, the dealer will utilize the services of a repo broker who will find the desired collateral and arrange the transaction for a fee.

Repo/Reverse to Maturity

One important type of repo is a repo/reverse to maturity. A repo/reverse to maturity is one where the term of the repurchase agreement coincides with the maturity date of the collateral and the repurchase price equals the proceeds of the collateral. As before, whether the transaction is a repo or reverse is viewed from the dealer's perspective. This type of transaction is driven primarily for accounting/tax reasons. For example, suppose a dealer has a customer has bond in their portfolio that they would like to sell but the bond is trading below its carrying value. Further suppose the customer does have any gains to offset the loss. In this case, the customer might consider a repo to maturity as an alternative to selling the bond. By doing so, the customer is using the bonds as collateral for a loan and gains access to funds without selling the bond outright.

Buy/Sell-Back

Another securities lending arrangement that is functionally equivalent to a repurchase agreement is a *buy/sell-back agreement.* A buy/sell-back agreement separates a securities lending transaction into separate buy and sell trades that are entered into simultaneously. The security borrower buys the security in question and agrees to return the borrowed security (that is, sell back) at some future date for an agreed upon forward price. The forward price is usually derived using a repo rate. A buy/sell-back agreement differs from a repurchase agreement in that the security borrower receives legal title and beneficial ownership of the securities for the length of the agreement. Moreover, the security borrower retains any accrued interest and coupon payments until the security is returned to the lender. Nevertheless, the price on the termination date reflects the fact that the economic benefits of the coupon interest being transferred back to the seller.

Repo Market Structures

Structured repurchase agreements have developed in recent years mainly in the U.S. market where repo is widely accepted as a money market instrument. Following the introduction of new repo types it is also possible now to transact them in other liquid markets.

LIBOR Financed Treasury Repo

As the name implies, a London Interbank Offered Rate (LIBOR) financed Treasury repurchase agreement differs from a traditional repo in that the repo rate is tied to three-month LIBOR rather than the overnight Federal funds rates. The repo rate is reset quarterly according to movements in the level of three-month LIBOR. Accordingly, unlike a traditional repo, the repo rate over the term of the agreement is uncertain.

Cross-Currency Repo

A cross-currency repo is an agreement in which the cash lent and securities used as collateral are denominated in different currencies say, borrow U.S. dollars with U.K. gilts used as collateral. Of course, fluctuating foreign exchange rates mean that it is likely that the transaction will need to be marked-to-market frequently in order to ensure that cash or securities remain fully collateralized.

Callable Repo

In a callable repo arrangement, the lender of cash in a term fixed-rate repo has the option to terminate the repo early. In other words, the repo transaction has an embedded interest rate option which benefits the lender of cash if rates rise during the repo's term. If rates rise, the lender may exercise the option to call back the cash and reinvest at a higher rate. For this reason, a callable repo will trade at a lower repo rate than an otherwise similar conventional repo.

Whole Loan Repo

A whole loan repo structure developed in the U.S. market as a response to investor demand for higher yields in a falling interest rate environment. Whole loan repo trades at a higher rate than conventional repo because a lower quality collateral is used in the transaction. There are generally two types: mortgage whole loans and consumer whole loans. Both are unsecuritized loans or interest receivables. The loans can also be credit card payments and other types of consumer loans. Lenders in a whole loan repo are not only exposed to credit risk but prepayment risk as well. This is the risk that the loan package is paid off prior to the maturity date which is often the case with consumer loans. For these reasons, the yield on a whole loan repo is higher than conventional repo collateralized by say U.S. Treasuries, trading at around 20 to 30 basis points over LIBOR.

Total Return Swap

A total return swap structure, also known as a "total rate of return swap," is economically identical to a repo. The main difference between a total return swap and a repo is that the former is governed by the International Swap Dealers Association (ISDA) agreement as opposed to a repo agreement. This difference is largely due to the way the transaction is reflected on the balance sheet in that a total return swap is recorded as an off-balance-sheet transaction. This is one of the main motivations for entering into this type of contract. The transaction works as follows:

1. The institution sells the security at the market price
2. The institution executes a swap transaction for a fixed term, exchanging the security's total return for an agreed rate on the relevant cash amount
3. On the swap's maturity date the institution repurchases the security for the market price

In theory, each leg of the transaction can be executed separately with different counterparties; in practice, the trade is bundled together and so is economically identical to a repo.

DOLLAR ROLLS

Dollar rolls resemble repurchase agreements on a number of dimensions. For example, a dollar roll is a collateralized loan that calls for the sale and repurchase of a pass-through security on different settlement dates. However, unlike a repurchase agreement, the dealer who borrows pass-through securities need only return "substantially identical securities." Although we will discuss this in more detail shortly, for now, "substantially identical securities" returned by the dealer must match certain criteria such as the coupon rate and security type (that is, issuer, e.g., Ginnie Mae) and mortgage collateral (e.g., 30-year fixed rate). These are the same general trade parameters that buyer and seller would agree to when trading pass-throughs on a to-be-announced (TBA) basis. This feature provides valuable flexibility to dealers for either covering short positions or obtaining pass-throughs to collateralize a *collateralized mortgage obligation* (CMO) deal. In order to obtain this flexibility, the dealer provides the security lender (that is, the investor) provides 100% financing—no overcollateralization or margin required. The financing cost may also be cheaper (sometimes considerably so) because of this flexibility. Lastly, recall that with a repurchase agreement, there is no transfer of security's cash flows. The original owner continues to receive any principal and coupon interest. Not so with a dollar roll, the dealer borrowing the pass-through security keeps the coupon interest and any principal paydown during the length of the agreement.

Background Information on Agency Pass-throughs

A mortgage *pass-through* security (henceforth, pass-through) is created when one or more mortgage holders form a collection (pool) of mortgages and sell shares or participation certificates in the pool. The cash flow of a pass-through depends on the cash flow of the underlying mortgages. It consists of monthly mortgage payments representing interest, the scheduled repayment of principal, and any prepayments.

Payments are made to security holders each month. Neither the amount nor the timing, however, of the cash flow from the mortgage pool is identical to that of the cash flow passed through to investors. The monthly cash flow for a pass-through is less than the monthly cash flow of the underlying mortgages by an amount equal to servicing and other fees. The other fees are those charged by the issuer or guarantor of the pass-through for guaranteeing the issue. The coupon rate on a pass-through is less than the mortgage rate on the underlying pool of mortgage loans by an amount equal to the servicing and guaranteeing fees.

The timing of the cash flows is also different. The monthly mortgage payment is due from each mortgagor on the first day of each month. There is then a delay in passing through the corresponding monthly cash to the security holders, which varies by the type of pass-through. Because of prepayments, the cash flow of a pass-through is not known with certainty.

There are three major types of pass-throughs guaranteed by the following organizations: Government National Mortgage Association ("Ginnie Mae"), Fannie Mae, and Freddie Mac. These are called *agency pass-throughs.* Ginnie Mae pass-throughs are backed primarily by Federal Housing Authority (FHA) insured or Veterans Administration (VA) guaranteed mortgage loans. Correspondingly, Fannie Mae and Freddie Mac securitize conforming mortgage loans. Agency pass-throughs are identified by a pool prefix and pool number provided by the agency. The prefix indicates the type of pass-through. The pool number indicates specific mortgages underlying the pass-through as well as the pass-through's issuer.

The trading and settlement of mortgage-backed securities is governed by rules established by the Bond Market Association. We limit our discussion in this section to agency pass-through securities. Many trades of pass-through securities are derived from mortgage pools that have yet to be specified. As a result, no pool information is available at the time of the trade. Such a trade is denoted as a *TBA* trade (which stands for *to be announced*). In a TBA trade, the buy and seller agree on the issuer, type of program, coupon rate, face value, the price, and the settlement date. The actual pools underlying the pass-through are not specified. This information is provided by the seller to the buyer before delivery. There are also specified pool trades wherein the actual pool numbers to be delivered are specified.

Agency pass-throughs usually trade on a forward basis and settlement occurs once month. Each pass-through is assigned a settlement day during the month based on the issuer and type of collateral. This system of forward settlement is crucial to the MBS market for two reasons. First, forward settlement allows the originators of mortgages to sell pass-throughs forward before creating mortgage pools. Accordingly, originators can hedge the mortgage rates at which they are lending. Second, forward settlement also facilitates CMO production as the collateral for CMOs is agency pass-throughs. The settlement of CMO deals is usually one month from the pricing date. Thus, issuers of CMOs are active players in the one-month forward market (see Davidson and Ching, 2005). Moreover, it is the demand for newly minted pass-throughs needed for CMO collateral that gives rise to the existence of the dollar roll market.

Determination of the Financing Cost

The process for determining the dollar roll's financing cost is not as straightforward as that of a repurchase agreement. The key elements in determining a dollar roll's financing cost assuming that the dealer is borrowing securities/lending cash are:

1. The sale price and the repurchase price.
2. The amount of the coupon payment.
3. The amount of scheduled principal payments.
4. The projected prepayments of the security sold to the dealer.
5. The attributes of the substantially identical security returned by the dealer.
6. The amount of under- or overdelivery permitted.

Let us consider each of these elements. The repurchase price is usually less than the sale price in a dollar roll. At first blush, this may seem counterintuitive. After all, the repurchase price is always greater than the sale price where the difference represents repo interest. In a dollar roll, the reason the repurchase price is less than the sales price is because of the second element—the investor surrenders any coupon payments they would have received had they simply held the securities during the length of the dollar roll agreement. Thus, the financing costs of a dollar roll depend on the difference between what the investor gives up in terms of forgone coupon interest and what the investor gives back in the form of a lower repurchase price. Specifically, when the yield curve is positively sloped (that is, long-term interest rates exceed short-term interest rates), the coupon rates of newly minted pass-throughs will exceed short-term collateralized borrowing rates. The greater the slope of the yield curve, the lower the repurchase price must be to offset the forgone coupon interest, other things equal.

The third and fourth elements involve principal payments. There are two types of principal payments—scheduled and prepayments. Scheduled principal payments are predictable and are due to loan amortization. Prepayments occur because the homeowner's option to make principal payments in excess of the scheduled amount (in whole or in part) at any time prior to the mortgage's maturity date usually at no cost. As with the coupon payments, the investor forfeits any principal payments during the length of the agreement, A gain will be realized by the dealer on any principal payments if the security is purchased by the dealer at a discount and a loss if purchased at a premium. Because of prepayments, the principal paydown over the life of the agreement is unknown so the investor's borrowing rate is not known with certainty. This uncertainty represents another difference between dollar rolls and repurchase agreements. In a repurchase agreement, the lender of securities/borrower of funds borrows at a known financing rate. Conversely, with a dollar roll, the financing rate is unknown at the outset of the agreement and can only be projected based on an assumed prepayment rate.

The fifth element is another risk since the effective financing cost will depend on the attributes of the substantially identical security that the dealer returns to the lender. Note this differs from a repurchase agreement in that the security borrower must return securities that are identical to those pledged as collateral. A dealer that borrows mortgage pass-throughs will almost never return the identical securities (that is, pass-throughs derived from the same mortgage pools) to the investor. Instead the dealer is only required mortgage pass-throughs that met certain criteria. The American Institute of Certified Public Accountants, Statement of Position 90-3 requires substantially identical securities met the following criteria:

1. Be collateralized by similar mortgages.
2. Be issued by the same agency and be a part of the same program.
3. Have the same original stated maturity.
4. Have identical coupon rates.
5. Be priced at similar market yields.
6. Satisfy delivery requirements; that is, the aggregate principal amounts of the securities delivered and received back must be within 0.1% of the initial amount delivered.

There are literally hundreds if not thousands of pass-through securities that meet these criteria at any given time. However, these pass-throughs differ in that they are securitized by different mortgage pools. As a result, even among substantially identical securities, some pools perform worse than others.

The last element is the amount of under- or overdelivery permitted. Specifically, the BMA (Bond Market Association) delivery standards permit under- or overdelivery of up to 0.01%. In a dollar roll agreement, both the investor and the dealer have the option to under- or overdeliver: the investor when delivering the securities at the outset of the transaction and the dealer when returning the securities at the end of the agreement.

Illustrations of Dollar Roll Agreements

The decision of whether a mortgage-backed securities investor will participate in a dollar roll agreement depends on a number of factors. These factors include the size of the difference between the sale price and the repurchase price (called the drop or the forward drop), prepayment speeds of the collateral underlying the securities, and available reinvestment rates. In this section, we present two illustrations using discount and premium pass-throughs highlighting how these factors impact the investor's decision to roll their securities.

Dollar Roll with Discount Pass-throughs

Using some information obtained from Bloomberg, consider some Fannie Mae pass-throughs that carry a 5% coupon and a principal balance of $1 million. The payment delay is 54 days and the settlement date is October 12, 2006. Suppose that an investor enters into an agreement with a dealer in which it agrees to sell $1 million par value (that is, unpaid aggregate balance) of these Fannie Mae 5s at $96^{13}/_{32}$ and repurchase substantially identical securities one month later at $96^{12}/_{32}$ (the repurchase price). (In market parlance, a trader would say "buy $1 million of the October/November roll.") Note that the difference in the sales price and the repurchase price is $^{1}/_{32}$ and is called the *forward drop*. The key question that the investor faces is whether she should roll the pass-throughs versus simply holding them over the same time period.

If the investor chooses to roll the pass-throughs, she will receive the sale price of $96^{12}/_{32}$ or $964,062.50 for a $1 million principal balance on the settlement date of October 12, 2006. In addition, the investor will receive 11 days accrued interest of $1,805.56 ($1 million × 5% × ($^{11}/_{360}$)) because interest starts accruing October 1. The total amount received on the settlement date is $965,590.28. The investor assumes the proceeds of the dollar roll will be reinvested for the length of the agreement from October 12 to November 13 or 32 days. Suppose the relevant reinvestment rate is the rate on a repurchase agreement collateralized by Treasury securities over the same period, which is 5.16%. The reinvestment income generated over the length of the agreement is the repo interest of $4,428.84 (5.16% × $965,590.28 × ($^{32}/_{360}$)). The investor will have $949,135.71 at the end of the agreement November 13, 2006. This number must be compared to the number of dollars generated by simply holding the pass-through securities over the same period of time.

If the investor holds the pass-throughs, she will receive a cash flow on November 25 for the month of October because of the 54 day stated payment delay. The cash flows consist of coupon interest and principal paydown—both scheduled and prepayments. The interest for the month of October is $4,166.67 (5% × $1 million × $^{1}/_{12}$). While the scheduled principal payments are known, prepayments must be forecasted using an assumed prepayment rate which is say 162 PSA. (The prepayment rate, referred to as the *prepayment speed*, is measured using a standardized benchmark developed by the Bond Market Association (BMA), now the Securities Industry and Financial Markets Association (SIFMA). This prepayment rate is referred to as the PSA speed.) The October principal payments are projected to be $3,567.62. The total cash flow to be received by the investor on November 25 is $7,734.29. We must compute the present value of this payment to find how much they will be worth on November 15 (the end of the dollar roll agreement). Using the reinvestment rate of 5.16%, we discount $7,734.29 back 12 days to obtain $7,720.98.

The remaining principal is worth $996,432.38 using the dollar roll repurchase price of $96^{12}/_{32}$. In addition, there will 12 days accrued interest for the month of November of $1,660.72 (5% × $996,432.38 × $^{12}/_{360}$). The total number of dollars generated by continuing to hold the pass-through securities on November 13 is $969,693.41. Comparing this number to the future value generated by the dollar roll of $970,019.12 indicates there is a $325.71 gain per $1 million of principal for rolling the pass-throughs.

From the investor's perspective, engaging in a dollar roll is tantamount to financing the pass-throughs using a repurchase agreement. As such, it is possible to compute a breakeven reinvestment rate that would make dollar advantage of rolling the securities equal to zero all else equal. In this example, the breakeven rate is 4.781%. When the investor's reinvestment rate is higher than this, there is an advantage to rolling the pass-throughs all else equal. In comparing financing costs, it is important that the dollar amount of the cost be compared to the amounts borrowed. Moreover, it is not proper to compare financing costs of other alternatives without recognizing the risks associated with a dollar roll.

Dollar Roll with Premium Pass-throughs

Now suppose an investor is contemplating a dollar roll with $1 million Fannie Mae 6% pass-throughs at a price of $100^{16}/_{32}$ for settlement on October 12, 2006. The forward drop is $^{1}/_{32}$, and the repurchase price on November 13 (the end of the dollar roll agreement) is $100^{15}/_{32}$. Using an assumed prepayment assumption of 309 PSA, the breakeven reinvestment rate is 5.386%. If the investor uses the one-month repo rate as a proxy for their reinvestment rate of 5.16%, the investor would not want to roll the pass-throughs. Specifically, using these assumptions, there is a $202.02 (per $1 million) advantage for holding rather than rolling these premium pass-throughs.

Risks in a Dollar Roll from the Investor's Perspective

Because of the unusual nature of the dollar roll transaction as a collateralized borrowing vehicle, it is only possible to estimate the financing cost (that is, the breakeven reinvestment rate). The reason being that the speed of prepayments will affect the financing rate the investor pays by rolling the pass-throughs. In our illustration, since the pass-throughs are trading at a discount, faster prepayments will benefit whoever holds the securities. Thus, the investor's financing rate obtained via a dollar roll will be directly related to the prepayment speed. An investor can perform a sensitivity analysis to determine the effect of varying prepayment speeds on the financing rate. If the pass-throughs are trading at a premium, the investor's financing rate will be inversely related to the prepayment speed.

In addition to the uncertainty about the prepayment speed, there is another risk that involves the substantially identical securities returned by the dealer at the end of the dollar roll. As noted earlier, even among substantially identical securities, some pools perform worse than others. The risk is that the dealer will deliver securities from pools that perform poorly.

SUMMARY

A repurchase agreement is the sale of a security with a commitment by the seller to buy the same security back from the purchaser at a specified price at a designated future date. They serve a means to finance bond positions and borrow securities. The liquidity of a bond market is enhanced by an active repo market. In this chapter, we discussed the mechanics of a repurchase agreement and the determinants of repo interest rates. We described a buy/sell-back agreement as well as a structured repo agreement, which is becoming more widely traded. We also discuss a specialized reverse repurchase with pass-through securities serving as collateral known as a dollar roll. After briefly reviewing some background information about pass-through securities and their trading/settlement process, we detailed the mechanics of dollar roll agreements with particular attention to the determination of the financing costs. Finally, the risks in a dollar roll from the investor's perspective were examined.

REFERENCES

Comotto, R. (2005). The European repo market. In F. J. Fabozzi and S. V. Mann (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 234–240). Hoboken, NJ: John Wiley & Sons.

Davidson, A., and Ching, A. (2005). Agency mortgage-backed securities. In F. J. Fabozzi (ed.), *The Handbook of Fixed-Income Securities*, 7th edition (pp. 513–540). New York: McGraw Hill.

Fabozzi, F. J., and Mann, S. V. (2005a). Repurchase and reverse repurchase agreements. In F. J. Fabozzi and S. V. Mann (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 221–240). Hoboken, NJ: John Wiley & Sons.

Fabozzi, F. J. and Mann, S. V. (2005b). Dollar rolls. In F. J. Fabozzi and S. V. Mann (eds.), *Securities Finance: Securities Lending and Repurchase Agreements* (pp. 283–298). Hoboken, NJ: John Wiley & Sons.

Fleming, M. J. (2000). The benchmark U.S. Treasury market: Recent performance and possible alternatives. *FRBNY Economic Policy Review*, April: 129–145.

Index