*Handbooks in Finance*

# HANDBOOK of CORPORATE FINANCE

## EMPIRICAL CORPORATE FINANCE

## Volume 1

Editor: B. Espen Eckbo

North-Holland

North-Holland is an imprint of Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

First edition 2007
Reprinted 2008



**British Library Cataloguing in Publication Data**
A catalogue record for this book is available from the British Library

**Library of Congress Cataloging-in-Publication Data**
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-444-50898-0
ISSN: 1873-1503

For information on all North-Holland publications
visit our website at www.elsevierdirect.com

Printed and bound in *Hungary*

08 09 10 10 9 8 7 6 5 4 3 2

# INTRODUCTION TO THE SERIES



The Handbooks in Finance are intended to be a definitive source for comprehensive and accessible information in the field of finance. Each individual volume in the series presents an accurate self-contained survey of a sub-field of finance, suitable for use by finance and economics professors and lecturers, professional researchers, graduate students and as a teaching supplement. The goal is to have a broad group of outstanding volumes in various areas of finance.

William T. Ziemba
University of British Columbia

Publisher's Note
For a complete overview of the Handbooks in Finance Series, please refer to the listing at the end of this volume.

# PREFACE: EMPIRICAL CORPORATE FINANCE

Judging by the sheer number of papers reviewed in this Handbook, the empirical analysis of firms' financing and investment decisions—empirical corporate finance—has become a dominant field in financial economics. The growing interest in everything "corporate" is fueled by a healthy combination of fundamental theoretical developments and recent widespread access to large transactional data bases. A less scientific—but nevertheless important—source of inspiration is a growing awareness of the important social implications of corporate behavior and governance. This Handbook takes stock of the main empirical findings to date across the entire spectrum of corporate finance issues, ranging from econometric methodology, to raising capital and capital structure choice, and to managerial incentives and corporate investment behavior. The surveys are written by leading empirical researchers that remain active in their respective areas of interest. With few exceptions, the writing style makes the chapters accessible to industry practitioners. For doctoral students and seasoned academics, the surveys offer dense roadmaps into the empirical research landscape and provide suggestions for future work.

### Part 1 (Volume 1): Econometric Issues and Methodological Trends

The empirical corporate finance literature is progressing through a combination of large-sample data descriptions, informal hypothesis testing, as well as structural tests of theory. Researchers are employing a wide spectrum of econometric techniques, institutional settings, and market structures in order to distill the central message in the data. Part 1 of Volume 1 begins by reviewing econometric issues surrounding event studies, and proceeds to explain the econometrics of self-selection. It then explains and illustrates methodological issues associated with the growing use of auction theory, and it ends with a discussion of key elements of the corporate finance evidence from a behavioral perspective.

In Chapter 1, "Econometrics of event studies", S.P. Kothari and Jerold Warner review the power of the event-study method; the most successful empirical technique to date for isolating the price impact of the information content of corporate actions. The usefulness of event studies arises from the fact that the magnitude of abnormal performance at the time of an event provides a measure of the (unanticipated) impact of this type of event on the wealth of the firms' claimholders. Thus, event studies focusing on announcement effects over short horizons around an event provide evidence relevant for understanding corporate policy decisions. Long-horizon event studies also serve an important purpose in capital market research as a way of examining market efficiency. The survey discusses sampling distributions and test statistics typically used in event studies, as well as criteria for reliability, specification and power. While much is known about the statistical properties of short-horizon event studies, the survey provides a critical review of potential pitfalls of long-horizon abnormal return estimates. Serious challenges related to model specification, skewness and cross-correlation remain. As they also point out, events are likely to be associated with return-variance increases, which are equivalent to abnormal returns varying across sample securities. Misspecification induced by variance increases can cause the null hypothesis to be rejected too often unless the test statistic is adjusted to reflect the variance shift. Moreover, the authors emphasize the importance of paying close attention to specification issues for nonrandom samples of corporate events.

Self-selection is endemic to voluntary corporate events. In Chapter 2, "Self-selection models in corporate finance", Kai Li and Nagpurnanand Prabhala review the relevant econometric issues with applications in corporate finance. The statistical issue raised by self-selection is the wedge between the population distribution and the distribution within a selected sample, which renders standard linear (OLS/GLS) estimators biased and inconsistent. This issue is particularly relevant when drawing inferences about the determinants of event-induced abnormal stock returns from multivariate regressions, a technique used by most event studies today. These regressions are typically run using samples that exclude non-event firms. The standard solution is to include a scaled estimate of the event probability—the inverse Mills ratio (the expected value of the true but unobservable regression error term)—as an additional variable in the regression. Interestingly, testing for the significance of the inverse Mills ratio is equivalent to testing whether the sample firms use private information when they self-select to undertake the event. Conversely, if one believes that the particular event being studied is induced by or reflects private information (market overpricing of equity, arrival of new investment projects, merger opportunities, etc.), then consistent estimation of the parameters in the cross-sectional regression requires the appropriate control for self-selection. What is "appropriate" generally depends on the specific application and should ideally be guided by economic theory. The survey also provides a useful overview of related econometric techniques—including matching (treatment effect) models, panel data with fixed effects, and Bayesian self-selection models—with specific applications.

In Chapter 3, "Auctions in corporate finance", Sudipto Dasgupta and Robert Hansen introduce auction theory and discuss applications in corporate finance. The authors explain theoretical issues relating to pricing, efficiency of allocation (the conditions under which the asset is transferred to the most efficient buyer), differential information, collusion among buyers, risk aversion, and the effects of alternative auctions designs (sealed-bid versus open auction, seller reserve price, entry fees, etc.). It is important for empirical research in corporate finance to be informed of auction theory for at least two reasons. First, when sampling a certain transaction type that in fact takes place across a variety of transactional settings, auction theory help identify observable characteristics that are likely to help explain the cross-sectional distribution of things like transaction/bid prices, expected seller revenues, valuation effects, and economic efficiency. This is perhaps most obvious in studies of corporate takeovers (negotiation versus auction, strategic bidding behavior, etc.) and in public security offerings (role of intermediaries, degree and role of initial underpricing, long-run pricing effects, etc.). Second, auction theory provides solutions to the problem of optimal selling mechanism design. This is highly relevant in debates over the efficiency of the market for corporate control (negotiations versus auction, desirability of target defensive mechanisms, the role of the board), optimality of a bankruptcy system (auctions versus court-supervised negotiations, allocation of control during bankruptcy, prospects for fire-sales, risk-shifting incentives, etc.), and the choice of selling mechanism when floating new securities (rights offer, underwritten offering, fixed-price, auction, etc.).

In Chapter 4, "Behavioral corporate finance", Malcolm Baker, Richard Ruback and Jeffery Wurgler survey several aspects of corporate finance and discuss the scope for competing behavioral and rational interpretations of the evidence. The idea that inherent behavioral biases of CEOs—and their perception of investor bias—may affect corporate decisions is both intuitive and compelling. A key methodological concern is how to structure tests with the requisite power to discriminate between behavioral explanations and classical hypotheses based on rationality. The "bad model" problem—the absence of clearly empirically testable predictions—is a challenge for *both* rational and behavioral models. For example, this is evident when using a scaled-price ratio such as the market-to-book ratio (B/M), and where the book value is treated as a fundamental asset value. A high value of B/M may be interpreted as "overvaluation" (behavioral) or, alternatively, as B poorly reflecting economic fundamentals (rational). Both points of view are consistent with the observed inverse relation between B/M and expected returns (possibly with the exception of situations with severe short-selling constraints). Also, measures of "abnormal" performance following some corporate event necessarily condition on the model generating expected return. The authors carefully discuss these issues and how researchers have tried to reduce the joint model problem, e.g. by considering cross-sectional interactions with firm-characteristics such as measures of firm-specific financing constraints. The survey concludes that behavioral approaches help explain a number of important financing and investment patterns, and it offers a number of open questions for future research.

## Part 2 (Volume 1): Banking, Public Offerings, and Private Sources of Capital

In Part 2, the Handbook turns to investment banking and the capital acquisition process. Raising capital is the lifeline of any corporation, and the efficiency of various sources of capital, including banks, private equity and various primary markets for new securities is an important determinant of the firm's cost of capital.

In Chapter 5, "Banks in capital markets", Steven Drucker and Manju Puri review empirical work on the dual role of banks as lenders and as collectors of firm-specific private information through the screening and monitoring of loans. Until the late 1990s, U.S. commercial banks were prohibited from underwriting public security offerings for fear that these banks might misuse their private information about issuers (underwriting a low quality issuer and market it as high quality). Following the repeal of the Glass–Steagall Act in the late 1990s, researchers have examined the effect on underwriter fees of the emerging competition between commercial and investment banks. Commercial banks have emerged as strong competitors: in both debt and equity offerings, borrowers receive lower underwriting fees when they use their lending bank as underwriter. The evidence also shows that having a lending relationship constitutes a significant competitive advantage for the commercial banks in terms of winning underwriting mandates. In response, investment banks have started to develop lending units, prompting renewed concern with conflicts of interest in underwriting. Overall, the survey concludes that there are positive effects from the interaction between commercial banks' lending activities and the capital markets, in part because the existence of a bank lending relationship reduces the costs of information acquisition for capital market participants.

In Chapter 6, "Security offerings", Espen Eckbo, Ronald Masulis and Øyvind Norli review studies of primary markets for new issues, and they extend and update evidence on issue frequencies and long-run stock return performance. This survey covers all of the key security types (straight and convertible debt, common stock, preferred stock, ADR) and the most frequently observed flotation methods (IPO, private placement, rights offering with or without standby underwriting, firm commitment underwritten offering). The authors review relevant aspects of securities regulations, empirical determinants of underwriter fees and the choice of flotation method, market reaction to security issue announcements internationally, and long-run performance of U.S. issuers. They confirm that the relative frequency of *public* offerings of seasoned equity (SEOs) is low and thus consistent with a financial pecking order based on adverse selection costs. They also report that the strongly negative announcement effect of SEOs in the U.S. is somewhat unique to U.S. issuers. Equity issues in other countries are often met with a significantly positive market reaction, possibly reflecting a combination of the greater ownership concentration and different selling mechanisms in smaller stock markets. They conclude from this evidence that information asymmetries have a first-order effect on the choice of which security to issue as well as by which method. Their large-sample estimates of post-issue long-run abnormal performance, which covers a wide range of security types, overwhelmingly reject the hypothesis that the performance is 'abnormal'. Rather, the long-run performance is commensurable with issuing firms' exposures to commonly accepted definitions of pervasive risk factors. They conclude that the long-run evidence fails to support hypotheses which hold that issuers systematically time the market, or hypotheses which maintain that the market systematically over- or under-reacts to the information in the issue announcement.

The cost of going public is an important determinant of financial development and growth of the corporate sector. In Chapter 7, "IPO underpricing", Alexander Ljungqvist surveys the evidence on one significant component of this cost: IPO underpricing, commonly defined as the closing price on the IPO day relative to the IPO price. He classifies theories of underpricing under four broad headings: 'asymmetric information' (between the issuing firm, the underwriter, and outside investors), 'institutional' (focusing on litigation risk, effects of price stabilization, and taxes), 'control' (how the IPO affects ownership structure, agency costs and monitoring), and 'behavioral' (where irrational investors bid up the price of IPO shares beyond true value). From an empirical perspective, these theories are not necessarily mutually exclusive, and several may work to successfully explain the relatively modest level of underpricing (averaging about 15%) observed before the height of the technology-sector offerings in 1999–2000. Greater controversy surrounds the level of underpricing observed in 1999–2000, where the dollar value of issuers' underpricing cost ('money left on the table') averaged more than four times the typical 7% investment banking fee. Two interesting—and mutually exclusive—candidate explanations for this unusual period focus on inefficient selling method design (failure of the fix-priced book-building procedure to properly account for the expected rise in retail investor demand) and investor irrationality (post-offering pricing 'bubble'). Additional work on the use and effect of IPO auctions, and on the uniquely identifying characteristics of a pricing 'bubble', is needed to resolve this issue.

Multidivisional (conglomerate) firms may exist in part to take advantage of internal capital markets. However, in apparent contradiction of this argument, the early literature on conglomerate firms identified a 'conglomerate discount' relative to pure-play (single-plant) firms. In Chapter 8, "Conglomerate firms and internal capital markets", Vojislav Maksimovic and Gordon Phillips present a comprehensive review of how the literature on the conglomerate discount has evolved to produce a deeper economic understanding of the early discount evidence. They argue that issues raised by the data sources used to define the proper equivalent 'pure-play' firm, econometric issues arising from firms self-selecting the conglomerate form, and explicit model-based tests derived from classical profit-maximizing behavior, combine to explain the discount without invoking agency costs and investment inefficiencies. As they explain, a firm that chooses to diversify is a different type of firm than one which stays with a single segment—but either type may be value-maximizing. They conclude that, on balance, internal capital markets in conglomerate firms appear to be efficient in reallocating resources.

After reviewing internal capital markets, bank financing, and public securities markets, Volume 1 ends with the survey "Venture capital" in Chapter 9. Here, Paul Gompers defines venture capital as "independent and professionally managed, dedicated pools of capital that focus on equity or equity-linked investments in privately held, high-growth companies". The venture capital industry fuels innovation by channeling funds to start-up firms and, while relatively small compared to the public markets, has likely had a disproportionately positive impact on economic growth in the United States where the industry is most developed. The empirical literature on venture capital describes key features of the financial contract (typically convertible preferred stock), staging of the investment, active monitoring and advice, exit strategies, etc., all of which affect the relationship between the venture capitalist and the entrepreneur. While data sources are relatively scarce, there is also growing evidence on the risk and return of venture capital investments. Paul Gompers highlights the need for further research on assessing venture capital as a financial asset, and on the internationalization of venture capital.

## Part 3 (Volume 2): Dividends, Capital Structure, and Financial Distress

The first half of Volume 2 is devoted to the classical issue of capital structure choice. This includes the effect of taxes, expected bankruptcy costs, agency costs, and the costs of adverse selection in issue markets on the firm's choice of financial leverage and dividend policy. More recent empirical work also links debt policy to competition in product markets and to the firm's interaction with its customers and suppliers. There is also substantial empirical work on the effect on expected bankruptcy- and distress costs of the design of the bankruptcy code, where claim renegotiation under court supervision (such as under Chapter 11 of the U.S. code) and auctions in bankruptcy (such as in Sweden) are major alternatives being studied.

In Chapter 10, "Payout policy", Avner Kalay and Michael Lemmon refer to payout policy as "the ways in which firms return capital to their equity investors". Classical dividend puzzles include why firms keep paying cash dividends in the presence of a tax-disadvantage relative to capital gains, and why dividend changes have information contents. In contrast to increases in debt interest payments, dividend increases are not contractually binding and therefore easily reversible. So, where is the commitment to maintain the increased level of dividends? While there is strong evidence of a positive information effect of unanticipated dividend increases, they argue that available signaling models are unlikely to capture this empirical phenomenon. Moreover, there is little evidence that dividend yields help explain the cross-section of expected stock returns—which fails to reveal a tax effect of dividend policy. Recent surveys indicate that managers today appear to consider dividends as a second order concern after investment and liquidity needs are met, and to an increased reliance on stock repurchase as an alternative to cash payouts.

In Chapter 11, "Taxes and corporate finance", John Graham reviews research specifically relating corporate and personal taxes to firms' choice of payout policy, capital structure, compensation policy, pensions, corporate forms, and a host of other financing arrangements. This research often finds that taxes do appear to affect corporate decisions, but the economic magnitude of the tax effect is often uncertain. There is cross-sectional evidence that high-tax rate firms use debt more intensively than do low-tax rate firms, but time-series evidence concerning whether firm-specific changes in tax status affect debt policy is sparse. Many firms appear to be "underleveraged" in the sense that they could capture additional tax-related benefits of debt at a low cost—but refrain from doing so. Conclusions concerning "underleverage" are, however, contingent on a model of the equilibrium pricing implications of the personal tax-disadvantage of interest over equity income, a topic that has been relatively little researched. Graham also points to the need for a total tax-planning view (as opposed to studying tax issues one by one) to increase the power of tests designed to detect overall tax effects on firm value.

In Chapter 12, "Tradeoff and pecking order theories of debt", Murray Frank and Vidhan Goyal review the empirical evidence on firms capital structure choice more generally. Under the classical tradeoff theory, the firm finds the optimal debt level at the point where the marginal tax benefit of another dollar of debt equals the marginal increase in expected bankruptcy costs. This theory is somewhat challenged by the evidence of underleverage surveyed by Graham. However, corporate leverage ratios appears to be mean-reverting over long time horizons, which is consistent with firms trying to maintain target leverage ratios. This target may reflect transaction costs of issuing securities, agency costs, and information asymmetries as well as taxes and bankruptcy costs, and the available evidence does not indicate which factors are the dominant ones. They report several stylized facts about firms leverage policies. In the aggregate for large firms (but not for small firms), capital expenditures track closely internal funds, and the "financing deficit" (the difference between investments and internal funds) track closely debt issues. This is as predicted by the "pecking order" hypothesis, under which debt is preferred over equity as a source of external finance. For small firms, however, the deficit tracks closely equity issues, which reverses the prediction of the pecking order. The authors conclude that "no currently available model appears capable of simultaneously accounting for the stylized facts".

In Chapter 13, "Capital structure and corporate strategy", Chris Parsons and Sheridan Titman survey arguments and evidence that link firms' leverage policies to structural characteristics of product markets. Capital structure may affect how the firm chooses to interact with its non-financial stakeholders (customers, workers, and suppliers concerned with the firm's survival) as well as with competitors. To account for endogeneity problems that commonly arise in this setting, most papers in this survey analyze firms' responses to a "shock", whether it be a sharp (and hopefully unanticipated) leverage change, an unexpected realization of a macroeconomic variable, or a surprising regulatory change. This approach often allows the researcher to isolate the effect of leverage on a firm's corporate strategy, and in some cases, makes it possible to pinpoint the specific channel (for example, whether a financially distressed firm lowers prices in response to predation by competitors or by making concessions to its customers). There is evidence that debt increases a firm's employment sensitivity to demand shocks (perhaps perpetuating recessions), but can also protect shareholder wealth by moderating union wage demands. Excessive leverage can also inhibit a firm's ability to compete in the product market, as measured by prices and market shares. Firms that depend crucially on non-fungible investments from stakeholders are most sensitive to these losses, and choose more conservative capital structures as a result.

To avoid formal bankruptcy, financially distressed firms engage in asset sales, equity issues and debt renegotiations. In Chapter 14, "Bankruptcy and resolution of financial distress", Edith Hotchkiss, Kose John, Robert Mooradian and Karin Thorburn survey empirical work on the costs, benefits, and effectiveness of out-of-court debt workouts and of formal "one size fits all" bankruptcy procedures. Failing to renegotiate their debt claims out of court, the firm files for bankruptcy, where it is either liquidated piecemeal or restructured as a going concern under court protection. For reasons that are poorly understood, different bankruptcy systems have evolved in different countries, with a trend toward the structured bargaining process characterizing Chapter 11 of the U.S. code. The U.S. code substantially restricts the liquidation rights of creditors as filing triggers automatic stay of debt payments, prevents repossession of collateral, and allows the bankrupt firm to raise new debt with super-priority (debtor-in-possession financing). In contrast, UK bankruptcy is akin to a contract-driven receivership system where creditor rights are enforced almost to the letter. Here, assets pledged as collateral can be repossessed even if they are vital for the firm, and there is no stay of debt claims. This makes it difficult to continue to operate the distressed firm under receivership, even if the bankrupt firm is economically viable. A third system is found in Sweden where the filing firm is automatically turned over to a court-appointed trustee who arranges an open auction (while all debt claims are stayed). The authors survey the international evidence on bankruptcies (which also includes France, Germany, and Japan). They conclude that it remains an open question whether Chapter 11 in the U.S.—with its uniquely strong protection of the incumbent management team—represents an optimal bankruptcy reorganization procedure.

## Part 4 (Volume 2): Takeovers, Restructurings, and Managerial Incentives

Modern corporate finance theory holds that in a world with incomplete contracting, financial structure affects corporate investment behavior and therefore firm value. The Handbook ends with comprehensive discussions of the value-implications of major corporate investment and restructuring decisions (outside of bankruptcy) and of the role of pay-for-performance type of executive compensation contracts on managerial incentives and risk taking behavior.

In Chapter 15, "Corporate takeovers", Sandra Betton, Espen Eckbo and Karin Thorburn review and extend the evidence on mergers and tender offers. They focus in particular on the bidding process as it evolves sequentially from the first bid through bid revision(s) and towards the final bid outcome. Central issues include bid financing, strategic bidding, agency issues and the impact of statutory and regulatory restrictions. The strategic arsenal of the initial bidder includes approaching the target with a tender offer or a merger bid, acquiring a toehold to gain an advantage over potential competitors, offering a payment method (cash or stock) which signals a high bidder valuation of the target, and/or simply bid high (a preemptive strike). The survey provides new evidence on the magnitude of successive bid jumps, and on the speed of rival firm entry and the time between the first and the final bids in multi-bidder contests. The survey confirms that the average abnormal return to bidders is insignificantly different from zero, and that the sum of the abnormal returns to targets and bidders is positive, suggesting that takeovers improve the overall efficiency of resource allocation. Takeover bids also tend to generate positive abnormal returns throughout the industry of the target, in part because they increase the likelihood that industry rivals may become targets themselves (industry "in-play" effect). The evidence strongly rejects the hypothesis that horizontal mergers reduce consumer welfare through increased market power—even when the merger-induced change in industry concentration is non-trivial. However, some input suppliers suffer losses following downstream mergers that increase the downstream industry's bargaining power. The survey ends with a discussion of merger waves.

In Chapter 16, "Corporate restructurings", Espen Eckbo and Karin Thorburn review a number of financial and asset restructuring techniques—other than corporate takeovers and bankruptcy reorganizations. They distinguish between transactions that securitize corporate divisions from those that recapitalize the entire firm. Forms of divisional securitization include spinoff, splitoff, divestiture, equity carveout and tracking stock. Forms of recapitalizations of the entire firm include leveraged recapitalization, leveraged buyout (LBO), demutualization, going-private transactions, and state privatizations. They show transaction frequency, describe the financing technique, discuss regulatory and tax issues, and review evidence on the associated valuation effects. Announcement-induced abnormal stock returns are generally reported to be positive. Potential sources of this wealth creation include improved alignment of management and shareholder incentives through post-transaction compensation contracts that include divisional stock grants, the elimination of negative synergies, improved governance systems through the disciplinary effect of leverage, the avoidance of underinvestment costs, wealth transfers from old bondholders experiencing claim dilution and risk increase following new debt issues, and an "in-play" effect as divisional securitization increases the probability that the division will become a future acquisition target. Unbundling corporate assets and allowing public trade of securities issued by individual divisions also leads to a general welfare increase from increased market completeness and analyst following. The evidence indicates improved operating performance following spinoffs and LBOs, and increased takeover activity after spinoffs and carveouts, and that a minority of LBO firms goes public within five years of the going-private transaction.

Delegation of corporate control to managers gives rise to costly agency conflicts as the personal interests of managers and owners diverge. The literature on executive compensation seeks to identify the form of the employment contract that minimizes agency costs. In Chapter 17, "Executive compensation and incentives", Rajesh Aggarwal surveys the empirical findings of this literature over the past two decades, focusing in particular on evidence concerning stock options and restricted stock grants. The optimal provision of incentives in managerial compensation contracts depends on factors such as executive risk and effort aversion, managerial productivity, and information asymmetries. A key limitation on incentive provision appears to be the need to share risk between managers and shareholders. Also, while optimal contracting theory implies that firm performance should be evaluated relative to an industry or market wide benchmark, relative performance provisions (e.g. by indexing the exercise price of a stock option to the market) are rarely observed. This puzzle may be explained in part by accounting and tax rules, and in part by the cost to shareholders of indexed options (relative to other forms of compensation) when managers are risk averse. Observed compensation practices may also reflect a governance problem if the CEO has undue influence over the determination of her own level of pay. Some researchers argue that rent extraction by the CEO is a major issue of concern for shareholders, an issue that remains controversial.

For a given compensation contract, risk-averse managers have a personal incentive to limit risk exposure by lowering the volatility of the firm's cash flow ex post. If unchecked, this incentive may lead to value-reducing overinvestment in risk-reducing technologies and projects. However, as reviewed by Clifford Smith in Chapter 18, “Managing corporate risk”, it is widely accepted that active cash flow risk management can also lead to increased shareholder value. For example, if hedging alters the timing of taxable cash flows, there may be a net tax benefit. Hedging may also reduce expected costs of financial distress which in turn may allow the firm to capture additional benefits from leverage. Hedging opportunities (using various forms of derivatives and hybrid instruments) have increased substantially over the past decade, and their costs have decreased. As a result, today some form of hedging activity is common among large publicly traded firms. The evidence indicates that smaller firms—with greater default risk—tend to hedge a larger percentage of their exposures than larger firms. However, Smith points to several data problems that limit the power of the empirical research in this area.

I would like to thank all the contributors for their hard work and patience in seeing this Handbook to fruition. A special thank goes to the Series Editor William T. Ziemba for his enthusiasm for this project.

B. Espen Eckbo
Dartmouth College, 2007

*Chapter 1*

# ECONOMETRICS OF EVENT STUDIES*

S.P. KOTHARI

*Sloan School of Management, E52-325, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142, USA*
*e-mail: kothari@mit.edu*

JEROLD B. WARNER

*William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, NY 14627-0107, USA*
*e-mail: warner@simon.rochester.edu*

## Contents



* We thank Espen Eckbo, Jon Lewellen, Adam Kolasinski, and Jay Ritter for insightful comments, and Irfan Safdar and Alan Wancier for research assistance.

*Handbook of Corporate Finance, Volume 1*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01001-4*

## Abstract

The number of published event studies exceeds 500, and the literature continues to grow. We provide an overview of event study methods. Short-horizon methods are quite reliable. While long-horizon methods have improved, serious limitations remain. A challenge is to continue to refine long-horizon methods. We present new evidence illustrating that properties of event study methods can vary by calendar time period and can depend on event sample firm characteristics such as volatility. This reinforces the importance of using stratified samples to examine event study statistical properties.

## Keywords

event study, abnormal returns, short-horizon tests, long-horizon tests, cross-sectional tests, risk adjustment

## 1. Introduction and background

This chapter focuses on the design and statistical properties of event study methods. Event studies examine the behavior of firms' stock prices around corporate events.[1] A vast literature on event studies written over the past several decades has become an important part of financial economics. Prior to that time, "there was little evidence on the central issues of corporate finance. Now we are overwhelmed with results, mostly from event studies" (Fama, 1991, p. 1600). In a corporate context, the usefulness of event studies arises from the fact that the magnitude of abnormal performance at the time of an event provides a measure of the (unanticipated) impact of this type of event on the wealth of the firms' claimholders. Thus, event studies focusing on announcement effects for a short-horizon around an event provide evidence relevant for understanding corporate policy decisions.

Event studies also serve an important purpose in capital market research as a way of testing market efficiency. Systematically nonzero abnormal security returns that persist after a particular type of corporate event are inconsistent with market efficiency. Accordingly, event studies focusing on long-horizons following an event can provide key evidence on market efficiency (Brown and Warner, 1980; Fama, 1991).

Beyond financial economics, event studies are useful in related areas. For example, in the accounting literature, the effect of earnings announcements on stock prices has received much attention. In the field of law and economics, event studies are used to examine the effect of regulation, as well as to assess damages in legal liability cases.

The number of published event studies easily exceeds 500 (see Section 2), and continues to grow. A second and parallel literature, which concentrates on the methodology of event studies, began in the 1980s. Dozens of papers have now explicitly studied statistical properties of event study methods. Both literatures are mature.

From the methodology papers, much is known about how to do—and how not to do—an event study. While the profession's thinking about event study methods has evolved over time, there seems to be relatively little controversy about statistical properties of event study methods. The conditions under which event studies provide information and permit reliable inferences are well-understood.

This chapter highlights key econometric issues in event study methods, and summarizes what we know about the statistical design and the interpretation of event study experiments. Based on the theoretical and empirical findings of the methodology literature, we provide clear guidelines both for producers and consumers of event studies. Rather than provide a comprehensive survey of event study methods, we seek to sift through and synthesize existing work on the subject. We provide many references and

[1] We discuss event studies that focus only on the mean stock price effects. Many other types of event studies also appear in the literature, including event studies that examine return variances (e.g., Beaver, 1968, and Patell, 1976), trading volume (e.g., Beaver, 1968, and Campbell and Wasley, 1996), operating (accounting) performance (e.g., Barber and Lyon, 1996), and earnings management via discretionary accruals (e.g., Dechow, Sloan and Sweeney, 1995, and Kothari, Leone, and Wasley, 2005).

borrow heavily from the contributions of published papers. Two early papers that cover a wide range of issues are by Brown and Warner (1980, 1985). More recently, an excellent chapter in the textbook of Campbell, Lo, and MacKinlay (1997) is a careful and broad outline of key research design issues. These standard references are recommended reading, but predate important advances in our understanding of event study methods, in particular on long horizon methods. We provide an updated and much needed overview, and include a bit of new evidence as well.

Although much emphasis will be on the statistical issues, we do not view our mission as narrowly technical. As financial economists, our ultimate interest is in how to best specify and test interesting economic hypotheses using event studies. Thus, the econometric and economic issues are interrelated, and we will try to keep sight of the interrelation.

In Section 2, we briefly review the event study literature and describe the changes in event study methodology over time. In Section 3 we discuss how to use events studies to test economic hypotheses. We also characterize the properties of the event study tests and how these properties depend on variables such as security volatility, sample size, horizon length, and the process generating abnormal returns. Section 4 is devoted to issues most likely encountered when conducting long-horizon event studies. The main issues are risk adjustment, cross-correlation in returns, and changes in volatility during the event period.

## 2. The event study literature: basic facts

### 2.1. The stock and flow of event studies

To quantify the enormity of the event study literature, we conducted a census of event studies published in 5 leading journals: the *Journal of Business* (JB), *Journal of Finance* (JF), *Journal of Financial Economics* (JFE), *Journal of Financial and Quantitative Analysis* (JFQA), and the *Review of Financial Studies* (RFS). We began in 1974, the first year the JFE was published.

Table 1 reports the results for the years 1974 through 2000. The total number of papers reporting event study results is 565. Since many academic and practitioner-oriented journals are excluded, these figures provide a lower bound on the size of the literature. The number of papers published per year increased in the 1980s, and the flow of papers has since been stable. The peak years are 1983 (38 papers), 1990 (37 papers), and 2000 (37 papers). All five journals have significant representation. The JFE and JF lead, with over 200 papers each.

Table 1 makes no distinction between long horizon and short horizon studies. While the exact definition of "long horizon" is arbitrary, it generally applies to event windows of 1 year or more. Approximately 200 of the 565 event studies listed in Table 1 use a maximum window length of 12 months or more, with no obvious time trend in the year by year proportion of studies reporting a long-horizon result.

Table 1
Event studies, by year and journal. For each journal, all papers that contain an event study are included. Survey and methodological papers are excluded

| Year | *Journal of Business* | *Journal of Finance* | *Journal of Financial Economics* | *Journal of Financial and Quant. Analysis* | *Review of Financial Studies* | Grand total |
|---|---|---|---|---|---|---|
| 1974 | 2 | | 2 | 1 | | 5 |
| 1975 | | 2 | 2 | 1 | | 5 |
| 1976 | | 5 | 1 | 1 | | 7 |
| 1977 | | 5 | 5 | 1 | | 11 |
| 1978 | 1 | 5 | 4 | 1 | | 11 |
| 1979 | | 7 | | 2 | | 9 |
| 1980 | 3 | 4 | 2 | 2 | | 11 |
| 1981 | 1 | 3 | 4 | 2 | | 10 |
| 1982 | 1 | 6 | 2 | 1 | | 10 |
| 1983 | 2 | 14 | 18 | 4 | | 38 |
| 1984 | | 5 | 5 | 1 | | 11 |
| 1985 | 2 | 4 | 7 | 2 | | 15 |
| 1986 | 2 | 7 | 14 | 4 | | 27 |
| 1987 | | 7 | 18 | 1 | | 26 |
| 1988 | 1 | 4 | 7 | 5 | 1 | 18 |
| 1989 | | 11 | 11 | 1 | 1 | 24 |
| 1990 | 5 | 17 | 7 | 6 | 2 | 37 |
| 1991 | 5 | 17 | 2 | 4 | 1 | 29 |
| 1992 | 4 | 13 | 9 | 4 | 1 | 31 |
| 1993 | 5 | 7 | 5 | 5 | 3 | 25 |
| 1994 | 1 | 10 | 10 | 5 | | 26 |
| 1995 | 1 | 8 | 14 | 11 | 2 | 36 |
| 1996 | 1 | 7 | 10 | 5 | 3 | 26 |
| 1997 | 3 | 8 | 12 | 3 | | 26 |
| 1998 | 1 | 14 | 11 | 3 | | 29 |
| 1999 | 1 | 7 | 12 | 1 | 4 | 25 |
| 2000 | 2 | 15 | 13 | 5 | 2 | 37 |
| Totals | 44 | 212 | 207 | 82 | 20 | 565 |

No survey of these 565 event study papers is attempted here. For the interested reader, the following are some examples of event study surveys. MacKinlay (1997) and Campbell, Lo, and MacKinlay (1997) document the origins and breadth of event studies. The relation of event studies to tests of market efficiency receives considerable attention in Fama (1991), and in recent summaries of long-horizon tests in Kothari and Warner (1997) and Fama (1998). Smith (1986) presents reviews of event studies of financing decisions. Jensen and Ruback (1983), Jensen and Warner (1988), and Jarrell, Brickley, and Netter (1988) survey corporate control events. Recently, Kothari (2001) reviews event studies in the accounting literature.

### 2.2. Changes in event study methods: the big picture

Even the most cursory perusal of event studies done over the past 30 years reveals a striking fact: the basic statistical format of event studies has not changed over time. It is still based on the table layout in the classic stock split event study of Fama et al. (1969). The key focus is still on measuring the sample securities' mean and cumulative mean abnormal return around the time of an event.

Two main changes in methodology have taken place, however. First, the use of daily (and sometimes intraday) rather than monthly security return data has become prevalent, which permits more precise measurement of abnormal returns and more informative studies of announcement effects. Second, the methods used to estimate abnormal returns and calibrate their statistical significance have become more sophisticated. This second change is of particular importance for long-horizon event studies. The changes in long-horizon event study methods reflect new findings in the late 1990s on the statistical properties of long-horizon security returns. The change also parallels developments in the asset pricing literature, particularly the Fama–French 3-factor model.

While long-horizon methods have improved, serious limitations of long-horizon methods have been brought to light and still remain. We now know that inferences from long-horizon tests "require extreme caution" (Kothari and Warner, 1997, p. 301) and even using the best methods "the analysis of long-run abnormal returns is treacherous" (Lyon, Barber, and Tsai, 1999, p. 165). These developments underscore and dramatically strengthen earlier warnings (e.g., Brown and Warner, 1980, p. 225) about the reliability—or lack of reliability—of long-horizon methods. This contrasts with short-horizon methods, which are relatively straightforward and trouble-free. As a result, we can have more confidence and put more weight on the results of short-horizon tests than long-horizon tests. Short-horizon tests represent the "cleanest evidence we have on efficiency" (Fama, 1991, p. 1602), but the interpretation of long-horizon results is problematic. As discussed later, long-horizon tests are highly susceptible to the joint-test problem, and have low power.

Of course these statements about properties of event study tests are very general. To provide a meaningful basis for assessing the usefulness of event studies—both short- and long-horizon—it is necessary to have a framework that specifies: (i) the economic and statistical hypotheses in an event study, and (ii) an objective basis for measuring and comparing the performance of event study methods. Section 3 lays out this framework, and summarizes general conclusions from the methodology literature. In the remainder of the chapter, additional issues and problems are considered with more specificity.

## 3. Characterizing event study methods

### 3.1. An event study: the model

An event study typically tries to examine return behavior for a sample of firms experiencing a common type of event (e.g., a stock split). The event might take place at different points in calendar time or it might be clustered at a particular date (e.g., a regulatory event affecting an industry or a subset of the population of firms). Let $t = 0$ represent the time of the event. For each sample security $i$, the return on the security for time period $t$ relative to the event, $R_{it}$, is:

$$R_{it} = K_{it} + e_{it}, \tag{1}$$

where $K_{it}$ is the "normal" (i.e., expected or predicted return given a particular model of expected returns), and $e_{it}$ is the component of returns which is abnormal or unexpected.[2] Given this return decomposition, the abnormal return, $e_{it}$, is the difference between the observed return and the predicted return:

$$e_{it} = R_{it} - K_{it}. \tag{2}$$

Equivalently, $e_{it}$ is the difference between the return conditional on the event and the expected return unconditional on the event. Thus, the abnormal return is a direct measure of the (unexpected) change in securityholder wealth associated with the event. The security is typically a common stock, although some event studies look at wealth changes for firms' preferred or debt claims.

A model of normal returns (i.e., expected returns unconditional on the event but conditional on other information) must be specified before an abnormal return can be defined. A variety of expected return models (e.g., market model, constant expected returns model, capital asset pricing model) have been used in event studies.[3] Across alternative methods, both the bias and precision of the expected return measure can differ, affecting the properties of the abnormal return measures. Properties of different methods have been studied extensively, and are discussed later.

### *3.2. Statistical and economic hypotheses*

#### *3.2.1. Cross-sectional aggregation*

An event study seeks to establish whether the cross-sectional distribution of returns at the time of an event is abnormal (i.e., systematically different from predicted). Such an exercise can be conducted in many ways. One could, for example, examine the entire distribution of abnormal returns. This is equivalent comparing the distributions of actual with the distribution of predicted returns and asking whether the distributions are the same. In the event study literature, the focus almost always is on the mean of the distribution of abnormal returns. Typically, the specific null hypothesis to be tested is whether the mean abnormal return (sometimes referred to as the average residual, AR) at time $t$ is equal to zero. Other parameters of the cross-sectional distribution (e.g., median, variance) and determinants of the cross-sectional variation in abnormal returns are

[2] This framework is from Brown and Warner (1980) and Campbell, Lo, and MacKinlay (1997).

[3] For descriptions of each of these models, see Brown and Warner (1985) or Campbell, Lo, and MacKinlay (1997).

sometimes studied as well. The focus on mean effects, i.e., the first moment of the return distribution, makes sense if one wants to understand whether the event is, on average, associated with a change in security holder wealth, and if one is testing economic models and alternative hypotheses that predict the sign of the average effect. For a sample of $N$ securities, the cross-sectional mean abnormal return for any period $t$ is:

$$\mathrm{AR}_t = \frac{1}{N}\sum_{i=1}^{N} e_{it}. \tag{3}$$

### 3.2.2. *Time-series aggregation*

It is also of interest to examine whether mean abnormal returns for periods around the event are equal to zero. First, if the event is partially anticipated, some of the abnormal return behavior related to the event should show up in the pre-event period. Second, in testing market efficiency, the speed of adjustment to the information revealed at the time of the event is an empirical question. Thus, examination of post-event returns provides information on market efficiency.

In estimating the performance measure over any multi-period interval (e.g., time 0 through +6), there are a number of methods for time-series aggregation over the period of interest. The cumulative average residual method (CAR) uses as the abnormal performance measure the sum of each month's average abnormal performance. Later, we also consider the buy-and-hold method, which first compounds each security's abnormal returns and then uses the mean compounded abnormal return as the performance measure. The CAR starting at time $t_1$ through time $t_2$ (i.e., horizon length $L = t_2 - t_1 + 1$) is defined as:

$$\mathrm{CAR}(t_1, t_2) = \sum_{t=t_1}^{t_2} \mathrm{AR}_t. \tag{4}$$

Both CAR and buy-and-hold methods test the null hypothesis that mean abnormal performance is equal to zero. Under each method, the abnormal return measured is the same as the returns to a trading rule that buys sample securities at the beginning of the first period, and holds through the end of the last period. CARs and buy-and-hold abnormal returns correspond to security holder wealth changes around an event. Further, when applied to post-event periods, tests using these measures provide information about market efficiency, since systematically nonzero abnormal returns following an event are inconsistent with efficiency and imply a profitable trading rule (ignoring trading costs).

## 3.3. *Sampling distributions of test statistics*

For a given performance measure, such as the CAR, a test statistic is typically computed and compared to its assumed distribution under the null hypothesis that mean abnormal performance equals zero.[4] The null hypothesis is rejected if the test statistic exceeds a critical value, typically corresponding to the 5% or 1% tail region (i.e., the test level or size of the test is 0.05 or 0.01). The test statistic is a random variable because abnormal returns are measured with error. Two factors contribute to this error. First, predictions about securities' unconditional expected returns are imprecise. Second, individual firms' realized returns at the time of an event are affected for reasons unrelated to the event, and this component of the abnormal return does not average to literally zero in the cross-section.

For the CAR shown in equation (4), a standard test statistic is the CAR divided by an estimate of its standard deviation.[5] Many alternative ways to estimate this standard deviation have been examined in the literature (see, for example, Campbell, Lo, and MacKinlay, 1997). The test statistic is given by:

$$\frac{\mathrm{CAR}(t_1, t_2)}{[\sigma^2(t_1, t_2)]^{1/2}}, \tag{5}$$

where

$$\sigma^2(t_1, t_2) = L\sigma^2(\mathrm{AR}_t) \tag{6}$$

and $\sigma^2(\mathrm{AR}_t)$ is the variance of the one-period mean abnormal return. Equation (6) simply says that the CAR has a higher variance the longer is $L$, and assumes time-series independence of the one-period mean abnormal return. The test statistic is typically assumed unit normal in the absence of abnormal performance. This is only an approximation, however, since estimates of the standard deviation are used.

The test statistic in equation (5) is well-specified provided the variance of one-period mean abnormal return is estimated correctly. Event-time clustering renders the independence assumption for the abnormal returns in the cross-section incorrect (see Collins and Dent, 1984, Bernard, 1987, and Petersen, 2005, and more detailed discussion in Section 4 below). This would bias the estimated standard deviation downward and the test statistic given in equation (5) upward. To address the bias, the significance of the event-period average abnormal return can be and often is gauged using the variability of the time series of event portfolio returns in the period preceding or after the event date. For example, the researcher can construct a portfolio of event firms and obtain a time series of daily abnormal returns on the portfolio for a number of days (e.g., 180 days) around the event date. The standard deviation of the portfolio returns can be used to assess the significance of the event-window average abnormal return. The cross-sectional

[4] Standard tests are "classical" rather than "Bayesian". A Bayesian treatment of event studies is beyond the scope of this chapter.

[5] An alternative would be a test statistic that aggregates standardized abnormal returns, which means each observation is weighted in inverse proportion of the standard deviation of the estimated abnormal return. The standard deviation of abnormal returns is estimated using time-series return data on each firm. While a test using standardized abnormal returns is in principle superior under certain conditions, empirically in short-horizon event studies it typically makes little difference (see Brown and Warner, 1980, 1985).

dependence is accounted for because the variability of the portfolio returns through time incorporates whatever cross-dependence that exists among the returns on individual event securities.

The portfolio return approach has a drawback, however. To the extent the event period is associated with increased uncertainty, i.e., greater return variability, the use of historical or post-event time-series variability might understate the true variability of the event-period abnormal performance. An increase in event-period return variability is economically intuitive. The event might have been triggered by uncertainty-increasing factors and/or the event itself causes uncertainty in the economic environment for the firm. In either case, the event-period return variability is likely to exceed that during other time periods for the event firms. Therefore, the statistical significance of the event-window abnormal performance would be overstated if it is evaluated on the basis of historical variability of the event-firm portfolio returns (Brown and Warner, 1980, 1985; Collins and Dent, 1984). One means of estimating the likely increase in the variability of event-period returns is to estimate the cross-sectional variability of returns during the event and non-event periods. The ratio of the variances during the event period and non-event periods might serve as an estimate of the degree of increase in the variability of returns during the event period, which can be used to adjust for the bias in the test statistic calculated ignoring the increased event-period uncertainty.[6]

### *3.4. Criteria for "reliable" event study tests*

Using the test statistics, errors of inference are of two types. A Type I error occurs when the null hypothesis is falsely rejected. A Type II error occurs when the null is falsely accepted. Accordingly, two key properties of event study tests have been investigated. The first is whether the test statistic is correctly specified. A correctly-specified test statistic yields a Type I error probability equal to the assumed size of the test. The second concern is power, i.e., a test's ability to detect abnormal performance when it is present. Power can be measured as one minus the probability of a Type II error. Alternatively, it can be measured as the probability that the null hypothesis will be rejected given a level of Type I error and level of abnormal performance. When comparing tests that are well-specified, those with higher power are preferred.

### *3.5. Determining specification and power*

#### *3.5.1. The joint-test problem*

While the specification and power of a test can be statistically determined, economic interpretation is not straightforward because all tests are joint tests. That is, event study

[6] Use of non-parametric tests of significance, as suggested in Corrado (1989), might also be effective in performing well-specified tests in the presence of increased event-period uncertainty.

tests are well-specified only to the extent that the assumptions underlying their estimation are correct. This poses a significant challenge because event study tests are joint tests of whether abnormal returns are zero and of whether the assumed model of expected returns (i.e., the CAPM, market model, etc.) is correct. Moreover, an additional set of assumptions concerning the statistical properties of the abnormal return measures must also be correct. For example, a standard $t$-test for mean abnormal performance assumes, among other things, that the mean abnormal performance for the cross-section of securities is normally distributed. Depending on the specific $t$-test, there may be additional assumptions that the abnormal return data are independent in time-series or cross-section. The validity of these assumptions is often an empirical question. This is particularly true for small samples, where one cannot rely on asymptotic results or the central limit theorem.

### *3.5.2. Brown–Warner simulation*

To directly address the issue of event study properties, the standard tool in event study methodology research is simulation procedures that use actual security return data. The motivation and specific research design is initially laid out in Brown and Warner (1980, 1985), and has been followed in almost all subsequent methodology research.

Much of what is known about general properties of event study tests comes from such large-scale simulations. The basic idea behind the event study simulations is simple and intuitive.[7] Different event study methods are simulated by repeated application of each method to samples that have been constructed through a random (or stratified random) selection of securities and random selection of an event date to each. If performance is measured correctly, these samples should show no abnormal performance, on average. This makes it possible to study test statistic specification, that is, the probability of rejecting the null hypothesis when it is known to be true. Further, various levels of abnormal performance can be artificially introduced into the samples. This permits direct study of the power of event study tests, that is, the ability to detect a given level of abnormal performance.

### *3.5.3. Analytical methods*

Simulation methods seem both natural and necessary to determine whether event study test statistics are well-specified. Once it has been established using simulation methods that a particular test statistic is well-specified, analytical procedures have also been used to complement simulation procedures. Although deriving a power function analytically for different levels of abnormal performance requires additional distributional assumptions, the evidence in Brown and Warner (1985, p. 13) is that analytical and simulation methods yield similar power functions for a well-specified test statistic. As illustrated below, these analytical procedures provide a quick and simple way to study power.

[7] This characterization of simulation is from Brown and Warner (1985, p. 4).

### *3.6. A quick summary of our knowledge*

#### *3.6.1. Qualitative properties*

Table 2 highlights, in qualitative terms, what is known about the properties of event study tests. The table shows the characteristics of event study methods along three dimensions: specification, power against specific types of alternative hypotheses, and the sensitivity of specification to assumptions about the return generating process. The table also shows how these properties can differ sharply for short and long horizon studies. Much of the remainder of the chapter deals with the full details of this table.

From Table 2, horizon length has a big impact on event study test properties. First, short-horizon event study methods are generally well-specified, but long-horizon methods are sometimes very poorly specified. While much is understood about how to reduce misspecification in long horizon studies (see Section 4), no procedure in whose specification researchers can have complete confidence has yet been developed. Second, short-horizon methods are quite powerful if (but only if) the abnormal performance is concentrated in the event window. For example, a precise event date is known for earnings announcements, but insider trading events might be known to have occurred

Table 2
General characterization of properties of event study test methods

| Criterion | Length of event window | |
|---|---|---|
| | Short (<12 months) | Long (12 months or more) |
| Specification | Good | Poor/Moderate |
| *Power when abnormal performance is*: | | |
| Concentrated in event window | High | Low |
| Not concentrated in event window | Low | Low |
| *Sensitivity of test statistic specification to assumptions about the return generating process*: | | |
| Expected returns, unconditional on event | Low | High |
| Cross-sectional and time-series dependence of sample abnormal returns | Low/Moderate | Moderate/High |
| Variance of abnormal returns, conditional on event | High | High |
| *Sensitivity of power to*: | | |
| Sample size | High | High |
| Firm characteristics (e.g., size, industry) | High | High |

only sometime during a one-month window. In contrast to the short-horizon tests, long-horizon event studies (even when they are well-specified) generally have low power to detect abnormal performance, both when it is concentrated in the event window and when it is not. That power to detect a given level of abnormal performance is decreasing in horizon length is not surprising, but the empirical magnitudes are dramatic (see below). Third, with short-horizon methods the test statistic specification is not highly sensitive to the benchmark model of normal returns or assumptions about the cross-sectional or time-series dependence of abnormal returns. This contrasts with long-horizon methods, where specification is quite sensitive to assumptions about the return generating process.

Along several lines, however, short- and long-horizon tests show similarities, and these results are easy to show using either simulation or analytical procedures. First, a common problem shared by both short- and long-horizon studies is that when the variance of a security's abnormal returns conditional on the event increases, test statistics can easily be misspecified, and reject the null hypothesis too often. This problem was first brought to light and has been studied mainly in the context of short-horizon studies (Brown and Warner, 1985, and Corrado, 1989). A variance increase is indistinguishable from abnormal returns differing across sample securities at the time of an event, and would be expected for an event. Thus, this issue is likely to be empirically relevant both in a short- and long-horizon context as well. Second, power is higher with increasing sample size, regardless of horizon length. Third, power depends on the characteristics of firms in the event study sample. In particular, firms experiencing a particular event can have nonrandom size and industry characteristics. This is relevant because individual security variances (and abnormal return variances) exhibit an inverse relation to firm size and can vary systematically by industry. Power is inversely related to sample security variance: the noisier the returns, the harder to extract a given signal. As shown below, differences in power by sample type can be dramatic.

### *3.6.2. Quantitative results*

To provide additional texture on Table 2, below we show specific quantitative estimates of power. We do so using the test statistic shown previously in equations (5) and (6), using two-tailed tests at the 0.05 significance level.[8] Since this test statistic is well-specified, at least at short horizons, the power functions are generated using analytic (rather than simulation) procedures. The estimates are for illustrative purposes only, however, and only represent "back of the envelope" estimates. The figures and the test statistic on which they are based assume independence of the returns (both through time and in the cross-section), and that all securities within a sample have the same standard

[8] This format for displaying power functions is similar to Campbell, Lo, and MacKinlay (1997, pp. 168–172). Our test statistic and procedures are the same as for their test statistic J1, but as discussed below we use updated variance inputs.

deviation. The power functions also assume that return and abnormal return variances are the same (i.e., the model of abnormal returns is the "mean-adjusted returns" model of Brown and Warner, 1980).

### *3.6.3. Volatility*

In calculating the test statistic in an event study, a key input required here is the individual security return (or abnormal return) variance (or standard deviation). To determine a reasonable range of standard deviations, we estimate daily standard deviations for all CRSP listed firms from 1990 to 2002. Specifically, for each year, we: (i) calculate each stock's standard deviation, and (ii) assign firms to deciles ranked by standard deviation. From each decile, the averages of each year's mean and median values are reported in Table 3. The mean daily standard deviation for all firms is 0.053. This is somewhat higher than the value of 0.026 reported by Brown and Warner (1985, p. 9) for NYSE/AMEX firms and the value of 0.035 reported by Campbell and Wasley (1993, p. 79) for NASDAQ firms. The differences reflect that individual stocks have become more volatile over time (Campbell et al., 2001). This is highly relevant because it suggests that the power to detect abnormal performance for events over 1990–2002 is lower than for earlier periods. From Table 3, there is wide variation across the deciles. Firms in decile 1 have a mean daily standard deviation of 0.014, compared to 0.118 for decile 10. The figure of 0.118 for decile 10 seems very high, although this is likely to represent both very small firms and those with low stock prices. Further, there is a strong negative

Table 3
Standard deviation of daily returns on individual securities using all CRSP common-stock securities from 1990–2002. For each year, firms are ranked by their estimated daily standard deviation. Firms with missing observations are excluded. The numbers under mean and median columns represent the average of the annual mean and median values for the firms in each decile and for all firms. The number of firms in each decile ranges from 504 in 2002 to 673 in 1997

| Decile | Standard deviation | |
|---|---|---|
| | Mean | Median |
| 1 | 0.014 | 0.014 |
| 2 | 0.019 | 0.019 |
| 3 | 0.023 | 0.023 |
| 4 | 0.028 | 0.028 |
| 5 | 0.033 | 0.033 |
| 6 | 0.039 | 0.039 |
| 7 | 0.046 | 0.046 |
| 8 | 0.055 | 0.055 |
| 9 | 0.069 | 0.068 |
| 10 | 0.118 | 0.098 |
| All firms | 0.053 | 0.053 |

empirical relation between volatility and size. Our qualitative results apply if ranking is by firm size, so Table 3 is not simply picking up measurement error in volatility.

### *3.6.4. Results*

Figure 1 shows how, for a sample comprised of securities of average risk and 10% abnormal performance, the power to detect abnormal performance falls with horizon length. This level of abnormal performance seems economically highly significant. If the abnormal performance is concentrated entirely in one day (and the day in known with certainty), a sample of only six stocks detects this level of abnormal performance 100% of the time. In contrast, if the same abnormal performance occurs over six months, a sample size of 200 is required to detect the abnormal performance even 65% of the time. These various rejection frequencies are lower than those using pre-1990 volatilities (not reported), although this is not surprising.

Figure 2(a)–(c) show related results using a one-day horizon for samples whose individual security standard deviations correspond to the average standard deviation for: the lowest decile (Figure 2(a)); all firms (Figure 2(b)); and the highest decile (Figure 2(c)). For decile 1 firms, with 1% abnormal performance a 90% rejection rate requires only 21 stocks. For firms in decile 10, even with 5% abnormal performance a 90% rejection rate requires 60 stocks. These comparisons may distort the differences in actual power if high variance firms are less closely followed and events are bigger surprises. When the effect of events differs cross-sectionally, analysis of test properties (i.e., power and specification) is more complicated.

Collectively, our results illustrate that power against alternative hypotheses can be sensitive to calendar time period and sample firm characteristics, and highlight the importance already recognized in the profession of studying test statistic properties for

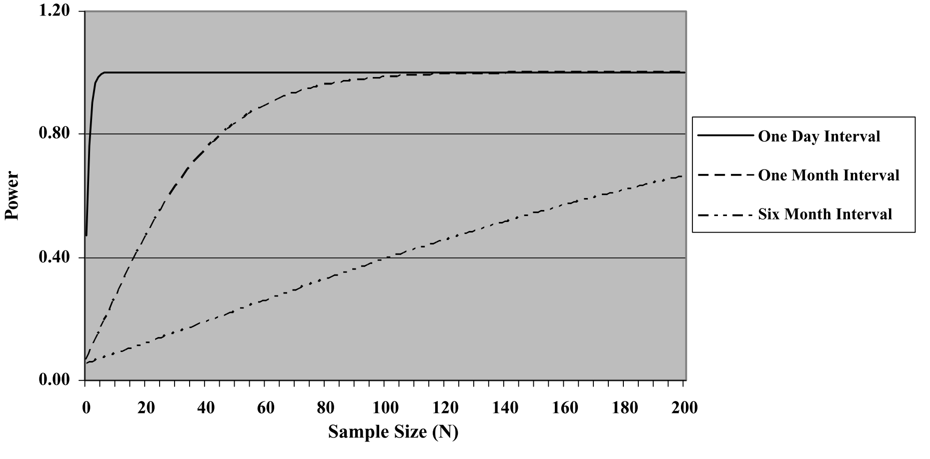
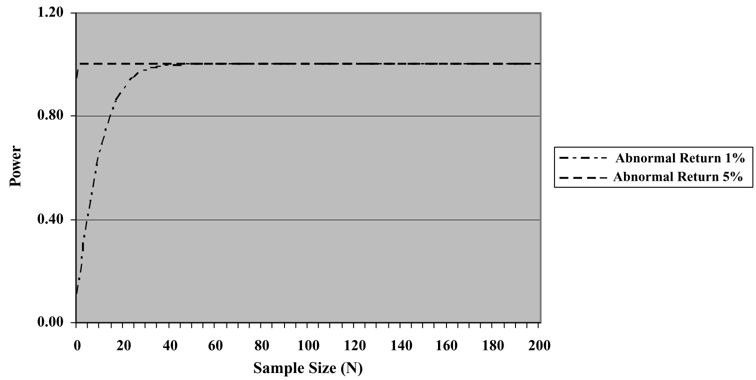
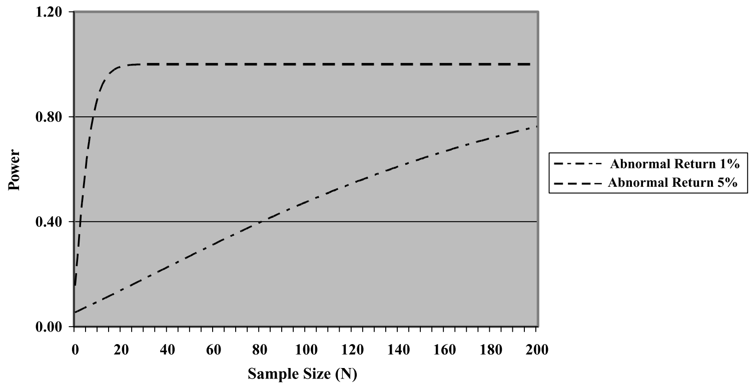


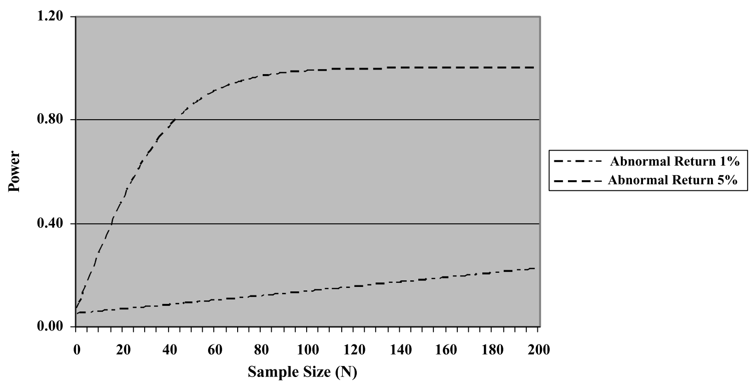

Fig. 1. Power of event study test statistic when abnormal return is 10%.




(a)




(b)




(c)

Fig. 2. (a) Power of event study for firms in the lowest volatility decile. (b) Power of event study for firms with average volatility. (c) Power of event study for firms in the highest volatility decile.

samples stratified by firm characteristics. A complete analysis of these issues would focus on abnormal return (rather than return) volatility, and study how specification (and abnormal return distributional properties such as skewness) varies across time and firm characteristics.

### *3.7. Cross-sectional tests*

This section's focus thus far has been event study tests for mean stock price effects. These tests represent the best understood class of event study tests. To provide a more complete picture of event-related tests, we briefly call attention to cross-sectional tests. These tests examine how the stock price effects of an event are related to firm characteristics. For a cross-section of firms, abnormal returns are compared to (e.g., regressed against) firm characteristics. This provides evidence to discriminate among various economic hypotheses.

Cross-sectional tests are a standard part of almost every event study. They are relevant even when the mean stock price effect of an event is zero. In addition, they are applicable regardless of horizon length. They are simple to do, but as discussed below, "one must be careful in interpreting the results" (Campbell, Lo, and MacKinlay, 1997, p. 174).

One reason that abnormal returns vary cross-sectionally is that the economic effect of the event differs by firm. For such a situation, Sefcik and Thompson (1986) examine the statistical properties of cross-sectional regressions. They are concerned with the effects of cross-sectionally correlated abnormal returns and heteroscedasticity in the abnormal returns. They argue that accounting for each appears to be potentially important for inferences, and they suggest procedures to deal with these issues.

Abnormal returns also vary cross-sectionally because the degree to which the event is anticipated differs by firm. For example, for firms that are more closely followed (e.g., more analysts), events should be more predictable, all else equal. Further, events are endogenous, reflecting a firm's self selection to choose the event, which in turn reflects insiders' information. Recognizing these factors, and recognizing that it is the unexpected information provided by an event that determines the stock price effect, has numerous consequences. For example, standard estimates of cross-sectional coefficients can be biased (Eckbo, Maksimovic, and Williams, 1990). Appropriate procedures for treating self-selection and partial anticipation issues is the subject of an entire chapter by Li and Prabhala (2007) (Chapter 2 in this volume).

Quite apart from the issues discussed in the context of Li and Prabhala, there are several additional dimensions where our understanding of cross-sectional tests is incomplete, and where additional work is potentially fruitful. One area concerns the power of cross-sectional procedures. While specification of cross-sectional regression methods (i.e., biases in regression coefficients) has received much attention, the power of alternative procedures to detect underlying cross-sectional effects has received less study.

A related point is that a simple type of cross-sectional procedure is to form portfolios based on firm characteristics, and compare portfolio abnormal returns. Such procedures are common, but methodological comparisons to cross-sectional regressions would prove useful. Portfolio procedures seem less amenable to multivariate comparisons than do regression procedures, but the relative empirical merits of each in an event-study context have not been investigated.

We also note that some studies focus not on the stock price effect of an event, but on predicting a corporate event (e.g., management turnover, or a security issue of a particular type), sometimes using past stock prices as one explanatory variable. These tests use cross-sectional methods in the sense that the cross-section includes both event and non-event firms. Typically, discrete choice models (e.g., probit or logit model) relate whether or not the event occurred to firm-specific characteristics. This seems intuitive, since we would like to know what factors led the firm to have the event. These methods complement standard event study methods. Methodological work on prediction models could enhance our understanding of how to best to use information about events to test economic hypotheses about firm behavior.

Finally, additional important issues to consider in an event study are: (i) whether the event was partially anticipated by market participants (e.g., a governance-related regulation might be anticipated following corporate scandals or CEO turnover is likely in the case of a firm experiencing steep stock-price decline and poor accounting performance), and (ii) whether the partial anticipation is expected to vary cross-sectionally in a predictable fashion (e.g., market participants might anticipate that managers of firms experiencing high price run-ups are likely to make value-destroying stock acquisitions, but the negative announcement effect of an actual merger announcement might have been largely anticipated for the firms who have experienced relatively high prior price run up). These issues arising from the nature of information arrival, partial anticipation of events, and cross-sectional variation in the degree of anticipation are also beyond the scope of this chapter. Interested readers will find treatments in Malatesta and Thompson (1985), Eckbo, Maksimovic, and Williams (1990), and, especially, Thompson (1995) of considerable interest.

## 4. Long-horizon event studies

All event studies, regardless of horizon length, must deal with several basic issues. These include risk adjustment and expected/abnormal return modeling (Section 4.2), the aggregation of security-specific abnormal returns (Section 4.3), and the calibration of the statistical significance of abnormal returns (Section 4.4). These issues become critically important with long horizons. The remainder of this chapter focuses on efforts in the long-horizon literature to deal with the issues.

### *4.1. Background*

Long-horizon event studies have a long history, including the original stock split event study by Fama et al. (1969). As evidence inconsistent with the efficient markets hypothesis started to accumulate in the late seventies and early eighties, interest in long-horizon studies intensified. Evidence on the post-earnings announcement effect (Ball and Brown, 1968; Jones and Litzenberger, 1970), size effect (Banz, 1981), and earnings yield effect (Basu, 1977, 1983) contributed to skepticism about the CAPM as well as market efficiency. This evidence prompted researchers to develop hypotheses about market inefficiency stemming from investors' information processing biases (DeBondt and Thaler, 1985, 1987) and limits to arbitrage (De Long et al., 1990a, 1990b; Shliefer and Vishny, 1997).

The "anomalies" literature and the attempts to model the anomalies as market inefficiencies has led to a burgeoning field known as behavioral finance. Research in this field formalizes (and tests) the security pricing implications of investors' information processing biases.[9] Because the behavioral biases might be persistent and arbitrage forces might take a long time to correct the mispricing, a vast body of literature hypothesizes and studies abnormal performance over long horizons of one-to-five years following a wide range of corporate events. The events might be one-time (unpredictable) phenomena like an initial public offering or a seasoned equity offering, or they may be recurring events such as earnings announcements.

Many long-horizon studies document apparent abnormal returns spread over long horizons. The literature on long-horizon security price performance following corporate events is summarized extensively in many studies, including Fama (1998), Kothari and Warner (1997), Schwert (2001), and Kothari (2001). Whether the apparent abnormal returns are due to mispricing, or simply the result of measurement problems, is a contentious and unresolved issue among financial economists. The methodological research in the area is important because it demonstrates how easy it is to conclude there is abnormal performance when none exists. Before questions on mispricing can be answered, better methods than currently exist are required.

We summarize some of the salient difficulties and the state-of-the-art event study methods for estimating long-horizon security price performance. More detailed discussions appear in Barber and Lyon (1997), Kothari and Warner (1997), Fama (1998), Brav (2000), Lyon, Barber, and Tsai (1999), Mitchell and Stafford (2000), Jegadeesh and Karceski (2004), Viswanathan and Wei (2004), Eckbo, Masulis, and Norli (2006) and Petersen (2005).

### *4.2. Risk adjustment and expected returns*

In long-horizon tests, appropriate adjustment for risk is critical in calculating abnormal price performance. This is in sharp contrast to short-horizon tests in which risk adjustment is straightforward and typically unimportant. The error in calculating abnormal performance due to errors in adjusting for risk in a short-horizon test is likely to be small. Daily expected returns are about 0.05% (i.e., annualized about 12–13%). Therefore, even if the event firm portfolio's beta risk is misestimated by 50% (e.g., estimated

[9] See Shleifer (2000), Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), Daniel, Hirshleifer, and Teoh (2002), Hirshleifer (2001), and Hong and Stein (1999).

beta risk of 1.0 when true beta risk is 1.5), the error in the estimated abnormal error is small relative to the abnormal return of 1% or more that is typically documented in short-window event studies. Not surprisingly, Brown and Warner (1985) conclude that simple risk-adjustment approaches to conducting short-window event studies are quite effective in detecting abnormal performance.

In multi-year long-horizon tests, risk-adjusted return measurement is the Achilles heel for at least two reasons. First, even a small error in risk adjustment can make an economically large difference when calculating abnormal returns over horizons of one year or longer, whereas such errors make little difference for short horizons. Thus, the precision of the risk adjustment becomes far more important in long-horizon event studies. Second, it is unclear which expected return model is correct, and therefore estimates of abnormal returns over long horizons are highly sensitive to model choice. We now discuss each of these problems in turn.

### *4.2.1. Errors in risk adjustment*

Such errors can make an economically non-trivial difference in measured abnormal performance over one-year or longer periods. The problem of risk adjustment error is exacerbated in long-horizon event studies because the potential for such error is greater for longer horizons. In many event studies, (i) the event follows unusual prior performance (e.g., stock splits follow good performance), or (ii) the event sample consists of firms with extreme (economic) characteristics (e.g., low market capitalization stocks, low-priced stocks, or extreme book-to-market stocks), or (iii) the event is defined on the basis of unusual prior performance (e.g., contrarian investment strategies in DeBondt and Thaler, 1985, and Lakonishok, Shleifer, and Vishny, 1994). Under these circumstances, accurate risk estimation is difficult, with historical estimates being notoriously biased because prior economic performance negatively impacts the risk of a security. Therefore, in long-horizon event studies, it is crucial that abnormal-performance measurement be on the basis of post-event, not historical risk estimates (Ball and Kothari, 1989; Chan, 1988; Ball, Kothari, and Shanken, 1995; Chopra, Lakonishok, and Ritter, 1992). However, how the post-event risk should be estimated is itself a subject of considerable debate, which we summarize below in an attempt to offer guidance to researchers.

### *4.2.2. Model for expected returns*

The question of which model of expected returns is appropriate remains an unresolved issue. As noted earlier, event studies are joint tests of market efficiency and a model of expected returns (e.g., Fama, 1970). On a somewhat depressing note, Fama (1998, p. 291) concludes that "all models for expected returns are incomplete descriptions of the systematic patterns in average returns", which can lead to spurious indications of abnormal performance in an event study. With the CAPM as a model of expected returns being thoroughly discredited as a result of the voluminous anomalies evidence, a quest for a better-and-improved model began. The search culminated in the Fama and French (1993) three-factor model, further modified by Carhart (1997) to incorporate the momentum factor.[10] However, absent a sound economic rationale motivating the inclusion of the size, book-to-market, and momentum factors, whether these factors represent equilibrium compensation for risk or they are an indication of market inefficiency has not been satisfactorily resolved in the literature (see, e.g., Brav and Gompers, 1997). Fortunately, from the standpoint of event study analysis, this flaw is not fatal. Regardless of whether the size, book-to-market, and momentum factors proxy for risk or indicate inefficiency, it is essential to use them when measuring abnormal performance. The purpose of an event study is to isolate the incremental impact of an event on security price performance. Since the price performance associated with the size, book-to-market, and momentum characteristics is applicable to all stocks sharing those characteristics, not just the sample of firms experiencing the event (e.g., a stock split), the performance associated with the event itself must be distinguished from that associated with other known determinants of performance, such as the aforementioned four factors.[11]

### *4.3. Approaches to abnormal performance measurement*

While post-event risk-adjusted performance measurement is crucial in long-horizon tests, actual measurement is not straightforward. Two main methods for assessing and calibrating post-event risk-adjusted performance are used: characteristic-based matching approach and the Jensen's alpha approach, which is also known as the calendar-time portfolio approach (Fama, 1998; Eckbo, Masulis, and Norli, 2000; Mitchell and Stafford, 2000). Analysis and comparison of the methods is detailed below. Despite an extensive literature, there is still no clear winner in a horse race. Both have low power against economically interesting null hypotheses, and neither is immune to misspecification.

#### *4.3.1. BHAR approach*

In recent years, following the works of Ikenberry, Lakonishok, and Vermaelen (1995), Barber and Lyon (1997), Lyon, Barber, and Tsai (1999), the characteristic-based matching approach (or also known as the buy-and-hold abnormal returns, BHAR) has been widely used. Mitchell and Stafford (2000, p. 296) describe BHAR returns as "the average multiyear return from a strategy of investing in all firms that complete an event and selling at the end of a prespecified holding period versus a comparable strategy

[10] More recently, considerable evidence suggests the importance of a liquidity factor in determining expected returns (Brennan and Subrahmanyam, 1996; Pastor and Stambaugh, 2003; Sadka, 2006). However, still others have begun to question the usefulness of the liquidity factor (see Chordia et al., 2006, and Ng, Rusticus, and Verdi, 2006).

[11] See Kothari, Leone, and Wasley (2005) for an extended discussion.

using otherwise similar nonevent firms". An appealing feature of using BHAR is that buy-and-hold returns better resemble investors' actual investment experience than periodic (monthly) rebalancing entailed in other approaches to measuring risk-adjusted performance.[12] The joint-test problem remains in that any inference on the basis of BHAR hinges on the validity of the assumption that event firms differ from the "otherwise similar nonevent firms" only in that they experience the event. The researcher implicitly assumes an expected return model in which the matched characteristics (e.g., size and book-to-market) perfectly proxy for the expected return on a security. Since corporate events themselves are unlikely to be random occurrences, i.e., they are unlikely to be exogenous with respect to past performance and expected returns, there is a danger that the event and nonevent samples differ systematically in their expected returns notwithstanding the matching on certain firm characteristics. This makes matching on (unobservable) expected returns more difficult, especially in the case of event firms experiencing extreme prior performance.

Once a matching firm or portfolio is identified, BHAR calculation is straightforward. A $T$-month BHAR for event firm $i$ is defined as:

$$\mathrm{BHAR}_i(t, T) = \prod_{t=1 \text{ to } T} (1 + R_{i,t}) - \prod_{t=1 \text{ to } T} (1 + R_{\mathrm{B},t}), \tag{7}$$

where $R_\mathrm{B}$ is the return on either a non-event firm that is matched to the event firm $i$, or it is the return on a matched (benchmark) portfolio.[13] If the researcher believes that the Carhart (1997) four-factor model is an adequate description of expected returns, then firm-specific matching might entail identifying a non-event firm that is closest to an event firm on the basis of firm size (i.e., market capitalization of equity), book-to-market ratio, and past one-year return. Alternatively, characteristic portfolio matching would identify the portfolio of all non-event stocks that share the same quintile ranking on size, book-to-market, and momentum as the event firm (see Daniel et al., 1997, or Lyon, Barber, and Tsai, 1999, for details of benchmark portfolio construction). The return on the matched portfolio is the benchmark portfolio return, $R_\mathrm{B}$. For the sample of event firms, the mean BHAR is calculated as the (equal- or value-weighted) average of the individual firm BHARs.

### *4.3.2. Jensen-alpha approach*

The Jensen-alpha approach (or the calendar-time portfolio approach) to estimating risk-adjusted abnormal performance is an alternative to the BHAR calculation using

[12] Apart from similarity with the actual investment experience, the BHAR approach also avoids biases arising from security microstructure issues when portfolio performance is measured with frequent rebalancing (see Blume and Stambaugh, 1983; Roll, 1983; Ball, Kothari, and Shanken, 1995). The latter biases are also reduced if value-weight portfolio performance is examined.

[13] See Mitchell and Stafford (2000) for details.

a matched-firm approach to risk adjustment. Jaffe (1974) and Mandelker (1974) introduced a calendar time methodology to the financial-economics literature, and it has since been advocated by many, including Fama (1998) and Mitchell and Stafford (2000).[14] The distinguishing feature of the most recent variants of the approach is to calculate calendar-time portfolio returns for firms experiencing an event, and calibrate whether they are abnormal in a multifactor (e.g., CAPM or Fama–French three factor) regression. The estimated intercept from the regression of portfolio returns against factor returns is the post-event abnormal performance of the sample of event firms.

To implement the Jensen-alpha approach, assume a sample of firms experiences a corporate event (e.g., an IPO or an SEO).[15] The event might be spread over several years or even many decades (the sample period). Also assume that the researcher seeks to estimate price performance over two years ($T = 24$ months) following the event for each sample firm. In each calendar month over the entire sample period, a portfolio is constructed comprising all firms experiencing the event within the previous $T$ months. Because the number of event firms is not uniformly distributed over the sample period, the number of firms included in a portfolio is not constant through time. As a result, some new firms are added each month and some firms exit each month. Accordingly, the portfolios are rebalanced each month and an equal or value-weighted portfolio excess return is calculated. The resulting time series of monthly excess returns is regressed on the CAPM market factor, or the three Fama and French (1993) factors, or the four Carhart (1997) factors as follows:

$$\begin{aligned} R_{pt} - R &= a_p + b_p(R_{mt} - R) \\ &\qquad + s_p\,\mathrm{SMB}_t + h_p\,\mathrm{HML}_t + m_p\,\mathrm{UMD}_t + e_{pt}, \end{aligned} \tag{8}$$

where

- $R_{pt}$ is the equal or value-weighted return for calendar month $t$ for the portfolio of event firms that experienced the event within the previous $T$ months;
- $R_{ft}$ is the risk-free rate;
- $R_{mt}$ is the return on the CRSP value-weight market portfolio;
- $\mathrm{SMB}_{pt}$ is the difference between the return on the portfolio of "small" stocks and "big" stocks;
- $\mathrm{HML}_{pt}$ is the difference between the return on the portfolio of "high" and "low" book-to-market stocks;
- $\mathrm{UMD}_{pt}$ is the difference between the return on the portfolio of past one-year "winners" and "losers";
- $a_p$ is the average monthly abnormal return (Jensen alpha) on the portfolio of event firms over the $T$-month post-event period,
- $b_p$, $s_p$, $h_p$, and $m_p$ are sensitivities (betas) of the event portfolio to the four factors.

[14] For a variation of the Jensen-alpha approach, see Ibbotson (1975) *returns across time and securities* (RATS) methodology, which is used in Ball and Kothari (1989) and others.

[15] The description here is based on Mitchell and Stafford (2000).

Inferences about the abnormal performance are on the basis of the estimated $a_p$ and its statistical significance. Since $a_p$ is the average monthly abnormal performance over the $T$-month post-event period, it can be used to calculate annualized post-event abnormal performance.

Recent work on the implications of using the Jensen-alpha approach is mixed. For example, Mitchell and Stafford (2000) and Brav and Gompers (1997) favor the Jensen-alpha approach. However, Loughran and Ritter (2000) argue against using the Jensen-alpha approach because it might be biased toward finding results consistent with market efficiency. Their rationale is that corporate executives time the events to exploit mispricing, but the Jensen-alpha approach, by forming calendar-time portfolios, under-weights managers' timing decisions and over-weights other observations. In the words of Loughran and Ritter (2000, p. 362): "If there are time-varying misvaluations that firms capitalize on by taking some action (a supply response), there will be more events involving larger misvaluations in some periods than in others... In general, tests that weight firms equally should have more power than tests that weight each time period equally". Since the Jensen-alpha (i.e., calendar-time) approach weights each period equally, it has lower power to detect abnormal performance if managers time corporate events to coincide with misvaluations. As a means of addressing the problem, Fama (1998) advocates weighting calendar months by their statistical precision, which varies with sample size. Countering the criticism of Loughran and Ritter (2000), Eckbo, Masulis, and Norli (2000) point out another problem with the buy-and-hold abnormal return methods. The latter is not a feasible portfolio strategy because the total number of securities is not known in advance.[16]

### *4.4. Significance tests for BHAR and Jensen-alpha measures*

The choice between the matched-firm BHAR approach to abnormal return measurement and the calendar time Jensen-alpha approach (also known as the calendar-time portfolio approach) hinges on the researcher's ability to accurately gauge the statistical significance of the estimated abnormal performance using the two approaches. Unbiased standard errors for the distribution of the event-portfolio abnormal returns are not easy to calculate, which leads to test misspecification. Assessing the statistical significance of the event portfolio's BHAR has been particularly difficult because (i) long-horizon returns depart from the normality assumption that underlies many statistical tests; (ii) long-horizon returns exhibit considerable cross-correlation because the

[16] The BHAR approach is also criticized for "pseudo-timing" because BHAR mechanically produces underperformance following a clustering of issues experiencing a common event, e.g., an IPO, in an up or down market (Schultz, 2003; Eckbo and Norli, 2005). The criticism assumes that those seeking to exploit the event-related market inefficiency do not have market-timing ability. The question of pseudo-timing and return predictability is a topic of intense current interest and appears currently unresolved (Baker, Talliaferro, and Wurgler, 2004, 2006; Goyal and Welch, 2003, 2005; Boudoukh, Richardson, and Whitelaw, 2006; Cochrane, 2006).

return horizons of many event firms overlap and also because many event firms are drawn from a few industries; and (iii) volatility of the event firm returns exceeds that of matched firms because of event-induced volatility. We summarize below the econometric inferential issues encountered in performing long-horizon tests and some of the remedies put forward in recent studies.

#### *4.4.1. Skewness*

Long-horizon buy-and-hold returns, even after adjusting for the performance of a matched firm (or portfolio), tend to be right skewed. The right skewness of buy-and-hold returns is not surprising because the lower bound is $-100\%$ and returns are unbounded on the upside. Skewness in abnormal returns imparts a skewness bias to long-horizon abnormal performance test statistics (see Barber and Lyon, 1997). Brav (2000, p. 1981) concludes that "with a skewed-right distribution of abnormal returns, the Student $t$-distribution is asymmetric with a mean smaller than the zero null". While the right-skewness of individual firms' long-horizon returns is undoubtedly true, the extent of skewness bias in the test statistic for the hypothesis that mean abnormal performance for the portfolio of event firms is zero is expected to decline with sample size.[17] Fortunately, the sample size in long-horizon event studies is often several hundred observations (e.g., Teoh, Welch, and Wong, 1998, and Byun and Rozeff, 2003). Therefore, if the BHAR observations for the sample firms are truly independent, as assumed in using a $t$-test, the Central Limit Theorem's implication that "the sum of a large number of independent random variables has a distribution that is approximately normal" should apply (Ross, 1976, p. 252). The right-skewness of the distribution of long-horizon abnormal returns on event *portfolios*, as documented in, for example, Brav (2000) and Mitchell and Stafford, 2000, appears to be due largely to the lack of independence arising from overlapping long-horizon return observations in event portfolios. That is, skewness in portfolio returns is in part a by-product of cross-correlated data rather than a direct consequence of skewed firm-level buy-and-hold abnormal (or raw) returns.

#### *4.4.2. Cross-correlation*

*4.4.2.1. The issue* Specification bias arising due to cross-correlation in returns is a serious problem in long-horizon tests of price performance. Brav (2000, p. 1979) attributes the misspecification to the fact that researchers conducting long-horizon tests typically "maintain the standard assumptions that abnormal returns are independent and normally distributed although these assumptions fail to hold even approximately

[17] Simulation evidence in Barber and Lyon (1997) on skewness bias is based on samples consisting of 50 firms and early concern over skewness bias as examined in Neyman and Pearson (1928) and Pearson (1929a, 1929b) also refers to skewness bias in small samples.

at long horizons".[18] The notion that economy-wide and industry-specific factors would generate contemporaneous co-movements in security returns is the cornerstone of portfolio theory and is economically intuitive and empirically compelling. Interestingly, the cross-dependence, although muted, is also observed in risk-adjusted returns.[19] The degree of cross-dependence decreases in the effectiveness of the risk-adjustment approach and increases in the homogeneity of the sample firms examined (e.g., sample firms clustered in one industry). Cross-correlation in abnormal returns is largely irrelevant in short-window event studies when the event is not clustered in calendar time. However, in long-horizon event studies, even if the event is not clustered in calendar time, cross-correlation in abnormal returns cannot be ignored (Brav, 2000; Mitchell and Stafford, 2000; Jegadeesh and Karceski, 2004). Long-horizon abnormal returns tend to be cross-correlated because: (i) abnormal returns for subsets of the sample firms are likely to share a common calendar period due to the long measurement period; (ii) corporate events like mergers and share repurchases exhibit waves (for rational economic reasons as well as opportunistic actions on the part of the shareholders and/or management); and (iii) some industries might be over-represented in the event sample (e.g., merger activity among technology stocks).

If the test statistic in an event study is calculated ignoring cross-dependence in data, even a fairly small amount of cross-correlation in data will lead to serious misspecification of the test. In particular, the test will reject the null of no effect far more often than the size of the test (Collins and Dent, 1984; Bernard, 1987; Mitchell and Stafford, 2000). The overrejection is caused by the downward biased estimate of the standard deviation of the cross-sectional distribution of buy-and-hold abnormal returns for the event sample of firms.

*4.4.2.2. Magnitude of bias* To get an idea of approximate magnitude of the bias, we begin with the cross-sectional standard deviation of the event firms' abnormal returns, AR, assuming equal variances and pairwise covariances across all sample firms' abnormal returns:

$$\sigma_{\mathrm{AR}} = \left[\frac{1}{N}\sigma^2 + \frac{N-1}{N}\rho_{i,j}\sigma^2\right]^{1/2}, \tag{9}$$

where $N$ is the number of sample firms, $\sigma^2$ is the variance of abnormal returns, which is assumed to be the same for all firms; and $\rho_{i,j}$ is the correlation between firm $i$ and $j$'s abnormal returns, which is also assumed to be the same across all firms. The second term in the square brackets in equation (9) is due to the cross-dependence in the data, and it would be absent if the standard deviation is calculated assuming independence

[18] Also see Barber and Lyon (1997), Kothari and Warner (1997), Fama (1998), Lyon, Barber, and Tsai (1999), Mitchell and Stafford (2000), and Jegadeesh and Karceski (2004).

[19] See Schipper and Thompson (1983), Collins and Dent (1984), Sefcik and Thompson (1986), Bernard (1987), Mitchell and Stafford (2000), Brav (2000), and Jegadeesh and Karceski (2004).

in the data. The bias in the standard deviation assuming independence is given by the ratio of the "true" standard deviation allowing for dependence to the standard deviation assuming independence:

$$\frac{\sigma_{\mathrm{AR}}\ \text{(Dependence)}}{\sigma_{\mathrm{AR}}\ \text{(Independence)}} = \left[1 + (N-1)\rho_{i,j}\right]^{1/2}. \tag{10}$$

The ratio in equation (10) is the factor by which the standard error in a test for the significance of abnormal performance is understated and therefore the factor by which the test statistic (e.g., $t$-statistic) itself is overstated. The ratio is increasing in the pairwise cross-correlation, $\rho_{i,j}$. Empirical estimates of the average pairwise correlation between annual BHARs of event firms are about 0.02 to 0.03 (see Mitchell and Stafford, 2000). The average pairwise correlation in multi-year BHARs is likely to be greater than that for annual returns because Bernard (1987, Table 1) reports that the average cross-correlations increase with return horizon. Assuming the average pairwise cross-sectional correlation to be only 0.02, for a sample of 100, the ratio in equation (4) is 1.73, and it increases with both sample size and the degree of cross-correlation. Since the sample size in many long-horizon event studies is a few hundred securities, and the BHAR horizon is three-to-five years, even a modest degree of average cross-correlation in the data can inflate the test statistics by a factor of two or more. Therefore, accounting for cross-correlation in abnormal returns is crucial to drawing accurate statistical inferences in long-horizon event studies. Naturally, this has been a subject of intense interest among researchers.

*4.4.2.3. Potential solutions* One simple solution to the potential bias due to cross-correlation is to use the Jensen-alpha approach. It is immune to the bias arising from cross-correlated (abnormal) returns because of the use of calendar-time portfolios. Whatever the correlation among security returns, the event portfolio's time series of returns in calendar time accounts for that correlation. That is, the variability of portfolio returns is influenced by the cross-correlation in the data. The statistical significance of the Jensen alpha is based on the time-series variability of the portfolio return residuals. Since returns in an efficient market are serially uncorrelated (absent nontrading), on this basis the independence assumption in calculating the standard error and the $t$-statistic for the regression intercept (i.e., the Jensen alpha) seems quite appropriate. However, the evidence is that this method is misspecified in nonrandom samples (Lyon, Barber, and Tsai, 1999, Table 10). This is unfortunate, given that the method seems simple and direct. The reasons for the misspecification are unclear (see Lyon, Barber, and Tsai, 1999). Appropriate calibration under calendar time methods probably warrants further investigation.

In the BHAR approach, estimating standard errors that account for the cross-correlation in long-horizon abnormal returns is not straightforward. As detailed below, there has been much discussion, and some interesting progress. Statistically precise estimates of pairwise cross-correlations are difficult to come by for the lack of availability of many time-series observations of long-horizon returns to accurately estimate the correlations (see Bernard, 1987). The difficulty is exacerbated by the fact that only a portion of the post-event-period might overlap with other firms. Researchers have developed bootstrap and pseudoportfolio-based statistical tests that might account for the cross-correlations and lead to accurate inferences.

*4.4.2.4. Cross-correlation and skewness* Lyon, Barber, and Tsai (1999) develop a bootstrapped skewness-adjusted $t$-statistic to address the cross-correlation and skewness biases. The first step in the calculation is the skewness-adjusted $t$-statistic (see Johnson, 1978). This statistic adjusts the usual $t$-statistic by two terms that are a function of the skewness of the distribution of abnormal returns (see equation (5) in Lyon, Barber, and Tsai, 1999, p. 174). Notwithstanding the skewness adjustment, the adjusted $t$-statistic indicates overrejection of the null and thus warrants a further refinement. The second step, therefore, is to construct a bootstrapped distribution of the skewness-adjusted $t$-statistic (Sutton, 1993; Lyon, Barber, and Tsai, 1999). To bootstrap the distribution, a researcher must draw a large number (e.g., 1,000) of resamples from the original sample of abnormal returns and calculate the skewness-adjusted $t$-statistic using each resample. The resulting empirical distribution of the test statistics is used to ascertain whether the skewness-adjusted $t$-statistic for the original event sample falls in the $\alpha\%$ tails of the distribution to reject the null hypothesis of zero abnormal performance.

The pseudoportfolio-based statistical tests infer statistical significance of the event sample's abnormal performance by calibrating against an empirical distribution of abnormal performance constructed using repeatedly-sampled pseudoportfolios.[20] The empirical distribution of average abnormal returns on the pseudoportfolios is under the null hypothesis of zero abnormal performance. The empirical distribution is generated by repeatedly constructing matched firm samples with replacement. The matching is on the basis of characteristics thought to be correlated with the expected rate of return. Following the Fama and French (1993) three-factor model, matching on size and book-to-market as expected return determinants is quite common (e.g., Lyon, Barber, and Tsai, 1999, Byun and Rozeff, 2003, and Gompers and Lerner, 2003). For each matched-sample portfolio, an average buy-and-hold abnormal performance is calculated as the raw return minus the benchmark portfolio return. It's quite common to use 1,000 to 5,000 resampled portfolios to construct the empirical distribution of the average abnormal returns on the matched-firm samples. This distribution yields empirical 5 and 95% cut-off probabilities against which the event-firm sample's performance is calibrated to infer whether or not the event-firm portfolio buy-and-hold abnormal return is statistically significant.

Unfortunately, the two approaches described above, which are aimed at correcting the bias in standard errors due to cross-correlated data, are not quite successful in their intended objective. Lyon et al. find pervasive test misspecification in non-random samples.

[20] See, for example, Brock, Lakonishok, and LeBaron (1992), Ikenberry, Lakonishok, and Vermaelen (1995), Ikenberry, Rankine and Stice (1996), Lee (1997), Lyon, Barber, and Tsai (1999), Mitchell and Stafford (2000), and Byun and Rozeff (2003).

Because the sample of firms experiencing a corporate event is not selected randomly by the researcher, correcting for the bias in the standard errors stemming from the non-randomness of the event sample selection is not easy. In a strident criticism of the use of bootstrap- and pseudoportfolio-based tests, Mitchell and Stafford (2000, p. 307) conclude that long-term event studies often incorrectly "claim that bootstrapping solves all dependence problems. However, that claim is not valid. Event samples are clearly different from random samples. Event firms have chosen to participate in a major corporate action, while nonevent firms have chosen to abstain from the action. An empirical distribution created by randomly selecting firms with similar size-BE/ME characteristics does not replicate the covariance structure underlying the original event sample. In fact, the typical bootstrapping approach does not even capture the cross-sectional correlation structure related to industry effects...". Jegadeesh and Karceski (2004, pp. 1–2) also note that the Lyon, Barber, and Tsai (1999) approach is misspecified because it "assumes that the observations are cross-sectionally uncorrelated. This assumption holds in random samples of event firms, but is violated in nonrandom samples. In nonrandom samples where the returns for event firms are positively correlated, the variability of the test statistics is larger than in a random sample. Therefore, if the empiricist calibrates the distribution of the test statistics in random samples and uses the empirical cutoff points for nonrandom samples, the tests reject the null hypothesis of no abnormal performance too often".

*4.4.2.5. Autocorrelation* To overcome the weaknesses in prior tests, Jegadeesh and Karceski (2004) propose a correlation and heteroskedasticity-consistent test. The key innovation in their approach is to estimate the cross-correlations using a monthly time-series of portfolio long-horizon returns (see Jegadeesh and Karceski, 2004, Section II.A for details). Because the series is monthly, but the monthly observations contain long-horizon returns, the time-series exhibits autocorrelation that is due to overlapping return data. The autocorrelation is, of course, due to cross-correlation in return data. The autocorrelation is expected to be positive for $H - 1$ lags, where $H$ is the number of months in the long horizon. The length of the time-series of monthly observations depends on the sample period during which corporate events being examined take place. Because of autocorrelation in the time series of monthly observations, the usual $t$-statistic that is a ratio of the average abnormal return to the standard deviation of the time series of the monthly observations would be understated. To obtain an unbiased $t$-statistic, the covariances (i.e., the variance–covariance matrix) should be taken into account. Jegadeesh and Karceski (2004) use the Hansen and Hodrick (1980) estimator of the variance–covariance matrix assuming homoskedasticity. They also use a heteroskedasticity-consistent estimator that "generalizes White's heteroskedasticity-consistent estimator and allows for serial covariances to be non-zero" (p. 8). In both random and non-random (industry) samples the Jegadeesh and Karceski (2004) tests perform quite well, and we believe these might be the most appropriate to reduce misspecification in tests of long-horizon event studies.

#### *4.4.3. The bottom line*

Despite positive developments in BHAR calibration methods, two general long-horizon problems remain. The first concerns power. Jegadeesh and Karceski (2004) report that their tests show no increase in power relative to that of the test employed in previous research, which already had low power. For example, even with seemingly huge cumulative abnormal performance (25% over 5 years) in a sample of 200 firms, the rejection rate of the null is typically under 50% (see their Table 6).

Second, specification issues remain. For example, as discussed earlier (Section 3.6), events are generally likely to be associated with variance increases, which are equivalent to abnormal returns varying across sample securities. Previous literature shows that variance increases induce misspecification, and can cause the null hypothesis to be rejected far too often. Thus, whether a high level of measured abnormal performance is due to chance or mispricing (or a bad model) is still difficult to empirically determine, unless the test statistic is adjusted downward to reflect the variance shift. Solutions to the variance shift issue include such intuitive procedures as forming subsamples with common characteristics related to the level of abnormal performance (e.g., earnings increase vs. decrease subsamples). With smaller subsamples, however, specification issues unrelated to variance shifts become more relevant. Moreover, the importance of examining specification for nonrandom samples cannot be overemphasized.

Given the various power and specification issues, a challenge that remains for the profession is to continue to refine long-horizon methods. Whether calendar time, BHAR methods or some combination can best address long-horizon issues remains an open question.

## References

Baker, M., Talliaferro, R., Wurgler, J., 2006. Predicting returns with managerial decision variables: Is there a large-sample bias? Journal of Finance, in press.

Baker, M., Talliaferro, R., Wurgler, J., 2004. Pseudo market timing and predictive regression. Working Paper #10823, National Bureau of Economic Research.

Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. Journal of Accounting Research 6, 159–177.

Ball, R., Kothari, S., 1989. Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns. Journal of Financial Economics 25, 51–74.

Ball, R., Kothari, S., Shanken, J., 1995. Problems in measuring portfolio performance: An application to contrarian investment strategies. Journal of Financial Economics 38, 79–107.

Banz, R., 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9, 3–18.

Barber, B., Lyon, J., 1996. Detecting abnormal operating performance: The empirical power and specification of test statistics. Journal of Financial Economics 41, 359–399.

Barber, B., Lyon, J., 1997. Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. Journal of Financial Economics 43, 341–372.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

Basu, S., 1977. The investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient markets hypothesis. Journal of Finance 32, 663–682.

Basu, S., 1983. The relationship between earnings yield, market value, and returns for NYSE common stocks: Further evidence. Journal of Financial Economics 12, 129–156.

Beaver, W., 1968. The information content of annual earnings announcements. Journal of Accounting Research Supplement 6, 67–92.

Bernard, V., 1987. Cross-sectional dependence and problems in inference in market-based accounting research. Journal of Accounting Research 25, 1–48.

Blume, M., Stambaugh, R., 1983. Biases in computed returns: an application to the size effect. Journal of Financial Economics 12, 387–404.

Boudoukh, J., Richardson, M., Whitelaw, R., 2006. The myth of long-horizon predictability. Working Paper, New York University.

Brav, A., 2000. Inference in long-horizon event studies: A Bayesian approach with application to initial public offerings. Journal of Finance 55, 1979–2016.

Brav, A., Gompers, P., 1997. Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies. Journal of Finance 52, 1791–1821.

Brennan, M., Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of Financial Economics 41, 441–464.

Brock, W., Lakonishok, J., LeBaron, B., 1992. Simple trading rules and the stochastic properties of stock returns. Journal of Finance 47, 1731–1764.

Brown, S., Warner, J., 1980. Measuring security price performance. Journal of Financial Economics 8, 205–258.

Brown, S., Warner, J., 1985. Using daily stock returns: The case of event studies. Journal of Financial Economics 14, 3–31.

Byun, J., Rozeff, M., 2003. Long-run performance after stock splits: 1927 to 1996. Journal of Finance 58, 1063–1085.

Campbell, C., Wasley, C., 1996. Measuring abnormal trading volume for samples of NYSE/ASE and NASDAQ securities using parametric and nonparametric test statistics. Review of Quantitative Finance and Accounting.

Campbell, C., Wasley, C., 1993. Measuring security price performance using daily NASDAQ returns. Journal of Financial Economics 33, 73–92.

Campbell, J., Lettau, M., Malkiel, B., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. Journal of Finance 56, 1–43.

Campbell, J., Lo, A., MacKinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ.

Carhart, M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Chan, K., 1988. On the contrarian investment strategy. Journal of Business 61, 147–163.

Chopra, N., Lakonishok, J., Ritter, J., 1992. Measuring abnormal performance: Does the market overreact. Journal of Financial Economics 32, 235–268.

Chordia, T., Goyal, A., Sadka, G., Sadka, R., Shivakumar, L., 2006. Liquidity and the post-earnings-announcement drift. Working Paper, LBS.

Cochrane, J., 2006. The dog that did not bark: A defense of return predictability. Working Paper, University of Chicago and NBER.

Collins, D., Dent, W., 1984. A comparison of alternative testing methodologies used in capital market research. Journal of Accounting Research 22, 48–84.

Corrado, C., 1989. A nonparametric test for abnormal security-price performance in event studies. Journal of Financial Economics 23, 385–395.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035–1058.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53, 1839–1885.

Daniel, K., Hirshleifer, D., Teoh, S., 2002. Investor psychology in capital markets: Evidence and policy implications. Journal of Monetary Economics 49, 139–209.
De Long, J., Shleifer, A., Vishny, R., Waldman, R., 1990a. Noise trader risk in financial markets. Journal of Political Economy 98, 703–738.
De Long, J., Shleifer, A., Vishny, R., Waldman, R., 1990b. Positive feedback investment strategies and destabilizing rational speculation. Journal of Finance 45, 375–395.
DeBondt, W., Thaler, R., 1985. Does the stock market overreact? Journal of Finance 40, 793–805.
DeBondt, W., Thaler, R., 1987. Further evidence of investor overreaction and stock market seasonality. Journal of Finance 42, 557–581.
Dechow, P., Sloan, R., Sweeney, A., 1995. Detecting earnings management. The Accounting Review 70, 3–42.
Eckbo, B., Maksimovic, V., Williams, J., 1990. Consistent estimation of cross-sectional models in event studies. Review of Financial Studies 3, 343–365.
Eckbo, B., Masulis, R., Norli, O., 2000. Seasoned public offerings: Resolution of the 'New Issues Puzzle'. Journal of Financial Economics 56, 251–291.
Eckbo, B., Masulis, R., Norli, O., 2006. Security offerings. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. In: Handbook in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 6 (this volume).
Eckbo, B., Norli, O., 2005. Liquidity risk, leverage and long-run IPO returns. Journal of Corporate Finance 11, 1–35.
Fama, E., 1970. Efficient capital markets: A review of theory and empirical work. Journal of Finance 25, 383–417.
Fama, E., 1991. Efficient capital markets: II. Journal of Finance 46, 1575–1617.
Fama, E., 1998. Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics 49, 283–306.
Fama, E., Fisher, L., Jensen, M., Roll, R., 1969. The adjustment of stock prices to new information. International Economic Review 10, 1–21.
Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.
Gompers, P., Lerner, J., 2003. The really long-run performance of initial public offerings: The pre-Nasdaq evidence. Journal of Finance 58, 1355–1392.
Goyal, A., Welch, I., 2003. Predicting the equity premium with dividend ratios. Management Science 49, 639–654.
Goyal, A., Welch, I., 2005. A comprehensive look at the empirical performance of equity premium prediction. Working Paper, Brown University.
Hansen, L., Hodrick, R., 1980. Forward rates as optimal predictors of future spot rates: An econometric analysis. Journal of Political Economy 88, 829–853.
Hirshleifer, D., 2001. Investor psychology and asset pricing. Journal of Finance 56, 1533–1598.
Hong, H., Stein, J., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54, 2143–2184.
Ibbotson, R., 1975. Price performance of common stock new issues. Journal of Financial Economics 2, 235–272.
Ikenberry, D., Lakonishok, J., Vermaelen, T., 1995. The underreaction to open market share repurchases. Journal of Financial Economics 39, 181–208.
Ikenberry, D., Rankine, G., Stice, E., 1996. What do stock splits really signal? Journal of Financial and Quantitative Analysis 31, 357–376.
Jaffe, J., 1974. Special information and insider trading. Journal of Business 47, 411–428.
Jarrell, G., Brickley, J., Netter, J., 1988. The market for corporate control—The empirical evidence since 1980. Journal of Economic Perspectives 2, 49–68.
Jegadeesh, N., Karceski, J., 2004. Long-run performance evaluation: Correlation and heteroskedasticity-consistent tests. Working Paper, Emory University.

Jensen, M., Ruback, R., 1983. The market for corporate control—The scientific evidence. Journal of Financial Economics 11, 5–50.
Jensen, M., Warner, J., 1988. The distribution of power among corporate managers, shareholders, and directors. Journal of Financial Economics 20, 3–24.
Johnson, N., 1978. Modified *t* tests and confidence intervals for symmetrical populations. Journal of American Statistical Association 73, 536–544.
Jones, C., Litzenberger, R., 1970. Quarterly earnings reports and intermediate stock price trends. Journal of Finance 25, 143–148.
Kothari, S., 2001. Capital markets research in accounting. Journal of Accounting and Economics 31, 105–231.
Kothari, S., Leone, A., Wasley, C., 2005. Performance-matched discretionary accruals. Journal of Accounting and Economics 39, 163–197.
Kothari, S., Warner, J., 1997. Measuring long-horizon security price performance. Journal of Financial Economics 43, 301–339.
Lakonishok, J., Shleifer, A., Vishny, R., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541–1578.
Lee, I., 1997. Do firms knowingly sell overvalued equity? Journal of Finance 52, 1439–1466.
Li, K., Prabhala, N., 2007. Self-selection models in corporate finance. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. Elsevier/North-Holland, Amsterdam, Chapter 2 (this volume).
Loughran, T., Ritter, J., 2000. Uniformly least powerful tests of market efficiency. Journal of Financial Economics 55, 361–389.
Lyon, J., Barber, B., Tsai, C., 1999. Improved methods of tests of long-horizon abnormal stock returns. Journal of Finance 54, 165–201.
Malatesta, P., Thompson, R., 1985. Partially anticipated events: A model of stock price reactions with an application to corporate acquisitions. Journal of Financial Economics 14, 237–250.
MacKinlay, A.C., 1997. Event studies in economics and finance. Journal of Economic Literature 35, 13–39.
Mandelker, G., 1974. Risk and return: The case of merging firms. Journal of Financial Economics 1, 303–335.
Mitchell, M., Stafford, E., 2000. Managerial decisions and long-term stock price performance. Journal of Business 73, 287–329.
Neyman, J., Pearson, E., 1928. On the use and interpretation of certain test criteria for purposes of statistical inference, part I. Biometrika 20A, 175–240.
Ng, J., Rusticus, T., Verdi, R., 2006. Implications of transaction costs for the post-earnings-announcement drift. Working Paper, MIT Sloan School of Management.
Pastor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.
Patell, J., 1976. Corporate forecasts of earnings per share and stock price behavior: Empirical tests. Journal of Accounting Research 14, 246–276.
Pearson, E., 1929a. The distribution of frequency constants in small samples from symmetrical distributions. Biometrika 21, 356–360.
Pearson, E., 1929b. The distribution of frequency constants in small samples from non-normal symmetrical and skew populations. Biometrika 21, 259–286.
Petersen, M., 2005. Estimating standard errors in finance panel data sets: Comparing approaches. Working Paper, Northwestern University.
Roll, R., 1983. On computing mean returns and the small firm premium. Journal of Financial Economics 12, 371–386.
Ross, S., 1976. A First Course in Probability. Macmillan, New York.
Sadka, R., 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. Journal of Financial Economics 80, 309–349.
Schipper, K., Thompson, R., 1983. The impact of merger-related relationships on the shareholders of acquiring firms. Journal of Accounting Research 21, 184–221.
Schultz, P., 2003. Pseudo market timing and the long-run underperformance of IPOs. Journal of Finance 58, 483–517.

Schwert, G.W., 2001. Anomalies and market efficiency. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. North-Holland, Amsterdam, pp. 939–974.
Sefcik, S., Thompson, R., 1986. An approach to statistical inference in cross-sectional models with security abnormal returns as dependent variables. Journal of Accounting Research 24, 316–334.
Shleifer, A., 2000. Inefficient Markets: An Introduction to Behavioral Finance. Oxford University Press, London.
Shliefer, A., Vishny, R., 1997. The limits of arbitrage. Journal of Finance 52, 35–55.
Smith, C., 1986. Investment banking and the capital acquisition process. Journal of Financial Economics 15, 3–29.
Sutton, C., 1993. Computer-intensive methods for tests about the mean of an asymmetrical distribution. Journal of American Statistical Association 88, 802–808.
Teoh, S., Welch, I., Wong, T., 1998. Earnings management and the long-run performance of initial public offerings. Journal of Finance 53, 1935–1974.
Thompson, R., 1995. Empirical methods of event studies in corporate finance. In: Jarrow, R., Maksimovic, V., Ziemba, B. (Eds.), In: Finance. Series of Handbooks in Operations Research and Management Science. North-Holland, Amsterdam, pp. 1017–1072.
Viswanathan, S., Wei, B., 2004. Endogenous event and long run returns. Working Paper, Duke University.

*Chapter 2*

# SELF-SELECTION MODELS IN CORPORATE FINANCE*

KAI LI

*Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC, V6T 1Z2 Canada*
*e-mail: kai.li@sander.ubc.ca*

NAGPURNANAND R. PRABHALA

*Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA*
*e-mail: prabhala@rhsmith.umd.edu*

## Contents



* We thank N.K. Chidambaran, Craig Doidge, Espen Eckbo, Andrew Karolyi, Gordon Phillips, Vojislav Maksimovic, Jeffrey Smith, and Xinlei Zhao without implicating them for any errors or omissions, which remain ours. Li acknowledges the financial support from the Social Sciences and Humanities Research Council of Canada, and the W.M. Young Chair in Finance from the Sauder School of Business at UBC. Li also wishes to thank the support and hospitality of the MIT Sloan School of Management where she completed most of her work on this chapter.

*Handbook of Corporate Finance, Volume 1*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01002-6*

**Abstract**

Corporate finance decisions are not made at random, but are usually deliberate decisions by firms or their managers to *self-select* into their preferred choices. This chapter reviews econometric models of self-selection. The review is organized into two parts. The first part reviews econometric models of self-selection, focusing on the key assumptions of different models and the types of applications they may be best suited for. Part two reviews empirical applications of selection models in the areas of corporate investment, financing, and financial intermediation. We find that self-selection is a rapidly growing area in corporate finance, partly reflecting its recognition as a pervasive feature of corporate finance decisions, but more importantly, the increasing recognition of selection models as unique tools for understanding, modeling, and testing the role of private information in corporate finance.

**Keywords**

selection, private information, switching regression, treatment effect, matching, propensity score, Bayesian selection methods, panel data, event study, underwriting, investment banking, diversification

## Introduction

Corporate finance concerns the financing and investment choices made by firms and a broad swathe of decisions within these broad choices. For instance, firms pick their target capital structure, and to achieve the target, must make several choices including issue timing of security issues, structural features of the securities issued, the investment bank chosen to underwrite it, and so on. These choices are not usually random, but are deliberate decisions by firms or their managers to *self-select* into their preferred choices. This chapter reviews econometric models of self-selection. We review the approaches used to model self-selection in corporate finance and the substantive findings obtained by implementing selection methods.

Self-selection has a rather mixed history in corporate finance. The fact that there is self-selection is probably not news; indeed, many papers at least implicitly acknowledge its existence. However, the literature differs on whether to account for self-selection using formal econometric methods, and why one should do so. One view of self-selection is that it is an errant nuisance, a "correction" that must be made to prevent other parameter estimates from being biased. Selection is itself of little economic interest under this view. In other applications, self-selection is itself of central economic interest, because models of self-selection represent one way of incorporating and controlling for unobservable private information that influences corporate finance decisions. Both perspectives find expression in the literature, although an increasing emphasis in recent work reflects the positive view in which selection models are used to construct interesting tests for private information.

Our review is organized into two parts. Part I focuses on econometric models of self-selection. We approach selection models from the viewpoint of a corporate finance researcher who is implementing selection models in an empirical application. We formalize the notion of self-selection and overview several approaches towards modeling it, including reduced form models, structural approaches, matching methods, fixed effect estimators, and Bayesian methods. As the discussion clarifies, the notion of selection is not monolithic. No single model universally models or accounts for all forms of selection, so there is no one "fix" for selection. Instead, there are a variety of approaches, each of which makes its own economic and statistical assumptions. We focus on the substantive economic assumptions underlying the different approaches to illustrate what each can and cannot do and the type of applications a given approach may be best suited for. We do not say much on estimation, asymptotic inference, or computational issues, but refer the reader to excellent texts and articles on these matters.

Part II of our review examines corporate finance applications of self-selection models. We cover a range of topics such as mergers and acquisitions, stock splits, equity offerings, underwriting, analyst behavior, share repurchases, and venture capital. Our objective is to illustrate the wide range of corporate finance settings in which selection arises and the different econometric approaches employed in modeling it. Here, we focus on applications published in the last decade or so, and on articles in which self-selection is a major component of the overall results.[1]

## I. MODELING SELF-SELECTION

This portion of our review discusses econometric models of self-selection. Our intention is not to summarize the entire range of available models and their estimation. Rather, we narrow our focus to models that have been applied in the corporate finance literature, and within these models, we focus on the substantive assumptions made by each specification. From the viewpoint of the empirical researcher, this is the first order issue in deciding what approach suits a given application in corporate finance. We do not touch upon asymptotic theory, estimation, and computation. These important issues are well covered in excellent textbooks.[2]

We proceed as follows. Section 1 describes the statistical issue raised by self-selection, the wedge between the population distribution and the distribution within a selected sample. Sections 2–6 develop the econometric models that can address selection. Section 2 discusses a baseline model for self-selection, the "Heckman" selection model analyzed in Heckman (1979), a popular modeling choice in corporate finance.[3] We discuss identification issues related to the model, which are important but not frequently discussed or justified explicitly in corporate finance applications. Because the Heckman setting is so familiar in corporate finance, we use it to develop a key point of this survey, the analogy between econometric models of self-selection and private information models in corporate finance. Section 3 considers switching regressions and structural self-selection models. While these models generalize the Heckman selection model in some ways, they also bring additional baggage in terms of economic and statistical assumptions that we discuss.

We then turn to other approaches towards modeling selection. Section 4 discusses matching models, which are methods *du jour* in the most recent applications. The popularity of matching models can be attributed to their relative simplicity, easy interpretation of coefficients, and minimal structure with regard to specification. However, these gains come at a price. Matching models make the strong economic assumption that unobservable private information is irrelevant. This assumption may not be realistic in many corporate finance applications. In contrast, selection models explicitly model and incorporate private information. A second point we develop is that while matching

[1] Our attempt is to capture the overall flavor of self-selection models as they stand in corporate finance as of the writing. We apologize to any authors whose work we have overlooked: no slight is intended.

[2] The venerable reference, Maddala (1983), continues to be remarkably useful, though its notation is often (and annoyingly, to the empirical researcher) different from that used in other articles and software packages. Newer material is covered in Wooldridge (2002) and Greene (2003).

[3] Labeling any one model as "the" Heckman model surely does disservice to the many other contributions of James Heckman. We choose this label following common usage in the literature.

methods are often motivated by the fact that they yield easily interpretable *treatment effects*, selection methods also estimate treatment effects with equal ease. Our review of methodology closes by briefly touching upon fixed effect models in Section 5 and Bayesian approaches to selection in Section 6.

## 1. Self-selection: The statistical issue

To set up the self-selection issue, assume that we wish to estimate parameters $\beta$ of the regression

$$Y_i = X_i \beta + \epsilon_i \tag{1}$$

for a population of firms. In equation (1), $Y_i$ is the dependent variable, which is typically an *outcome* such as profitability or return. The variables explaining outcomes are $X_i$, and the error term is $\epsilon_i$. If $\epsilon_i$ satisfies usual classical regression conditions, standard OLS/GLS procedures consistently estimate $\beta$.

Now consider a sub-sample of firms who self-select choice $E$. For this sub-sample, equation (1) can be written as

$$Y_i | E = X_i \beta + \epsilon_i | E. \tag{2}$$

The difference between equations (2) and (1) is at the heart of the self-selection problem. Equation (1) is a specification written for the population but equation (2) is written for a subset of firms, those that self-select choice $E$. If self-selecting firms are not random subsets of the population, the usual OLS/GLS estimators applied to equation (2), are no longer consistent estimators of $\beta$.

Accounting for self-selection consists of two steps. Step 1 specifies a model for self-selection, using economic theory to model why some firms select $E$ while others do not. While this specification step is not often discussed extensively in applications, it is critical because the assumptions involved ultimately dictate what econometric model should be used in the empirical application. Step 2 ties the random variable(s) driving self-selection to the outcome variable $Y$.

## 2. The baseline Heckman selection model

### *2.1. The econometric model*

Early corporate finance applications of self-selection are based on the model analyzed in Heckman (1979). We spend some time developing this model because most other specifications used in the finance literature can be viewed as extensions of the Heckman model in various directions.

In the conventional perspective of self-selection, the key issue is that we have a regression such as equation (1) that is well specified for a population but it must be estimated using sub-samples of firms that self-select into choice $E$. To estimate population parameters from self-selected subsamples, we first specify a self-selection mechanism. This usually takes the form of a probit model in which firm $i$ chooses $E$ if the net benefit from doing so, a scalar $W_i$, is positive. Writing the selection variable $W_i$ as a function of explanatory variables $Z_i$, which are assumed for now to be exogenous,[4] we have the system

$$C = E \equiv W_i = Z_i\gamma + \eta_i > 0, \tag{3}$$

$$C = NE \equiv W_i = Z_i\gamma + \eta_i \leqslant 0, \tag{4}$$

$$Y_i = X_i\beta + \epsilon_i, \tag{5}$$

where $Z_i$ denotes publicly known information influencing a firm's choice, $\gamma$ is a vector of probit coefficients, and $\eta_i$ is orthogonal to public variables $Z_i$. In the standard model, $Y_i$ is observed only when a firm picks one of $E$ or $NE$ (but not both), so equation (5) would require the appropriate conditioning. Assuming that $\eta_i$ and $\epsilon_i$ are bivariate normal, the likelihood function and the maximum likelihood estimators for equations (3)–(5) follow, although a simpler two-step procedure (Heckman, 1979, and Greene, 1981) is commonly used for estimation. Virtually all applied work is based on the bivariate normal structure discussed above.

### *2.2. Self-selection and private information*

In the above setup, self-selection is a nuisance problem. We model it because not doing so leads to inconsistent estimates of parameters $\beta$ in regression (1). Self-selection is, by itself, of little interest. However, this situation is frequently reversed in corporate finance, because tests for self-selection can be viewed as tests of private information theories. We develop this point in the context of the Heckman (1979) model outlined above, but we emphasize that this private information interpretation is more general.

We proceed as follows. Following a well-established tradition in econometrics, Section 2.2.1 presents selection as an omitted variable problem. Section 2.2.2 interprets the omitted variable as a proxy for unobserved private information. Thus, including the omitted self-selection variable controls for and tests for the significance of private information in explaining ex-post outcomes of corporate finance choices.

#### *2.2.1. Selection: An omitted variable problem*

Suppose that firm $i$ self-selects choice $E$. For firm $i$, we can take expectations of equation (5) and write

[4] Thus, we preclude for now the possibility that $Z$ includes the outcome variable $Y$. This restriction can be relaxed at a cost, as we show in later sections.

$$Y_i|E = X_i\beta + (\epsilon_i|Z_i\gamma + \eta_i > 0) \tag{6}$$
$$= X_i\beta + \pi(\eta_i|Z_i\gamma + \eta_i > 0) + \nu_i. \tag{7}$$

Equation (7) follows from the standard result that $\epsilon_i|\eta_i = \pi\eta_i + \nu_i$ where $\pi$ is the coefficient in the regression of $\epsilon_i$ on $\eta_i$, and $\nu_i$ is an orthogonal zero-mean error term.[5] Given the orthogonality and zero-mean properties of $\nu_i$, we can take expectations of equation (7) and obtain the regression model

$$E(Y_i|E) = X_i\beta + \pi E(\eta_i|Z_i\gamma + \eta_i > 0) \tag{8}$$

and a similar model for firms choosing not to announce $E$,

$$E(Y_i|NE) = X_i\beta + \pi E(\eta_i|Z_i\gamma + \eta_i \leqslant 0). \tag{9}$$

Equations (8) and (9) can be compactly rewritten as

$$E(Y_i|C) = X_i\beta + \pi\lambda_C(Z_i\gamma) \tag{10}$$

where $C \in \{E, NE\}$ and $\lambda_C(.)$ is the conditional expectation of $\eta_i$ given $C$. In particular, if $\eta$ and $\epsilon$ are bivariate normal, as is standard in the bulk of the applied work, $\lambda_E(.) = \frac{\phi(.)}{\Phi(.)}$ and $\lambda_{NE}(.) = -\frac{\phi(.)}{1-\Phi(.)}$ (Greene, 2003, p. 759).

A comparison of equations (1) and (10) clarifies why self-selection is an omitted variable problem. In the population regression in equation (1), regressing outcome $Y$ on $X$ consistently estimates $\beta$. However, in self-selected samples, consistent estimation requires that we include an additional variable, the inverse Mills ratio $\lambda_C(.)$. Thus, the process of correction for self-selection can be viewed as including an omitted variable.

### 2.2.2. *The omitted variable as private information*

In the probit model (3) and (4), $\eta_i$ is the part of $W_i$ not explained by public variables $Z_i$. Thus, $\eta_i$ can be viewed as the private information driving the corporate financing decision being modeled. The ex-ante expectation of $\eta_i$ should be zero, and it is so, given that it has been defined as an error term in the probit model.

Ex-post after firm $i$ selects $C \in \{E, NE\}$, the expectations of $\eta_i$ can be updated. The revised expectation, $E(\eta_i|C)$, is thus an updated estimate of the firm's private information. If we wished to test whether the private information in a firm's choice affected post-choice outcomes, we would regress outcome $Y$ on $E(\eta_i|C)$. But $E(\eta_i|C) = \lambda_C(.)$ is the inverse Mills ratio term that we add anyway to adjust for self-selection. Thus, correcting for self-selection is equivalent to testing for private information. The omitted variable used to correct for self-selection, $\lambda_C(.)$, is an estimate of the private information

[5] Note that $\pi = \rho_{\eta\epsilon}\sigma_\epsilon$ where $\rho_{\eta\epsilon}$ is the correlation between $\epsilon$ and $\eta$, and $\sigma_\epsilon^2$ is the variance of $\epsilon$.

underlying a firm's choice and testing its significance is a test of whether private information possessed by a firm explains ex-post outcomes. In fact, a two-step procedure most commonly used to estimate selection models follows this logic.[6]

Our main purpose of incorporating the above discussion of the Heckman model is to highlight the dual nature of self-selection "corrections". One can think of them as a way of accounting for a statistical problem. There is nothing wrong with this view. Alternatively, one can interpret self-selection models as a way of testing private information hypotheses, which is perhaps an economically more useful perspective of selection models in corporate finance. Selection models are clearly useful if private information is one's primary focus, but even if not, the models are useful as means of controlling for potential private information effects.

### *2.3. Specification issues*

Implementing selection models in practice poses two key specification issues: the need for exclusion restrictions and the assumption that error terms are bivariate normal. While seemingly innocuous, these issues, particularly the exclusion question, are often important in empirical applications, and deserve some comment.

#### *2.3.1. Exclusion restrictions*

In estimating equations (3)–(5), researchers must specify two sets of variables: those determining selection ($Z$) and those determining the outcomes ($X$). An issue that comes up frequently is whether the two sets of variables can be identical. This knotty issue often crops up in practice. For instance, consider the self-selection event $E$ in equations (3) and (4) as the decision to acquire a target and suppose that the outcome variable in equation (5) is post-diversification productivity. Variables such as firm size or the relatedness of the acquirer and the target could explain the acquisition decision. The same variables could also plausibly explain the ex-post productivity gains from the acquisition. Thus, these variables could be part of both $Z$ and $X$ in equations (3)–(5). Similar arguments can be made for several other explanatory variables: they drive firms' decision to self-select into diversification and the productivity gains after diversification. Do we need exclusion restrictions so that there is at least one variable driving selection, an instrument in $Z$ that is not part of $X$?

Strictly speaking, exclusion restrictions are not necessary in the Heckman selection model because the model is identified by non-linearity. The selection-adjusted outcome regression (10) regresses $Y$ on $X$ and $\lambda_C(Z'\gamma)$. If $\lambda_C(.)$ were a linear function of $Z$, we would clearly need some variables in $Z$ that are not part of $X$ or the regressors

[6] Step 1 estimates the probit model (3) and (4) to yield estimates of $\gamma$, say $\hat{\gamma}$, and hence the private information function $\lambda_C(Z_i\hat{\gamma})$. In step 2, we substitute the estimated private information in lieu of its true value in equation (10) and estimate it by OLS. Standard errors must be corrected for the fact that $\gamma$ is estimated in the second step, along the lines of Heckman (1979), Greene (1981), and Murphy and Topel (1985).

would be collinear.[7] However, under the assumption of bivariate normal errors, $\lambda_C(.)$ is a non-linear function. As Heckman and Navarro-Lozano (2004) note, collinearity between the outcome regression function (here and usually the linear function $X_i\beta$) and the selection "control" function $\lambda_C(.)$ is not a generic feature, so some degree of non-linearity will probably allow the specification to be estimated even when there are no exclusion restrictions.

In practice, the identification issue is less clear cut. The problem is that while $\lambda_C(.)$ is a non-linear function, it is roughly linear in parts of its domain. Hence, it is entirely possible that $\lambda_C(Z'\gamma)$ has very little variation relative to the remaining variables in equation (10), i.e., $X$. This issue can clearly arise when the selection variables $Z$ and outcome variables $X$ are identical. However, it is important to realize that merely having extra instruments in $Z$ may not solve the problem. The quality of the instruments also matters. Near-multicollinearity could still arise when the extra instruments in $Z$ are weak and have limited explanatory power.

What should one do if there appears to be a multicollinearity issue? It is tempting to recommend that the researcher impose additional exclusion restrictions so that self-selection instruments $Z$ contain unique variables not spanned by outcome variables $X$. Matters are, of course, a little more delicate. Either the exclusions make sense, in which case these should have been imposed in the first place. Alternatively, the restrictions are not reasonable, in which case it hardly makes sense to force them on a model merely to make it estimable. In any event, as a practical matter, it seems reasonable to always run diagnostics for multicollinearity while estimating selection models whether one imposes exclusion restrictions or not.

The data often offer one degree of freedom that can be used to work around particularly thorny cases of collinearity. Recall that the identification issue arises mainly because of the $1/0$ nature of the selection variable $W_i$, which implies that we do not observe the error term $\eta_i$ and we must take its expectation, which is the inverse Mills ratio term. However, if we could observe the *magnitude* of the selection variable $W_i$, we would introduce an independent source of variation in the selection correction term and in effect observe the private information $\eta_i$ itself and use it in the regression in lieu of the inverse Mills ratio. Exclusion restrictions are no longer needed. This is often more than just a theoretical possibility. For instance, in analyzing a sample of firms that have received a bank loan, we do observe the bank loan amount conditional on a loan being made. Likewise, in analyzing equity offerings, we observe the fact that a firm made an equity offering and also the size of the offer. In hedging, we do observe (an estimate of) the extent of hedging given that a firm has hedged. This introduces an independent source of variation into the private information variable, freeing one from the reliance on non-linearity for identification.

[7] In this case, having a variable in $X$ that is not part of $Z$ does not help matters. If $\lambda_C(.)$ is indeed linear, it is spanned by $X$ whenever $Z$ is spanned by $X$. Thus, we require extra variables that explain the decision to self-select but are unrelated to the outcomes following self-selection.

### *2.3.2. Bivariate normality*

A second specification issue is that the baseline Heckman model assumes that errors are bivariate normal. In principle, deviations from normality could introduce biases in selection models, and these could sometimes be serious (for an early illustration, see Goldberger, 1983). If non-normality is an issue, one alternative is to assume some specific non-normal distribution (Lee, 1983, and Maddala, 1983, Chapter 9.3). The problem is that theory rarely specifies a particular alternative distribution that is more appropriate. Thus, whether one uses a non-normal distribution and the type of the distribution should be used are often driven by empirical features of the data. One approach that works around the need to specify parametric structures is to use semi-parametric methods (e.g., Newey, Powell and Walker, 1990). Here, exclusion restrictions are necessary for identification.

Finance applications of non-normal selection models remain scarce, so it is hard at this point of time to say whether non-normality is a first order issue deserving particular attention in finance. In one application to calls of convertible bonds (Scruggs, 2006), the data were found to be non-normal, but non-normality made little difference to the major conclusions.

## 3. Extensions

We review two extensions of the baseline Heckman self-selection model, switching regressions and structural selection models. The first allows some generality in specifying regression coefficients across alternatives, while the second allows bidirectional simultaneity between self-selection and post-selection outcomes.[8] Each of these extensions generalizes the Heckman model by allowing some flexibility in specification. However, it should be emphasized that the additional flexibility that is gained does not come for free. The price is that the alternative approaches place additional demands on the data or require more stringent economic assumptions. The plausibility and feasibility of these extra requirements should be carefully considered before selecting any alternative to the Heckman model for a given empirical application.

### *3.1. Switching regressions*

As in Section 2, a probit model based on exogenous variables drives firms' self-selection decisions. The difference is that the outcome is now specified separately for firms selecting *E* and *NE*, so the single outcome regression (5) in system (3)–(5) is now replaced

[8] For instance, in modeling corporate diversification as a decision involving self-selection, structural models would allow self-selection to determine post-diversification productivity changes, as in the standard setup, but also allow anticipated productivity changes to impact the self-selection decision.

by two regressions. The complete model is as follows:

$$C = E \equiv Z_i\gamma + \eta_i > 0, \tag{11}$$

$$C = NE \equiv Z_i\gamma + \eta_i \leqslant 0, \tag{12}$$

$$Y_{E,i} = X_{E,i}\beta_E + \epsilon_{E,i}, \tag{13}$$

$$Y_{NE,i} = X_{NE,i}\beta_{NE} + \epsilon_{NE,i}, \tag{14}$$

where $C \in \{E, NE\}$. Along with separate outcome regression parameter vectors $\beta_E$ and $\beta_{NE}$, there are also two covariance coefficients for the impact of private information on outcomes, the covariance between private information $\eta$ and $\epsilon_E$ and that between $\eta$ and $\epsilon_{NE}$. Two-step estimation is again straightforward, and is usually implemented assuming that the errors $\{\eta_i, \epsilon_{E,i}, \epsilon_{NE,i}\}$ are trivariate normal.[9]

Given the apparent flexibility in specifying two outcome regressions (13) and (14) compared to the one outcome regression in the standard selection model, it is natural to ask why we do not always use the switching regression specification. There are three issues involved. First, theory should say whether there is a single population regression whose LHS and RHS variables are observed conditional on selection, as in the Heckman model, or whether we have two regimes in the population and the selection mechanism dictates which of the two we observe. In some applications, the switching regression is inappropriate: for instance, it is not consistent with the equilibrium modeled in Acharya (1988). A second issue is that the switching regression model requires us to observe outcomes of firms' choices in both regimes. This may not always be feasible because we only observe outcomes of firms self-selecting $E$ but have little data on firms that choose not to self-select. For instance, if we were analyzing stock market responses to merger announcements as in Eckbo, Maksimovic and Williams (1990), implementing switching models literally requires us to obtain a sample of would-be acquirers that had never announced to the market and the market reaction on the dates that the markets realize that there is no merger forthcoming. These data may not always be available (Prabhala, 1997).[10] A final consideration is statistical power: imposing restrictions such as equality of coefficients $\{\beta, \pi\}$ for $E$ and $NE$ firms (when valid), lead to greater statistical power.

A key advantage of the switching regression framework is that we obtain more useful estimates of (unobserved) counterfactual outcomes. Specifically, if firm $i$ chooses $E$, we observe outcome $Y_{E,i}$. However, we can ask what the outcome might have been had

[9] Write equations (13) and (14) in regression form as

$$Y_{C,i} = X_{C,i}\beta_C + \pi_C\lambda_C(Z_i\gamma), \tag{15}$$

where $C \in \{E, NE\}$. The two-step estimator follows: the probit model (11) and (12) gives estimates of $\gamma$ and hence the inverse Mills ratio $\lambda_C(.)$, which is fed into regression (15) to give parameters $\{\beta_E, \beta_{NE}, \pi_E, \pi_{NE}\}$. As before, standard errors in the second step regression require adjustment because $\lambda_C(Z\hat{\gamma})$ is a generated regressor (Maddala, 1983, pp. 226–227).

[10] Li and McNally (2004) and Scruggs (2006) describe how we can use Bayesian methods to update priors on counterfactuals. More details on their approach are given in Section 6.

firm $i$ chosen *NE*, the unobserved counterfactual, and what the gain is from firm $i$'s having made choice $E$ rather than *NE*. The switching regression framework provides an estimate. The net benefit from choosing $E$ is the outcome of choosing $E$ less the counterfactual had it chosen *NE*, i.e., $Y_{E,i} - Y_{NE,i} = Y_{E,i} - X_i\beta_{NE} - \pi_{NE}\lambda_{NE}(Z_i\gamma)$. The *expected* gain for firm $i$ is $X_i(\beta_E - \beta_{NE}) + (\pi_E\lambda_E(.) - \pi_{NE}\lambda_{NE}(.))$.[11] We return to the counterfactuals issue when we deal with treatment effects and propensity scores. We make this point at this stage only to emphasize that selection models do estimate treatment effects. This fact is often not apparent when reading empirical applications, especially those employing matching methods.

### 3.2. Simultaneity in self-selection models

The models considered thus far presume that the variables $Z$ explaining the self-selection decision (equations (3) and (4) or equations (11) and (12)) are exogenous. In particular, the bite of this assumption is to preclude the possibility that the decision to self-select choice $C$ does not directly depend on the anticipated outcome from choosing $C$. This assumption is sometimes too strong in corporate finance applications. For instance, suppose we are interested in studying the diversification decision and that the outcome variable to be studied is firm productivity. The preceding models would assume that post-merger productivity does *not* influence the decision to diversify. If firms' decisions to diversify depend on their anticipated productivity changes, as theory might suggest (Maksimovic and Phillips, 2002), the assumption that $Z$ is exogenous is incorrect.

The dependence of the decision to self-select on outcomes and the dependence of outcomes on the self-selection decision is essentially a problem of simultaneity. Structural selection models can account for simultaneity. We review two modeling choices. The Roy (1951) model places few demands on the data but it places tighter restrictions on the mechanism by which self-selection occurs. More elaborate models are less stringent on the self-selection mechanism, but they demand more of the data, specifically instruments, exactly as in conventional simultaneous equations models.

#### 3.2.1. The Roy model

The Roy model hard-wires the dependence of self-selection on post-selection outcomes. Firms self-select $E$ or *NE* depending on which of the two alternatives yields the higher outcome. Thus, $\{E, Y_E\}$ is observed for firm $i$ if $Y_{E,i} > Y_{NE,i}$. If, on the other hand,

[11] This expression stands in contrast to the basic Heckman setup. There, in equation (9), $\beta_E = \beta_{NE}$ and $\pi_E = \pi_{NE}$, so the expected difference is $\pi(\lambda_E(.) - \lambda_{NE}(.))$. There, the sign of the expected difference is fixed: it *must* equal to the sign of $\pi$ because $(\lambda_E(.) - \lambda_{NE}(.)) > 0$. Additionally, the expected difference in the setup of Section 2 does not vary with $\beta$ or variables $X$ that are not part of $Z$: here, it does. In short, the counterfactual choices that could be made but were not are less constrained in the switching regression setup.

$Y_{NE,i} \geqslant Y_{E,i}$, we observe $\{NE, Y_{NE,i}\}$. The full model is

$$C = E \equiv Y_{E,i} > Y_{NE,i}, \tag{16}$$

$$C = NE \equiv Y_{E,i} \leqslant Y_{NE,i}, \tag{17}$$

$$Y_{E,i} = X_i \beta_E + \epsilon_{E,i}, \tag{18}$$

$$Y_{NE,i} = X_i \beta_{NE} + \epsilon_{NE,i}, \tag{19}$$

where the $\epsilon$'s are (as usual) assumed to be bivariate normal. The Roy model is no more demanding of the data than standard selection models. Two-step estimation is again fairly straightforward (Maddala, 1983, Chapter 9.1).

The Roy selection mechanism is rather tightly specified on two dimensions. One, the model exogenously imposes the restriction that firms selecting $E$ would experience worse outcomes had they chosen *NE* and vice versa. This is often plausible. However, it is unclear whether this should be a hypothesis that one wants to test or a restriction that one imposes on the data. Two, the outcome differential is the *only* driver of the self-selection decision in the Roy setup. Additional flexibility can be introduced by loosening the model of self-selection. This extra flexibility is allowed in models to be described next, but it comes at the price of requiring additional exclusion restrictions for model identification.

### 3.2.2. *Structural self-selection models*

In the standard Heckman and switching regression models, the explanatory variables in the selection equation are exogenous. At the other end of the spectrum is the Roy model of Section 3.2.1, in which self-selection is driven solely by the endogenous variable. The interim case is one where selection is driven by both exogenous and outcome variables. This specification is

$$C = E \equiv Z_i \gamma + \delta(Y_{E,i} - Y_{NE,i}) + \eta_i > 0, \tag{20}$$

$$C = NE \equiv Z_i \gamma + \delta(Y_{E,i} - Y_{NE,i}) + \eta_i \leqslant 0, \tag{21}$$

$$Y_{E,i} = X_i \beta_E + \epsilon_{E,i}, \tag{22}$$

$$Y_{NE,i} = X_i \beta_{NE} + \epsilon_{NE,i}. \tag{23}$$

The structural model generalizes the switching regression model of Section 3.1, by incorporating the extra explanatory variable $Y_{E,i} - Y_{NE,i}$, the net outcome gain from choosing $E$ over *NE*, in the selection decision, and generalizes the Roy model by permitting exogenous variables $Z_i$ to enter the selection equation. Estimation of the system (20)–(23) follows the route one typically treads in simultaneous equations systems estimation—reduced form estimation followed by a step in which we replace the dependent variables appearing in the RHS by their fitted projections. A trivariate normal assumption is standard (Maddala, 1983, pp. 223–239). While structural self-selection models have been around for a while in the labor economics literature, particularly those studying unionism and the returns to education (see Maddala, 1983, Chapter 8), applications in finance are of very recent origin.

The structural self-selection model clearly generalizes every type of selection model considered before. The question is why one should not always use it. Equivalently, what additional restrictions or demands does it place on the data? Because it is a type of the switching regression model, it comes with all the baggage and informational requirements of the switching regression. As in simultaneous equations systems, instruments must be specified to identify the model. In the diversification example at the beginning of this section, the identification requirement demands that we have at least one instrument that determines whether a firm diversifies but does not determine the ex-post productivity of the diversifying firm. The quality of one's estimates depends on the strength of the instrument, and all the caveats and discussion of Section 2.3.1 apply here.

## 4. Matching models and self-selection

This section reviews matching models, primarily those based on propensity scores. Matching models are becoming increasingly common in applied work. They represent an attractive means of inference because they are simple to implement and yield readily interpretable estimates of "treatment effects." However, matching models are based on fundamentally different set of assumptions relative to selection models. Matching models assume that unobserved private information is irrelevant to outcomes. In contrast, unobserved private information is the essence of self-selection models. We discuss these differences between selection and matching models as well as specific techniques used to implement matching models.

To clarify the issues, consider the switching regression selection model of Section 3.1, but relabel the choices to be consistent with the matching literature. Accordingly, firms are *treated* and belong to group $E$ or *untreated* and belong to group $NE$. This assignment occurs according to the probit model

$$pr(E|Z) = pr(Z\gamma + \eta) > 0, \tag{24}$$

where $Z$ denotes explanatory variables, $\gamma$ is a vector of parameters and we drop firm subscript $i$ for notational convenience. The probability of being untreated is $1-pr(E|Z)$. We write post-selection outcomes as $Y_E$ for treated firms and $Y_{NE}$ for untreated firms, and for convenience, write

$$Y_E = X_E\beta_E + \epsilon_E, \tag{25}$$

$$Y_{NE} = X_{NE}\beta_{NE} + \epsilon_{NE}, \tag{26}$$

where (again suppressing subscript $i$) $\epsilon_C$ denotes error terms, $X_C$ denotes explanatory variables, $\beta_C$ denotes parameter vectors, and $C \in \{E, NE\}$. We emphasize that the basic setup is identical to that of a switching regression of Section 3.1.

### 4.1. Treatment effects

Matching models focus on estimating *treatment effects*. A treatment effect, loosely speaking, is the value added or the difference in outcome when a firm undergoes treatment $E$ relative to not undergoing treatment, i.e., choosing *NE*. Selection models such as the switching regression specification (equations (11)–(14)) estimate treatment effects. Their approach is indirect. In selection models, we estimate a vector of parameters and covariances in the selection equations and use these parameters to estimate treatment effects. In contrast, matching models go directly to treatment effect estimation, setting aside the step of estimating parameters of regression structures specified in selection models.

The key question in the matching literature is whether treatment effects are significant. In the system of equations (24)–(26), this question can be posed statistically in a number of ways.

- At the level of an individual firm $i$, the effectiveness of a treatment can be judged by asking whether $E(Y_{E,i} - Y_{NE,i}) = 0$.
- For the group of treated firms, the effectiveness of the treatment for treated firms is assessed by testing whether the *treatment effect on treated* (TT), equals zero, i.e., whether $E[(Y_E - Y_{NE})|C = E] = 0$.
- For the population as a whole whether treated or not, we test the significance of the *average treatment effect* (ATE) by examining whether $E(Y_E - Y_{NE}) = 0$.

The main issue in calculating any of the treatment effects discussed above, whether by selection or matching models, is the fact that unchosen counterfactuals are not observed. If a firm $i$ chooses $E$, we observe outcome of its choice $Y_{E,i}$. However, because firm $i$ chose $E$, we do not explicitly observe the outcome $Y_{NE,i}$ that would occur had the firm instead made the counterfactual choice *NE*. Thus, the difference $Y_{E,i} - Y_{NE,i}$ is never directly observed for any particular firm $i$, so its expectation—whether at the firm level, or across treated firms, or across treated and untreated firms—cannot be calculated directly. Treatment effects can, however, be obtained via selection models or by matching models, using different identifying assumptions. We discuss selection methods first and then turn to matching methods.

### 4.2. Treatment effects from selection models

Self-selection models obtain treatment effects by first estimating parameters of the system of equations (24)–(26). Given the parameter estimates, it is straightforward to estimate treatment effects described in Section 4.1, as illustrated, e.g., in Section 3.1 for the switching regression model. The key identifying assumption in selection models is the specification of the variables entering selection and outcome equations, i.e., variables $X$ and $Z$ in equations (24)–(26).

Two points deserve emphasis. The first is that the entire range of selection models discussed in Section 2 through Section 3.2 can be used to estimate treatment effects. This point deserves special mention because in received corporate finance applications, the tendency has been to report estimates of matching models and as a robustness check, an accompanying estimate of a selection model. With virtually no exception, the selection model chosen for the robustness exercise is the Heckman model of Section 2. However, there is no a priori reason to impose this restriction—any other model, including the switching regression models or the structural models, can be used, and perhaps ought to at least get a hearing. The second point worth mentioning is that unlike matching models, selection models always explicitly test for and incorporate the effect of unobservable private information, through the inverse Mills ratio term, or more generally, through *control functions* that model private information (Heckman and Navarro-Lozano, 2004).

### *4.3. Treatment effects from matching models*

In contrast to selection models, matching models begin by assuming that private information is irrelevant to outcomes.[12] Roughly speaking, this is equivalent to imposing zero correlation between private information $\eta$ and outcome $Y_E$ in equations (24)–(26).

Is irrelevance of private information a reasonable assumption? It clearly depends on the specific application. The assumption is quite plausible if the decision to obtain treatment $E$ is done through an exogenous randomization process. It becomes less plausible when the decision to choose $E$ is an endogenous choice of the decision-maker, which is probably close to many corporate finance applications except perhaps for exogenous shocks such as regulatory changes.[13] If private information can be ignored, matching methods offer two routes to estimate treatment effects: dimension-by-dimension matching and propensity score matching.

#### *4.3.1. Dimension-by-dimension matching*

If private information can be ignored, the differences in firms undergoing treatment $E$ and untreated *NE* firms only depend on observable attributes $X$. Thus, the treatment effect for any firm $i$ equals the difference between its outcome and the outcome for a firm $j(i)$ that matches it on all observable dimensions, Formally, the treatment effect equals $Y_{i,E} - Y_{j(i),NE}$, where $j(i)$ is such that $X_{i,k} = X_{j(i),k}$ for all $K$ relevant dimensions, i.e., $\forall k$, $k = 1, 2, \ldots, K$. Other measures such as TT and ATE defined in Section 4.1 follow immediately.[14]

Dimension-by-dimension matching methods have a long history of usage in empirical corporate finance, as explained in Chapter 1 (Kothari and Warner, 2007) in this book.

[12] See, e.g., Wooldridge (2002) for formal expressions of this condition.

[13] Of course, even here, if unobservable information guides company responses to such shocks, irrelevance of unobservables is still not a good assumption.

[14] One could legitimately ask why we need to match dimension by dimension when we have a regression structure such as (25) and (26). The reason is that dimension-by-dimension matching is still consistent when the data come from the regressions, but dimension-by-dimension matching is also consistent with other data generating mechanisms. If one is willing to specify equations (25) and (26), the treatment effect is immediately obtained as the difference between the fitted values in the two equations.

Virtually all studies routinely match on size, industry, the book-to-market ratio, and so on. The "treatment effect" is the matched-pair difference in outcomes. There is nothing inherently wrong with these methods. They involve the same economic assumptions as other matching methods based on propensity scores used in recent applications. In fact, dimension-by-dimension matching imposes less structure and probably represents a reasonable first line of attack in typical corporate finance applications.

Matching on all dimensions and estimating the matched-pair differences in outcomes poses two difficulties. One is that characteristics are not always exactly matched in corporate finance applications. For instance, we often match firm size or book-to-market ratios with 30% calipers. When matches are inexact, substantial biases could build up as we traverse different characteristics being matched. A second issue that proponents of matching methods frequently mention is dimensionality. When the number of dimensions to be matched goes up and the matching calipers become fine (e.g., size and prior performance matched within 5% rather than 30%, and 4-digit rather than 2-digit SIC matches), finding matches becomes difficult or even infeasible. When dimension-by-dimension matching is not feasible, a convenient alternative is methods based on propensity scores. We turn to these next.

### *4.3.2. Propensity score (PS) matching*

Propensity score (PS) matching methods handle the problems caused by dimension-by-dimension matching by reducing it to a problem of matching on a single one: the probability of undergoing treatment $E$. The probability of treatment is called the *propensity score*. Given a probability model such as equation (24), the treatment effect equals the outcome for the treated firm minus the outcome for an untreated firm with equal treatment probability. The simplicity of the estimator and its straightforward interpretation makes the propensity score estimator attractive.

It is useful to review the key assumptions underlying the propensity score method. Following Rosenbaum and Rubin (1983), suppose that the probability model in equation (24) satisfies

- PS1: $0 < pr(E|Z) < 1$.
- PS2: Given $Z$, outcomes $Y_E$, $Y_{NE}$ do not depend on whether the firm is in group $E$ ($NE$).

Assumption (PS1) requires that at each level of the explanatory variable $Z$, some firms should pick $E$ and others pick $NE$. This constraint is frequently imposed in empirical applications by requiring that treated and untreated firms have common support.

Assumption (PS2) is the *strong ignorability* or conditional independence condition. It requires that unobserved private information should not explain outcome differentials between firms choosing $E$ and those choosing $NE$. This is a crucial assumption. As Heckman and Navarro-Lozano (2004) show, even fairly mild departures can trigger substantial biases in treatment effect estimates.

Given assumptions (PS1) and (PS2), Rosenbaum and Rubin (1983) show that the treatment effect is the difference between outcomes of treated and untreated firms having identical treatment probabilities (or propensity scores). Averaging across different treatment probabilities gives the average treatment effect across the population.[15]

### *4.3.3. Implementation of PS methods*

In light of Rosenbaum and Rubin (1983), the treatment effect is the difference between outcomes of treated and untreated firms with identical propensity scores. One issue in implementing matching is that we need to know propensity scores, i.e., the treatment probability $pr(E|Z)$. This quantity is not ex-ante known but it must be estimated from the data, using, for instance, probit, logit, or other less parametrically specified approaches. The corresponding treatment effects are also estimated with error and the literature develops standard error estimates (e.g., Heckman, Ichimura and Todd, 1998; Dehejia and Wahba, 1999; Wooldridge, 2002, Chapter 18).

A second implementation issue immediately follows. What variables do we include in estimating the probability of treatment? While self-selection models differentiate between variables determining outcomes and variables determining probability of being treated ($X$ and $Z$, respectively, in equations (24)–(26)), matching models make no such distinction. Roughly speaking, either a variable determines the treatment probability, in which case it should be used in estimating treatment probability, or it does not, in which case it should be randomly distributed across treated and untreated firms and is averaged out in computing treatment effects. Thus, for matching models, the prescription is to use all relevant variables in estimating propensity scores.[16]

A third issue is estimation error. In principle, matching demands that treated firms be compared to untreated firms with the same treatment probability. However, treatment probabilities must be estimated, so exact matching based on the true treatment probability is usually infeasible. A popular approach, following Dehejia and Wahba (1999), divides the data into several probability bins. The treatment effect is estimated as the average difference between the outcomes of $E$ and $NE$ firms within each bin. Heckman, Ichimura and Todd (1998) suggest taking the weighted average of untreated firms, with weights declining inversely in proportion to the distance between the treated and untreated firms. For statistical reasons, Abadie and Imbens (2004) suggest that the counterfactual outcomes should be estimated not as the actual outcomes for a matched untreated firm, but as the fitted value in a regression of outcomes on explanatory variables.[17]

[15] This discussion points to another distinction between PS and selection methods. The finest level to which PS methods can go is the propensity score or the probability of treatment. Because many firms can have the same propensity score, PS methods do not estimate treatment effects at the level of the individual firm, while selection methods can do so.

[16] This statement is not, of course, a recommendation to engage in data snooping. For instance, in fitting models to estimate propensity scores, using quality of fit as a model selection criterion leads to difficulties, as pointed out by Heckman and Navarro-Lozano (2004).

[17] The statistical properties of different estimators has been extensively discussed in the econometrics literature, most recently in a review issue devoted to the topic (*Symposium on the Econometrics of Matching*, Review of Economics and Statistics 86 (1), 2004).

## 5. Panel data with fixed effects

In self-selection models, the central issue is that unobserved attributes that lead firms to self-select could explain variation in outcomes. In panel data settings, we have multiple observations on the same firm over different periods. If the unobservable attributes are fixed over time, we can control for them by including firm fixed effects. Applications of fixed effect models in corporate finance include Himmelberg, Hubbard and Palia (1999), Palia (2001), Schoar (2002), Bertrand and Mullainathan (2003), and Çolak and Whited (2005). There are undoubtedly many more. One question is whether the use of such fixed effect models alleviates self-selection issues. Not necessarily, as we discuss next.

There are two main issues with using firm fixed effects to rule out unobservables. One is that the unobservables should be time invariant. When time invariant effects exist and ought to be controlled for, fixed effect models are effective. However, time invariance is unlikely to be an appropriate modeling choice for corporate events where unobservables are not only time varying but also related to the event under consideration. Furthermore, unobservables often have a causal role in precipitating the corporate finance event being studied. For instance, in the framework of Maksimovic and Phillips (2002), firms diversify or focus because they receive an unobserved shock that alters the optimal scope of the firm. Thus, in studying conglomerate diversification or spinoffs, the central unobservable of importance is the scope-altering shock. It is time varying and it leads to the event of interest—diversification. Including time-invariant firm fixed effects does nothing to address such event-related unobservable shocks. This point also applies to the difference-in-difference methods related to fixed effects. They do not account for event-related self-selection. Such methods are just not designed to capture time-varying and event-related unobservables, which are, in contrast, the central focus of selection models.[18]

A second issue with fixed effect models is statistical power. Models with fixed effects rely on time variation in RHS variables and LHS outcomes for a given firm. Thus, fixed effect models often have limited power when the underlying variables vary slowly over time. In this scenario, causal effects, if any, are primarily manifested in the cross-section rather than time series. Zhou (2001) presents an argument on these lines with an empirical application. Thus, it appears especially important to take a more careful look at the lack of power as an explanation for insignificant results when using fixed effects. It should also be pointed out that the regression $R^2$ in fixed effects regressions could easily lead to misleading impressions of the strength of an economic relation.[19]

[18] A related issue is the use of period-by-period estimates of Heckman-style selection models in panel data. Imposing such a structure imposes the assumption that the period-by-period disturbances are pairwise uncorrelated with next-period disturbances, which may not necessarily be realistic.

[19] Most cross-sectional studies in corporate finance with reasonable sample sizes report a modest $R^2$ when there are no fixed effects. However when one adds fixed effects, there is often an impressive improvement in

## 6. Bayesian self-selection models

Thus far, our discussion covered inference via classical statistical methods. An alternative approach towards estimating selection models involves Bayesian methods. These techniques often represent an elegant way of handling selection models that are computationally too burdensome to be practical for classical methods. We review the Bayesian approach briefly and illustrate their potential value by discussing a class of selection models based on Markov Chain Monte Carlo (MCMC) simulations (see Poirier (1995) for a more in-depth comparison between Bayesian and classical statistical inferences).

### *6.1. Bayesian methods*

The Bayesian approach begins by specifying a prior distribution over parameters that must be estimated. The prior reflects the information known to the researcher without reference to the dataset on which the model is estimated. In time series context, a prior can be formed by looking at out of sample historical data. In most empirical corporate finance applications, which are cross-sectional in nature, researchers tend to be agnostic and use non-informative diffuse priors.

Denote the parameters to be estimated by $\theta$ and the prior beliefs about these parameters by the density $p(\theta)$. If the observed sample is $y$, the posterior density of $\theta$ given the sample can be written as

$$p(\theta|y) = \frac{P(y|\theta)p(\theta)}{p(y)}, \tag{27}$$

where $p(y|\theta)$ denotes the likelihood function of the econometric model being estimated. Given the prior and the econometric model, equation (27) employs Bayes rule to generate the posterior distribution $p(\theta|y)$ about parameter $\theta$. The posterior density $p(\theta|y)$ summarizes what one learns about $\theta$ after seeing the data. It is the central object of interest that Bayesian approaches wish to estimate.

A key difficulty in implementing Bayesian methods is the computation of the posterior. Except for a limited class of priors and models, posteriors do not have closed-form analytic expressions, which poses computational difficulties in implementing Bayesian models. However, recent advances in computational technology and more importantly, the advent of the Gibbs sampler and the Metropolis–Hastings algorithm, which are specific implementations of MCMC methods, simplify implementation of fairly complex Bayesian models. In some cases, it even provides a viable route for model estimation where classical methods prove to be computationally intractable. Chib and Greenberg (1996) and Koop (2003) provide more detailed discussions of these issues.

$R^2$ (see, e.g., Campa and Kedia, 2002, and Villalonga, 2004, for interesting illustrations of this point). The high $R^2$ should not be misattributed to the explanatory power of the included variables, because they often arise due to the (ultimately unexplained) fixed effects.

### *6.2. Bayesian methods for selection models*

To illustrate the implementation of the Bayesian approach to selection models, consider the switching regression model of Section 3.1. For notational convenience, rewrite this model as the system of equations

$$I = 1_{Z_i\gamma+\eta_i>0}, \tag{28}$$

$$Y_{E,i} = X_{E,i}\beta_E + \epsilon_{E,i}, \tag{29}$$

$$Y_{NE,i} = X_{NE,i}\beta_{NE} + \epsilon_{NE,i}, \tag{30}$$

where $1_{\{.\}}$ denotes the indicator function and the other notation follows that in Section 3.1. As before, assume that the errors are trivariate normal with the probit error variance in equation (28) normalized to unity.

The critical unobservability issue, as discussed in Section 4, is that if a firm self-selects $E$, we observe the outcome $Y_{E,i}$. However, we do not observe the counterfactual $Y_{NE,i}$ that would have occurred had firm $i$ chosen $NE$ instead of $E$. Following Tanner and Wong (1987), a Bayesian estimation approach generates counterfactuals by augmenting the observed data with simulated observations of the unobservables through a "data augmentation" step. When augmented data are generated in a manner consistent with the structure of the model, the distribution of the augmented data converges to the distribution of the observed data. The likelihood of both the observed and the augmented data can be used as a proxy for the likelihood of the observed data. Conditional on the observed and augmented data and given a prior on parameters $\gamma$, $\beta$ and the error covariances, approximate posteriors for the model parameters can be obtained by using standard simulation methods. The additional uncertainty introduced by simulating unobserved data can then be integrated out (Gelfand and Smith, 1990) to obtain posteriors conditional on only the observed data.

Explicitly modeling the unobserved counterfactuals offers advantages in the context of selection models. The counterfactuals that are critical in estimating treatment effects are merely the augmented data that are anyway employed in Bayesian estimation. The augmented data also reveal deficiencies in the model that are not identified by simple tests for the existence of selectivity bias. In addition, one can obtain exact small sample distributions of parameter estimates that are particularly useful when sample sizes are small to moderate, such as self-selection involving relatively infrequent events. Finally, we can impose parameter constraints without compromising estimation. In later sections, we review empirical applications that employ the Bayesian approach towards estimating counterfactuals (Li and McNally, 2004; Scruggs, 2006). We also illustrate an application to a matching problem (Sørensen, 2005) in which the tractability of the conditional distributions given subsets of parameters leads to computationally feasible estimators in a problem where conventional maximum likelihood estimators are relatively intractable.

# II. EMPIRICAL APPLICATIONS

This part reviews empirical applications of self-selection models in corporate finance. We limit our scope to papers in which self-selection is an important element of the econometric approach or substantive findings. We begin with applications in event-studies. Here, the specifications are related to but differ from standard selection models. We then review applications in security offerings and financial intermediation, where more conventional selection models are used to characterize how private information affects debt issue pricing. We then turn to the diversification discount literature, where a range of methods have been used to address self-selection issues. The remaining sections include a collection of empirical applications based on selection and propensity score based matching methods. A last section covers Bayesian techniques. As will be clear from the review, most applications are relatively recent and involve a reasonably broad spectrum of approaches. In most cases, the model estimates suggest that unobserved private information is an important determinant of corporate finance choices.

## 7. Event studies

Event studies are a staple of empirical corporate finance. Hundreds of studies routinely report the stock market reactions to announcements such as mergers, stock splits, dividend announcements, equity issues, etc. Evidence in these studies has been used as a basis for testing and generating a wealth of theories, policies, and regulations. Chapter 1 in this volume (Kothari and Warner, 2007) overviews the literature.

Self-selection entered the event-study literature relatively recently. Its main use has been as a tool to model private information revealed in events. The basic idea is that when firms announce events, they reveal some latent "private" information. If the private information has value, it should explain the announcement effects associated with an event. Selection models are convenient tools to model the information revelation process and estimate "conditional" announcement effects.

### *7.1. Conditional announcement effects:* Acharya (1988)

Acharya (1988) introduces the self-selection theme to event-studies, using a version of the standard Heckman specification to model calls of convertible bonds. In Acharya's model, firms decide whether to call an outstanding convertible bond (event $E$) or not (event $NE$) according to a probit model, viz.,

$$E \quad \text{if } W_i = Z_i\gamma + \eta_i > 0, \tag{31}$$

$$NE \quad \text{if } W_i = Z_i\gamma + \eta_i \leqslant 0, \tag{32}$$

where $Z$ denotes known observables and $\eta$, the probit error term, is private information. Ex-ante, private information has zero mean, but ex-post, once the firm has announced

$E$ or $NE$, markets update expectations. If the private information affects stock prices, the stock price reaction $y$ to the firm's choice should be related to the updated value of private information. Assuming that $(\eta, y)$ are bivariate normal with mean, variances, and correlation equal to $(0, 0, 1, \sigma_y^2, \rho)$, we can write

$$E(y|E) = \pi E(\eta_i|\eta_i > -Z_i'\gamma) = \pi\lambda_E(Z_i'\gamma), \tag{33}$$

where $\pi = \rho\sigma_\epsilon$ and $\lambda_E(Z_i'\gamma) = \pi\phi(Z_i'\gamma)/\Phi(Z_i'\gamma)$, the inverse Mills ratio. Equation (33) gives the *conditional* announcement effect associated with event $E$. It is a specialized version of the Heckman (1979) model (e.g., equation (10)) in which there are no regressors other than the inverse Mills ratio.[20]

The empirical application in Acharya (1988) is conversion-forcing calls of convertible bonds (event $E$) while $NE$ denotes the decision to delay forced conversion. Acharya finds that the coefficient $\pi$ in equation (33) is statistically significant, suggesting that the markets do react to the private information revealed in the call. The coefficient is negative, consistent with the Harris and Raviv (1985) signaling model. A legitimate question is whether testing for the significance of unconditional announcement effects and running a linear regression on characteristics $Z$ could yield inferences equivalent to those from Acharya's model. Acharya (1993) offers simulation evidence and the question is formally analyzed in Prabhala (1997). Self-selection models add most value when one has samples of firms that chose *not* to announce $E$ because these methods offer a natural way of exploiting the information in samples of silent non-announcers.

### 7.2. *Two announcements on the same date:* Nayak and Prabhala (2001)

In the Acharya model, there is one announcement on an event-date. Nayak and Prabhala (2001) analyze a specification in which two announcements are made on the same date. They present a model to recover the individual impact of each announcement from the observed announcement effects, which reflect the combined impact of both announcements made on one date.

The empirical application in Nayak and Prabhala is to stock splits, 80% of which are announced jointly with dividends. Nayak and Prabhala model the joint decisions about whether to split a stock and whether to increase dividends using a bivariate probit model, which can be specified as

$$SPL_i = \gamma_s Z_{si} + \psi_{si}, \tag{34}$$

$$DIV_i = \gamma_d Z_{di} + \psi_{di}. \tag{35}$$

If $SPL_i$ exceeds zero, a firm splits, and if $DIV_i$ exceeds zero, it increases dividends. The private information components of these two latent variables are $\psi_{si}$ and $\psi_{di}$, and these have potentially non-zero correlation $\rho_{sd}$. The announcement effect from the two

[20] The absence of other regressors is dictated by the condition that announcement effects should not be related to ex-ante variables under the efficient markets hypothesis.

decisions is

$$E(AR_{sdi}) = \gamma_{sd} + \beta_d E(\psi_{di}|C, S) + \beta_s E(\psi_{si}|C, S). \tag{36}$$

The question of substantive interest is to decompose the joint split-dividend announcement effect into a portion due to the dividend information implicit in a split and the portion unrelated to the dividend information in the split. This decomposition cannot be inferred directly from equation (36) because the term relating to splits ($\beta_d E(\psi_{di}|C, S)$) incorporates both the dividend and the non-dividend portion of the information in splits. However, this decomposition is facilitated by writing the split information $\psi_{si}$ into dividend and non-dividend components. Accordingly, write $\psi_{si} = \rho_{sd}\psi_{di} + \psi_{s-d,i}$, in which case the joint announcement effect is

$$E(AR_{sdi}|C, S) = \gamma_{sd} + (\alpha_d - \rho_{sd}\alpha_{s-d})E(\psi_{di}|C, S) + \alpha_{s-d}E(\psi_{si}|C, S), \tag{37}$$

where $\alpha_d$ and $\alpha_{s-d}$ denote the reaction to the dividend and pure split components of the information in splits. Given these, Nayak and Prabhala show that the market's reaction to a hypothetical "pure" split unaccompanied by a dividend is

$$E(AR_{si}) = \left(1 - \rho_{sd}^2\right)\alpha_{s-d}\psi_{si} + \rho_{sd}\alpha_d\psi_{si}. \tag{38}$$

The first component in equation (38) represents the market's reaction to pure split information orthogonal to dividends and the second represents the reaction to the dividend information implied by a split. Estimating the model is carried out using a two-step procedure.[21] Using a sample of splits made between 1975 and 1994 divided into two sub-samples of ten years each, Nayak and Prabhala report that about 46% of split announcement effects are due to information unrelated to the dividend information in splits.

The Nayak and Prabhala analysis has interesting implications for sample selection in event studies. In many cases, an event is announced together with secondary information releases. For instance, capital expenditure, management, or compensation announcements may be made together with earnings releases, creating noisy samples. The conventional remedy for this problem is to pick samples in which the primary announcement of interest is not accompanied by a secondary announcements by firms. However, the analysis in Nayak and Prabhala suggests that this remedy may not cure the ill, since markets can form expectations about and price secondary announcements even when they are not explicitly announced on the event date. A different approach is to model both announcements and extract the information content of each. Selection methods are useful tools in this regard because they explicitly model and incorporate the latent information from multiple announcements.

[21] The parameter $\rho_{sd}$ is obtained as the correlation coefficient in the bivariate probit model (34) and (35). The inverse Mills ratios for equation (37) follow (they require modification from standard expressions to incorporate non-zero correlation between bivariate latent variables). The other coefficients can be estimated from regression (37).

### 7.3. *Takeovers:* Eckbo, Maksimovic and Williams (1990)

Eckbo, Maksimovic and Williams (1990)—henceforth EMW—propose variants of the "truncated regression" specification, rather than the Heckman selection model used in Acharya (1988) model to analyze announcement effects. Their empirical application is to takeovers, the subject of Chapter 15 (Betton, Eckbo and Thorburn, 2007).

EMW develop two models for announcement effects. In both models, managers announce event $E$ if the stock market gain, $y_i = x_i\gamma + \eta_i$ is positive. As before, $\eta_i$ is private information, normally distributed with mean zero and variance $\omega^2$ and $x_i$ denotes publicly known variables. In model 1, event $E$ completely surprises the capital markets. In this case, the bidder's announcement effect is

$$\begin{aligned} F(x_i) &= E(y_i|y_i = x_i\gamma + \eta_i > 0) \\ &= x_i\gamma + \omega\frac{\phi(x_i\gamma/\omega)}{\Phi(x_i\gamma/\omega)}. \end{aligned} \tag{39}$$

In model 2, the market learns about the impending takeover on a prior rumor date. The probability that the takeover will be announced is the probability that the takeover gain is positive, i.e., $Pr(x_i\gamma+\eta_i > 0) = \Phi(x_i\gamma/\omega)$. If the takeover occurs, the gain is $F(x_i)$, while the absence of the takeover is assumed to lead to zero gain. Thus, the expected stock return on the rumor date is $F(x_i)\Phi(x_i\gamma/\omega)$. On the actual merger announcement date, the takeover probability rises to 1 and the announcement effect is

$$G(x_i) = \left[x_i\gamma + \omega\frac{\phi(x_i\gamma/\omega)}{\Phi(x_i\gamma/\omega)}\right]\bigl[1 - \Phi(x_i\gamma/\omega)\bigr]. \tag{40}$$

The EMW expression in equation (40) is different from the Acharya model because EMW assume that private information has value only conditional on the takeover $E$, but has no value if there is no takeover. Thus, EMW model the real gains specific to mergers rather than non-specific information modeled by Acharya. In the actual empirical application. EMW find that bidder gains decrease with the size of the bidder relative to the target, the concentration of firms in the industry, and the number of previous takeovers in the industry. As a model diagnostic, they show that OLS estimates differ from those of the non-linear model (40), which is supported by the Vuong (1989) test statistics. EMW also report that $\omega^2$ is significant, indicating that bidders' private information is valued by capital markets.

The EMW framework has been widely applied in other event-studies with cross-sectional regressions. Eckbo (1990) examines the valuation effects of greenmail prohibitions and finds that the precommittment not to pay greenmail is value enhancing. Maksimovic and Unal (1993) estimate the after-market price performance of public offers in thrift conversions recognizing that management's choice of issue size reflects the value of growth opportunities and conflicts of interest between managers and investors. Servaes (1994) relates takeover announcement effects to excess capital expenditure. Hubbard and Palia (1995) find an increasing and then decreasing relation between merger announcement effects and managerial ownership levels. Bohren, Eckbo and

Michalsen (1997) use it to explain why rights flotations are not favored over public offerings despite the greater direct costs of the latter. Li and McNally (2006) apply the EMW method to open market share repurchases in Canada and find evidence supporting a signaling interpretation of repurchase announcement effects. We study one particular extension of EMW, Eckbo (1992), in greater detail next.

### 7.4. *Takeover deterrence:* Eckbo (1992)

Eckbo (1992) extends the EMW framework to account for the fact that regulatory challenges and court decisions on these could affect merger gains. To the extent these decisions also involved unobserved private information, they introduce additional selection bias terms into the final specification. Eckbo develops these models and applies them to horizontal mergers and price effects of rivals not involved in takeovers.

As in EMW, horizontal mergers occur if the acquirer's share of the synergy gains, $y_j = x_j\gamma + \eta_j > 0$. Under the EMW assumptions, the model for the announcement effects is equation (40). Additionally, regulators can choose whether to initiate anti-trust actions or not, and subsequently courts can decide whether to stop a merger or not. These actions are modeled using additional probit models.

$$R = x_i\phi_r + \eta_r > 0, \tag{41}$$

$$C = x_i\phi_c + \eta_c > 0. \tag{42}$$

Merger gains are realized if mergers are not challenged or they are challenged but challenges are unsuccessful. Assuming that challenges have a cost $c$ proportional to merger gains, conditional announcement effects of merger announcements can be written as

$$\begin{aligned} E(AR_i|E) = {} & \left[(1 - p_{ri}p_{ci})\left(x_i\gamma + \omega\frac{\phi(x_i\gamma/\omega)}{\Phi(x_i\gamma/\omega)}\right) - p_{ri}c\right] \\ & \times \left[1 - \Phi(x_i\gamma/\omega)\right]. \end{aligned} \tag{43}$$

Eckbo applies the truncated regression models to U.S. and Canadian data. For Canadian data, Eckbo uses the EMW models (39) and (40) because there is no regulatory overhang. He uses equation (43) in U.S. horizontal mergers where regulatory overhang exists. The explanatory variables include the ratio of the market values of the bidder and target firms, the number of non-merging rival firms in the industry of the horizontal merger, the pre-merger level of and merger-induced change in industry concentration. Eckbo finds that bidder gains are positively related to the pre-merger concentration ratio and are negatively related to the merger-induced changes in the concentration ratio. These do not support the collusion explanation for merger gains. In an interesting innovation, Eckbo also estimates the models for non-merging rivals. He reports similar and even sharper findings in challenged deals where court documents identify rivals more precisely. Changes in concentration are negatively related to rival gains in the regulatory overhang free environment in Canada, further refuting the collusion hypothesis.

## 8. The pricing of public debt offerings

Companies making a debt issue must make several decisions related to the offering such as the terms and structure of the offering, the type of the underwriter for the issue. Private information held by the issuer or the intermediaries participating in the offering could affect the choices made by firms. If such information has value, it affects the prices at which issues can be sold. A fairly wide range of self-selection models have been used to address the existence of private information and its effect on the pricing of debt issues. We review some of the applications and the key findings that emerge.

### 8.1. *Bank underwritings and the Glass–Steagall Act:* Puri (1996)

The choice of an underwriter is an area that has been extensively analyzed using self-selection models. An early application is Puri (1996), who investigates the information in a firm's choice between commercial banks and investment banks as underwriters of public debt offerings. Commercial banks are often thought to possess private information about their borrowers. If they use the private information positively, commercial bank underwritten offerings should be priced higher (the "certification" hypothesis). Alternatively, banks could use their information negatively to palm off their lemons to the market, in which case the markets should discount commercial bank underwritten offerings (the "conflicts of interest" hypothesis). Selection models are natural avenues to examine the nature of these private information effects.

Puri models the private information in the underwriter choice using a probit model, viz.,

$$C = CB \equiv W_i = Z_i\gamma + \eta_i > 0, \tag{44}$$

$$C = IB \equiv W_i = Z_i\gamma + \eta_i \leqslant 0, \tag{45}$$

where *CB* denotes a commercial bank, *IB* denotes an investment bank, and $\eta_i$ is the private information in offering $i$. Markets price issue $i$ at yield $y_i$ where

$$y_i = x_i\beta + \epsilon_i, \tag{46}$$

$$E(y_i|C) = X_i\beta + \pi\lambda_C(Z_i\gamma). \tag{47}$$

Equation (47) follows from equation (46) and the assumption that $\epsilon$ and $\eta$ are bivariate normal. The above system is, of course, the standard Heckman model of Section 2, so the sign of the covariance coefficient $\pi$ determines the impact of private information on offer yields. If $\pi > 0$, markets demand higher yield for CB offerings, consistent with a conflicts of interest hypothesis, while $\pi < 0$ supports the certification hypothesis.

The data in Puri (1996) are debt and preferred stock issues prior to the passage of the 1933 Glass–Steagall Act. She includes issue size, credit rating, syndicate size, whether the security is exchange listed, whether it is collateralized, and the age of the issuer as determinants of the offer yield. She finds that $\pi < 0$, consistent with the certification hypothesis. Additionally, $\pi$ is more negative for information sensitive securities, where the conflicts of interest hypothesis predicts the more *positive* coefficient.[22] Gande et al. (1997), and Gande, Puri and Saunders (1999) report similar findings for debt issues offered after the 1989 relaxation of the Glass–Steagall Act. Underwritings by commercial banks convey positive information that improves the prices at which debt offerings can be sold.[23]

### 8.2. *Underwriting syndicate structure:* Song (2004)

Song (2004) analyzes debt offerings as in Puri (1996), Gande et al. (1997), and Gande, Puri and Saunders (1999) but there are some important differences in her specifications. Song uses a switching regression instead of the Heckman model. Second, she focuses on the effect of the syndicate structure rather than the commercial/investment banking dichotomy on debt issue spreads.

In Song's model, commercial banks could enter as lead underwriters or be part of a hybrid syndicate with investment banks. Alternatively, issues could be underwritten by a pure investment bank syndicate. For each outcome, we observe the yield of the debt offering, which is modeled as a function of public information and (implicitly) the private information conveyed in the firm's choice of a syndicate structure. The resulting specification is a variant of the switching regression model of Section 3.1, and can be written as

$$A_i = 1 \quad \text{if } (-Z_{Ai}\gamma_A + \eta_{Ai}) > 0, \tag{48}$$

$$B_i = 1 \quad \text{if } (-Z_{Bi}\gamma_B + \eta_{Bi}) > 0, \tag{49}$$

$$C_i = 1 \quad \text{if } (-Z_{Ci}\gamma_C + \eta_{Ci}) > 0, \tag{50}$$

$$Y_{1i} = X_{1i}\beta_1 + \eta_{1i}, \tag{51}$$

$$Y_{2i} = X_{2i}\beta_2 + \eta_{2i}, \tag{52}$$

$$Y_{3i} = X_{3i}\beta_3 + \eta_{3i}, \tag{53}$$

where we have adapted Song's notation for consistency with the rest of this chapter.[24] In equations (48)–(50), the counterfactuals are $A = 0$, $B = 0$, and $C = 0$, respectively.

In Song's model $A_i = 1$ if a lead investment bank invites a commercial bank to participate in the syndicate. $B_i = 1$ if the commercial bank joins the syndicate, and zero otherwise. $C_i = 1$ if a commercial bank led syndicate is chosen and $C_i = 0$

[22] Of course, it is possible that investors paid more for bank underwritten issues but were fooled into doing so. Puri (1994) rules out this hypothesis by showing that bank underwritten offerings defaulted less than non-bank issues.

[23] Chiappori and Salanie (2000) use similar methods to analyze the role of private information in insurance markets. Liu and Malatesta (2006) is a recent application of self-selection models to seasoned equity offerings. They analyze the availability of a credit rating on the underpricing and announcement effects of SEOs.

[24] Song's usage of signs for coefficients and error terms illustrates some confusing notation in the limited dependent variable literature. Her notation follows Maddala (1983) where the selection criterion is often written as $Z\gamma - \eta > 0$, while the more modern textbook convention is to use $Z\gamma + \eta > 0$.

if a pure investment bank syndicate is chosen. Thus, a hybrid syndicate is observed (regime 1) when $A_i = 1$ and $B_i = 1$; a pure investment bank syndicate (regime 2) is observed when $A_i = 0$ and $C_i = 0$, while a commercial bank led syndicate (regime 3) is observed when $B_i = 0$ and $C_i = 1$.[25] Song assumes that the latent errors $\eta$ are i.i.d. normal, correlated with yields $Y$ with regression coefficients $\sigma_{\omega j}$ where $\omega \in \{A, B, C\}$ and $j \in \{1, 2, 3\}$. The yields in each regime can be expressed in regression form as

$$E(y_{1i}|A_i = 1, B_i = 1) = X_{1i}\beta_1 + \sigma_{A1}\frac{\phi(Z_{Ai}\gamma_A)}{1 - \Phi(Z_{Ai}\gamma_A)} + \sigma_{B1}\frac{\phi(Z_{Bi}\gamma_B)}{1 - \Phi(Z_{Bi}\gamma_B)}, \tag{54}$$

$$E(y_{2i}|A_i = 0, C_i = 0) = X_{2i}\beta_2 - \sigma_{A2}\frac{\phi(Z_{Ai}\gamma_A)}{\Phi(Z_{Ai}\gamma_A)} - \sigma_{C2}\frac{\phi(Z_{Ci}\gamma_C)}{\Phi(Z_{Ci}\gamma_C)}, \tag{55}$$

$$E(y_{3i}|B_i = 0, C_i = 1) = X_{3i}\beta_3 - \sigma_{B3}\frac{\phi(Z_{Bi}\gamma_B)}{\Phi(Z_{Bi}\gamma_B)} + \sigma_{C3}\frac{\phi(Z_{Ci}\gamma_C)}{1 - \Phi(Z_{Ci}\gamma_C)}. \tag{56}$$

Song's sample comprises 2,345 bond issues offered between January 1991 and December 1996. In the first step probit estimates, Song reports that compared to pure investment bank syndicates, hybrid syndicates underwrite small firms that have made smaller debt issues in the past, have low S&P stock rankings, invest less, and use more bank debt. These findings are reminiscent of those in Gande et al. (1997) and Gande, Puri and Saunders (1999) that commercial banks underwrite informationally sensitive companies. Compared to commercial bank led syndicates, hybrid syndicates underwrite smaller firms with lower stock rankings that issue to refinance debt and lower ranked firms, consistent with the claim that these underwritings potentially alleviate conflicts of interest with commercial banks. Only two out of six private information coefficients in equations (54)–(56) are significant. Pricing benefits are seen in pure investment banking syndicates (equation (55)) where excluding a commercial bank leads to higher yields, consistent with a certification hypothesis. On the other hand, picking an investment bank to run the syndicate increases yields, because the coefficient $\sigma_{C2}$ in the same equation (55) is positive. Thus, the ex-ante effect of awarding a syndicate to an investment bank cannot be a priori signed.

Relative to prior work, Song (2004) has very different sample, sample period, and explanatory variables, not to mention the changes in underwriter classification, which is based on syndicate structure rather than on classification into commercial/investment bank or on bank reputation. Thus, it is hard to pinpoint the specific value added by her elaborate selection model. In addition, absent additional diagnostics, it is also difficult to interpret whether the general insignificance of most selection terms reflects coefficients that are truly zero, the lack of power, perhaps due to collinearity, or perhaps an unmodeled correlation between errors in equations (48)–(50) that are assumed to be

[25] Song does not explicitly write out the extensive form of the model she estimates. It is unclear whether pure investment bank syndicates should also include the node at which an investment bank is awarded the mandate and chooses to invite a commercial bank but the bank declines to join.

i.i.d. for the purposes of estimation. As Song points out, additional data may not help shed light on interpretation or robustness because there have been structural changes in the banking industry since 1996, due to several mergers and further relaxation of the Glass–Steagall Act.

### 8.3. *Underwriter reputation:* Fang (2005)

Like the other papers reviewed in this section, Fang (2005) also studies the role of underwriter choice in explaining at-issue bond spreads. Unlike the other papers in the section, however, Fang draws on an earlier literature and classifies underwriters by reputation rather than by organization into commercial or investment banks. Fang examines whether the information in the choice of a reputed underwriter impacts underwriting spreads and yields.

Fang uses a probit specification to model underwriter-issuer matching. If issue $i$ is underwritten by a reputed underwriter, the yield is $Y_{E,i}$ and if not, the yield is $Y_{NE,i}$. Yields are specified as a function of regressors $x_i$ with different regression coefficients across the two choices. Thus, Fang's model is exactly the switching regression of Section 3.1. Fang also estimates an auxiliary regression where the dependent variable is gross spread rather than offer yield.

Fang finds that reputed underwriters underwrite higher grade, less risky issues of large and frequent issuers, and are more likely to be associated with longer maturity callable issues that she interprets as being more complex. The self-selection term in the yield equation is negative. Thus, the unobserved information that leads firms to choose reputed underwriters leads to lower bond yields or better offer prices. In the specification analyzing gross spreads, Fang finds that issue size increases fees more rapidly but risk variables matter less for reputed underwriters, indicating greater marginal costs and superior risk bearing capacity of reputed underwriters. Most importantly, the coefficient for the inverse Mills ratio in the gross spread equation is positive, suggesting that reputed underwriters charge greater fees to issuers.

Taken together, the yield and gross spread specifications show that reputed underwriters charge issuers greater fees and lower the offer yields (i.e., increase the offer price) to borrowers. Fang shows that the benefit of lowered debt yields typically outweighs the higher commissions paid by issuers. The pattern of results is shown to strengthen in lower yield bonds, so that reputation matters more for more informationally sensitive issues.

### 8.4. *Debt covenants:* Goyal (2005)

While the papers reviewed in this section study and model information in underwriter choice, Goyal (2005) examines the information in the choice of covenants attached to debt issues. Goyal argues that commercial banks often enjoy franchise value because of regulations that deter free entry. Banks with more valuable franchises are less likely to engage in excessive risk taking, so they should have less need to include covenants in their debt issues. This incentive is recognized and priced by the market, and the pricing differential again feeds back into firms' decisions about whether to include covenants. In other words, the decision to include covenants influences and is influenced by the expected pricing benefits from doing so. Goyal implements the structural self-selection model of Section 3.2 to model the simultaneity.

Goyal estimates the structural model on a sample of 415 subordinated debt issues made by firms between 1975 and 1994. He finds that yields are negatively related to franchise value. This finding is consistent with the hypothesis that banks with greater franchise value have less incentives to take risk, latent information that is recognized and priced by financial markets. The inverse Mills ratio term is significant in the no-covenant sub-sample but not in the sample with restrictive covenants. In the equation explaining whether firms use covenants or not, the coefficient for the yield differential with/without covenants is significant in explaining covenant choice, suggesting that anticipated pricing benefits do influence whether firms select covenants or not in their debt issues. Many of Goyal's results are more prominent in the 1981–1988 sub-period, when the risk-taking activity in the U.S. was more elevated.[26]

### *8.5. Discussion*

The public debt issue pricing area is interesting for the wide range of selection models employed. One issue, however, is that it is a little difficult to place the literature in perspective because the sources of self-selection modeled vary across papers. An additional issue is, of course, that there is probably self-selection on other dimensions as well, such as maturity, collateral, or the callability of an issue, not speaking of the decision to issue debt in the first place. This raises another thorny question, one that probably has no easy answer. What dimensions of self-selection should one control for in a given empirical application? Modeling every source of selection seems infeasible, while studying some sources of bias while ignoring others also seems a little ad-hoc. Embarking on a purely empirical search for sources of selection that matter is certainly undesirable, smacking of data snooping. A happy middle way is likely to emerge as the literature matures.

## 9. Other investment banking applications

### *9.1. Underwriter compensation in IPOs:* Dunbar (1995)

Dunbar (1995) presents an interesting application of a Roy (1951) style self-selection model to the study of underwriter compensation. Some IPO issuers offer warrants to compensate their underwriters while other issuers do not. Dunbar examines the role

[26] Reisel (2004) provides an interesting extension, a structural self-selection model applied to debt covenants included in industrial bonds.

of self-selection in explaining this choice, and in particular, whether firms choose the alternative that minimizes their underwriting costs.

Let $W$ denote the decision to use warrants to compensate underwriters and $N$ if not, subscripts $w$ and $n$ denote the costs if warrants are used or not, respectively, $U$ denote underpricing costs and $C$ the other costs of going public. If firm $i$ chooses underwriter warrant compensation, we observe the pair $\{U_{wi}, C_{wi}\}$ while we observe $\{U_{ni}, C_{ni}\}$ if it chooses just straight cash compensation. The key self-selection issue is that we observe the choice made by firm $i$ but not the costs of the alternative not chosen by firm $i$. Without knowing the unchosen counterfactuals, we cannot tell how much a company saved by choosing to include or exclude warrants to compensate its underwriters.

Dunbar models the decision to use warrants using a probit model

$$W = \xi(U_{ni} + C_{ni} - U_{wi} - C_{wi}) - \varepsilon_i > 0, \tag{57}$$

$$N = \xi(U_{ni} + C_{ni} - U_{wi} - C_{wi}) - \varepsilon_i \leqslant 0. \tag{58}$$

The expression in parentheses in equation (57) is the reduction in offering costs if warrants are used as compensation instead of straight cash compensation. Each component of costs is written as a function of observables and unobservables as follows:

$$U_{ni} = X_{ni}\beta_n + \varepsilon_{uni}, \tag{59}$$

$$U_{wi} = X_{wi}\beta_w + \varepsilon_{uwi}, \tag{60}$$

$$C_{ni} = Z_{ni}\gamma_n + \varepsilon_{cni}, \tag{61}$$

$$C_{wi} = Z_{wi}\gamma_w + \varepsilon_{cwi}. \tag{62}$$

Assuming that the errors in equations (59)–(62) are i.i.d. normal but potentially correlated with the probit error term, Dunbar's system is a version of the Roy (1951) self-selection model.

Dunbar reports that variables such as offering size, underwriter reputation, and a hot issue dummy explain underpricing in the warrant and cash compensation samples. The self-selection term is significant in the non-warrant sample but not in the warrant compensation sample. Most interesting are Dunbar's estimates of unobserved counterfactuals. For firms that do not use warrants, underpricing (other costs) would be 11.6% (19.2%) on average had warrants been used compared to actual underpricing (other costs) of 12.8% (9.8%). For firms that do use warrants, underpricing (other costs) would be 36.4% (14.6%) if warrants had not been used, compared to actual costs of 23.3% (23.9%). While warrants are associated with high underpricing in reduced form cross-sectional regressions, it is incorrect to conclude that warrants result in higher underpricing. Estimates of the self-selection model indicates that the use of warrants actually *reduces* underpricing compared to what it would be without warrants. Firms appear to use warrants to reduce underpricing costs.

### 9.2. *Analyst coverage:* Ljungqvist, Marston and Wilhelm (2006)

Ljungqvist, Marston and Wilhelm (2006) examine the relation between the decision to award an underwriting mandate to a bank and the coverage offered by the bank's analyst. The self-selection issue in Ljungqvist et al. is that banks self-select on whether they cover a stock or not. If the bank covers a stock, we observe the nature of the stock recommendation and we can tie it to the decision to award an underwriting mandate. However, if a bank does not elect to cover a stock, we do not know what the nature of its recommendation might have been had it chosen to cover the stock. Ljungqvist et al. model this source of self-selection in testing whether a firm with more positive coverage of a firm is more likely to win the firm's underwriting mandates.

Ljungqvist et al. model the probability that bank $j$ covers firm $i$'s stock as a probit model

$$\begin{aligned} y_C &= 1 \quad \text{if } y_C^* = X_C\beta_C + u_C > 0, \\ y_C &= 0 \quad \text{if } y_C^* = X_C\beta_C + u_C \leqslant 0, \end{aligned} \tag{63}$$

where all subscripts are suppressed for notational convenience. If there is coverage, the tie between coverage and the award of an underwriting mandate is established by the equations

$$\left.\begin{aligned} y_A &= \beta_A X_A + u_A \\ y_L &= I_{\beta_L X_L + \delta_L y_A + u_L > 0} \end{aligned}\right\} \quad \text{if } y_C^* > 0. \tag{64}$$

If there is no coverage, we have

$$\left.\begin{aligned} y_A &= 0 \\ y_L &= I_{\beta_{LNC} X_L + u_{LNC} > 0} \end{aligned}\right\} \quad \text{if } y_C^* \leqslant 0, \tag{65}$$

where $y_A$ is the nature of an analyst's recommendation, $y_L$ is a $1/0$ dummy for whether an underwriting mandate is awarded to a bank, $I$ is the $1/0$ indicator function, and $X$'s are explanatory variables. Equations (63)–(65) represent a switching regression system, similar to the type analyzed in Section 3.1. The difference here is that we have two recursive equations observed in each regime instead of just one regression in Section 3.1.

Ljungqvist et al. find that the decision to cover a stock is positively related to the type of coverage offered by an analyst for debt underwriting transactions. Prior relationships in the underwriting and loan markets are the other most significant explanatory variables. There is no evidence that the type of coverage influences the decision to award equity underwriting mandates. Even when it is significant, the coefficient for analyst recommendation $\beta_A$ in equation (64) is negative. Ljungqvist et al. interpret this finding as evidence that even if analysts are overly biased, issuers refrain from using them for underwriting.

The analysis of Ljungqvist et al. has appealing features. The choice of instruments is carefully motivated, with both economic intuition and tests for instrument strength suggested by Staiger and Stock (1997). Their analysis also suggests some natural extensions. One issue is that the very decision to cover a stock—rather than the type of coverage—might affect the probability of winning an underwriting mandate. A second and perhaps more difficult issue is that of cross-sectional correlation. The 16,000+ transactions in the Ljungqvist et al. sample occur over overlapping periods, which leads to commonality across transactions and potential cross-sectional correlation in the disturbance terms.

## 10. Diversification discount

The scope of the firm is an issue that has occupied economists since Coase (1933). One issue in this literature has been whether firms should diversify or not. While the question can be examined from several perspectives, a now well developed literature in finance investigates the diversification question from a valuation perspective. Does diversification impact firm value, and if so, in what direction, and why does diversification have this effect? Our review of this literature focuses on self-selection explanations for diversification. Chapter 8 (Maksimovic and Phillips, 2007) provides a more complete review of the now vast literature on diversification.

The recent finance literature on diversification begins with the empirical observation that diversified firms trade below their imputed value, which is the weighted average value of stand-alone firms in the same businesses as the divisions of the diversified firm (see, e.g., Lang and Stulz, 1994, Berger and Ofek, 1995, and Servaes, 1996). The difference between the actual and imputed values is called the diversification discount. The existence of a diversification discount is frequently interpreted as a value destroying consequence of diversification, although there is no consensus on the issue (e.g., Chevalier, 2000, and Graham, Lemmon and Wolf, 2002). We review three papers that discuss the role of self-selection in explaining the source of diversification discount.

### *10.1. Unobservables and the diversification discount:* Campa and Kedia (2002)

Campa and Kedia (2002) argue that firms self-select into becoming diversified and that self-selection explains the diversification discount. They model the decision to become diversified using a probit model

$$D_{it} = 1 \quad \text{if } Z_{it}\gamma + \eta_{it} > 0, \tag{66}$$

$$D_{it} = 0 \quad \text{if } Z_{it}\gamma + \eta_{it} \leqslant 0, \tag{67}$$

where $D_{it}$ is a diversification dummy that takes the value of 1 if the firm operates in more than one segment, and 0 otherwise, and $Z_{it}$ is a set of explanatory variables. The notations are adapted to match that in Section 2. Excess value $V_{it}$ is specified as

$$V_{it} = d_0 + d_1 X_{it} + d_2 D_{it} + \epsilon_{it}, \tag{68}$$

where $X_{it}$ is a set of exogenous observable characteristics of firm $i$ at time $t$. Coefficient $d_2$ is the key parameter of interest. If it is negative, becoming diversified causes the diversification discount. If not, the diversification discount could not be due to diversification. Under the assumption that the error terms in equations (67) and (68) are bivariate normal, the system is akin to and is estimated just like the basic Heckman selection model.[27]

In the empirical application, Campa and Kedia measure the LHS variable in equation (68), as the difference between the actual value of the firm and the sum of the imputed value of each of its segments. Segment imputed values are estimated using multipliers based on market value to sales or market value to book value of assets of peer firms. The explanatory variables for equation (68) include profitability, size, capital expenditure, and leverage. The additional instruments used in the probit specification equations (66) and (67) include industry instruments such as the fraction of firms (or their sales) in an industry that are diversified, time instruments, macroeconomic indicators such as the overall economic growth and economic expansion/contraction, and firm specific instruments such as being listed on a major exchange or being included in a stock index. Campa and Kedia extensively discuss their choices for instruments.

Campa and Kedia show that in OLS specifications, $d_2$ is negative, so that diversified firms do appear to trade at a discount. However, once they include the inverse Mills ratio to correct for self-selection, the coefficient $d_2$ becomes positive. The negative sign seen in OLS estimates is soaked up by the coefficient for the inverse Mills ratio. This indicates that diversified firms possess private information that makes them self-select into being diversified. The information is negatively associated with value and leads to the diversification discount. After accounting for unobserved private information, there is no diversification discount: in fact, there is a premium, implying that diversification may well be in shareholders' best interests.

The flip in the sign of $d_2$ when the selection term is introduced does raise the question of robustness of results, particularly with respect to potential collinearity between the dummy variable for diversification and the inverse Mills ratio that corrects for selection. Campa and Kedia address this issue by reporting several other models, including a simultaneous questions system that instruments the diversification dummy $D_{it}$ and evidence based on a sample of refocusing firms. The main results are robust: there is indeed a diversification discount as found by Lang and Stulz (1994) or Berger and Ofek (1995) when using OLS estimates. However, this discount is not due to diversification, but by private information that leads firms to become diversified. In fact, the Campa and Kedia self-selection estimates suggest that diversified firms trade at a premium relative to their value had they not diversified.

[27] Compared to the standard Heckman model, there is one additional variable in the second stage equation (68), specifically, the dummy variable $D_{it}$. The Heckman model with the additional dummy variable is called a "treatment effects" model. The panel data setting also requires the additional assumption that the unobserved errors be i.i.d. period by period. Campa and Kedia estimate fixed effects models as an alternative to Heckman-style selection models to handle the panel structure of the data.

### *10.2. Observables and the discount:* Villalonga (2004)

While Campa and Kedia (2002) attribute the diversification discount to unobservables causing firms to diversify, Villalonga (2004) offers an explanation based on differences in observables. Villalonga uses a longitudinal rather than cross-sectional analysis, focusing on *changes* in excess value around diversification rather than the level of the excess value itself.

Villalonga's main sample comprises 167 cases where firms move from being one segment to two segments. She tracks the changes in the diversification discount around the diversification event compared to a control group of non-diversifying firms, using propensity score (PS) based matching to construct matched control firms. Following the methods discussed in Section 4.3.2, Villalonga fits a probit model to estimate the probability that a given firm will diversify using variables similar to those in Campa and Kedia (2002). She matches each diversifying firm with a non-diversifying firm with a similar propensity score, i.e., diversifying probability. Her final sample has five quintiles of firms based on their estimated propensity scores and having a common support.

Villalonga estimates the "treatment effect" caused by diversification as the difference between the change in excess value of a diversifying firm and the excess value change of a comparable non-diversifying firms with the closest propensity score. She reports that while the treatment effect is negative, it is not significant whether she uses the Dehejia and Wahba (1999) or the Abadie and Imbens (2004) technique for estimation. Villalonga also reports similar findings when using a Heckman correction, presumably a treatment effect model on the lines of Campa and Kedia (2002).[28]

Two aspects of Villalonga's results deserve comment. One issue is perhaps semantic, the use of the term *causal inference*. In reading the work, one could easily come away with the impression that matching methods somehow disentangle causality from correlation. This is incorrect. Matching methods rule out correlation by arbitrary fiat: causality is an *assumption* rather than a statistically tested output of these methods. This fact is indeed acknowledged by Villalonga but easy to overlook given the prominence attached to the term "causal inference" in the paper.

A second issue is that some point estimates of treatment effects are insignificant but not very different in economic magnitude from those in Lang and Stulz (1994) and Berger and Ofek (1995)—and indeed, from the baseline industry-adjusted estimates that Villalonga herself reports. Thus, in fairness to Lang and Stulz and Berger and Ofek, Villalonga's results do not necessarily refute their earlier work. Nevertheless, Villalonga's

[28] In reviewing applications, we often found references to "the" Heckman model or "standard" Heckman models to be quite confusing. Campa and Kedia (2002) and Çolak and Whited (2005) use it to denote a treatment effects model, and focus on the coefficient for the diversification dummy variable. However, the Heckman (1979) model is not a treatment effects model. Also, it is not clear from the papers whether the coefficient of interest is the coefficient for the dummy variable in a treatment effects model or for the inverse Mills ratio term. It is perhaps a better practice not to use labels but instead describe fully the specification being estimated.

work does make an important point. Specifically, the statistical significance of discount based on industry/size matching methods is not a given fact, but is an open question in light of her results.

*10.3. Refocusing and the discount:* Çolak and Whited (2005)

If one accepts the diversification discount as a fact, then the question is what causes the discount. One view is that conglomerates (i.e., diversified firms) follow inefficient investment policies, subsidizing inefficient divisions with cash flow from the efficient divisions. Çolak and Whited (2005) evaluate the efficiency of investment in conglomerate and non-conglomerate firms by comparing investments made by focusing firms with those made by firms that do not focus. The focusing sample in Çolak and Whited (2005) consists of 267 divestitures and 154 spinoffs between 1981 and 1996. Control non-focusing firms are multi-segment firms in similar businesses that do not focus in years $-3$ through $+3$ where year 0 is the focusing event for a sample point.

The main specification used in Çolak and Whited (2005) employs propensity scores to match focusing and non-focusing firms. As in standard propensity score method implementations, Çolak and Whited (2005) estimate the propensity score as the probability that a given firm will focus in the period ahead. The probit estimates broadly indicate that firms are more likely to focus if they are larger, have less debt, diversity in segments (entropy), and have had recent profit shocks.

The central issue in Çolak and Whited is, of course, on change in investment efficiency after a focusing activity. Çolak and Whited use several measures of change in investment efficiency, including investment $Q$-sensitivity, the difference in adjusted investment to sales ratio between high and low growth segments, and the relative value added, which is akin to weighted investment in high minus low $Q$ segments. Çolak and Whited find that the changes in these measures are not significant relative to changes in firms that do not focus and that have similar propensity scores, using the Dehejia and Wahba (1999) matching procedure and the Abadie and Imbens (2004) implementation. There is no evidence that post-spinoff efficiency improves once the focusing firms are matched by propensity score to the non-focusing firms.

For robustness, Çolak and Whited also report estimates of a treatment effects model, equation (68) of Campa and Kedia (2002). There is little evidence for efficiency gains, except for one case in which the investment efficiency has a significance level of 10% for focusing firms. This could, however, arise due to pure chance given the wide number of dependent variables and specifications examined. While the paper does not report the coefficient for the inverse Mills ratio in the treatment effects model, Toni Whited confirms to us in private communication, that this selection term *is* significant. This suggests that self-selection is the main explanation for why firms experience efficiency gains after focusing. The unobserved private information that leads firms to focus explains post-focusing improvements in efficiency; controlling for self-selection, there is little evidence of any additional efficiency gains.

### *10.4. Discussion*

A key advantage of the diversification discount literature is that it has reasonably similar datasets, so it is easier to see the changes due to different econometric approaches. By the same token, it becomes easier to raise additional questions on model choice. We raise these questions here for expositional convenience, but emphasize that the questions are general in nature and not particular to the diversification discount literature.

One issue is statistical power. The diversification discount is significant using conventional industry-size matching but it is insignificant using PS based matching methods. Is this because the latter lack power? Çolak and Whited offer some welcome Monte Carlo evidence with respect to their application, simulating data with sample sizes, means, covariance matrix, and covariates with third and fourth moments equal to that observed in the actual data. They confirm that their tests have appropriate size, and at the level of the treatment effects in the sample, there is a better than 20% chance of detecting the observed treatment effect. More on these lines would probably be useful.

A second issue is the use of PS based matching methods as primary means of inference about treatment effects. There are good reasons to be uncomfortable with such an approach. The main issue is that propensity score methods assume that private information is irrelevant. However, this assumption is probably violated to at least some degree in most corporate finance applications. In fact, in the diversification literature, private information does empirically matter. Thus, using PS methods as the primary specification seems inappropriate without strong arguments as to why firms' private information is irrelevant. Heckman and Navarro-Lozano (2004) stress and show explicitly that even small deviations from this assumption can introduce significant bias. Thus, the practice followed in the finance literature of reporting private information specifications in conjunction with matching models is probably appropriate, although more full discussion on reconciling the results from different approaches would be useful.

A final comment is about the self-selection specifications used to control for private information. While the literature has used versions of the baseline Heckman (1979) model, we emphasize that this restriction is neither necessary nor desirable. Other models, such as switching regressions and structural models are viable alternatives for modeling self-selection and private information. Because these models come with their own additional requirements, it is not clear that they would always be useful, but these issues are ultimately empirical.

## 11. Other applications of selection models

### *11.1. Accounting for R&D:* Shehata (1991)

Shehata (1991) applies self-selection models to analyze the accounting treatment of research and development (R&D) expenditures chosen by firms during the period of the introduction of FASB ruling SFAS No. 2. This ruling pushed firms to expense rather than defer R&D expenditures. Other studies examined the issue by comparing observed changes in R&D expenditures for a sample of capitalizing firms with those of expensing firms. If firms self-select into the choice they prefer, it is inappropriate to treat the choice as exogenous and assess its impact by comparing differences between capitalizers and expensers. Shehata uses a switching regression instead.

Shehata uses a probit specification to model how firms choose an accounting method, and two regressions to determine the level of the R&D expenditure, one for each accounting choice. This is, of course, the switching regression system of Section 3.1. Shehata estimates the system using standard two-step methods. As discussed in Section 3.1, one useful feature of the system is the estimation of counterfactuals: what the R&D spending would be for firms that expensed had they elected to defer and vice-versa. Shehata reports that capitalizers are small, highly leveraged, have high volatility of R&D expenditures, more variable earnings, and spend a significant portion of their income on R&D activities. The second stage regression shows that the two groups of firms behave differently with respect to R&D spending. For instance, R&D is a non-linear function of size and is related to the availability of internally generated funds for capitalizers but the size relation is linear and internally generated funds do not matter for expensers. Thus, it is more appropriate to use a switching regression specification rather than the Heckman (1979) setup to model selection.

The inverse Mills ratio that corrects for self-selection matters in the second stage regression for both groups. Thus, standard OLS estimates tend to understate the impact of SFAS No. 2 on R&D expenditures. Finally, Shehata (1991) reports predictions of the expected values of R&D expenditures for both expensing and capitalizing samples had they elected to be in the other group. The mean value of R&D for each group is lower under the unchosen alternative. The decline is more pronounced for the capitalizing group, where it declines from \$ 0.69 mm to \$ 0.37 mm, while the decline is from \$ 0.85 mm to \$ 0.79 mm for the expensing group.

### *11.2. Bankruptcy costs:* Bris, Zhu and Welch (2006)

Bris, Zhu and Welch (2006) analyze the relative costs of bankruptcy under the Chapter 11 and Chapter 7 procedures in the U.S., codes that are discussed more fully in Chapter 14 (John et al., 2007). The sample consists of close to 300 bankruptcy filings in Arizona and Southern New York, the largest sample in the literature as of this writing.

The specification is the basic Heckman model of Section 2, with treatment effects in some specifications. Step 1 is a probit specification that models the choice between Chapter 11 and Chapter 7, conditional on deciding to file for bankruptcy. Bris et al. show that the procedural choice is related to firm characteristics such as size, managerial ownership, and the structure of debt including variables such as the number of creditors, whether the debt is secured or not, and the presence of banks as a company creditor. Step 2 involves modeling the costs of bankruptcy. Bris et al. analyze four metrics to specify the LHS dependent variable: the change in value of the estate during bankruptcy; the time spent in bankruptcy; the expenses submitted to and approved by the bankruptcy court; and the recovery rates of creditors. These are modeled as a function of a comprehensive set of regressors that include linear and non-linear functions of firm size, various proxies for the structure of the filing firm and managerial ownership. Because the variables in the two stages are similar, the study essentially relies on non-linearity for identification.

Bris et al. find no evidence that firms that were more likely to self-select into Chapter 11 were any faster or slower in completing the bankruptcy process. Controlling for self-selection, Chapter 11 cases consumed more fees, not because Chapter 11 is intrinsically the more expensive procedure, but because of intrinsic differences in firms that choose to reorganize under this code. After controlling for self-selection, Chapter 11 emerges as the cheaper mechanism, and Bris et al. report that self-selection explains about half of the variation in bankruptcy expenses. With self-selection controls, Chapter 11 cases had higher recovery rates than Chapter 7 cases. In sum, selection has a significant impact on estimates of reorganization costs under different bankruptcy codes. After controlling for selection, Chapter 7 takes almost as long, consumes no less and probably more in professional fees, and creditors rarely receive as much, so there is little evidence that it is more efficient than Chapter 11 reorganizations.

### *11.3. Family ownership and value:* Villalonga and Amit (2006)

Villalonga and Amit (2006) examine the effect of family ownership, control, and management on value for a sample of Fortune 500 firms from 1994 to 2000. The specification is a standard Heckman style selection model of Section 2 with a treatment effect.

The first step is a probit specification that models whether a firm remains family owned or not. Family ownership is defined as firms in which the founding family owns at least 5% of shares or holds the CEO position. In the second step, value, proxied by Tobin's $Q$, is regressed on a dummy variable for family ownership, industry dummy variables, the Gompers, Ishii and Metrick (2003) shareholder rights index, firm-specific variables from COMPUSTAT, outside block ownership and proportion of non-family outside directors, and, of course, the inverse Mills ratio that corrects for self-selection. To assist in identification, Villalonga and Amit include two additional instruments in the selection equation lagged $Q$ and idiosyncratic risk. Idiosyncratic risk is presumably related to family ownership but not to $Q$ if only systematic risk is priced by the market.

Villalonga and Amit report that family ownership has a positive effect on value in the overall sample and in sub-samples in which the founder is the CEO. Interestingly, the sign is negative when the founder is not the CEO. Villalonga and Amit interpret their findings as evidence that the benefits of family ownership are lost when the family retains control in the post-founder generation. Their results strengthen when they incorporate a control for self-selection. In the self-selection specification, the inverse Mills ratio is significant and negative in the overall specification and sub-samples in which the CEO is the founder. In these samples, family ownership appears to be associated with unobserved attributes that are negatively related to value. These unobserved attributes positively impact value if the founder is not the CEO.[29]

## 12. Other applications of matching methods

### *12.1. Bank debt versus bonds:* Bharath (2004)

Debt financing by a corporation gives rise to conflicts of interest between creditors and shareholders that can reduce the value of the firm. Such conflicts are limited more effectively in bank loans than in public debt issues if banks monitor. Bharath (2004) measures the size of agency costs by calculating the yield spread between corporate bonds and bank loans (the Bond-Bank spread) of the same firm at the same point in time. To quantify the difference, Bharath needs to match bonds with bank loans of the same firm at the same point in time and having substantively identical terms. The matching problem is complicated by the fact that bank loans and public bonds are contractually very different on multiple dimensions such as credit rating, seniority, maturity, and collateral.

Bharath argues that because bank loans and bonds are matched at the same point of time and for the same firm, matching based on observables should adequately control for differences between bank debt and public debt. Thus, propensity score based matching methods are appropriate tools to control for differences between bank loans and public debt. Bharath uses the propensity score matched difference between bank and bond credit spreads as the treatment effect, or the value added by banks. The spread can be interpreted as the value added by banks in enforcing better investment policies, or more generally, as the price of the "specialness" of banks due to their ability to monitor, generate information, or better renegotiate loans, or even perhaps other explanations such as monopoly rents.

Using a sample of over 15,000 yield observations, Bharath finds that the Bond-Bank spread is negative for high credit quality firms and positive for low credit quality firms. He interprets his findings as being consistent with the view that for high quality firms, the benefits of bank monitoring are outweighed by the costs of bank hold-up. This causes the spread to be negative, indicating that bank debt offers few benefits for high quality firms. For low quality firms, the opposite is true, causing the spread to be positive. The magnitude of the potential agency costs mitigated by banks is more important for poor quality firms, justifying the decision to borrow from banks.

[29] An interesting question raised by this study is survivorship (e.g., Brown, Goetzmann and Ross, 1995). Perhaps family owned firms that survived and made it to Fortune 500 status are of better quality, and hence these firms are valued more. This question can perhaps be resolved by looking at broader samples that incorporate smaller firms outside the Fortune 500 universe. Bennedsen et al. (2006) take a step in this direction.

### *12.2. Matching and long-run performance:* Cheng (2003), Li and Zhao (2006)

A vast literature on market efficiency examines the long-run stock return after events such as IPOs, SEOs, share repurchases, listing changes, etc. The semi-strong version of the efficient markets hypothesis predicts that long-run returns should be zero on average. However, several papers report empirical evidence against the efficiency hypothesis (Fama, 1998). In most studies, post-event buy-and-hold returns are systematically positive or negative relative to benchmarks over periods of three to five years. Chapter 1 (Kothari and Warner, 2007) offers an overview of this literature. We focus on applications of matching models to assess long-run performance.

To test whether abnormal returns are zero or not, one needs a model of benchmark returns. As discussed in Chapter 1, the standard approach, is to match an event firm with a non-event firm on between two and four characteristics that include size, book-to-market, past returns, and perhaps industry. This method runs into difficulties when the number of dimensions becomes large and the calipers become fine, when it becomes difficult to generate matching firms. Propensity score (PS) based matching methods reviewed in Section 4.3.2 are potentially useful alternatives in this scenario. Two recent papers, Cheng (2003) and Li and Zhao (2006) use PS methods to reexamine the long-term performance of stock returns after SEOs. Both papers find that while characteristic-by-characteristic matching results in significant long-term abnormal returns after SEOs, abnormal returns are insignificant if one uses propensity score based matching methods instead.

Cheng (2003) studies SEOs offered between 1970 and 1997 for which necessary COMPUSTAT data are available on firm characteristics. She finds significant buy-and-hold abnormal returns of between $-6\%$ and $-14\%$ over three to five years in the full sample and various sub-samples when matches are constructed on size, industry and book-to-market. She then uses three logit models, one for each decade, to predict the probability of issuance. Several firm characteristics such as size, book-to-market, industry, R&D, exchange, as well as 11-month past returns predict the issuance decision. Cheng matches each issuer with a non-issuer in the SEO year with a similar propensity score (i.e., predicted probability). She finds little evidence of significant abnormal returns except for one sub-sample in the 1970s.

Li and Zhao undertake an exercise similar to that in Cheng (2003) for issuers from 1986 to 1997. They show that characteristic-by-characteristic matching produces inadequate matches between issuers and non-issuers in terms of average size.[30] They estimate propensity scores with size, book-to-market, and past returns in three quarters prior to issuance, one model per year, and add interaction terms for better predictions and delete firms as necessary to have a common support. In their final sample, conventional matching gives average three-year buy-and-hold abnormal returns of $-16\%$, but this drops to an insignificant $-4\%$ with PS matching.

[30] Medians are not reported, so it is hard to assess the role of outliers.

Cheng (2003) and Li and Zhao (2006) emphasize that PS methods are merely substitutes for characteristic-by-characteristic matching of observables. This perspective is probably appropriate. The main issue in these applications is the data driven nature of the exercise in fitting probit models. Characteristics and interaction terms are added as needed to achieve balance in characteristics and propensity scores. While we recognize that a reasonable probit model seems necessary to place faith in treatment effect estimates, the search required to achieve balance, however transparent, nevertheless raises data dredging concerns and even inconsistency of estimates (Heckman and Navarro-Lozano, 2004). The general use of PS methods in studies of long-term stock return or operating performance as an alternative to methods studied in Barber and Lyon (1996, 1997), Barber, Lyon and Tsai (1999), and Kothari and Warner (1997) remains an open question.

## 13. Bayesian methods

### *13.1. Matching:* Sørensen (2005)

Investors differ in their abilities to select good investments, and in their ability to take a given investment and monitor and manage it so as to add value to what they invest in. A key question in the venture capital literature is the differentiation of selection from value-addition. To what extent are better performing venture capitalists more successful because of their ability to select good investments rather than their ability to value-add to their investments? Sørensen (2005) employs a matching-selection model to separate these two influences, using Bayesian MCMC (Markov Chain Monte Carlo) methods to estimate it.

In Sørensen's model, there is a set of venture capital investors indexed by $i$. Each investor evaluates a set of potential investments indexed by $j$ and ultimately invests (i.e., becomes the lead investor) in a subset of these. Once an investment occurs, its outcome is specified as the variable *IPO* which equals one if the investment results in a public offering and zero otherwise. In Sørensen's model, feasible investments for each investor are partly determined by the characteristics of the other agents in the market. These characteristics are related to the investment decision but unrelated to the investment outcome, so they provide the exogenous variation used for identification of the model. On the other hand, this type of sorting also causes interaction between investment decisions by different venture capitalists, which leads to a dimensionality problem and considerable numerical difficulties in estimation. Bayesian methods offer feasible routes for estimation.

Sørensen specifies normally distributed and diffuse prior beliefs with prior variances that are over 300 times the posterior variance. He assumes that error terms for different deals are independent. There are three sets of exogenous variables. The characteristics of the company includes the stage of development of the company and industry dummies. The characteristics of the venture capital investor include his experience and amount of capital he has available. The characteristic of the market is the year of the investment. There are two parameters of central interest. One is the access of better venture capitalists to deal flow, which is captured by the experience of the venture capitalist. The other is the synergy between venture capitalists and their target investments or the value added by venture capitalists, which is captured by the correlation between the private information in the decision to invest and the probability of going public.

Sørensen's final sample includes 1,666 investments made by 75 venture capitalists between 1975 and 1995 in the states of California and Massachusetts. Experience is proxied by the total and stage-of-life-cycle-specific number of deals done since 1975. Sørensen reports a number of interesting findings. He finds evidence for sorting. Experienced investors are more likely to have access to the better deals whose probability of going public (and doing so faster) increases by about two-thirds. This type of sorting explains about 60% of the increased probability of success, leaving about 40% for the synergies, or the value added by venture capital investors. Sørensen explains why one might get different results from estimating a standard selection model compared to one with sorting.

### *13.2. Switching regressions:* Li and McNally (2004), Scruggs (2006)

Li and McNally (2004) and Scruggs (2006) offer interesting applications of Bayesian methods to estimate switching regression models of self-selection. Both papers emphasize that the value of the Bayesian approach is not merely the difference in philosophy or technique; rather, the techniques offer insights not readily available through classical methods. The application in Li and McNally (2004) is the choice of a mechanism to effect share repurchases, while the application in Scruggs relates to whether convertibles are called with or without standby underwriting arrangements. For convenience, we focus on Li and McNally, but substantially similar insights on methodology are offered in the work by Scruggs.[31]

Share repurchases started becoming popular in the 1980s as a way to return excess cash to shareholders in lieu of dividends. Repurchases tend to be more flexible in timing and quantity relative to the fixed cash flow stream expected by markets when companies raise dividends. Share repurchases can be implemented in practice as a direct tender offer or more open-ended open market repurchases. Li and McNally (2004) investigate the choice between the two mechanisms and their impact on share price reactions to announcements of repurchases using Bayesian self-selection methods.

Li and McNally propose the following system of equations to analyze the choice of a repurchase mechanism

$$I^* = Z_i \gamma + \eta_i, \tag{69}$$

[31] Wald and Long (2006) present an application of switching regression using classical estimation methods. They analyze the effect of state laws on capital structure.

$$p_1^* = X_1\beta_1 + \epsilon_1, \tag{70}$$

$$p_2^* = X_2\beta_2 + \epsilon_2, \tag{71}$$

$$y_1^* = W_1\alpha_1 + v_1, \tag{72}$$

$$R_1^* = V_1\theta_1 + u_1, \tag{73}$$

$$R_2^* = V_2\theta_2 + u_2, \tag{74}$$

where $I^*$ is an unobserved latent variable representing the incremental utility of tender offers over open market repurchases, $p_1^*$, $y_1^*$, $R_1^*$ are the percentage of shares sought, tender premium and announcement effects under the tender offer regime, and $p_2^*$, $R_2^*$ are the proportion sought and announcement effects in an open market repurchase regime. The error terms in equations (69)–(74) are assumed to have a multivariate normal distribution.

The system of equations (69)–(74) represents a switching regression system discussed in Section 3.1, but with more than one regression in each regime. The key issue in estimating the system is the lack of information on unobserved counterfactuals. We observe outcomes in the repurchase technique actually chosen by a firm but do not explicitly observe what would happen if the firm had chosen the alternative technique instead. Li and McNally employ MCMC methods that generate counterfactuals as a natural by-product of the estimation procedure. This approach involves a data augmentation step in which the observed data are supplemented with counterfactuals generated consistent with the model structure. The priors about parameters are updated and posteriors obtained using standard simulation methods after which the additional uncertainty due to the data augmentation step can be integrated out. Observations on counterfactual choices and outcomes are generated as part of the estimation procedure. These can be directly used to examine the impact of choosing a given type of repurchase mechanism not just in isolation, but also relative to the impact of choosing the unchosen alternative.

The sample in Li and McNally comprises 330 fixed price tender offers, 72 Dutch auction tender offers, and 1,197 open market repurchases covering time periods from 1962 to 1988. In terms of findings, Li and McNally report that firms choose the tender offer mechanism when they have financial slack and large shareholders that monitor management. Firms prefer the open market repurchase in times of market turbulence or weak business conditions. Unobserved private information affects both the type of the repurchase program and the repurchase terms and is reflected in the stock market announcement effects. The estimates of counterfactuals are quite interesting. For instance, if the open market repurchasers had opted for tender offers, the proportion of shares sought would have been 36% (versus actual of about 7%) and the tender premium would have been 33% compared to 0% actuals, and the five-day announcement effect would be 16% compared to the actual announcement effect of 2.2%. Likewise, tender offer firms would have repurchased 10.6% (actual = 19.7%) and experienced announcement effects of 3.7% (actual = 10.2%). Firms appear to have a comparative advantage in their chosen repurchase mechanisms.

## 14. Conclusions

Our review suggests that self-selection is a growth area in empirical corporate finance. The rapidly expanding number of applications undoubtedly reflects the growing recognition in the finance profession that self-selection is an important and pervasive feature of corporate finance decisions. The range of econometric models in use is also growing as techniques diffuse from the econometrics literature to finance. However, the key issue in implementing self-selection models still remains the choice of specification, particularly the economic assumptions that make one model or another more appropriate for a given application. One size does not fit all. Each self-selection model addresses a different kind of problems, places its own demands on the type of data needed, and more importantly, carries its own baggage of economic assumptions. The plausibility of these assumptions is perhaps the primary criterion to guide what is used in empirical applications.

## References


Abadie, A., Imbens, G., 2004. Simple and bias-corrected matching estimators for average treatment effects. Working Paper, Harvard University.

Acharya, S., 1988. A generalized econometric model and tests of a signaling hypothesis with two discrete signals. Journal of Finance 43, 413–429.

Acharya, S., 1993. Value of latent information: Alternative event study methods. Journal of Finance 48, 363–385.

Barber, B., Lyon, J., 1996. Detecting abnormal operating performance: The empirical power and specification of test statistics. Journal of Financial Economics 41, 359–400.

Barber, B., Lyon, J., 1997. Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. Journal of Financial Economics 43, 341–372.

Barber, B., Lyon, J., Tsai, C., 1999. Improved methods for tests of long-run abnormal stock returns. Journal of Finance 54, 165–201.

Bennedsen, M., Nielsen, K., Perez-Gonzalez, F., Wolfenzon, D., 2006. Inside the family firm: The role of family in succession decisions and performance. Working Paper, New York University.

Berger, P.G., Ofek, E., 1995. Diversification's effect on firm value. Journal of Financial Economics 37, 39–65.

Bertrand, M., Mullainathan, S., 2003. Enjoying the quiet life? Corporate governance and managerial preferences. Journal of Political Economy 111, 1043–1075.

Betton, S., Eckbo, E., Thorburn, K., 2007. Takeovers, restructurings, and corporate control. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 2. Elsevier/North-Holland, Amsterdam, Chapter 15, forthcoming.

Bharath, S., 2004. Agency costs, bank specialness and renegotiation. Working Paper, Stephen M. Ross School of Business, University of Michigan.

Bohren, O., Eckbo, B.E., Michalsen, D., 1997. Why underwrite rights offerings? Some new evidence. Journal of Financial Economics 46, 223–261.

Bris, A., Zhu, N., Welch, I., 2006. The cost of bankruptcy: Chapter 7 liquidation versus Chapter 11 reorganization. Journal of Finance 61, 1253–1303.

Brown, S., Goetzmann, W., Ross, S., 1995. Survival. Journal of Finance 50, 853–873.

Campa, J.M., Kedia, S., 2002. Explaining the diversification discount. Journal of Finance 57, 1731–1762.

Cheng, Y., 2003. Propensity score matching and the new issues puzzle. Working Paper, Florida State University.

Chevalier, J., 2000. Why do firms undertake diversifying mergers? An examination of the investment policies of merging firms. Working Paper, Yale School of Management.

Chiappori, P., Salanie, B., 2000. Testing for asymmetric information in insurance markets. Journal of Political Economy 108, 56–78.

Chib, S., Greenberg, E., 1996. Markov Chain Monte Carlo simulation methods in econometrics. Econometric Theory 12, 409–431.

Coase, R.H., 1933. The nature of the firm. Economica 4, 386–405.

Çolak, G., Whited, T.M., 2005. Spin-offs, divestitures, and conglomerate investment. Review of Financial Studies, in press.

Dehejia, R.H., Wahba, S., 1999. Casual effects in non-experimental studies: Re-evaluating the evaluation of training programs. Journal of American Statistical Association 94, 1053–1062.

Dunbar, C.G., 1995. The use of warrants as underwriter compensation in initial public offerings. Journal of Financial Economics 38, 59–78.

Eckbo, B.E., 1990. Valuation effects of greenmail prohibitions. Journal of Financial and Quantitative Analysis 25, 491–505.

Eckbo, B.E., 1992. Mergers and the value of antitrust deterrence. Journal of Finance 47, 1005–1029.

Eckbo, B.E., Maksimovic, V., Williams, J., 1990. Consistent estimation of cross-sectional models in event studies. Review of Financial Studies 3, 343–365.

Fama, E., 1998. Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics 49, 283–306.

Fang, L.H., 2005. Investment bank reputation and the price and quality of underwriting services. Journal of Finance 60, 2729–2761.

Gande, A., Puri, M., Saunders, A., 1999. Bank entry, competition, and the market for corporate securities underwriting. Journal of Financial Economics 54, 165–195.

Gande, A., Puri, M., Saunders, A., Walter, I., 1997. Bank underwriting of debt securities: Modern evidence. Review of Financial Studies 10, 1175–1202.

Gelfand, A.E., Smith, A.F.M., 1990. Sampling based approaches to calculating marginal densities. Journal of American Statistical Association 85, 398–409.

Goldberger, A., 1983. Abnormal selection bias. In: Karlin, S., Amemiya, T., Goodman, L. (Eds.), Studies in Econometrics, Time Series, and Multivariate Statistics. Academic Press, New York.

Gompers, P.A., Ishii, J.L., Metrick, A., 2003. Corporate governance and equity prices. Quarterly Journal of Economics 118, 107–155.

Goyal, V.K., 2005. Market discipline of bank risk: Evidence from subordinate debt contracts. Journal of Financial Intermediation 14, 318–350.

Graham, J.R., Lemmon, M.L., Wolf, J.G., 2002. Does corporate diversification destroy value? Journal of Finance 57, 695–720.

Greene, W., 1981. Sample selection bias as a specification error: Comment. Econometrica 49, 795–798.

Greene, W., 2003. Econometric Analysis, 5th edition. Prentice Hall, New York.

Harris, M., Raviv, A., 1985. A sequential signaling model of convertible debt call policy. Journal of Finance 40, 1263–1282.

Heckman, J.J., 1979. Sample selection as a specification error. Econometrica 47, 153–161.

Heckman, J.J., Ichimura, H., Todd, P., 1998. Matching as an econometric evaluation estimator. Review of Economic Studies 65, 261–294.

Heckman, J.J., Navarro-Lozano, S., 2004. Using matching, instrumental variables, and control functions to estimate economic choice models. Review of Economics and Statistics 86, 30–57.

Himmelberg, C., Hubbard, R., Palia, D., 1999. Understanding the determinants of managerial ownership and the link between ownership and performance. Journal of Financial Economics 53, 353–384.

Hubbard, R.G., Palia, D., 1995. Benefits of control, managerial ownership, and the stock returns of acquiring firms. RAND Journal of Economics 26, 782–792.

John, K., Hotchkiss, E., Mooradian, R.C., Thorburn, K., 2007. Bankruptcy and the resolution of financial distress. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 2. Elsevier/North-Holland, Amsterdam, Chapter 14, forthcoming.

Koop, G., 2003. Bayesian Econometrics. John Wiley & Sons, New York.

Kothari, S.P., Warner, J.B., 1997. Measuring long-horizon security price performance. Journal of Financial Economics 43, 301–339.

Kothari, S.P., Warner, J.B., 2007. Econometrics of event studies. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. Elsevier/North-Holland, Amsterdam, Chapter 1, this volume.

Lang, L.H.P., Stulz, R.M., 1994. Tobin's $Q$, corporate diversification, and firm performance. Journal of Political Economy 102, 1248–1280.

Lee, L.F., 1983. Generalized econometric models with selectivity. Econometrica 51, 507–512.

Li, K., McNally, W., 2004. Open market versus tender offer share repurchases: A conditional event study. Working Paper, Sauder School of Business UBC. Available at http://finance.sauder.ubc.ca/~kaili/buyback.pdf.

Li, K., McNally, W., 2006. The information content of Canadian open market repurchase announcements. Managerial Finance (A Special Issue on Payout Policy), in press.

Li, X., Zhao, X., 2006. Is there an SEO puzzle? Journal of Empirical Finance 13, 351–370.

Liu, Y., Malatesta, P., 2006. Credit ratings and the pricing of seasoned equity offerings. Working Paper, University of Washington.

Ljungqvist, A., Marston, F., Wilhelm Jr., W.J., 2006. Competing for securities underwriting mandates: Banking relationships and analyst recommendations. Journal of Finance 61, 301–340.

Maddala, G.S., 1983. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press, Cambridge.

Maksimovic, V., Phillips, G., 2002. Do conglomerate firms allocate resources efficiently? Journal of Finance 57, 721–767.

Maksimovic, V., Phillips, G., 2007. Conglomerate firms and internal capital markets. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. Elsevier/North-Holland, Amsterdam, Chapter 8, this volume.

Maksimovic, V., Unal, H., 1993. Issue size choice and "underpricing" in thrift mutual-to-stock conversions. Journal of Finance 48, 1659–1692.

Murphy, K., Topel, R., 1985. Estimation and inference in two step econometric models. Journal of Business and Economic Statistics 3, 370–379.

Nayak, S., Prabhala, N.R., 2001. Disentangling the dividend information in splits: A decomposition using conditional event-study methods. Review of Financial Studies 14, 1083–1116.

Newey, W., Powell, J., Walker, J., 1990. Semi-parametric estimation of selection models. American Economic Review 80, 324–328.

Palia, D., 2001. The endogeneity of managerial compensation in firm valuation: A solution. Review of Financial Studies 14, 735–764.

Poirier, D.J., 1995. Intermediate Statistics and Econometrics. MIT Press, Cambridge.

Prabhala, N.R., 1997. Conditional methods in event studies and an equilibrium justification for standard event-study procedures. Review of Financial Studies 10, 1–38.

Puri, M., 1994. The long-term default performance of bank underwritten securities issues. Journal of Banking and Finance 18, 397–418.

Puri, M., 1996. Commercial banks in investment banking: Conflict of interest or certification role? Journal of Financial Economics 40, 373–401.

Reisel, N., 2004. On the value of restrictive covenants: An empirical investigation of public bond issues. Working Paper, Rutgers University.

Rosenbaum, R., Rubin, D., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–55.

Roy, A.D., 1951. Some thoughts on the distribution of earnings. Oxford Economic Papers 3, 135–146.

Schoar, A., 2002. The effect of diversification on firm productivity. Journal of Finance 62, 2379–2403.

Scruggs, J.T., 2006. Estimating the cross-sectional market response to an endogenous event: Naked vs. underwritten calls of convertible bonds. Journal of Empirical Finance, in press.

Servaes, H., 1994. Do takeover targets overinvest? Review of Financial Studies 7, 253–277.
Servaes, H., 1996. The value of diversification during the conglomerate merger wave. Journal of Finance 51, 1201–1225.
Shehata, M., 1991. Self-selection bias and the economic consequences of accounting regulation: An application of two-stage switching regression to SFAS No. 2. Accounting Review 66, 768–787.
Song, W., 2004. Competition and coalition among underwriters: The decision to join a syndicate. Journal of Finance 59, 2421–2444.
Sørensen, M., 2005. How smart is smart money? An empirical two-sided matching model of venture capital. Journal of Finance, in press.
Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557–586.
Tanner, M.A., Wong, W.H., 1987. The calculation of posterior distributions by data augmentation. Journal of the American Statistical Association 82, 528–550.
Villalonga, B., 2004. Does diversification cause the "diversification discount"? Financial Management 33, 5–27.
Villalonga, B., Amit, R., 2006. How do family ownership, control, and management affect firm value? Journal of Financial Economics 80, 385–417.
Vuong, Q., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57, 307–333.
Wald, J., Long, M., 2006. The effect of state laws on capital structure. Journal of Financial Economics, in press.
Wooldridge, J., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge.
Zhou, X., 2001. Understanding the determinants of managerial ownership and the link between ownership and performance: Comment. Journal of Financial Economics 62, 559–571.

*Chapter 3*

# AUCTIONS IN CORPORATE FINANCE*

SUDIPTO DASGUPTA

*Department of Finance, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*
*e-mail: dasgupta@ust.hk*

ROBERT G. HANSEN

*Tuck School of Business, Dartmouth College, Hanover, NH 03755, USA*
*e-mail: robert.g.hansen@dartmouth.edu*

## Contents



* We thank Parimal K. Bag, Espen Eckbo and Mike Fishman for comments and suggestions.

*Handbook of Corporate Finance, Volume 1*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

## Abstract

This paper reviews the applications of auction theory to corporate finance. It starts with a review of the main auction theory frameworks and the major results. It then goes on to discuss how auction theory can be applied, in the context of the market for corporate control, not only to "inform" a company's board or regulators, but also to understand some of the observed empirical evidence on target and bidder returns. It then considers the role of preemptive bidding, stock versus cash offers, the effect of toeholds on bidding behavior, the effect of bidder heterogeneity and discrimination in auctions, merger waves, bankruptcy auctions, share repurchases and "Dutch" auctions, IPO auctions, and the role of debt in auctions. It concludes with a brief discussion of the econometrics of auction data.

## Keywords

bidders, targets, private value, common value, winner's curse, auctions, bidding, takeovers, mergers and acquisitions, toeholds, bankruptcy auctions, IPO auctions

## 1. Introduction

This paper reviews developments in auction theory, with a focus on applications to corporate finance. Auctions, viewed broadly, are economic mechanisms that transfer control of an asset and simultaneously determine a price for the transaction.[1] Auctions are ubiquitous across the world. Formal auctions are used to buy and sell goods and services from fish to mineral rights and from logging contracts to lawyers' services in class action lawsuits. In the world of finance, auctions are used to buy and sell entire firms (in bankruptcy and out of bankruptcy) as well as securities issued by governments and companies. In the most recent and public example of auctions in corporate finance, the internet search firm Google sold its shares via a Dutch auction method in its initial public offering.

Auction theory has developed to explore a variety of issues, with the most important ones relating to pricing, efficiency of the allocation, differential information, collusion, risk aversion, and of course a very large topic, the effects of different auction rules (sealed-bids versus open auctions, reserve prices, entry fees, etc.) on the revenue to the seller.[2] Concomitant with theoretical work, there has been significant work in applications of auction theory, with many of these being related in some way to corporate finance. On one level, application of auction theory to corporate finance is very natural, for corporate finance sometimes directly involves auctions (e.g., auctions in bankruptcy). At another level, though, auction theory should serve to inform corporate finance because the underlying primitive issues are the same: pricing of assets, exchange of control, uncertainty especially in regard to asset valuation, heterogeneity of agents, asymmetric or disparate information, and strategic behavior. Given this similarity in the underlying frameworks, one should expect auction theory to have significant influence, both direct and indirect, on corporate finance research. There has also been, particularly in recent years, much work in the estimation of auction models. The econometrics of this work is very sophisticated, utilizing structural estimation methods that can retrieve estimates of the underlying distribution of bidders' valuations from bid data. While these techniques have not yet been applied to data from finance-related auctions, there would seem to be room for application to, for instance, corporate bankruptcy auctions. The broad lesson from these econometric studies is also very relevant for empirical work in financial auctions: use the restrictions from the theory to learn more from the data than non-structural methods will reveal. For this reason, empirical finance researchers studying auctions should have a good knowledge of auction theory.

[1] Throughout this survey, we will normally consider auctions where an item is being sold. Reverse auctions, where an auction is used to purchase an item, can generally be modeled by simply reversing the direction of payment.

[2] Krishna (2002) provides an excellent, comprehensive review of all existing auction theory. Klemperer (2000) is a shorter, recent review of auction theory, while McAfee and McMillan (1987) is thorough but a bit dated by now.

A careful review of the literature shows that auction theory has had a significant but not overwhelming influence on corporate finance. Perhaps the more insightful applications have been in the context of corporate takeover bidding: pre-emptive bidding, means of payment (takeover auctions are not always financed with cash), bidder heterogeneity, and discrimination amongst bidders. Application of auction theory to these contexts has at times produced new insights. Overall, however, while the applications have extended our understanding of the inefficiencies that are due to the underlying primitive construct of private information, they have not changed that understanding in any fundamental way.

The survey proceeds as follows. Section 2 reviews the simplest auction setting, that of independent private values. Many key insights can be developed from this simplest model: the basic pricing result that an auction's expected price equals the expected second-highest value; general solution methodology; effects of more bidders' risk aversion, reserve prices; revenue equivalence of the different auction forms; revenue enhancement from ex-post means-of-payment; and the solution of auction models via the Revelation Principle. Section 3 considers the interdependence amongst bidders' valuations (including the special case of a "common value" for the object) and reviews Milgrom and Weber's (1982a, 1982b) generalized auction model. Critical insights in this section pertain to the effects of the winner's curse; that lack of disclosure by the seller can lower expected prices; and that the different auction forms are no longer revenue-equivalent. With the basic theory developed in these sections, Section 4 turns to the applications most relevant to corporate finance. Section 5 ends with some thoughts about future applications and further development of auction theory that would make it more relevant for corporate finance.

## 2. The most basic theory: Independent private values

### 2.1. Initial assumptions

Auction theory begins with assumptions on how bidders value the asset for sale; the model then shows how an auction converts valuations into a price and an exchange of control. Valuation assumptions are absolutely key to auction theory. However, as we will argue later, the existing paradigms are not complete as they do not consider certain sets of valuation assumptions that are particularly relevant in corporate finance.

Independent preference (sometimes called independent private values) assumptions are straightforward: each bidder is simply assumed to know her value for the asset. For bidder $i$, denote this value as $v_i$. While each bidder knows her own value, to make the situation realistic and interesting, we assume that a bidder does not know other bidders' values. To model this uncertainty, we assume that each bidder believes other bidders' values to be independent draws from a distribution $F(v)$. We have therefore introduced a degree of symmetry in the model, that of symmetric beliefs.[3]

Fix a particular bidder, and focus on the highest value among the remaining $N-1$ values from the other $N-1$ bidders, and denote this value as $v_2$. Since $v_2$ is the highest among $N-1$ independent draws from the same distribution, its probability distribution $G(v_2)$ (i.e., the probability that $N-1$ independent draws are less than a value $v_2$) is

$$G(v_2) = F(v_2)^{N-1}. \tag{1}$$

Notice that the distribution $G(v_2)$ has a density function $g(v_2) = (N-1)F(v_2)^{N-2} \times f(v_2)$. If $F(v)$ is uniform over the unit interval, i.e., $F(v) = v$ for $0 \leqslant v \leqslant 1$, then note that

$$G(v_2) = F(v_2)^{N-1} = v_2^{N-1}. \tag{2}$$

### 2.2. *First-price sealed-bid auctions*

We are now in a position to evaluate any specific set of auction rules. Turn first to the common first-price sealed-bid auction, where bidders submit sealed bids and the highest bidder wins and pays the amount of her bid (hence the "first-price" qualifier). For now we assume a zero reserve price (a price below which the seller will keep the asset rather than sell).

In placing a bid $b$, bidder $i$ has expected profit of

$$E(\pi_i) = \Pr(\text{win})(v_i - b), \tag{3}$$

where one can note that in the case that bidder $i$ loses, her profit is zero. While (3) does not make it explicit, Pr(win) will be a function of $b$, normally increasing. This creates the essential tension in selecting an optimal bid: increasing one's bid increases the chance of winning, but the gain upon winning is less.

To solve this model, we need just a bit more structure. Let us use an intuitive version of the so-called Revelation Principle. Fix a bid function $b(v)$, and think of bidder $i$ as choosing the $v$ she "reports" rather than choosing her actual bid. So long as $b(v)$ is properly behaved, we have not restricted bidder $i$'s choice in any way, for she could get to any bid $b$ desired by simply "reporting" the requisite $v$.

Looking ahead, we are searching for a symmetric Nash equilibrium in bidding strategies. In terms of our $b(v)$ function, symmetry means that all bidders use the same $b(v)$. Nash equilibrium requires that, given other bidders' strategies, bidder $i$'s bid strategy is optimal. In terms again of our $b(v)$ formulation, equilibrium requires each bidder to report $v = v_i$, i.e., "honest" reporting. Our requirement for Nash equilibrium will therefore be as follows. Suppose that the other bidders are using $b(v)$ and honestly reporting,

[3] Several papers examine the effects of asymmetric beliefs, for example, Maskin and Riley (2000a, 2000b). See also Krishna (2002).

so that bidder $j$'s bid is $b(v_j)$. If $b(v)$ represents a (symmetric) bidding equilibrium, then bidder $i$'s optimal decision will be to report $v = v_i$, so that her bid is $b(v_i)$.

In the situation where the other $N - 1$ bidders are both using $b(v)$ and reporting honestly, we can re-write (3) as

$$\begin{aligned} E(\pi_i) &= \Pr(\text{win})(v_i - b) \\ &= G(v)\big(v_i - b(v)\big), \end{aligned} \tag{4}$$

where we assume bidder $i$ is using $b(v)$ but not requiring $v = v_i$. Note that bidder $i$ wins if all other $N - 1$ values are less than the $v$ that bidder $i$ reports, hence the conversion of Pr(win) into $G(v)$, the distribution for the highest value among the remaining $N - 1$ values.

Now we simply require that bidder $i$'s optimum decision is also honest reporting. Taking the first derivative of (4) with respect to $v$, we have

$$\frac{dE(\pi_i)}{dv} = g(v)\big(v_i - b(v)\big) - G(v)\frac{db(v)}{dv} = 0. \tag{5}$$

The first term of (5) shows the marginal benefit of bidding higher while the second term shows the marginal cost. Re-arranging, we have

$$G(v)\frac{db(v)}{dv} = g(v)\big(v_i - b(v)\big). \tag{6}$$

For equilibrium, we require that (6) hold at $v = v_i$. Hence we get

$$G(v)\frac{db(v)}{dv} = g(v)\big(v - b(v)\big). \tag{7}$$

Equation (7) is a standard first-order differential equation that can be solved via integration-by-parts.[4] Doing this yields

$$b(v) = \frac{1}{G(v)}\int_0^v yg(y)\,dy. \tag{8}$$

Equation (8) can be easily interpreted. As $G(x)$ is the distribution for the highest value among the remaining $N - 1$ values, $g(x)/G(v)$ is the density of that value conditional on it being lower than $v$. Equation (8) tells a bidder to calculate the expected value of the highest value among the remaining $N - 1$ bidders, conditional on that value being less than bidder $i$'s own, and to bid that amount. This is about as far as intuition can take us: the expected value of the second-highest value is in some sense bidder $i$'s real competition, and equilibrium bidding calls for her to just meet that competition. (One other intuitive approach involves marginal revenue; we will turn to this view below.)

[4] Rewrite (7) as $G(y)\,db + b\,dG = y\,dG$ or $d(G(y)b(y)) = y\,dG$. Integrating, and using the fact that $G(0) = 0$, we get $b(v) = \int_0^v y\,dG/G(v) = \int_0^v yg(y)\,dy/G(v)$.

If beliefs on values are governed by the uniform distribution, then $G(v) = v^{N-1}$, $g(v) = (N-1)v^{N-2}$, and (8) becomes

$$\begin{aligned} b(v) &= \frac{1}{v^{N-1}} \int_0^v y(N-1)y^{N-2}\,dy \\ &= \frac{N-1}{v^{N-1}} \int_0^v y^{N-1}\,dy \\ &= \frac{N-1}{N} v. \end{aligned} \tag{9}$$

In the particular case when $N = 2$, (9) implies that equilibrium bidding calls for bidding half of one's value—a significant "shading" of one's bid beneath true value. Note that in this case, however, the lowest the competitor's value could be is zero. If the distribution of values was instead uniform over [8, 10], the equilibrium bid would be $(8 + v)/2$—halfway between the lower bound and one's own valuation.

To see that in general, there is bid shading, notice that we can write

$$\begin{aligned} b(v) &= \frac{1}{G(v)} \int_0^v yg(y)\,dy = \frac{1}{G(v)} \int_0^v y\,dG(y) \\ &= \frac{1}{G(v)} \left[ yG(y)\big|_0^v - \int_0^v G(y)\,dy \right] \\ &= v - \int_0^v \frac{G(y)}{G(v)}\,dy \\ &= v - \int_0^v \bigl[F(y)/F(v)\bigr]^{N-1}\,dy, \end{aligned}$$

where we have used integration-by-parts in the third line.[5] Notice that while $b(v) < v$, since $F(y) < F(v)$ within the integral, as $N \to \infty$, $b(v) \to v$. In other words, intense competition will cause bidders to bid very close to their true values, and be left with little surplus from winning.

How does the seller fare in this first-price auction? We can construct the seller's expected revenue by calculating the expected payment by one bidder and then multiplying that by $N$. Sticking to the uniform [0, 1] distribution for clarity, we have the expected payment by bidder $i$ as

$$\begin{aligned} E(\text{Payment}_i) &= \int_0^1 \Pr(\text{win}) b(y)\,dy \\ &= \int_0^1 y^{N-1} \frac{N-1}{N} y\,dy = \frac{N-1}{N(N+1)}. \end{aligned} \tag{10}$$

[5] Since $d(uv) = u\,dv + v\,du$, we can write $\int u\,dv = uv - \int v\,du$. This handy trick is used very commonly in the auction literature.

Multiplying this by $N$ gives the seller's expected revenue as

$$E(\text{Revenue to seller}) = \frac{N-1}{N+1}. \tag{11}$$

Intuitively, since each bidder is bidding her expectation of the highest value among the remaining $N-1$ bidders, conditional on her value being the highest, the expected payment received by the seller should be the unconditional expected value of the second-highest value. In general, the density for the second-highest value is, from the theory of order statistics,[6]

$$f_2(y_2) = N(N-1)\big(1-F(y_2)\big)F(y_2)^{N-2}f(y_2). \tag{12}$$

In the case of the uniform [0, 1] distribution, the expected value of the second-highest value is then

$$\begin{aligned} E(y_2) &= \int_0^1 xN(N-1)(1-x)x^{N-2}\,dx \\ &= \frac{N-1}{N+1} \end{aligned} \tag{13}$$

as expected.[7] To reiterate and emphasize, the seller's expected revenue from the auction is exactly the expected value of the second-highest value. This result, of course, extends beyond the uniform distribution.

### 2.3. *Open and second-price sealed-bid auctions*

As compared to the first-price sealed-bid auction, the open auction and second-price sealed-bid auctions are considerably easier to solve. For this reason, they are often chosen to model any kind of auction mechanism; the Revenue Equivalence Theorem discussed below ensures that, in many cases, the results for one auction form extend to others.

In an open auction, bidders cry out higher and higher bids until only one bidder, the winner, remains. It is easy to see that "staying in the auction" until the bid exceeds one's

[6] To see this, first note that the distribution function $F_2(y_2)$ of the second-highest value $y_2$ is the probability that either: (a) all $N$ values are less than or equal to $y_2$ or (b) *any* $N-1$ values are less than $y_2$ and the remaining value is greater than $y_2$. Note that this latter event can happen in $N$ possible ways. Thus, the probability is $F_2(y_2) = F^N(y_2) + NF^{N-1}(y_2)(1-F(y_2))$. Differentiating this expression with respect to $y_2$, we get the expression for $f_2(y_2)$.

[7] This interpretation of the expected revenue holds for any distribution. Notice that the expected payment from any bidder is $\int_0^{\bar{v}}[\text{Prob(win)} \cdot \text{Amount Bid}]f(v)\,dv = \int_0^{\bar{v}} G(v)b(v)f(v)\,dv = \int_0^{\bar{v}}(\int_0^{v} yg(y)\,dy)\,dF(v)$ from (8). Integrating by parts, this expression becomes $\int_0^{\bar{v}} yg(y)\,dy - \int_0^{\bar{v}} F(v)vg(v)\,dv = \int_0^{\bar{v}} yg(y)\,dy - \int_0^{\bar{v}} F(y)yg(y)\,dy = \int_0^{\bar{v}} y(1-F(y))g(y)\,dy = \int_0^{\bar{v}} y(1-F(y))(N-1)F^{N-2}(y)f(y)\,dy$. $N$ times this expression is the expected revenue to the seller in the auction, and is exactly the expected value of the second-highest valuation.

value is a dominant strategy.[8] Staying in the auction beyond the point of the bid equaling one's value cannot be rational. Likewise, if the item is about to be won by someone else at a bid less than $v_i$, then bidder $i$ should be willing to bid a bit higher than the current bid, for if such a bid wins the auction it will yield a profit.

An open auction therefore will quite easily find the second-highest valuation and establish that as the price—for bidding will cease once the bidder with the second-highest valuation is no longer willing to bid more.[9] The expected price in the open auction is therefore the expected value of the second-highest valuation, the same as for the first-price sealed-bid auction. This result is an implication of the Revenue Equivalence Theorem; we return to a more general statement of that below.

Turn now to a second-price sealed-bid auction: in such an auction, sealed bids are submitted and rules call for the highest bid to win but that the price paid will be the second-highest bid submitted. With these rules it is again a dominant strategy to submit a bid equal to one's valuation. Bidding more than one's value would mean possibly winning at a price in excess of value. Bidding less than one's value will mean possibly forgoing an opportunity to buy the object at a price less than value. The key to understanding the second-price auction is to note that the linkage between one's bid and the price one pays has been severed; bidding equal to value to maximize the probability of profitable wins becomes optimal. Thus, as is the case for the open auction, the second-price sealed bid auction will also yield as a price the second-highest value out of the $N$ values held by the bidders.[10]

All three auctions therefore yield the same expected price.[11] Note, however, that the first-price and second-price (including the open auction as essentially a second-price auction) auctions have equilibrium strategies that are easy to compute for both the modeller and the bidder. Note also that the first-price auction gives a different (and less volatile) price for any given set of bidders. It is also important that all three auctions are efficient in that the bidder with the highest valuation is the winner. Auctions can be seen as accomplishing two distinct tasks: reallocating ownership of an asset and determining a price for the transfer of ownership. Efficiency is an important characteristic of any sales procedure, and auctions under private value assumptions should get the asset

[8] A (weakly) dominant strategy in game theory is a strategy which does at least as well as any other strategy no matter what strategies other agents use.

[9] This neglects effects (usually unimportant) of a minimum bid increment.

[10] Therefore, the second-price and the open (also called English or Ascending) auctions are "equivalent" in the sense that they lead the bidders to bid or drop out at their private value for the object. However, this "equivalence" holds only in the private values setting. If the other bidders' signals or valuations are relevant for a given bidder's valuation of the object, this equivalence breaks down, as the open bids by the other bidders conveys additional information.

[11] The first-price auction is "strategically equivalent" to yet another auction known as the Dutch auction, which an open descending price auction in which the auctioneer starts with a high price and then gradually lowers the price until some bidder accepts the price. Provided that the object has not been sold yet, a bidder will accept an asking price that equals her bid in the first-price auction. Strategic equivalence is a stronger notion than the equivalence between the open and the second-price auctions.

to its most highly-valued use. Reserve prices, considered below, may hamper this efficient transfer. Efficiency of auctions under asymmetric beliefs is also not assured (see Krishna, 2002, for further discussion).

### 2.4. Revenue equivalence

The result that the second-price auctions and the first-price auction yield the same expected revenue to the seller is a consequence of the so-called "Revenue Equivalence Theorem". What is fascinating about the revenue equivalence of these two auctions is that such sophisticated models confirm a result which is really quite intuitive: different mechanisms all yield what is really a "competitive" price, that being the second-highest valuation. The seller cannot, under these standard auction rules, extract any more revenue than the valuation of the second-highest bidder.

The revenue equivalence result in this independent private value context can be generalized—not only to encompass a broader class of auctions, but also a more general value environment. Suppose that each bidder $i$ privately observes an informational variable $x_i$. To simplify notation, we assume $N = 2$. Assume that $x_1$ and $x_2$ are independently and identically distributed with a distribution function $F(x_i)$ and density $f(x_i)$ over $[0, \bar{x}]$ for $i = 1, 2$. Let $v_i = v(x_i, x_j)$ denote the value of the object to bidder $i$, $i = 1, 2$ and $i \neq j$.

Consider a class of auctions in which the equilibrium bid function is symmetric and increasing in the bidder's signal, and let $A$ denote a particular auction form. Let $\Pi_i^A(z, x)$ denote the expected payoff to bidder $i$ when she receives signal $x_i = x$ and bids as if she received signal $z$. Then

$$\Pi_i^A(z, x) = \int_0^z v(x, y) f(y)\, dy - P_i^A(z),$$

where $P_i^A(z)$ denotes the expected payment conditional on bidding as if the signal were $z$, and we have used the assumption that the bidders have symmetric and increasing bid functions, so that $i$ wins if and only if $x_j < z$. Differentiating with respect to $z$, we get:

$$\frac{\partial \Pi_i^A(z, x)}{\partial z} = v(x, z) f(z) - \frac{d P_i^A(z)}{dz}.$$

In equilibrium, $\frac{\partial \Pi_i^A(z,x)}{\partial z} = 0$ at $z = x$, and hence

$$\frac{d P_i^A(y)}{dy} = v(y, y) f(y).$$

Integrating, we get

$$P_i^A(x) = P_i^A(0) + \int_0^x v(y, y) f(y)\, dy.$$

Notice that $P_i^A(0)$ is the expected payment made by bidder $i$ with the lowest draw of the signal. Since the seller's expected revenue is simply 2 times $\int_0^{\bar{x}} P^A(x) f(x)\, dx$, it follows that all auctions in which the bid functions are symmetric and increasing, and in which the bidder drawing the lowest possible value of the signal pays zero in expected value, are "revenue equivalent".[12]

The model considered here is one in which the values of the bidders are "interdependent" in the sense that one bidder's signal affects the value (estimate) of the other bidders. The signals themselves, however, are statistically independent. An example of the value function we considered here would be, for example, $v_1 = \alpha x_1 + (1 - \alpha)x_2$ and $v_2 = \alpha x_2 + (1 - \alpha)x_1$, where $1 \geqslant \alpha \geqslant 0$. Clearly, the independent private values model is a special case, in which $\alpha = 1$. The case of $\alpha = 1/2$ corresponds to a case of the "pure common value" model, for which $v(x, y) = v(y, x)$, i.e., the bidders have identical valuations of the object as a function of both bidders' signals.

### 2.5. Reserve prices

As reserve prices have figured in some of the corporate finance literature, it is worthwhile to consider analysis of reserve prices in auctions. Sticking with independent private values, consider an open auction with two bidders. Suppose that bidder 1 has valuation $v_1 > 0$ and bidder 2 has valuation $v_2 = 0$. Then the open auction will yield a price of zero. Better in this case would be for the seller to have a reserve price set in-between 0 and $v_1$ so that bidder 1 would still win but pay the reserve. Of course, the problem with a reserve price is that if it is set above $v_1$ no sale will result.

To understand how the reserve price is chosen,[13] let us return to the independent private values model with $N$ bidders. Consider any auction form $A$ in the class of auctions with symmetric increasing bid functions. As above, denote by $P^A(z)$ the expected payment by a given bidder in auction $A$ when she bids $b^A(z)$. If the bidder's private value of the object is $v$, her expected profit is

$$\Pi^A(z, v) = G(z)v - P^A(z),$$

where $G(z) = F^{N-1}(z)$. As above, in equilibrium, it must be optimal for the bidder with valuation $v$ to bid $b(v)$, which requires that $\Pi^A(z, v)$ is maximized at $z = v$. This implies that

$$g(y)y = \frac{dP^A(y)}{dy}. \tag{14}$$

Let us suppose now that a bidder with private value $v^*$ is indifferent between bidding and not bidding. For such a bidder (known as the "marginal bidder"), by definition

[12] Absent reserve prices, the bidder drawing the lowest possible signal will typically be indifferent between bidding and not bidding.

[13] Our treatment of the problem here follows that in Riley and Samuelson (1981).

$\Pi^A(v^*, v^*) = G(v^*)v^* - P^A(v^*) = 0$. Now from (14), integrating, we get for $v \geqslant v^*$

$$\begin{aligned} P^A(v) &= P^A(v^*) + \int_{v^*}^{v} y g(y)\, dy \\ &= G(v^*)v^* + \int_{v^*}^{v} y\, dG(y) \\ &= vG(v) - \int_{v^*}^{v} G(y)\, dy, \end{aligned} \tag{15}$$

where in the last step, we used integration by parts.

The expected revenue for the seller from a single bidder is $\int_0^{\bar{v}} P^A(v) f(v)\, dv$. Again, using integration by parts, this can be written as

$$\begin{aligned} E(R_i^A) &= \int_0^{\bar{v}} P^A(v) f(v)\, dv \\ &= \int_{v^*}^{\bar{v}} P^A(v) f(v)\, dv \\ &= \int_{v^*}^{\bar{v}} vG(v) f(v)\, dv - \int_{v^*}^{\bar{v}} \left[ \int_{v^*}^{v} G(y)\, dy \right] dF \\ &= \int_{v^*}^{\bar{v}} vG(v) f(v)\, dv - \int_{v^*}^{\bar{v}} G(y)\, dy + \int_{v^*}^{\bar{v}} F(v)G(v)\, dv \\ &= \int_{v^*}^{\bar{v}} \big[ vf(v) - \big(1 - F(v)\big)\big] G(v)\, dv. \end{aligned} \tag{16}$$

Given equal treatment of all $N$ buyers, the expected revenue to the seller is simply $N$ times the above expression.

Notice that what we have shown is that all auction forms in the class of auctions being considered must provide the seller with the same expected revenue if the marginal bidder is the same. The reserve price will determine the marginal bidder. If no bidder has a valuation above that of the marginal bidder, the seller keeps the object. Assume that the seller values the object at $v_0$. Then for any auction, the seller should choose the marginal bidder to maximize

$$\int_{v^*}^{\bar{v}} \big[ vf(v) - \big(1 - F(v)\big)\big] G(v)\, dv + F(v^*) v_0.$$

From the first-order condition with respect to $v^*$, we get

$$v^* = v_0 + \frac{1 - F(v^*)}{f(v^*)}. \tag{17}$$

Since the optimal marginal bidder is the same in all auctions—all auctions in the class of auctions we are considering provide the seller with the same expected *profit* as well as revenue. The revenue equivalence result survives when a reserve price is introduced.

It remains to characterize the reserve prices in different auction settings. Suppose the reserve price is $r$. Notice that in both the first-price and the second-price auctions, no bidder with a value less than $r$ can make any positive profit, as they have to bid at least $r$ to win the object. On the other hand, the profit of a bidder with value greater than $r$ must be strictly positive (in the second price auction, if no other bidder bids higher than $r$, the bidder pays $r$). Thus, by continuity, the marginal bidder must have a value $v^* = r$.

Note from (17) that the optimal reserve price exceeds the seller's own valuation and is independent of the number of bidders. This latter point makes sense given that the optimal reserve price is only aimed at making the high bidder pay more in the instance when all other valuations are beneath the reserve price. Note also that a reserve price destroys the assurance of an efficient allocation; in the case where the highest valuation among the bidders is less than $v^*$ but greater than $v_0$, the seller will retain possession even though one of the bidders has a valuation greater than the seller.

Notice also that *entry fees* are an alternative way of implementing a positive reserve price. By setting an entry fee equal to the expected profit of a bidder with value $r$ when the reserve price is 0,[14] the seller can ensure that a bidder participates if and only if her value exceeds $r$.

### *2.6. Optimal selling mechanisms*

Auctions are best thought of as "selling mechanisms"—ways to sell an object when the seller does not know exactly how the potential buyers value the object. There is obviously a very large number of ways in which an object could be sold in such a situation: for example, the seller could simply post a price and pick one bidder randomly if more than one buyer is willing to pay that price; post a price and then negotiate; use any one of the common auctions; use any of the *less* common forms of auction such as an "all pay" auction in which all bidders pay their bids but only the highest bidder gets the object; impose non-refundable entry fees; use a "matching auction" in which one bidder bids first and the other bidder is given the object if he matches the first bidder's bid, and so on. The search for an optimal selling scheme in a possibly infinite class of selling schemes would indeed seem like a daunting task. The major breakthrough, however, was the insight that without loss of generality, one could restrict attention to selling mechanisms in which each buyer is induced to report her valuation (often called "type") truthfully. This is the so-called "Revelation Principle" (Myerson, 1981; Dasgupta, Hammond and Maskin, 1979; Harris and Raviv, 1981), and it greatly simplified the formulation of the problem.

Armed with the Revelation Principle, one can attack the problem in a more general setting than we have discussed so far. While we will still remain within the confines of the independent private values framework,[15] we can dispense with the assumption

[14] From (15), this is $\int_0^r G(y)\,dy$.

[15] Myerson's (1981) framework is slightly more general in that he allows the value estimate of a bidder as well as the seller to depend on the signals of all other bidders, i.e., his model is one in which the signals are independent and private, but the valuations are interdependent.

that all bidders' valuations are drawn from identical distributions, i.e., one can accommodate asymmetries among bidders. Asymmetries are important in many real world situations—for example, in procurement, when both domestic and foreign bidders participate, and especially in corporate finance, in the context of takeover bidding.

Before proceeding further, however, we need to introduce some notation. Let $\mathbf{v} = (v_1, v_2, \ldots, v_N)$ denote the set of valuations for bidders $1, \ldots, N$ and let $\mathbf{v} \in \mathbf{V} \equiv (\times V_i)_{i=1}^N$, where $V_i$ is some interval $[0, \bar{v_i}]$. Likewise, let $\mathbf{v_{-i}} = (v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_N)$, and $\mathbf{v_{-i}} \in \mathbf{V_{-i}} \equiv (\times V_i)_{j=1, j\neq i}^N$. Let $f(\mathbf{v})$ denote the joint density of the values; since the values are independently drawn, we have $f(\mathbf{v}) = f_1(v_1) \times f_2(v_2) \times \cdots \times f_N(v_N)$, and $f_{-i}(\mathbf{v_{-i}}) = f_1(v_1) \times \cdots \times f_{i-1}(v_{i-1}) \times f_{i+1}(v_{i+1}) \times \cdots \times f_N(v_N)$ is similarly defined.

The seller picks a mechanism, i.e., an allocation rule that assigns the object to the bidders depending on messages sent by the latter. By appealing to the Revelation Principle, we can restrict attention to *direct* mechanisms, i.e., mechanisms that ask the bidders to report their values $v_i$. Thus, the mechanism consists of a pair of functions $\langle Q_i(\mathbf{v}'), P_i(\mathbf{v}')\rangle_{i=1}^N$ for each $i$ which states the probability $Q_i$ with which the object would go to bidder $i$ and the expected payment $P_i$ that bidder $i$ would have to make for any vector of *reported* values of the bidder valuations. Of course, the mechanism has to satisfy two conditions: (i) it must be Incentive Compatible, i.e., it must be (weakly) optimal for each bidder to report her value truthfully given that all others are doing the same, and (ii) it must be Individually Rational, i.e., the bidders must be at least as well off participating in the selling process than from not participating.

Thus, the probability that bidder $i$ gets the object when she reports her value to be $z_i$ and all other bidders report truthfully is

$$q_i(z_i) = \int_{\mathbf{V_{-i}}} Q_i(z_i, \mathbf{v_{-i}}) f_{-i}(\mathbf{v_{-i}})\, d\mathbf{v_{-i}},$$

and the expected payment he makes is

$$p_i(z_i) = \int_{\mathbf{V_{-i}}} P_i(z_i, \mathbf{v_{-i}}) f_{-i}(\mathbf{v_{-i}})\, d\mathbf{v_{-i}}.$$

It can be shown[16] that (i) Incentive Compatibility is equivalent to the requirement that the $q_i(v_i)$ functions are non-decreasing, i.e., the probability that a bidder gets the object is non-decreasing in her reported value of the object, and (ii) Individual Rationality is equivalent to the requirement that the $p_i(v_i)$ functions satisfy $p_i(0) \leqslant 0$, i.e., the bidder with zero value has non-positive expected payment. It can also be shown that in the optimal selling mechanism, the $Q_i(\mathbf{v})$ need to be chosen to maximize the following

[16] For details, please see Myerson (1981) or Krishna (2002).

expression:

$$\sum_{i=1}^{N} p_i(0) + \sum_{i=1}^{N} \int_{\mathbf{V}} \left( v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right) Q_i(\mathbf{v}) f(\mathbf{v})\, d\mathbf{v}$$
$$= \sum_{i=1}^{N} p_i(0) + \int_{\mathbf{V}} \left( \sum_{i=1}^{N} J_i(v_i) Q_i(v_i) \right) f(\mathbf{v})\, d\mathbf{v} \tag{18}$$

and the payment made by bidder $i$ needs to satisfy

$$P_i(\mathbf{v}) = Q_i(\mathbf{v}) v_i - \int_0^{v_i} Q_i(z_i, \mathbf{v_{-i}})\, dz_i. \tag{19}$$

The quantities $J(v_i) = v_i - \frac{1-F_i(v_i)}{f_i(v_i)}$ are known as "virtual valuations" for reasons that will become clear below. Notice that $\frac{1-F_i(v_i)}{f_i(v_i)}$ is the inverse of the hazard rate $\frac{f(v_i)}{1-F_i(v_i)}$. If the hazard rate is increasing, then the virtual valuations are increasing in $v_i$. This is known as the "regular case" in the literature.

Ignoring the Incentive Compatibility and Individual Rationality constraints for the moment, it is clear that the objective function (18) is maximized pointwise if $Q_i(\mathbf{v})$ is set equal to the maximum value (i.e. 1, since it is a probability) when $J_i(v_i)$ is the highest for any realized **v**, and zero otherwise. Two implications immediately follow.

First, notice that the allocation rule implies that if the bidders are symmetric (i.e., the private values are drawn from the same distribution $F(v_i)$ for all $i$), then the bidder with the highest value gets the object with probability one. Moreover, from (19), any two selling procedures that have the same allocation rule must also result in the same expected payment made by the bidders and thus result in the same expected revenue for the seller. In particular, when the bidders are symmetric, all the standard auctions—since they result in the highest value bidder getting the object with probability 1—are optimal selling mechanisms and result in the same expected revenue for the seller.

Second, if the bidders are not symmetric, then the object need not go to the bidder with the highest $v_i$. For example, suppose $f_i(v_i) = \frac{1}{b_i - a_i}$. Then $J_i(v_i) = 2v_i - b_i$. Thus, $v_i > v_j \Rightarrow J_i(v_i) > J_j(v_j)$ if and only if $v_i - v_j > (b_i - b_j)/2$. In other words, the high-value bidder may not get the object if the upper bound on her value for the object is sufficiently high. The intuition is that the potential for such a bidder to under-represent her value is high; thus, by discriminating against her in terms of the likelihood of being awarded the object, the seller induces her to report truthfully when her valuation is high. The basic message here is of considerable importance, as we will see in more detail later: when bidders are asymmetric, it may pay to discriminate against the stronger bidder.[17]

[17] Notice that in the regular case, since the virtual valuations are non-decreasing, the $q_i$'s are non-decreasing as well. Moreover, it is easily checked that $P_i(0, \mathbf{v_{-i}}) = 0$ for all $\mathbf{v_{-i}}$; hence $p_i(0) = 0$ for all $i$. Thus, incentive compatibility and individual rationality conditions are satisfied.

*2.7. Interpreting the optimal auction: The marginal revenue view*

Bulow and Roberts (1989) provide an intuitive interpretation of the "virtual valuations" $J_i(v_i)$ according to which the object is allocated in the optimal selling scheme. Interpret $v_i$ as a "price" and $1 - F_i(v_i)$ as a demand curve: if a price $p$ is set as a take-it-or-leave-it price, $1 - F_i(p)$ gives the probability of a sale, i.e., the "quantity" $q(p)$ sold at price $p$. We can then calculate a marginal revenue curve in the usual way, but using $1 - F_i(v_i)$ as the demand curve:

$$\begin{aligned}\text{Total Revenue} &= v_i q(v_i)\\ \Rightarrow (\text{Marginal Revenue}) &= \frac{d(\text{Total Revenue})}{dq}\\ &= v_i + q(v_i)\frac{dv_i}{dq}\\ &= v_i + \bigl(1 - F_i(v_i)\bigr)\frac{1}{dq/dv_i}\\ &= v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}.\end{aligned} \tag{20}$$

Thus, the virtual valuations are marginal revenues, and the optimal mechanism awards the good to the bidder with the highest marginal revenue. Bulow and Roberts (1989) in fact provide the following "second marginal revenue" auction interpretation of the optimal selling scheme. Each bidder is asked to announce her value, and the value is converted into a marginal revenue. The object is awarded to the bidder with the highest marginal revenue ($M_1$), and the price she pays is the lowest value that she could have announced without losing the auction (i.e., $MR_1^{-1}(M_2)$).[18]

Why does the "second marginal revenue" auction call for the winner to pay the lowest value she could announce without losing the auction? This is, in fact, a property of the optimal selling mechanism discussed in the previous section. To see this, define $s_i(\mathbf{v_{-i}})$ as the smallest value (more precisely, the infimum) of $v_i$ for which $i$'s virtual valuation (marginal revenue) would be no less that the highest virtual valuation from the rest of the values. Clearly, $Q_i(z_i, \mathbf{v_{-i}}) = 1$ if $z_i > s_i(\mathbf{v_{-i}})$ and 0 otherwise. Thus, $Q_i(z_i, \mathbf{v_{-i}})$ is a step function, and this implies that $\int_0^{v_i} Q(z_i, \mathbf{v_{-i}})\,dz_i = v_i - s_i(\mathbf{v_{-i}})$ if $v_i > s_i(\mathbf{v_{-i}})$ and 0 otherwise. Since $v_i > s_i(\mathbf{v_{-i}})$ implies $Q_i(\cdot\,,\cdot) = 1$ and $\int_0^{v_i} Q(z_i, \mathbf{v_{-i}})\,dz_i = v_i - s_i(\mathbf{v_{-i}})$, from (19) we get $P_i(v_i, \mathbf{v_{-i}}) = s_i(\mathbf{v_{-i}})$ for the winning bidder. Thus, the bidder with the highest marginal revenue pays the lowest value that would win against all other values when the object is allocated according to the marginal revenue rule.

[18] If no bidder has positive marginal revenue, the seller keeps the object; if only one bidder has a positive marginal revenue, then she pays the price at which her marginal revenue is zero. It is easy to check that truthful reporting is a dominant strategy in this auction.

## 3. Common-value auctions

### 3.1. Common value assumptions

To this point we have mostly considered auctions where bidders' preferences were described by the independent private values assumptions. Clearly, in this framework, given their signals, bidders have complete information about the value of the object to themselves. We turn now to another class of models where each bidder has information that, if made public, would affect the remaining bidders' estimate of the value of the object. The general model could be described as each bidder having a value $V_i = v_i(t_1, t_2, \ldots, t_N)$, where $t_i$ represents bidder $i$'s signal. However, before we turn to the general model, it is useful to focus on a particularly important special case—the case of the "pure common value" model. In this scenario, every bidder has the same valuation for the item, hence the phrase "common value". In other words, we have

$$V_i = v(t_1, t_2, \ldots, t_N) \tag{21}$$

for each bidder $i$. Such an assumption is reasonable for auctions of many assets. The sale of a company, for instance, is sure to exhibit common-value characteristics, for the company's underlying cash flows will be uncertain but, at least to the first consideration, will be the same for all potential acquirers.[19]

Common or interdependent-value auctions involve a certain form of adverse selection, which if not accounted for by bidders, leads to what has been called the "winner's curse". Auctions are wonderful at selecting as winner the bidder with the highest valuation. However, the highest of several value estimates is itself a biased estimate, and this fact would cause the winner to adjust downward her estimate of the value of the object. For example, suppose that there are two bidders, the object is worth $v = t_1 + t_2$ to each, where each $t_i$ is an independent draw from the uniform [0, 1] distribution. Based on her signal alone, each bidder's estimate of the value is $t_i + 1/2$. However, if the bidders are symmetric, after learning that she is the winner in a first-price auction, bidder $i$'s estimate of the value will change to $t_i + E(t_j|t_j < t_i) = t_i + t_i/2 < t_i + 1/2$.

The point to emphasize here is that under almost any reasonable bidding scenario, the high bidder will be the one with the highest value estimate. While each bidder's estimate is an unbiased ex-ante estimate of the common value, the highest of those estimates is biased high. Or to put it another way, winning an auction gives a bidder information that they had the highest estimate of value. If one respects the fact that the other bidders are as good at estimating value as oneself, then the information that $N - 1$ other bidders thought the item is worth less should give one pause for reflection (and of course this pause should have been taken before the bid was submitted).

[19] The classroom "wallet game" mimics this particular common value auction model. In this game, two students are picked and each is asked to privately check the amount of money in his wallet. The teacher then announces that a prize equal to the combined amount of money in the wallets will be auctioned. The auction method is a standard ascending auction in which the price is gradually raised until one student drops out. The winner then gets the prize by paying that price. See Klemperer (1998).

### *3.2. Optimal bidding with a common value*

We begin with the illustrative example introduced above, and show how the principles apply.

Suppose there are only two bidders and the value to each bidder is given as

$$v = t_i + t_j, \tag{22}$$

where $t_i$ and $t_j$ are each bidder's privately known signals. We will suppose that the signals are independently distributed according to a uniform distribution on [0, 1].

Consider first a second-price auction. It is easy to show that in this auction, it is optimal for each bidder to bid $2t_i$. Suppose bidder $j$ is following this strategy, and bidder $i$ bids $b$. Then bidder $i$ wins the auction if $2t_j = b_j < b$, i.e., $t_j < b/2$. Her expected gain is $\int_0^{b/2}(t_i + \tilde{t} - 2\tilde{t})\, d\tilde{t} = t_i \frac{b}{2} - \frac{1}{2}\frac{b^2}{4}$. Maximizing with respect to $b$, one gets $b_i = 2t_i$, as claimed.

With two bidders, the second-price auction is equivalent to an ascending auction. Thus, it should be no surprise that the equilibrium bidding strategies in an ascending auction are identical to the one derived above. To see this, suppose bidder $j$ has a bidding strategy of $b_j = 2t_j$. If bidder $i$ continues to be in the auction at a price $b > 2t_i$, her profit if $j$ ended the auction by dropping out would be $t_i + b/2 - b = t_i - b/2 < 0$, and thus it cannot be optimal for her to be in the auction at that price. Similarly, if $b < 2t_i$ her profit if $j$ ends the auction would be $t_i + b/2 - b > 0$, and thus it cannot be optimal for her to quit at that price. Consequently, she must stay in the auction until the price reaches $2t_i$.

Notice that the bidders do take into account winner's curse in equilibrium. If the price reaches a level $b = 2t_i$, the value of the object is at least $t_i + b/2 = 2t_i$, since $j$ is still in the auction. Thus, the expected value is strictly higher than $2t_i$. However, $i$ would still quit at this price, because *if the auction had ended* at this price because $j$ quit, she would be breaking even. As we saw above, she would lose if the auction ends at any higher price and she is the winner.

A first-price auction is more complicated, but similar results hold. One can think of the optimal bid in a first-price auction as being the result of a two-stage process: first, adjust one's expected value for the bias associated with being the highest out of $N$ signals; and second, further lower the bid to account for the strategic nature of an auction.

### *3.3. Milgrom and Weber's* (1982a, 1982b) *generalized model*

#### *3.3.1. Core assumptions*

While both the independent private value and the pure common value model capture many key aspects of real auctions, they are obviously polar cases. Many real auctions will contain both private value and common value characteristics. In an auction of a company, for instance, the company's "core" cash flow will be a common value for all bidders, but synergies will likely differ across bidders and therefore contribute an element of independent private values. In a seminal paper, Milgrom and Weber (1982a, 1982b) developed analysis of a generalized valuation model for auctions. The key valuation assumption in Milgrom and Weber's general *symmetric* model is that the value of the item to bidder $i$ is given by

$$v_i = u(t_i, \mathbf{t_{-i}}). \tag{23}$$

In (23), $t_i$ is the signal privately observed by bidder $i$, and $\mathbf{t_{-i}}$ denotes the vector of signals $(t_i, t_2, \ldots, t_{i-1}, t_{i+1}, \ldots, t_N)$. The function $u(\cdot, \cdots)$ is non-decreasing in all its variables. The model is symmetric in the sense that interchanging the values of the components of $\mathbf{t_{-i}}$ does not change the value of the object to bidder $i$. In this symmetric model, note that both the private and pure common value models are special cases: if $v_i = u(t_i)$ for all $i$, we have the private value model, and if $v_i = u(t_1, t_2, \ldots, t_N)$ for all $i$ (i.e., $u(\cdot, \cdots)$ is symmetric in *all* the signals, then the model is a common value model. The interdependent values model with independent signals discussed earlier is also obviously a special case, in which the signals are i.i.d.

The symmetric model assumes that the joint density of the signals, denoted by $f(\cdot, \cdots, \cdot)$ is defined on $[0, \bar{t}\,]^N$, and is a symmetric function of its arguments. The density functions are also assumed to have a statistical property known as "affiliation", which is a generalized notion of positive correlation among the signals.

It will be convenient to work in terms of the expected value of the object to bidder $i$ conditional on her own signal $t_i$ and the highest among the remaining $N - 1$ signals. Without any loss of generality, we will focus on bidder 1, and accordingly, let us define

$$v(t, y) = E\big[v_i(\cdots)|t_1 = t, Y_1 = y\big], \tag{24}$$

where $Y_1$ is the highest signal among the remaining $N - 1$ signals of bidders $2, \ldots, N$. We will denote the distribution function of $Y_1$ by $G(y)$ and its density by $g(y)$. Notice that because of symmetry, it does not matter who among the remaining bidders has the highest signal, and moreover, by virtue of symmetry with respect to the way in which a bidder's own signal affects the value of the object to the bidder, the function is the same for all bidders. Because of affiliation, it follows that $v(\cdot, \cdot)$ is non-decreasing function in $t$ and $y$.

### 3.3.2. *Equilibrium bidding*

It is convenient to begin with the second-price auction. Generalizing the example in Section 3.2, we shall show that the symmetric equilibrium bid function is given by $v(t, t)$. Recall that the function $v(t, t)$ is the expected value of the bidder's valuation, conditional upon the bidder having signal $t$ and on the bidder with the second-highest signal also having signal $t$.

To see that $v(t, t)$ is the symmetric equilibrium bid function, notice that if bidder 1 bids $b_1$ assuming that all other bidders are following the proposed equilibrium bidding strategy, then her expected payoff is

$$\int_0^{b^{S^{-1}}(b_1)} \bigl(v(t,y) - v(y,y)\bigr) g(y|t)\, dy.$$

Differentiating, it is immediate that the first-order condition is satisfied if $b_1 = b^S(t)$ so that $b^{S^{-1}}(b_1) = t$.

Turning now to the ascending auction, suppose that the bidding is at a stage where all bidders are still active. Suppose bidder with signal $t$ has the strategy that she will remain in the bidding until the price $b^N(t) = u(t, t, \ldots, t)$ is established, provided no bidder has dropped out yet. If the first bidder to drop out does so at the price $p_N$, let $t_N$ be implicitly defined by $b^N(t_N) = p_N$. Then suppose every remaining bidder with signal $t$ has the strategy of staying until the price reaches $b^{N-1}(t, p_N) = u(t, t, \ldots, t, t_N)$. Let $p_{N-1}$ be the price at which the next bidder drops out. Then let $t_{N-1}$ be implicitly defined by $b^{N-1}(t_{N-1}, P_N) = p_{N-1}$. Now every remaining bidder has a strategy of remaining in the bidding until the price reaches $b^{N-2}(t, p_{N-1}, p_N) = u(t, t, \ldots, t_{N-1}, t_N)$. Proceeding in this manner, the bidding strategies of the bidders after each round can be written down until two bidders remain. Clearly, these strategies entail that each bidder drops out at that price at which, given the information revealed by the bidding up to that point, the expected value of the object would be exactly equal to the price if all remaining bidder except herself were to drop out all at once at that price.

We shall argue that these strategies constitute an equilibrium of the ascending auction. If bidder 1 wins the auction, then $t_1$ must exceed all other signals. Now, from the construction of the bidding strategies, it is clear that the bidder with highest signal among the remaining bidders quits at a price $u(y_1, y_1, y_2, y_3, \ldots, y_{N-1})$, where $y_i$ denotes the value of the $i$th highest signal among the rest of the bidders, i.e., excluding bidder 1. Thus, bidder 1 gets $u(t, y_1, y_2, y_3, \ldots, y_{N-1}) - u(y_1, y_1, y_2, y_3, \ldots, y_{N-1})$, which is strictly positive. Quitting earlier, she would have obtained zero, and any other strategy that makes her drop out after the bidder with signal $y_1$ cannot give her any higher payoff. Consider now a situation in which bidder 1 does not have the highest draw. For her to win the auction, she must have to pay $u(y_1, y_1, y_2, y_3, \ldots, y_{N-1})$; however, this exceeds the value of the object to her, which is $u(t, y_1, y_2, y_3, \ldots, y_{N-1})$. Thus, she cannot do better than drop out as prescribed by the equilibrium strategy.

To find the equilibrium bid in the first-price auction, assume that each of the other $N - 1$ bidders follow a bidding strategy $b^F(z)$, and that bidder 1 bids as though her private signal were $z$. Since the bids are increasing, the expected profit for bidder 1 whose signal is $t$ is

$$\Pi(z,t) = \int_0^{z} \bigl(v(t,y) - b^F(z)\bigr) g(y|t)\, dy.$$

The derivative of this expression with respect to $z$ is

$$\bigl(v(t,z) - b^F(z)\bigr) g(z|t) - b^{F'}(z) G(z|t),$$

which should be zero at $z = t$. Thus, we get

$$\big(v(t,t) - b^F(t)\big)\frac{g(t|t)}{G(t|t)} = b^{F'}(t). \tag{25}$$

Since $v(0,0) = 0$, we have the boundary condition $b^F(0) = 0$. The differential equation can then be solved[20]

$$b^F(t) = \int_0^t v(y,y)\,dL(y|t), \tag{26}$$

where $L(y|t) = \exp(-\int_y^t \frac{g(x|x)}{G(x|x)}\,dx)$.

It is easy to check that $L(\cdot|t)$ is in fact a probability distribution function on $[0,t]$, so that the expression for the equilibrium bid is an expected value with respect to some probability measure.

### *3.3.3. Revenue ranking and the linkage principle*

With affiliated signals, revenue equivalence no longer holds. The ascending auction generates at least as much expected revenue to the seller as the second-price auction, which in turn generates at least as much expected revenue as the first-price auction. While a direct comparison is possible, the so-called "Linkage Principle" provides a fundamental insight. Consider an auction $A$ in which a symmetric equilibrium exists, and suppose that all bidders are bidding in accordance with this symmetric equilibrium except possibly bidder 1, who has a signal $t$ but bids as though her signal were $z$ ($z$ could equal $t$). Suppose $W^A(z,t)$ denotes the expected price that is paid by that bidder if she is the winning bidder. Then the Linkage Principle says that of any two auctions $A$ and $B$ with $W^A(0,0) = W^B(0,0)$, the auction for which $W_2^i(t,t)$ (i.e., the partial derivate with respect to the second argument evaluated with both arguments at $t$) is higher will generate the higher expected revenue for the seller.

With the benefit of the Linkage Principle, it is easy to see why the first-price auction generates higher revenue than the second-price auction. In the first-price auction, a bidder with signal $t$ bidding as if the signal were $z$ would pay $b^F(z)$ conditional on winning, i.e., $W_2^F(z,t) = 0$ for all $t$ and $z$. On the other hand, in the second-price auction, the corresponding expected payment is $E[b^S(Y_1)|t_1 = t, Y_1 < z]$, where $Y_1$ is the highest signal among the other $N-1$ bidders. It can be shown that given that $b^S(\cdot)$ is an increasing function, affiliation implies that $E[b^S(Y_1)|t_1 = t, Y_1 < z]$ is increasing in $t$. Hence, the second-price auction generates higher expected revenue.

An important implication of the Linkage Principle—especially for corporate finance purposes—is that the seller can raise her expected price (revenue) by committing to release to all bidders any information relevant to valuations. More formally, if the

[20] The first-order condition is only a necessary condition. It can be shown that $P_i(z,t)$ is indeed maximized at $z = t$ if the signals are affiliated.

seller releases an informative variable that is affiliated with the other variables, then the expected equilibrium price (for all auction forms) is at least as high as when the information is not released.

### *3.4. Limitations of the common-value and general symmetric auctions*

For corporate finance situations especially, issues of information and efficiency in auctions should be important. Existing models do not allow for full consideration of some of these issues.

In the independent private values auction, efficiency has only one dimension: whether the item is sold to the bidder with the highest valuation. In the pure common value model, there is no real allocation problem so that from an efficiency standpoint, one might as well allocate the item randomly. While a random allocation may not provide optimal revenue for the seller, one should be suspicious of a model focused only on wealth-transfer and not efficiency considerations. One can imagine a variety of economic forces outside of the auction process itself that will tend to cause efficient processes to develop (competition between auctioneers, or even the law). Models that assume away any possibility of inefficiency may cause us to lose sight of the true economic issues in comparing alternative selling mechanisms.

The Milgrom and Weber (1982a, 1982b) model brings an allocation problem back into the picture, in that bidders' valuations differ, so there are efficiency implications of the allocation. On another level, though, this relatively general model still fails to permit a complete role for economic efficiency. As pointed out by Hirshleifer (1971), information can have both private and social value. For information to have social value, it must have the capability to affect the allocation of resources. One would expect that in an auction context, information would not only allow bidders to refine their estimates of value, but since the bidders do have inherently different valuations, one would also expect that information would possibly change relative valuations. That is, with one information set, bidder $i$ might have the highest expected value; but with a different information set, bidder $j$ might have the highest expected value.

The Milgrom and Weber model does not permit this kind of role for information. A simple example suffices to show this as well as to illustrate why it is important to allow information to play an efficiency role. Consider the following two-bidder, two-state model:

| | State | |
|---|---|---|
| | A | B |
| Bidder 1 | 100 | 200 |
| Bidder 2 | 200 | 100 |

In State A, the asset is worth 100 to bidder 1 and 200 to bidder 2, with the valuations reversing for State B. Recall that a major result from the Milgrom and Weber (1982a, 1982b) model is that the expected price increases upon the seller's release of additional information. In the example above this result does not hold. Consider an open auction, and let the information on state initially be diffuse, with each state believed to be equally likely. Then each bidder has an expected value of 150, and an open auction will yield a price of 150. Now let the seller release public information which discloses precisely which state prevails. In either state, an open auction will yield a sale price of only 100, the second-highest valuation. Release of information therefore lowers the expected price, contrary to the Milgrom and Weber findings. Interestingly, there is also now a tension between the seller's objective and economic efficiency: additional information improves efficiency by allocating the asset to its highest-valued use, but it lowers the seller's revenue. Little work has been done on the relative efficiency of auctions under circumstances such as this, but see Krishna (2002) for an excellent summary of efficiency in auctions. In corporate finance, it would seem that the issue of information and efficiency will be closely related: does additional information increase the efficiency of an auction (bearing in mind the cost of producing the information, possibly by multiple bidders) and does this create a conflict between revenue maximization and efficiency?

## 4. Applications of auction theory to corporate finance

### *4.1. Introduction*

We now turn to survey the more important applications of auction theory to corporate finance. We begin with the market for corporate control and auctions in bankruptcy, which are the two largest areas of application. Then we turn to share repurchases, IPOs, and a limited review of corporate finance issues in the Federal Communication Commission's auction of radio spectrum. We do not cover applications of auctions to capital markets finance, for instance to models of the stock trading process or to auctions of bonds by governments and companies. Our intent in this survey is to go beyond a simple review and to point out how well auction theory can actually be used to "inform" corporate finance.

### *4.2. Applications to the market for corporate control*

Auctions of one form or another typically occur in the market for corporate control. The field has proved fruitful for a variety of auction-based models to be constructed that explain many aspects of the market. One aspect is to explain the wealth gains to bidders and targets, as well as the combined wealth gains, on announcements of acquisitions.

#### *4.2.1. Returns to bidders and targets*

Many studies have documented the evidence on stock returns to bidders and targets in corporate acquisitions, and the overall evidence is that returns to targets are large and positive, while returns to acquirers are generally negative but statistically insignificant. Jarrell, Brikley and Netter (1988) provide evidence prior to 1988; Andrade, Mitchell and Stafford (2001) provide a recent update: over the period 1973–1998, with a database of

3688 acquisitions, the average two-day abnormal return around the announcement of an acquisition was 16% (statistically significant at the 5 percent level) for the target; $-0.7\%$ for the acquirer (statistically insignificant); and the combined gain was 1.8% (statistically significant at the 5% level). Boone and Mulherin (2003) further update the recent evidence; they find for a sample of acquisitions between 1989 and 1999 that target returns were on average 21.6%, and that the return to acquirers was an insignificant $-0.7\%$.

Further cuts on the data provide interesting results on the returns to bidders. Returns to bidders are generally more negative the more is the competition from other bidders (although see Boone and Mulherin, 2006b, discussed below). All-stock offers generally yield lower returns to bidders than do all-cash offers (see discussion below). Returns to bidders are generally more positive when the acquisition is large relative to the acquirer's size (Loderer and Martin, 1990; Eckbo and Thorburn, 2000; Moeller, Schlingemann and Stulz, 2004). One strong empirical regularity is that the total profit to bidders and targets (as measured by the event studies) is greater for auctions than for merger negotiations. This is true for both bidders and targets. This may point to a particular measurement problem: merger bids are often a more drawn-out and partially anticipated takeover process than auctions—which means profits in auctions are more easily measured. It is also possible that tender offers are more profitable because they tend to remove old management (to a greater extent than mergers).

The most recent evidence come from the large-sample studies of Betton, Eckbo and Thorburn (2005, 2006). They study more than 12,000 publicly traded targets of merger bids and tender offers over the period 1980–2004. Following the approach of Betton and Eckbo (2000), bids are organized sequentially to form contests for a given target, and they focus in particular on the first and on the winning bidder (which need not be the same). Since the surprise effect of the initial bid is greater than that of subsequent bids, and since the initial bidder starts the contest, studying abnormal returns to the initial bidder yields additional power to test hypotheses concerning the sign and magnitude of bidder gains. Moreover, since bids are studied sequentially in calendar time, they present a natural laboratory for testing auction-theoretic and strategic bidding propositions (toehold bidding, bid preemption, bid jumps, target defenses, etc.).

Initially, Betton, Eckbo and Thorburn follow the tradition and report average abnormal returns for samples of offer outcomes, including "successful" and "unsuccessful" bids. In the traditional analysis, abnormal returns to "success" ($AR_s$) is found by cumulating abnormal returns from the first bid announcement through completion of the takeover process which may take several months. The lengthy cumulation adds noise to this estimate of $AR_s$. Therefore, Betton–Eckbo–Thorburn also report ex-ante estimates of $AR_s$ using the more precisely measured market reaction to the initial bid announcement only. To illustrate, let $x$ denote a set of offer characteristics (e.g., bid premium, the payment method, toehold purchases), and $p(x)$ the probability that the bid will succeed as a function of $x$. The market reaction $\Gamma$ in response to the initial announcement of bid $i$ is

$$\Gamma_i(x_i) = AR_s p(x_i) + AR_u\bigl(1 - p(x_i)\bigr), \tag{27}$$

where $AR_u$ is the average abnormal return conditional on the offer being unsuccessful. Here, $AR_s$ and $AR_u$ are estimated as regression parameters in a cross-sectional regression involving *all* sample bids, whether ultimately successful or not.[21] Using the right-hand side of equation (27), they conclude that the expected value of the initial bid (conditional on $x$) is statistically indistinguishable from zero. As in the earlier literature, targets expected returns are positive and significant, as is the value of the sum of the gains to targets and bidders. Thus, the data do not support theories predicting value-destruction.

Betton, Eckbo and Thorburn also report that the magnitude and distribution of abnormal returns to bidders and targets depends significantly on whether they are private or publicly traded companies. Bidder gains are larger, and target premiums smaller, when the bidder is public but the target is a private firm. Moreover, private bidder firms have a significantly lower probability of succeeding with their bids for public targets. They also report that, in contests where no bids succeed, the target share price reverts back to the level where it was three calendar months prior to the initial bid in the contest. As noted by Bradley, Desai and Kim (1983) as well, this share price reversal is what one would expect if the market conditions the initial target stock price gain on a control change in fact taking pace (where control may be acquired by either the initial or some rival bidder).

Overall, the evidence suggests that auctions tend to yield great results for targets but that competition in the auction (or something else) tends to ensure that gains to bidders are at best minimal. From the standpoint of auction theory, this is surprising: certainly in a private values context, and even in a common value context, the strategic equilibrium of an auction should still yield an expected profit for the winning bidder. The fact that gains to bidders are minimal suggests that the pure auction models do not capture the richness of the process, and that other forces are likely at play. As Boone and Mulherin (2003) suggest, the evidence is in favor of two-stage models such as that of French and McCormick (1984) which analyze costly entry. While pure auction models imply an expected surplus for participating bidders, entry of additional bidders will cause that expected surplus to be dissipated through costly entry.

Roll (1986) first used the idea of the winner's curse to explain the empirical evidence that acquiring firms appear to over-bid for targets in that acquiring firms' stock prices fall (or stay at best constant) upon announcement of acquisitions. If bidders ignore the winner's curse, they may well over-pay (in a common value setting, which is not unreasonable in the corporate acquisition market). The problem, of course, is that equilibrium theory does not permit expected over-bidding, so Roll is relying upon acquirers making mistakes. Proponents of behavioral finance will find it quite convincing to think that bidders may not properly adjust their strategy for the pitfalls inherent in common value

[21] The estimation is in three steps: (1) estimate $AR_i$ using time series of returns to the bidder up to the first bid announcement, (2) estimate $p(x)$ using the cross-section of bids, and (3) run regression (27) to produce $AR_s$ and $AR_u$.

auctions, for avoidance of the winner's curse requires some careful analysis. Those inclined towards rational, equilibrium based models of behavior will be wary of models that assume incomplete strategic adjustment. Boone and Mulherin (2006b) use unique data that allows them to characterize sales of companies as either auctions or negotiations, and for the auctions, to say how many potential bidders were contacted in the sales process and how many actually submitted bids. Finding no relationship between bidder returns and these measures of competition, Boone and Mulherin conclude that their findings do not support the existence of a winner's curse.

A large literature attributes the acquirer wealth losses to managerial agency problems or "empire building" tendencies. For example, in a sample of 326 U.S. acquisitions between 1975 and 1987, Morck, Shleifer and Vishny (1990) find that three types of acquisitions have systematically lower and predominantly negative announcement period returns to bidding firms: diversifying acquisitions, acquisitions of rapidly growing targets, and acquisitions by firms whose managers performed poorly before the acquisition. The authors argue that these results are consistent with the view that managerial objectives may drive acquisitions that reduce bidding firms' values. Lang, Stulz and Walkling (1991) present related results. Jensen (2004) provides a new angle to this argument by hypothesizing that high market valuations increase managerial discretion, making it possible for managers to make poor acquisitions when they have run out of good ones.

Another recent approach to overbidding is based on the idea that when bidders own initial stakes or "toeholds" in the target firm, they are essentially wearing two hats—that of a buyer for the target's remaining shares, and that of a seller of their initial stakes to the rival bidder. We review the theory-based work in this area more fully below. For now, we note that in an independent private values model, Burkart (1995) and Singh (1998) show that a bidder with toehold will bid above her private value in a second-price auction. Similar results are also obtained in alternative value environments and under alternative auction procedures (Bulow, Huang and Klemperer, 1999; Dasgupta and Tsui, 2003). Evidence on the empirical relevance of toeholds, however, is mixed. In Jennings and Mazzeo's (1993) sample of 647 tender offers and mergers, the mean toehold is 3%, but only about 15% of the bidders own an initial stake. Betton and Eckbo (2000) study toeholds for initial and rival bidders in a sample of 1,250 tender offer contests over the period 1972–1991. They find that toeholds increase the probability of single-bid success and lower the price paid by the winning bidders.

Betton, Eckbo and Thorburn (2005) delve more deeply into the subtleties of various facts about toeholds. In their sample of 12,723 bids for control (3,156 tender offers and 9,034 mergers), 11% of the bids involved toeholds. The percentage was significantly higher for tender offers than for mergers, both for non-hostile targets (21% and 6%, respectively, for tender offers and mergers) and for hostile targets (62% and 31%, respectively). The mean and median toehold sizes conditional on being positive were 21% and 17%, respectively, for the overall sample. However, a majority of these toeholds were "long-term toeholds", i.e., acquired before 6 months prior to the bid. The percentage of bids involving short-term toeholds for the entire sample was only about 2%. Betton, Eckbo and Thorburn (2005) argue that since toeholds are likely to deter competition, the target might turn hostile if the bidder acquires toeholds when private negotiations might be going on. Thus, it is unclear to what extent toeholds are used strategically in bidding contexts. It is worth recalling in this context, however, Shleifer and Vishny's (1986) analysis of the role of large shareholders in the target firm: even when they are not bidders, the presence of large shareholders in the target firm who are willing to split the gains on their shares with a bidder has the same effect as the bidder having an initial stake in the target.

Another approach to reconciling the existing findings on loss of value to acquirers, the gains to targets, and joint value losses is presented by Jovanovic and Braguinsky (2004), even though their model is not explicitly auction-based. The model incorporates uncertainty over the skill of corporate managers, the value of projects that companies have, and the takeover market. In equilibrium, the takeover market facilitates the exchange of "good" projects from firms with "bad" managers to firms with "good" managers but "bad" projects. Ex-ante values of firms represent investors' knowledge of management type but uncertainty over project type. If a firm puts itself up for sale, which it does only if its project is good and its management is bad, then investors learn that the firm does have the property right to a good project and its value increases—hence the positive return to targets. A firm becomes an acquirer only if its own project is bad. Upon learning that a firm will be an acquirer, investors learn that the firm's own project is bad—hence the negative return to acquirers. For reasonable parameter values, including a cost incurred in the takeover process, joint values of the target and acquirer fall. Even so, the mergers in the model are welfare-enhancing.

### *4.2.2. The auction process in the market for corporate control*

As our previous discussion shows, takeover models help understand some of the observed empirical evidence on bidder and target returns. Another major role of auction theory, in so far as it facilitates our understanding of the takeover bidding process, has been to "inform" a company's board or regulators about the impact of selling processes or rules on shareholder wealth, efficiency and welfare. However, here, for the prescriptions to be useful, the auction models must at least reasonably mimic the takeover bidding environment. The question we address now is the extent to which this is the case.

First, it is important to note that auction theory has developed in the spirit of mechanism design, or the design of optimal selling schemes. Any auction model assumes a degree of commitment power on the part of the seller. There are clear "rules of the game" that the seller and the bidders are required to abide by. For example, in a first-price auction, in which bidders shade their bids, the losing bidders might want to submit a bid higher than the winning bid after the latter is disclosed. The seller must be able to commit not to entertain such bids. A similar argument applies to the reserve price. Casual observation, however, suggests that many bids (even when they are friendly) are not seller initiated. It might appear that many control contests are not really formal auctions, in which the seller is trying to secure the best price for the firm's shareholders by committing to a selling mechanism.

This perspective is misleading, for several reasons. First, the board has a formal responsibility to be an "auctioneer". Under Delaware law,[22] a company's board must act as "auctioneers charged with getting the best price for the stock-holders at a sale of the company". In several well-publicized cases, after potential bidders had indicated their interest in acquiring the company, the board of directors of the target company have conducted an auction.[23] Although procedures similar to the ascending auction are most commonly used, boards have also held single, and sometimes even multiple, rounds of sealed-bid auctions (e.g., in the well-documented case of RJR Nabisco).

The commitment issue discussed above may influence the board's choice of auction mechanism. For example, the board might have a preference for ascending auctions because, under alternative auction rules such as the sealed-bid auction, should a losing bidder offer a higher subsequent bid, it may be difficult to reject that bid if the board is required to obtain the "best price for the shareholders". In other words, it may be difficult to commit to a single round of bidding.

Legal scholars, however, have taken the view that whether or not it is feasible for the board to pursue a particular auction mechanism depends, ultimately, on how the courts view it. If, in a given context, the courts consider that a particular auction mechanism can generate higher revenue for the shareholders ex-ante than the more commonly used ones, there is no reason why a board cannot adopt it as a selling scheme. Further, if the shareholders do not perceive a particular selling scheme to be against their interests *ex-ante*, there is no reason why a board cannot secure shareholder approval prior to conducting a sale. It is exactly in this spirit that legal scholars have looked at alternative selling procedures (see, for example, Cramton and Schwartz, 1991). The focus of this literature has very much been on what one can learn from economic theory (in particular, auction theory) to "inform" takeover regulation or selling practices.

Second, the board's commitment power is sometimes underestimated. Boards can commit to awarding an object to a "winner" from a given round of bidding even when better bids might subsequently emerge—thereby undermining the auction—in a variety of ways. The most common practice is to enter into a lock-up arrangement[24] with the declared winner, together with an agreement to pay a break-up fee should the sale be

[22] The Delaware law is significant because many U.S. public companies are incorporated in Delaware.

[23] For example, in the takeover battle for Paramount between Viacom and QVC, the Paramount board eventually conducted an auction in an effort to "select the bidder providing the greatest value to shareholders".

[24] Lockups are "*agreements that give the acquirer the right to buy a significant division, subsidiary or other asset of the target at an agreed (and generally favorable) price when a competing bidder acquires a stated percentage of the target's shares*" (see Herzel and Shepro, 1990). They may also involve options to buy a block of target shares from the target that may make acquisition by a competing bidder more difficult. Lockup agreements are quite common in takeover contests. The legal status of lockups is unclear, as some courts have upheld them, while others have not. For an account of the legal literature on lockups, see Kahan and Klausner (1996a, 1996b).

terminated.[25] Another possibility is for the target board to refuse to rescind poison pills for any but the declared winning bidder. While it is unclear whether the courts will allow such poison pills to stand, the legal costs of challenging the poison pills and the possibility that the board might switch to an ascending auction (so that the challenger is by no means assured of winning the contest) may deter further challenge from a losing bidder.

Third, formal or informal auctions are much more common than is usually assumed. Boone and Mulherin (2006a) analyze a sample of 400 takeovers of U.S. corporations in the 1989–1999 period and find evidence consistent with the idea that boards act as auctioneers to get the best price for the shareholders in the sale of a company. Based on information from the SEC merger documents, the authors provide new information on the sale process. The most important evidence is that there is a significant *private* takeover market prior to the public announcement of a bid. The authors document that almost half of firms in their sample were auctioned among multiple bidding firms, and the rest conducted negotiations with a single bidder. A third of the firms in the former category went through a formal auction, in which the rules were clearly laid out. In all cases, the process usually began with the selling firm hiring an investment bank and preparing a list of potential bidders to contact. After the bidders agreed to sign a confidentiality/standstill agreement, they received non-public information. Subsequently, a subset of the bidders indicating preliminary interest was asked to submit sealed bids.[26]

Another issue relevant for the applicability of auction models to control contests concerns the complexity of the environments in which takeovers are conducted, compared to the standard auction environments. Auction models are nicely classified as belonging to different value environments, and results differ depending on which value environment is under consideration. The takeover environment is considerably more complicated. The motives for takeover bids could be varied. The early takeover models (e.g., Grossman and Hart, 1980) assumed that the benefit from a takeover comes from an improvement in the operational efficiency of the target company. As the authors showed, this could lead to a "free-rider" problem and the market for corporate control could fail. However, later models have focused on "merger synergies" as the source of gain from takeovers. If the synergies accrue to the bidding firm, then the standard auction environment is more applicable. Here, however, there are issues about whether "private values" or "common values" assumptions are more relevant. Since bidders are different and the synergies are likely to have idiosyncratic components, a private values model does not appear unreasonable. However, common value elements will also undoubtedly exist. Synergies can have common value components if their magnitudes

[25] For example, Viacom's initial offer for Paramount in 1993 was associated with (a) an option to buy 20% of Paramount's outstanding shares and (b) a termination fee of $150 million plus expenses, should the transaction not be concluded.

[26] Betton, Eckbo and Thorburn (2005) show that of the 12,000 contests, about 3,000 (25%) start out as tender offers (which subsequently turn into auctions). Some initial merger bids also end up in auctions, so the overall percentage of auctions maybe closer to 30%.

depend on the quality of the target's assets, or if the bidders plan to bundle these assets with other assets that they own and eventually sell these assets.[27,28]

Other complexities also arise when applying the auction framework to the analysis of takeover bidding. Bidders could bid for the company, or they could bid for a fraction of the company's shares. Different regulatory regimes permit different types of bids. Bids could be exclusionary, discriminatory, conditional, and so on. Bids can be in cash, or in shares of the target company. Bidders may have different toeholds, and they might have different degrees of expertise in the target industry (a factor that could affect the degree of information conveyed by their bids in common value environments). Finally, if the bidders are competing with each other in the same industries, then the outcome of the auction may impose externalities on the bidders. As we will argue below, while existing takeover models, drawing on auction theory, have evolved to deal with these many of these complexities, significant gaps still exist in the literature.

### *4.2.3. Auctions versus negotiations*

Several papers use auction theory to further refine our theoretical and empirical understanding of the auction process in corporate takeovers. Starting at the most basic level, Bulow and Klemperer (1996) show that in an English auction, it is always better to have $N+1$ bidders in a formal auction than to have $N$ bidders but with a follow-on (optimal) negotiation between the winning bidder and the seller. If $N = 1$, this shows that it is better to have an auction with two bidders than to sell by posting a reserve price.[29] Very simply, the auction process is extremely efficient at extracting value from the high bidder, more so than even an optimally conducted negotiation. This theoretical result does conflict with a stylized fact that companies do frequently avoid auctions and instead negotiate with just one buyer (Boone and Mulherin, 2006a).

### *4.2.4. Pre-emptive bidding*

Fishman (1988, 1989) considers models where one bidder has incentive to make a "pre-emptive" bid. In the main model of Fishman, a first bidder has incentive to put in a high bid that discourages the second bidder from bidding. The reason for this is that a high bid can signal a high valuation on the part of the first bidder, and a second bidder will

[27] Models of takeover bidding, when making common values assumption about the target's "true worth", have often tended to assume away the free-rider problem. If a bidder obtains a large majority of the shares, she may be able to "freeze out" the remaining minority shareholders. Also, the loss of liquidity on any remaining shares can have the same effect as "dilution" (see Grossman and Hart, 1980) that reduce the post-takeover value of the minority shares.

[28] Betton and Eckbo (2000) show that the average number of days from the initial tender offer bid to the second bid is 15 days (counting only auctions with two or more bids). They suggest that this very short period is evidence of correlated values. Of course, the vast majority of all cases develop a single bid only, which may be taken as evidence of private (uncorrelated) values, or preemptive bidding (see below).

[29] See Krishna (2002) for an analysis of this case.

then infer that the gain from participating in the auction is low (they are not likely to win in the final English auction). Since participation in the auction requires a bidder to spend resources to determine her own value, the second bidder can be discouraged from even entering the auction. Fishman (1989) extends this initial work by including the possibility of non-cash offers.

Fishman's (1988) model works as follows. The value of the target assets to the bidders depends in part on the realization of a state of nature which is observed only by the target. Conditional on the target's information, the value of the assets to the bidders is increasing in the bidders' independent private signals. The means-of-payment can be either cash or debt that is backed by the target's assets. Each bidder has to incur some cost to learn the private signal. Bidder 1 identifies a target by accident, and then incurs some cost to learn his signal (bidding is assumed to be not profitable if the true signal is unknown). If bidder 1 submits a bid, the target is "put in play", and a second bidder is aware of the target. This bidder then decides whether or not to compete for the target and incur the cost of learning her signal.

There is a stand-alone value of the target that is public information, and the target rejects all bids below this value. Since the bidders do not know completely how much the target assets are worth to them (recall that the target privately observes part of this information), bidders could end up overpaying for the target. Paying with debt mitigates the overpayment because the value of the debt is contingent on the value of the target assets (since the debt is backed by these assets). However, if bidder 1 draws a high private signal, a cash offer—though costly—will separate it from a bidder with low signal: the latter will prefer to pay only with debt since his own private signal is not sufficiently high. Thus, by bidding with cash, the first bidder can signal to the second bidder that the latter's likelihood of winning the ensuing auction is low: hence, the second bidder may decide not to incur the cost of learning her signal. This, then, is a "pre-emptive" bid. On the other hand, if the first bidder's signal is low, bidding high with a cash offer is too costly. Thus, such a bidder would decide not to preempt and instead bid with debt to mitigate the potential loss from buying a target with low synergy. Notice that one prediction of this model is that more competing offers should be forthcoming with non-cash offers than with cash offers.[30]

### *4.2.5. Modelling auctions of companies*

Hansen (2001) reviews the formal auction process used for selling private companies and divisions of public companies. The model explains the common practices of limiting the number of bidders and limiting the disclosure of information to bidders (Boone and Mulherin, 2006a), even though theory suggests that both practices would reduce

[30] Betton and Eckbo (2000) find that the average offer premium in successful single-bid contests is greater than the average offer premium in the *first* bid in multiple-bid contests. This what consistent with preemptive bidding.

prices. Hansen argues that some information in a corporate sale is competitive in nature, and that its broad release can destroy value in the selling company. The seller therefore faces a tradeoff between having many bidders and full disclosure versus protecting value by limiting disclosure, as well as the number of bidders. While not modeling negotiations formally, the analysis implies that negotiation with a single bidder may be optimal if the "competitive information cost" is high enough. The model also explains the practice of a two-stage auction, with a first stage calling for non-binding "indications of interest" (value estimates for the target) which are used to select bidders for the second round and giving them access to more information on the selling company. If the selling company uses the initial value estimates for the target to set a reserve price that is an increasing function of the estimates, bidders in the initial round will reveal their private valuations honestly and the selling company can select the most highly-valued bidders for the final, binding, round (see the discussion below on the process for pricing IPOs for an earlier similar finding).

### *4.3. Means-of-payment*

Hansen (1985a, 1985b, 1986) has considered the role of non-cash means of payment in the market for corporate control; this work has now been extended by DeMarzo, Kremer and Skrzypacz (2005). In one model, Hansen shows that ex-post means of payment can increase the seller's revenue beyond what cash payments can do. Take an independent private values context, where $v_i$ represents bidder $i$'s valuation of the target company. An ascending auction with cash as the means of payment will yield $v_2$—the second highest value—as the price. Consider, however, bidding using bidders' stock as the means-of-payment. Let each bidder have a common value, $v$, of her stand-alone equity. Then each bidder will be willing to bid up to $s_i$, where $s_i$ is the share of firm $i$ offered (implicitly through an offer of equity) and is defined to make the post-acquisition value of the bidder's remaining equity equal to its pre-acquisition value:

$$v = (v + v_i)(1 - s_i),$$

which implies

$$s_i = \frac{v_i}{v + v_i}. \tag{28}$$

The bidder with the highest valuation of the target will win this auction ($s_i$ is increasing in $v_i$) and she will have to offer a share defined by $v_2$, the valuation of the second-highest bidder. However, the value of this bid to the target will be

$$s_2 * (v + v_1) = \frac{v + v_1}{v + v_2} v_2 > v_2$$

since $v_1 > v_2$. The stock-based bidding therefore extracts more revenue from the high-bidder than does cash bidding. DeMarzo, Kremer and Skrzypacz (2005) generalize this result, showing that expected revenues are increasing in the "steepness" of the security design, where steepness refers, roughly, to the rate of change of a security's value in relation to the underlying true state. This paper also compares auction formats in a world where bids can be non-cash; it turns out that revenue equivalence does not always hold. Overall the paper concludes that the optimal auction is a first-price auction with call options as the means-of-payment.

Ex-post pricing mechanisms also yield benefits in common-value contexts. The reason for this follows from the return of the adverse selection problem inherent in the winner's curse: the problem arises because the price for the asset is being determined before the value of the asset is known. Any kind of pricing mechanism that determines all or part of the price ex-post can alleviate the problem. Using the acquiring firm's stock is an ex-post pricing mechanism, for that stock's value will depend upon the actual value of the target firm. Hansen (1986) builds on this insight and shows that stock and cash/stock offers can be used efficiently in mergers and acquisitions. However, in offering stock as the means-of-payment, acquiring firms bring in their own adverse selection problem—acquiring firms may offer stock when they have information that their own value is low. Taking into account both the ex-post pricing advantage of stock and the "reverse" adverse selection problem, it turns out that higher-valued acquirers will offer cash while low-valued acquirers offer stock. Fishman (1989) reaches a similar conclusion, in that non-cash offers induce the target firm to make more efficient sell/don't sell decisions, but that cash offers have an advantage in pre-empting other bids.[31]

Several studies on U.S. data show results consistent with Hansen and Fishman's work, that acquirers' returns are higher for cash offers than for stock offers (see Eckbo, Giammarino and Heinkel, 1990, for a brief summary). The first paper to explicitly model the choice of mixed offers is Eckbo, Giammarino and Heinkel (1990). These authors prove the existence of a fully separating equilibrium in which the market's revaluation of the bidder firm is increasing and convex in the proportion of the offer that is paid in cash. Since one can estimate the revaluation, and since the proportion paid in cash is observable, this theory is testable. Using over 250 Canadian takeovers (where tax issues do not confound the choice of payment method), the authors find empirical support for the "increasing" part but not for convexity.

### *4.4. Toeholds*

Recently, a number of theoretical papers have examined how toeholds affect takeover bidding. The main result that emerges from this literature is that the presence of makes bidders more aggressive, with the result that bidders can bid above the value of the object. The result holds for the second-price auction in both the independent private values as well as a common value environment.

[31] Rhodes-Kropf and Viswanathan (2000) consider a general model of non-cash auctions for a bankrupt firm. We discuss this model later.

Burkart (1995)[32] considers a two-bidder and independent private values model. The private values are best interpreted as synergies. The auction form is a second-price auction, which in this context is strategically equivalent to an ascending auction (Lemma 1 in the paper). From standard arguments, it follows that (i) it is a dominant strategy for the bidder with no toeholds to bid exactly her valuation, and (ii) it is a *dominated* strategy for the bidder with positive toehold to bid below her valuation. A general result is that any bidder with positive initial stake will bid strictly above her valuation. The model is then specialized to the case in which one bidder—call her bidder 1—has an initial stake of $\theta$ while the other bidder—bidder 2—has no initial stake.

Since bidder 2 will bid her value, we have $b_2(v_2) = v_2$. Thus, bidder 1's problem is to choose $b_1$ to maximize

$$\mathrm{Max}_{b_1} \Pi_1(v_1, b_1, \theta) = \int_0^{b_1} \big[v_1 - (1-\theta)v_2\big] f_2(v_2)\, dv_2 + \theta b_1\big(1 - F_2(b_1)\big). \quad (29)$$

The first-order condition is

$$\big(v_1 - (1-\theta)b_1\big) f_2(b_1) + \theta\big(1 - F_2(b_1)\big) - \theta b_1 f_2(b_1) = 0.$$

Re-arranging, we get

$$b_1 = v_1 + \theta \frac{1 - F_2(b_1)}{f_2(b_1)} > v_1. \quad (30)$$

If one assumed that the hazard function $\frac{f_2(\cdot)}{1-F_2(\cdot)}$ is increasing, then a number of results follows immediately. First, bidder 1's equilibrium bid is increasing in her valuation and the size of her toehold. Therefore, the probability that bidder 1 wins the auction is also increasing in her toehold. It is also clear that the auction outcome can be inefficient: since bidder one bids more aggressively than bidder 2, it is clearly possible that $v_1 < v_2 < b_1(v_1)$, i.e., bidder 1 has the lower valuation but wins the auction. This result is similar to the inefficiency in the standard auctions where the seller sets a reserve price. In fact, the intuition for the overbidding result is exactly that of an optimal reserve price from the point of view of a seller. Indeed, with a toehold, a bidder is a part-owner and we should not be surprised to find that she wants to "set a reserve price" in excess of her own value.

It is interesting to note that winning can be "bad news" for bidder 1. Suppose $v_1 = 0$ with probability 1. Then bidder 1 still bids a positive amount (equal to bidder 2's value) but since her bid exceeds the value of the synergy, she always overpays when she wins the auction. By continuity, the same conclusion holds for $\bar{v_1}$ (the upper bound of the support of the distribution of bidder 1's synergy) sufficiently small, and for bidder 2's valuation in some interval $[v_2', b_1(\bar{v_1})]$.[33]

[32] Singh (1998) has essentially similar results.

[33] Using Burkart's private value setting with two bidders, Betton, Eckbo and Thorburn (2005) also show optimal overbidding when the bidder has a lock-up agreement with the target. Moreover, they show optimal underbidding when the bidder has a breakup fee agreement with the target.

Bulow, Huang and Klemperer (1999) examine the effect of toeholds in a pure common value environment. They make a significant contribution to the literature on toeholds by deriving bid functions for both the second and first-price auctions when both bidders have positive toeholds. They examine how (for small positive toeholds) bidder asymmetry affects the takeover outcome in each auction, and compare expected revenues in the two auctions when the toeholds are symmetric as well as asymmetric. We first discuss their setup in some detail, before discussing the intuition for the main results.

Bulow, Huang and Klemperer (1999) consider a "pure common value" model with two bidders where each bidder draws an independent signal $t_i$ from a uniform $[0, 1]$ distribution. The value of the target to each bidder is $v(t_1, t_2)$. Bidder $i$ owns initial stake $\theta_i$ in the target, where $1/2 > \theta_i > 0$, for $i = 1, 2$. Each bidder bids for the remaining $1 - \theta_i$ fraction of the shares of the target.

In the second-price auction, bidder $i$'s problem is to choose $b_i$ to maximize

$$\mathrm{Max}_{b_i} \Pi_i(t_i, b_i) = \int_0^{b_j^{-1}(b_i)} \big[v(t_i, \alpha) - (1-\theta_i) b_j(\alpha)\big]\, d\alpha + \int_{b_j^{-1}(b_i)}^{1} \theta_i b_i \, d\alpha. \tag{31}$$

The first-order condition is

$$\frac{1}{b_j'}\big[v\big(t_i, b_j^{-1}(b_i)\big) - (1-\theta_i) b_j\big(b_j^{-1}(b_i)\big)\big] + \big[1 - b_j^{-1}(b_i)\big]\theta_i - \theta_i b_i \frac{1}{b_j'} = 0.$$

Let us now define $\phi_j(t_i) = b_j^{-1}(b_i(t_i))$, i.e., this defines the pair of signals for bidders $i$ and $j$ for which they have the same bid, since $b_j(\phi_j(t_i)) = b_i(t_i)$. Similarly, we can define $\phi_i(t_j) = b_i^{-1}(b_j(t_j))$. Using these definitions, we can rewrite the first-order condition as

$$b_j'\big(\phi_j(t_i)\big) = \frac{1}{\theta_i} \frac{1}{(1 - \phi_j(t_i))} \big[b_i(t_i) - v\big(t_i, \phi_j(t_i)\big)\big], \tag{32}$$

where we have replaced $t_j$ by $\phi_j(t_i)$.

The corresponding first-order condition for bidder $j$ is

$$b_i'\big(\phi_i(t_j)\big) = \frac{1}{\theta_j} \frac{1}{(1 - \phi_i(t_j))} \big[b_j(t_j) - v\big(\phi_i(t_j), t_j\big)\big], \tag{33}$$

where we have used the fact that $v(\phi_i(t_j), t_j) = v(t_j, \phi_i(t_j))$. Consider a pair of $t_i$ and $t_j$ that in equilibrium bid the same, then we must have $t_i = \phi_i(t_j)$ and $t_j = \phi_j(t_i)$. Using this, the last equation can be rewritten as

$$b_i'(t_i) = \frac{1}{\theta_j} \frac{1}{(1 - t_i)} \big[b_j\big(\phi_j(t_i)\big) - v\big(t_i, \phi_j(t_i)\big)\big]. \tag{34}$$

Since $b_j(\phi_j(t_i)) = b_i(t_i)$ and $b_i'(t_i) = b_j'(\phi_j(t_i))\phi_j'(t_i)$, dividing (34) by (32), we get

$$\phi_j'(t_i) = \frac{\theta_i}{\theta_j} \frac{1 - \phi_j(t_i)}{1 - t_i}. \tag{35}$$

Integrating, and using the boundary condition $b_i(0) = b_j(0)$ (see Bulow, Huang and Klemperer, 1999, for a proof), we get

$$\phi_j(t_i) = 1 - (1 - t_i)^{\theta_i/\theta_j}. \tag{36}$$

Since the probability that bidder $i$ wins the object is $\int_0^1 \int_0^{\phi_j(t_i)} dt\, dt_i = \frac{\theta_i}{\theta_i+\theta_j}$, it is clear that bidder $i$ is more likely to win the auction as her stake increases and that of bidder $j$ decreases. Remarkably, a bidder's probability of winning goes to 0 as her stake becomes arbitrarily small, given that the other bidder has a positive stake. The intuition for this result is that while bidder $i$ with zero stake has no incentive to bid above $v(t_i, \phi_j(t_i))$ given the equilibrium bidding strategy of $j$, as we shall see below, bidder $j$ with $t_j = \phi_j(t_i)$ and a positive stake will strictly bid above this value.[34]

Now, equation (34) can be integrated to give

$$b_i(t_i) = \frac{\int_{t_i}^1 v(t, \phi_j(t))(1 - t)^{\frac{1}{\theta_j} - 1}\, dt}{\int_{t_i}^1 (1 - t)^{\frac{1}{\theta_j} - 1}\, dt}, \tag{37}$$

where the boundary condition $b_i(1) = b_j(1) = v(1, 1)$ is used (see Bulow, Huang and Klemperer, 1999).

From (36), we then get

$$b_i(t_i) = \frac{\int_{t_i}^1 v(t, 1 - (1 - t)^{\theta_i/\theta_j})(1 - t)^{\frac{1}{\theta_j} - 1}\, dt}{\int_{t_i}^1 (1 - t)^{\frac{1}{\theta_j} - 1}\, dt}. \tag{38}$$

Bidder $j$'s bid function is derived similarly. From (37), it is clear that for $t_i < 1$, $b_i(t_i) > v(t_i, \phi_j(t_i))$. Thus, when bidder $i$ wins the auction, she is paying more than the target is worth to her. Moreover, bidder $i$'s bid is increasing in her stake $\theta_i$, i.e., a higher stake makes the bidder act more like a seller and causes her to bid higher.

Bulow, Huang and Klemperer (1999) extend the analysis in two main directions. First, they consider the effect of a more asymmetric distribution of the toeholds and find that subject to an overall constraint on the toeholds of the two bidders that is sufficiently small, a more uneven distribution of toeholds leads to lower expected sale price for

[34] Klemperer (1998) demonstrates in the context of the "Wallet Game" how a very small asymmetry in a common value model can give rise to very asymmetric equilibria. This is a consequence of the fact that in the standard Wallet Game, there are in fact a continuum of asymmetric equilibria. A small toehold—like a small bonus to one of the players in the Wallet Game—introduces a slight asymmetry that can have a major impact on the equilibrium, i.e., one of the bidders essentially having a zero probability of winning. With a slight advantage, the stronger player bids slightly more aggressively, but that increases the winner's curse on the weaker player. The latter then bids less aggressively, which reduces the winner's curse on the stronger player, who then bids still more aggressively, and so on. With slight entry or bidding costs, this prevents the weaker player/players from entering the auction, so that very low prices result. Klemperer (1998) provides several illustrative examples from Airwaves Auctions.

the target. This result is a consequence of the fact that as the toeholds become more asymmetric, the bidder with the higher toehold bids more aggressively, i.e., further away from the value. For the bidder with a smaller toehold, this implies that the target is worth less conditional on winning. Exposed to this "winner's curse", the bidder with the smaller toehold therefore bids lower. Since in the second-price auction the winner pays the lower of the two bids, the expected sale price is adversely affected when the toeholds become asymmetric.

Bulow, Huang and Klemperer (1999) next consider first-price auction and derive the equilibrium bid functions using methods similar to those described above for the second-price auction. In this case, we have $\tilde{\phi}_j(t_i) = t_i^{(1-\theta_i)/(1-\theta_j)}$. The probability that bidder $i$ with signal $t_i$ and toehold $\theta_i$ wins the auction in this case is given by $\frac{1-\theta_j}{(1-\theta_i)+(1-\theta_j)}$, which is increasing in $\theta_i$. It is easily checked that for $\theta_i < \theta_j$, the probability of bidder $i$ winning the auction is lower in the second-price auction than in the first-price auction. Since in both auctions the probability is exactly $1/2$ when $\theta_i = \theta_j$, this implies that the winning probability falls more steeply with a decrease in a bidder's toehold the second-price auction than in the first-price auction.

The incentive for bidders with toeholds to bid high in the first-price auction are not as strong as in the second-price auction. This is because in the in the first-price auction (unlike the second-price auction), bidding high does affect the bidder's cost, although a higher toehold does lower that cost since fewer shares need to be purchased.

Unlike the second-price auction, the expected sale price can increase in the first-price auction as the toeholds become more asymmetric. Revenue comparisons indicate that with *symmetric* toeholds, the expected sale price is higher in the second-price auction. This is because as the winner's curse problem is mitigated with symmetric toeholds, both bidders can bid more aggressively and essentially set a higher reserve price for their stakes in the second-price auction. With asymmetric toeholds, as we saw above, the second-price auction generates low expected sale prices due to the winner's curse.[35]

### *4.5. Bidder heterogeneity and discrimination in takeover auctions*

Bidder asymmetry is common in the context of corporate control contests and can take several forms. Asymmetry in initial stakes or toeholds, discussed in the previous section, is one form of bidder asymmetry. Bidder asymmetry can also arise when bidders draw their signals from different distributions, or when (in a common value environment) the bidder signals have asymmetric impact on the value function.

In Section 2.6, we saw that when bidders are asymmetric, the optimal mechanism may not allocate the object to the bidder with the highest valuation. For example, in the independent private value context, an allocation rule that discriminates against a

[35] The analysis of toeholds can be extended to models that include the private value model and the common value model of Bulow, Huang and Klemperer (1999) as special cases. Dasgupta and Tsui (2004) analyze auctions where bidding firms hold toeholds in each other in the context of such a model.

stronger bidder may provide a higher expected profit to the seller. Thus, standard auctions are no longer optimal in the presence of various forms of bidder heterogeneity.

To increase the expected sale price when bidders are asymmetric, the seller has essentially two alternative responses. Both involve "levelling the playing field". When the asymmetry is due to differences in toeholds or access to information, the target's board may decide to restore symmetry by allowing the disadvantaged bidder increase his toehold cheaply or provide access to additional information.[36] Alternatively, the board may decide to design the auction rules in a way that discriminates against the strong bidder.

An especially simple way to discriminate is to impose an order of moves on the bidders. Since bidding games are price-setting games, there is usually a "second-mover advantage" associated with bidding games (see Gal-Or, 1985, 1987). Thus, to discriminate against the strong bidder, the seller could ask this bidder to bid first. This bid could then be revealed to a second bidder, who wins the auction if she agrees to match the first bid. Otherwise, the first bidder wins. In the context of takeover bidding, this "matching auction" has been studied by Dasgupta and Tsui (2003), who note that since courts are more concerned about shareholder value than whether the playing field is level or not, it is unlikely that the matching auction will run into trouble because it does not treat the bidders symmetrically.[37]

To see that the matching auction can generate a higher expected sale price than the second-price auction in the independent private value setting, let us return to the private values model introduced in Section 4.4. Assume that the private values of both bidders are drawn from the uniform [0, 1] distribution. From equation (30), we get the bid of bidder 1 who has a toehold of $\theta$ to be

$$b_1(t_1) = \frac{t_1 + \theta}{1 + \theta}.$$

Thus, the expected bid from bidder 1 is $P_1 = \int_0^1 \int_0^{(t_1+\theta)/(1+\theta)} t_2\, dt_2\, dt_1 = \frac{1}{6}\frac{3\theta^2+3\theta+1}{(1+\theta)^2}$ and that from bidder 2 is $P_2 = \int_0^1 \int_0^{(1+\theta)t_2-\theta} \frac{t_1+\theta}{1+\theta}\, dt_1\, dt_2 = \frac{1}{6}\frac{1-2\theta^2+2\theta}{1+\theta}$. Thus, the expected sale price in the second-price auction is

$$P^S = P_1 + P_2 = \frac{(2\theta + 1)(2 + 2\theta - \theta^2)}{6(1 + \theta)^2}. \tag{39}$$

Now consider the matching auction. Given a bid $b_1$ from bidder 1, bidder 2 will match if and only if $t_2 > b_1$. Thus, bidder 1 chooses $b_1$ to maximize

$$\int_0^{b_1} \bigl(t_1 - (1 - \theta)b_1\bigr)\, dt_2 + \theta(1 - b_1)b_1.$$

[36] Betton and Eckbo (2000) note that when a rival (second) bidder enters the auction with a toehold, the toehold is of roughly the same magnitude as the initial bidder's toehold (about 5%). This is consistent with the "leveling the playing field" argument of Bulow, Huang and Klemperer (1999).

[37] Herzel and Shepro (1990) note: "Opinion in several cases in the Delaware Chancery court has noted that the duty and loyalty [of managers] runs to shareholders, not bidders. As a result, 'the board may tilt the playing field if it is in the shareholder interest to do so' ".

From the first-order condition, one readily gets $b_1(t_1) = (1/2)t_1 + (1/2)\theta$. Thus, the expected sale price in the matching auction is

$$P^M = \frac{1}{4} + \frac{1}{2}\theta. \tag{40}$$

Comparing (39) and (40), it can be verified that $P^M > P^S$ if and only if $\theta > 0.2899$. Thus, if the toeholds are sufficiently asymmetric, asking the strong bidder to move first increases the expected sale price.

The matching auction's properties in the context of a common value model with independent signals similar to Bulow, Huang and Klemperer (1999) have been explored by Dasgupta and Tsui (2003). The authors show that there exists a perfect Bayesian Nash equilibrium in which bidder 1 with stake $\theta_1$ bids

$$b_1(t_1) = v\big(t_1, F_2^{-1}(\theta_1)\big) \tag{41}$$

and bidder 2 matches if and only if $t_2 \geqslant F_2^{-1}(\theta_1)$.[38] Here, bidder $i$'s signal is drawn from the distribution $F_i(t_i)$. Notice that the expected sale price is then

$$P^M = E_{t_1}\big(v\big(t_1, F_2^{-1}(\theta_1)\big)\big). \tag{42}$$

Notice that (i) conditional on her bid, losing is better than winning for bidder 1, since her payoff in the former event is $\theta_1 v(t_1, F^{-1}(\theta_1))$, and her payoff in the latter event is at most $v(t_1, F^{-1}(\theta_1)) - (1 - \theta_1)v(t_1, F^{-1}(\theta_1))$, and (ii) as a consequence, winning is "bad news" for bidder 1, i.e., if she wins, there would be a negative effect on the stock price. In contrast, winning is always "good news" for the second bidder.

It is also immediate that the expected sale price increases in the first bidder's toehold. In contrast, bidder 2's stake has no effect on the expected sale price. The probability of bidder 1 winning the auction is $F_2^{-1}(\theta_1)$ and is therefore increasing in $\theta_1$. However, the common value feature of the model is apparent in that if bidder 1's toehold is 0, then her probability of winning is also 0; moreover, in this case, she bids $v(t_1, 0)$, i.e., the lowest possible value conditional on her own signal. This is because the bidder who moves first is subjected to an extreme winner's curse problem.

How can the matching auction improve the expected sale price compared to the standard auctions? Recall that in the second-price auction with asymmetric toeholds, the smaller toehold bidder is exposed to an extreme winner's curse problem. The matching auction is a way to shield the low toehold bidder from this extreme winner's curse by asking her to move second. This, of course, imposes a winner's curse on the first bidder. However, if the asymmetry is large, the first bidder with a higher toehold will act more like a seller, and this the sale price will not suffer as much. Dasgupta and Tsui (2003) show that, for the case of a value function that is symmetric and linear in the signals (i.e., $v(t_1, t_2) = t_1 + t_2$) that are drawn from the uniform distribution, the matching

[38] For a derivation and a complete characterization of the equilibrium, see Dasgupta and Tsui (2003).

auction generates a higher expected sale price than both the first- and the second-price auctions when the toeholds are sufficiently asymmetric and not too small.

Another type of bidder asymmetry arises in the common value framework if the value function is not symmetric, e.g., $v(t_1, t_2) = \alpha t_1 + (1 - \alpha) t_2$ and $\alpha > 1/2$. Dasgupta and Tsui (2003) show that with symmetric toeholds, the matching auction generates a higher expected sale price than the first-price auction if the value function is sufficiently asymmetric (i.e., $\alpha$ sufficiently close to 0 or 1); and it generates a higher expected sale price than the second-price auction if the value function is sufficiently asymmetric and the toeholds are not too large. Povel and Singh (2006) characterize the optimal selling mechanism for the zero toeholds case and show that discrimination against the strong bidder is optimal. However, to implement the optimal mechanism, the seller needs to know the precise value of $\alpha$ as well as the distribution of the signals. This is not required in the matching auction, for which only the identity of the stronger bidder is needed. In other words, the matching auction is a "detail-free" mechanism. This is an especially appealing property given that for sufficiently large asymmetry, the matching auction does almost as well as the optimal mechanism in extracting the surplus.

### *4.6. Merger waves*

There is no question that merger and acquisition activity goes in waves. Rhodes-Kropf and Viswanathan (2004) give the following perspective: in 1963–1964, there were 3,311 acquisition announcements while in 1968–1969 there were 10,569; during 1979 to 1980 and also from 1990 to 1991 there were only 4,000 announcements while in 1999 alone there were 9,278 announcements. The 1980s were generally a period of high merger and acquisition activity, and saw the emergence of the hostile takeover and corporate raiders, but activity dropped off in the early 90s only to rebound again late in the 90s. Holmstrom and Kaplan (2001) review the evidence on merger waves and offer a macro explanation based on changing regulatory and technological considerations which created a wedge between corporate performance and potential performance, along with developments in capital markets which gave institutional investors the incentives ands ability to discipline managers.

Rhodes-Kropf and Viswanathan (2004) offer an alternative explanation for merger waves based on an auction-theoretic model rich in informational assumptions. They note that periods of high merger activity tend to be periods of high market valuation, and the means of payment is generally stock. For example, the percentage of stock in acquisitions as a percentage of deal value was 24% in 1990, but 68% in 1998. They focus on mergers where stock is the means-of-payment. The essence of the argument is as follows: stock values of both targets and acquirers can become over-valued on a market-wide basis. These are economy-wide pricing errors that managers of neither targets nor acquirers have information on, but they do know they occur. Managers of targets know when their own stocks are overvalued; however, they do not know how much of that is due to economy-wide pricing errors and how much is firm-specific. When a stock offer is made *in an overvalued market*, target managers, knowing their own firms are overvalued but not knowing whether this is due to market-wide or firm specific factors, will overestimate potential synergies with acquirers. This is similar to search-based explanations of labor market unemployment, whereby workers think that a decrease in demand for their labor at one firm is firm-specific (when it is in fact business cycle related) and therefore accept unemployment, thinking that their economy-wide opportunities have not been affected. Thus, in times of economy-wide overvaluation, target firms will accept more bids, for they rationally infer that synergy with the bidder is high. Of course, with each merger, the market should rationally lower the price, taking the possibility of overvaluation into account. However, this does not rapidly lead to an end of a wave: if synergies are correlated, then merger waves can occur, because the market also revises upward the probability that synergies for *all* firms are high. Correlation of synergies can arise out of the sort of considerations that Holmstrom and Kaplan discuss, for example, changes in technology which increase the efficient scale of firms. Thus, a merger wave that begins when the market becomes overvalued may end only when the market realizes that the synergies that were anticipated are actually not there—i.e., the wave ends with a market crash.

Rhodes-Kropf and Viswanathan's model is one of an open auction with bidders offering shares of the combined firm, similar to that of Hansen (1985a, 1985b). Multiple bidders and cash offers are possible. High bids by other bidders imply more likely misvaluation in stock offers; however, since synergies are correlated, this does not cause the wave to end. Stock-based deals are also more likely than pure cash deals in times of economy-wide overvaluation because of the valuation errors that targets make given the information structure. Thus, the model explains not only merger waves but also the stylized fact that in times of intense merger activity, stock is more likely to be used as the means-of-payment.

Shleifer and Vishny (2003) propose a theory of mergers and acquisitions which has a similar flavor. They argue that merger activity is driven by the relative valuations of bidders and targets and perceptions of synergies from merger activity. Suppose that acquirer and target have $K_1$ and $K$ units of capital, respectively. The current market valuations per unit of capital are $Q_1$ and $Q$, respectively, where $Q_1 > Q$. The long-run value of all assets is $q$ per unit. If the two firms are combined, then the *short-run* value of the combined assets is $S(K + K_1)$, where $S$ is the "perceived synergy" from the merger. In other words, "$S$ is the story that the market consensus holds about the benefits of the merger. It could be a story about [the benefits of] diversification, or consolidation, or European integration". Suppose $P$ is the price paid to the target in a merger. If the means of payment is stock, it is easily checked that *long-run* benefit to the bidding firms' shareholders is $qK(1 - P/S)$ and that to the target shareholders is $qK(P/S - 1)$.[39] Thus, if $S > P > Q$, bidding firms' shareholders benefit in the long run but target

[39] Since the synergy is only in the mind of the beholder (the market), the long term benefit to the bidding firms' shareholders from a cash offer would be $q(K + K_1) - PK < qK_1$ since $P > Q > q$. On the other hand, if the synergy were real, a bidding firm would have no reason to prefer stock over cash. Thus, a large number of stock offer during a particular period should reveal to the market that the synergies are

shareholders benefit in the short run. Shleifer and Vishny (2003) argue that if target shareholders or managers have shorter horizons, they may be willing to trade off the short run benefits for the long run losses. For example, target management may be close to retirement or own illiquid stock and options.[40] Shleifer and Vishny (2003) argue that the example of family firms selling to conglomerates and entrepreneurial firms selling to firms such as Cisco and Intel in the 1990s fit this story very well. Alternatively, the bidding firm could simply "bribe" target management—Hartzell, Ofek and Yermack (2004) find that target management receive significant wealth gains in acquisitions, and acquisitions with higher wealth gains for target management are associated with lower takeover premia.

Overall, the theory predicts that cash offers will be made when perceived synergies are low but the target is undervalued ($Q < q$). This is likely to be a situation where the firm needs to be split up and/or incumbent management replaced to improve value, and will be associated with target management resistance and poor pre-acquisition target returns. In contrast, stock offers will be made when market valuations are high, but there is also significant dispersion in market values. Finally, for stock offers to succeed, there must be a widely accepted "story" about synergies, and target management must have shorter horizons. Notice that the model also predicts that the short term returns to bidders in stock offers would be negative if the synergies are not extremely high ($S > Q_1$, i.e., the bidder essentially has a money machine) and the long-run returns would also be negative. For cash offers, both short and long-term returns should be positive.

Shleifer and Vishny (2003) argue that the three most recent merger waves nicely fall into their framework. The conglomerate merger wave of the 1960s was fuelled by high market valuations and a story about the benefits of diversification through better management. The acquisition of firms in unrelated businesses might have been more attractive because target firms in the same industry would also have high market valuations. The targets were often family firms whose owners wished to cash out and retire. However, since there was really no synergy from diversification, the wave of the 1960s gave rise to the bust-up takeovers of the 1980s—acquisitions that were in cash, hostile, and of undervalued targets. Rising stock market prices ended this wave of takeover activity as undervalued targets became more difficult to find. The most recent wave of the 1990s was ushered in by the rising market valuations. The story of synergy was reinvented: technological synergies, the benefits of consolidation, and the European integration.

more apparent than real. It is precisely this kind of inference that is carefully modelled in Rhodes-Kropf and Viswanathan's (2004) model discussed above. Shleifer and Vishny (2003) brush aside these issues by assuming that the market is irrational.

[40] Cai and Vijh (2006) find that in the cross-section of all firms during 1993–2001, CEOs with higher illiquidity discount are more likely to get acquired. Further, in a sample of 250 completed acquisitions, target CEOs with higher illiquidity discount accept lower premium and are more likely to leave after acquisition. They also put up lower resistance and speed up the process.

### 4.7. Auctions in bankruptcy

One of the most fruitful areas for the application of auction theory in corporate finance is in the context of corporate bankruptcy. The theoretical efficiency of auctions in allocating assets to their most highly-valued use has led many scholars to propose auctions as a means to resolve some of the issues in bankruptcy. As an auction also yields a price for the corporation, the question of determining value (for the purpose of settling claims) is also solved. Unfortunately, the informational issues in bankruptcy are quite severe; so any complete auction-based model of the process which will yield predictions on total cost must include the cost of information acquired by bidders. There is also a fairly prevalent view that credit markets may not always allocate financing efficiently to potential buyers of bankrupt companies, so prices may be low because of a dearth of bidders. Some of these issues have been addressed empirically by examining the bankruptcy process in Sweden, where auctions of bankrupt companies are mandatory (see related discussion below).[41]

Baird (1986) was one of the first to point out that auctions may be preferable to the court-supervised reorganization process of the United States' Chapter 11 bankruptcy code. Baird, among others, used the auction processes and results of the corporate takeover market as an analogy to estimate the gains that may be achieved if auctions were used to transfer control of bankrupt companies' assets. Other researchers, Weiss (1990) in particular, turned to estimating the direct cost of Chapter 11 procedures—with those costs being estimated at between 2.8% and 7.5% of assets. Easterbrook (1990) argues against auctions, maintaining that the costs associated with the IPO process is a good analogy for estimating the costs of determining a firm's value, and calculates IPO direct costs at roughly 14% of proceeds. Hansen and Thomas (1998) argue that Easterbrook's figures need to be adjusted and put on a total asset, not proceeds, basis, and that the so-called "dealer's concession" built into IPO costs should also be subtracted as it is a cost of distribution, not of the auction process per se. Their resulting figure of 2.7% is then roughly equal to Weiss' estimates of the direct cost of bankruptcy. Thus, auctions and Chapter 11 would seem to have similar direct costs, leaving their relative efficiency to be determined by either theory or further empirical work.

On the empirical side, Thorburn (2000) has exploited the Swedish bankruptcy experience to draw important conclusions on the relative efficiency of cash auctions of bankrupt firms. In Sweden, the typical procedure has been for a bankrupt firm to be taken over by a court-appointed trustee who supervises a cash auction of the firm, either piecemeal or as an ongoing combination. These data therefore allow for direct examination of how auctions work in bankruptcy. Thorburn (2000) finds that three-quarters of the 263 bankrupt firms are auctioned as going-concerns, which compares favorably to Chapter 11 survival rates. As to cost, direct costs average 6.4% of pre-filing assets, with the one-third largest firms experiencing costs of only 3.7% of assets. As to debt recovery, the recovery rates are comparable to Chapter 11 reorganizations of much larger

[41] See Eckbo and Thorburn (2003, 2005).

firms: on average, creditors received 35% of their claims, with secured creditors receiving 69% and unsecured creditors only 25%. Thorburn finds that APR is maintained by the auction procedure.

Eckbo and Thorburn (2005) construct an auction-based model to examine the incentives of the main creditor bank in a bankruptcy auction. Their work addresses one fear of bankruptcy auctions, that credit market inefficiencies will sometimes limit credit and cause bankrupt companies to be sold at "fire-sale" prices, possibly to the benefit of the original owner/managers. Eckbo and Thorburn show that the main creditor bank has an incentive to provide financing to one bidder and to encourage that bidder to bid higher than would be in their private interest. The reason for this follows from the analysis of an optimal reserve price (see also the discussion of toeholds in Section 4.4) in an auction, for the main creditor bank is essentially a partial owner of the bankrupt company. Just as an optimal reserve price exceeds the seller's own valuation (see above), the optimal bid for a main-bank financed bidder exceeds that bidder's own valuation. The equation specifying the optimal bid in Eckbo and Thorburn is exactly analogous to the equation for an optimal reserve price. Eckbo and Thorburn, examining again the Swedish data, find strong results for the over-bidding theory and no evidence that auction prices are affected by industry-wide distress or business cycle downturns. They also demonstrate a surprising degree of competition in the automatic bankruptcy auctions, and that auction premiums are no lower when the firm is sold back to its own owners. Overall, their evidence—which is the first to exploit directly the cross-sectional variation in auction prices—fails to support either fire-sale arguments or the notion that salebacks are non-competitive transactions.

Auction theory has also been applied to study the question of optimal bankruptcy procedures. Hart et al. (1997) propose an ingenious three-stage auction process for bankrupt companies. The first stage solicits cash and non-cash bids for the firm, while the second and third stages determine prices and ownership of so-called "reorganization rights". Reorganization rights are new securities which consolidate all the various existing claims on the firm's assets. This proposal differs from Aghion, Hart and Moore (1992) in that there is a public auction (the third auction) for the reorganization rights. The purpose here is to reduce any inefficiencies caused by liquidity constraints in determining prices and allocations of the new securities which replace the old claims.

Rhodes-Kropf and Viswanathan (2000) extend the limited work done on non-cash bids in auctions discussed previously. While theory such as Hansen (1985a, 1985b) shows that non-cash bids such as equity can increase sales revenue, non-cash bids are themselves subject to uncertain valuation. Building on these basic insights, Rhodes-Kropf and Viswanathan show that in any separating equilibrium, a security auction (the means-of-payment is a security the value of which depends on the bidder's type) generates higher expected revenue to the seller than a cash auction. The reason for this is that in a security auction, the low types have a greater gain from mimicking the high types, so to separate, the high types have to bid more. However, some securities will not separate the bidders. The authors show that there is no incentive compatible separating equilibrium with stock alone. Debt bids, or a minimum debt requirement, can achieve separation in some cases; in others, cash payments or large non-pecuniary bankruptcy costs are needed to achieve separation (so that the highest value bidder can be identified). However, relative to cash bids, bids that involve debt or equity distort ex-post effort choices. Bids that involve high debt and low equity rank higher because they distort effort less. Convertibles can work better as they give the seller the option to affect the ex-post capital structure of the target firm. The model thus is capable of explaining why debt and convertibles are often part of reorganization plans, and why companies often end up more highly levered than when they were distressed (Gilson (1997)).

Hansen and Thomas (1998) apply the model of French and McCormick (1984) to argue that uncertainty surrounding a bankrupt firm's assets can cause auction prices to be low. Using the French and McCormick model, with free entry of bidders, the auction price will be $N^*C$ less than true value, where $N^*$ is the equilibrium number of bidders and $C$ is the pre-bid cost of entry (which they model as an information acquisition cost). Theoretically, then, the question is whether a court, by having to only obtain one (good) evaluation of the firm's assets, can hold costs below $N^*C$. They argue that the greater the uncertainty surrounding a firm's assets, the worse an auction will perform. By way of example, Reece (1978) shows that with high uncertainty, a common-value auction yields a price only 70% of true value.

### *4.8. Share repurchases*

Companies frequently buy back their shares through either fixed-price tender offers or Dutch auction mechanisms. In a Dutch auction repurchase, a company determines a quantity of shares to buy back and asks shareholders to submit bids specifying a price and quantity of shares that they are willing to sell. The bids are ordered according to price (low to high), creating a supply curve. As the Securities and Exchange Commission prohibits price discrimination, a uniform price is set corresponding to the lowest price that enables the firm to buy the pre-determined number of shares.

While there has been little formal modeling of the Dutch auction repurchase process itself (possibly because no real auction-theoretic issues are present) there is considerable empirical study, and their effects relative to fixed-price offers has been studied in a more traditional corporate finance setting. Bagwell (1992) studies 32 Dutch auction repurchases between 1988 and 1991. In one transaction, the highest bid was 14% above the pre-announcement market price, while the lowest bid was only 2% above. Such disparities in bids are documented for the entire sample, showing that the firms did face upward-sloping supply curves for their shares, contrary to naive ideas of a perfect capital market. Bagwell mentions several possible explanations, including differences in private valuations (for example, because of capital gains tax lock-ins), asymmetric information about a common value as in Milgrom and Weber (1982a, 1982b), or differences in opinion (Miller (1977)). While tax considerations could play a large role, it is certainly not a stretch to assume that shareholders will have different information on the value of a company (even though they share the public information embedded in the current price).

Other work has explored signaling aspects of Dutch auction repurchases relative to fixed price tender offers (Persons, 1994) and relative to paying dividends (Hausch and Seward, 1993). Persons (1994) considers a situation in which shareholders demand a premium (perhaps due to capital gains tax frictions) to tender their shares, but this premium varies across shareholders, resulting in an upward sloping supply curve. Repurchases are costly to existing shareholders because the tendering shareholders must be offered a premium. Importantly, the slope of the supply curve is random. In a fixed-price tender offer, the price is fixed, while the quantity of shares tendered adjusts to the random slope of the supply curve; in a Dutch auction, exactly the opposite is the case. If the manager intends to signal the true value by maximizing a weighted average of the intrinsic value and the market value of the shares (as in the dividend signaling model of Miller and Rock, 1985), fixed-price offers are more effective signals of the manager's private information; on the other hand, if the manager needs to buy back a specific number of shares to prevent a takeover threat, a Dutch auction is better as it guarantees that the required number of shares will be tendered.

### *4.9. Auction aspects of initial public offerings (IPOs)*

In the summer of 2004, the internet search firm Google completed the world's largest initial public offering to be conducted via an auction procedure. Google sold 19.6 million shares at an offering price of $85 each, for a total of $1.67 billion raised. The auction method used was a variant of the Wall Street Dutch auction, covered immediately above. Initial public offerings of equity shares would seem to be excellent candidates for an auction procedure: multiple units of the same item for sale, with uncertainty over value and ability of a seller to commit to a sales method.

Interestingly, however, the evidence suggests that formal auctions are not favored as a sales mechanism. Instead, the IPO procedure known as "bookbuilding" attracts most of the market in regions where multiple sales methods can legally exist (Sherman, 2005; Jagannathan and Sherman, 2006; Degeorge, Derrien and Womack, 2004). A fair amount of theoretical work has been done to explore differences in sales mechanisms for IPOs as well as issues within any one sales method. There is also a literature examining relative performance of auctions versus other sales methods, for in some countries we do have different sales methods co-existing.

In applying auction theory to IPOs, the place to start is the literature on uniform price, multiple unit auctions. The main initial contributions here are Wilson (1979), and Back and Zender (1993). A recent contribution is by Kremer and Nyborg (2004a, 2004b). The reason this literature is so important is that it shows how simple auction analysis yields the main underpricing result from IPO studies (that is, that the initial stock market returns immediately after setting the IPO price are overwhelmingly positive).[42] The auction models show in fact that uniform price, multiple unit auctions have a multitude

[42] For a detailed account of various theories of IPO underpricing, see Chapter 7 of Ljungqvist in this volume.

of equilibria with varying degrees of underpricing. The intuition of the underpricing result is quite simple: in a uniform price auction, bidders are asked to essentially submit demand schedules, specifying the number of shares they would be willing to buy at different prices. Wilson (1979) showed that instead of thinking of bidders as selecting a demand schedule to submit, a simple transformation allows us to model a bidder's decision as one of selecting the optimal "stop-out" price after subtracting other bidders' demands from the available supply. This makes each bidder a monopsonist over the residual supply and sets up the essential monopsonistic tension: a higher bid increases the quantity of shares purchased, but raises the price paid on all shares. Optimally, a bidder will submit a low stop-out bid, and as this will be the case for all bidders, a Nash equilibrium holds. Interestingly, the literature on underpricing in IPOs has not picked up on this simple explanation, relying instead on more complicated explanations.

While not relying on the Wilson/Back and Zender insights, Benveniste and Spindt (1989) nonetheless use an auction-based model to explain certain aspects of the IPO process. The basic idea is similar to that of Hansen (2001), as it involves conditions under which bidders reveal truthfully their information through bids that are non-binding "indications of interest". The model asks under what conditions an investor will reveal her information to the investment banker collecting demand information for an IPO. Under-pricing of the IPO guarantees a return to these investors; this is critical for otherwise there could be no incentive to honestly reveal information. Also, those investors who reveal high valuations must receive more of the under-valued shares, or again there would be no payoff from honestly revealing information (and there is a cost to honest revelation as it affects the offering price). Thus, this auction-based model explains two core features of the IPO process, under-pricing and differential allocations of shares.

Biais and Faugeron-Crouzet (2002) present a complex and quite general model of the IPO process that compares auctions to fixed-price offerings. Unfortunately, the authors' conclusion that the book-building approach dominates the auction method is clouded by the assumption that the auction method will induce collusion between the bidders. It is not at all clear why collusion, if profitable, will occur only in one auction method. This paper also shows why it is extremely difficult to use auction theory to convincingly show that one method is more efficient than another: to do this, one must introduce a myriad of assumptions, covering everything from valuations to costs of information collection. The validity of all these assumptions is difficult to evaluate, and the chances that the ranking of the sales methods would change, or become indeterminate, is high if some of the assumptions were changed.

Sherman (2005) compares bookbuilding to auctions under the very reasonable assumption that entry by bidders is an endogenous decision. Her model yields a result similar in spirit to a core result that emerges from comparing the basic auction methods that while the expected price is the same in sealed-bids versus second-price auctions, the variance of prices is greater for the second-price auction. This result comes about because in the first-price auction, bidders put in their bids using their *expectation* of what other bidders' values are, while in the second-price auction, the high-bid is dependent on the *actual* value of the second-highest valuation. Sherman focuses on the uncertainty in the number of bidders caused by a mixed strategy equilibrium in the game of entry into an IPO auction. If bidders are free to enter the IPO auction, then if there is some cost to entering and some classes of bidders are ex-ante identical, the equilibrium in the entry game has a probability of entry for at least some bidders; the result is uncertainty over the actual number of bidders. Sherman claims that this uncertainty over the actual number of bidders causes the IPO price to vary and in particular to vary in its relation to a "true" underlying value. Sherman observes that this additional uncertainty further worsens the "winners' curse" and considerably complicates the optimal "bid-shaving" calculation that is required when there is winners' curse. She also shows that each investor optimally collects less information in a uniform price rather than a discriminatory auction, because of the free rider (moral hazard) problem in the uniform price auction.[43]

Sherman assumes that in the bookbuilding process, the underwriter can select the number of investors to invite into an information-acquisition process; this makes the bookbuilding process more like the first-price auction in terms of the variance of its outcomes. Jagannathan and Sherman (2006) rely on this model to explain their findings of a worldwide abandonment of IPO auctions in favor of bookbuilding; they also support the theoretical model with evidence on the variance in number of participants for IPO auctions. The issues of number of bidders and information collection would seem to be key in an optimal IPO pricing/allocation mechanism. One wonders, however, if a slight twist on assumptions for the auction models—let the auctioneer control somehow the selection of bidders, à la Hansen (2001) would bring equivalence back to the two mechanisms. Sherman (2005, p. 619) does note that "If the term "auction" is interpreted in a broad sense, it is almost a tautology that an appropriate auction could be designed for IPOs". This exemplifies a general difficulty in building theoretical models of two different institutions to explain their empirical performances: one can capture the sense of institutional differences by making clear assumptions (e.g., the underwriter can select the number of potential investors for bookbuilding but not for auctions) but one is left wondering if the assumptions really do justice to what actually happens in practice.

In anther recent attempt at comparing bookbuilding to auctions, Degeorge, Derrien and Womack (2004) show that bookbuilding seems to dominate empirically (they look at France, where for a time auctions and bookbuilding had roughly equal market shares, but now auctions are virtually extinct) and they offer a justification for issuers' preference for the bookbuilding method that is based not on the price performance of bookbuilding but on the investment bankers' preference for the method. While one might understand why investment bankers prefer a method that creates more demand for their services, the link to issuers' interests is less clear. Degeorge et al. hypothe-

[43] In the uniform price auction, since the auction price is set by the actions of bidders who have already paid the information gathering and processing costs, there is an incentive for uninformed bidders to free ride and jump in with a high bid.

size that bankers agree to provide research coverage for issuers in return for using the bookbuilding method. What is left unstated is that issuers must be unable to buy such research coverage on the open market at prices similar to the costs paid by investment bankers: the authors agree that auctions would yield issuers a better price, so one must wonder why issuers put up with an inefficient procedure simply to get a tied service.

On the empirical side of the auctions/IPO issue, Kandel, Sarig and Wohl (1999) utilize a data set from Israel IPO auctions to document elasticity of demand and under-pricing. The under-pricing of Israeli IPOs is intriguing, for those IPOs had their prices set by an explicit auction mechanism. In the period 1993–1996, Israeli IPOs were conducted much like Dutch auction share repurchases: investors submitted sealed-bids specifying prices and quantities, a demand curve was determined, and a uniform price was set at the highest price for which demand equaled the supply of shares available. Kandel, Sarig and Wohl document some elasticity of demand for the reported bids: the average elasticity at the clearing price, based on the accumulated demand curves, was 37 (relatively elastic). Interestingly, even in these IPO auctions, there was under-pricing: the one-day return between the auction price and the market trading price was 4.5%. Another interesting feature of the Israeli auctions is that after the auction but before the first day of trading, the underwriters announce the market clearing price corresponding to the offered quantity, as well as the oversubscription at the minimum price stipulated in the auction. This essentially means that the investor can estimate the price elasticity of demand based on two points on the demand curve. The authors find that the abnormal return on the first day of trading is positively related to the estimate of the elasticity. The authors argue that this reflects greater homogeneity in the estimates of value on the part of the participants in the auction; this is "good news" either because it implies greater accuracy of information about future cash flows and thus leads to a lower risk-premium demanded by investors, or because it signifies greater "market depth" and hence greater future liquidity.

Kerins, Kutsuna and Smith (2003) examine IPOs in Japan in the period 1995–1997, a time when Japanese firms had to use a discriminatory (bidders pay the amount of their bid) auction to sell the first tranche of newly issued shares. This first tranche of shares would be relatively small, and the sale by auction was restricted to outside investors only, with further limitations on the amount that could be bought by any investor. These restrictions could be interpreted as limiting the informational advantages of any one bidder. Under that interpretation, it is not surprising that the authors find relatively little "underpricing" of the shares for the auction tranche: for all the issues, the auction proceeds were only 1.6% below what proceeds would have been at the final aftermarket price. The second stage of the Japanese process was a more traditional fixed-price offer, and there was considerable underpricing of shares at this stage. While this might suggest that the auction was a better choice of mechanism, one must recognize that costs of a larger auction (to sell the entire issue) could well be larger than costs of just the first tranche.

### *4.10. The spectrum auctions and the role of debt in auctions*

Beginning in 1994, the Federal Communications Commission in the United States auctioned licenses for the use of radio spectrum in designated areas. The licenses were auctioned using a novel auction format involving sequential rounds of sealed-bidding on numerous licenses simultaneously. At the end of each round, complete information on the level of bids for all licenses was revealed. The auction format was designed by economists, and at least in regard to the vast sums of money raised, was a great success. Numerous articles summarize all aspects of the auctions, including their design and performance: see, for example, McAfee and McMillan (1996), Milgrom (2000) and Salant (1997). For the empirical researcher, FCC auctions provide a wealth of information: for example, the FCC Web site (http://www.fcc.gov) lists all the bids in all the auctions. Moreover, many of the participating companies are publicly traded, so that company-specific information is also easily available. We focus here on one analysis which studied the effect of debt on the FCC auctions. Clayton and Ravid (2002) construct an auction model where bidders' debt induces lower bids than would otherwise be the case. In this model, bidders have outstanding debt that is large enough to induce bankruptcy if the auction is not won. Lower bids decrease the probability of winning, of course, but in this case guarantee some residual to the shareholders conditional on winning. In effect, in this model, pre-existing debt holders are "third parties" who have a prior claim of a part of the pie. Thus, pre-existing debt serves to reduce bidders' values and therefore reduces bids.[44] An empirical analysis of the FCC bidding data produces a negative but generally insignificant effect of a bidder's own debt on their bid but a negative and significant effect on a firm's bid of competitors' debt levels.

Che and Gale (1998) were the first to explicitly study the role of debt in auctions.[45] They have a result similar to Clayton and Ravid, although the models rely on different effects. In Che and Gale's framework, a second-price auction yields lower expected revenue than a first-price auction. To see how financial constraints affect revenue comparisons, suppose that due to budget constraints, bidders cannot bid more than a given budget, which is observed only by the bidder. The private valuations and budgetary endowments of each bidder are independently and identically distributed according to some joint distribution function. In this context, since bidders in the second-price auction bid their value, but in the first-price auction they bid *below* their value, bidding is more constrained in the second-price auction because of budget constraints, ceteris paribus. As a consequence, the first-price auction generates higher expected revenue. Che and Gale (1998) allow for financial constraints that are more general than we have considered here: for example, these could take the form of a marginal cost of borrowing that is increasing in the amount of the loan.

[44] On the role of debt holders as "third parties" in the context of bilateral bargaining, see Dasgupta and Sengupta (1993). On the role of "third party" shareholders in the context of bilateral bargaining, see Dasgupta and Tao (1998, 2000).

[45] For a recent contribution on the role of financing in auctions, see Rhodes-Kropf and Viswanathan (2005).

### 4.11. Advanced econometrics of auction data

There has been considerable progress in the application of econometric techniques to auction data. While the datasets used in these studies do not cover corporate finance directly, the techniques used should be of interest to corporate finance researchers, as they may be applicable to financial datasets and help resolve certain key issues. One broad topic that has been covered in empirical auction studies and that also appears in corporate finance are auctions with one informed bidder and numerous uninformed bidders. In corporate finance, such a situation could reasonably be assumed when current management is allowed to bid for a corporation, either in a takeover or bankruptcy context. Certainly in bankruptcy one concern has been that management, if allowed to bid in an auction, may be able to purchase the corporate assets at less than fair value. Hendricks and Porter (1988, 1992) have studied U.S. government oil lease auctions of so-called "drainage" tracts—tracts that have a neighboring tract currently under lease to one of the bidders. For these drainage tracts, it is reasonable to assume that the owner of the neighboring tract would have better information than other bidders. The authors of several studies have found this assumption, and the related equilibrium bidding theory, to be consistent with the data. The econometrics used relies heavily on the underlying auction theory. For example, equilibrium with one informed bidder imposes restrictions on the distributions of the informed bidder's bid distribution and the uninformed bidders' bid distributions. Note that a test of this type requires that data on all bids be available.

Structural models are also being used successfully to examine auction data. The most exciting approach here is to use equilibrium theory in conjunction with data on all bids to estimate the underlying probability distribution of the valuations of bidders. The essence of the idea here is that an equilibrium bid function maps a valuation to a bid. If data on bids are available, then with suitable econometrics one can recover the distribution of the underlying valuations from the bid data. Li, Perrigne and Vuong (2002) provide a step-by-step guide to structural estimation of the affiliated private value auction model. One aim of this work in the economics literature has been to estimate the optimal selling mechanism for a real auction. For example, if valuations are affiliated, then revenue equivalence no longer holds. Also, the optimal reserve price depends upon the underlying distribution of values, so if that distribution can be estimated, we can also get an estimate of the optimal reserve price. Researchers in empirical corporate finance should be aware of the progress made in structural estimation of auctions, for some of the issues at the heart of finance auctions may be resolved through structural estimation (and in some finance auctions, there should be data on all bids). For example, in the bankruptcy area, questions of reserve prices and informational rents abound, and these are two issues that structural estimation can get at.[46]

[46] In the context of takeover auctions, Betton and Eckbo (2000) pursue an interesting line of empirical research. A takeover contest typically associated with an "event tree" beginning with the initial bid, possibly

## 5. Conclusion

Upon reflection, the accomplishments of auction theory are really quite amazing. The "black box" of the Walrasian auctioneer has been opened, studied in depth, and its perfection questioned. We can now say a lot about the process of actual price formation in many real markets. While modelers have been able to explore theoretically important topics such as revenue comparisons across auctions, their work has also enabled economists to consult with governments on the design of optimal auctions to sell public assets. And with only a slight time lag, empirical work in auctions is following in the footsteps of theory, with structural estimation methods setting a new standard for creativity and rigor. Similar to the way that theoretical developments made their way into the real world of auction design, empirical work is focusing on real world auctions such as those found on Ebay and other online auctions. There are not too many topics in economics that allow researchers to cover such a broad swath of analytical territory, from the highly theoretical to the highly empirical and practical. In this way, auction theory resembles developments in financial asset pricing, where for instance the development of the option pricing model led to a surge in theoretical and empirical work while at the same time the model was applied in real markets.

The application of auction theory in corporate finance really needs to be seen as the intersection of two fields, that of auction theory and of information-based corporate finance theory. Nobody should have been surprised to see auction theory have a bit of a field day in being applied to topics in corporate finance, and as we think this survey shows, this is clearly what has happened and continues to happen. The question before us, however, must be: what have we learned in the process? That there has been considerable learning cannot be doubted, with the most significant learning being in interpreting the returns to bidders and targets in the market for corporate control, and in understanding the real institutional practices used in financial markets, such as underpricing in the IPO market, non-cash bids in takeover markets, and the role of asymmetries and discrimination against selected bidders. Perhaps the single best measure of auction theory's influence in corporate finance is that most PhD courses in corporate finance will include several papers, if not an entire module, on applications of auctions. As even a superficial study of auctions requires a fair amount of knowledge of game theory, the inclusion of auctions in PhD finance courses reinforces the study of games, itself a critical component of modern finance.

followed by the appearance of rival bidders, until the eventual success or failure of the initial bid. The market reaction to a bid (or indeed, at reaching any node) therefore represents the sum of the product of the probabilities of all subsequent events in the tree emanating from that node, and the associated payoffs. Since the probabilities and market reactions can be estimated, the payoff implications associated with the events (the "market prices") can be estimated. Betton and Eckbo (2000) find generally significant effects for the target, but less significant effects for the bidders.

While auction theory deserves much credit for its inroads into corporate finance, two areas of concern do emerge. First, there are some phenomena in corporate finance for which we still lack sufficient understanding, and where one might have expected auction theory to lead the way. Yes, we have increased our understanding of returns to bidders and targets in the market for corporate control, but why are acquirers' returns so small? Any auction with heterogeneity of valuations or information leads to strategic behavior and expected profits for inframarginal bidders. And why do acquirers seem to do better when acquiring private companies? There is still a huge question as to whether auctions in bankruptcy are better than a court-supervised valuation and division of assets. Why are toeholds so seldom taken, if they lead to a bidding advantage? If auctions really are so good, why are they used so infrequently in the initial public offering market, and why do some sellers of companies bypass an auction in favor of a one-on-one negotiation?

The second unsatisfactory aspect of auctions in corporate finance is simply that no new fundamental insights have emerged. We do understand better how information, values, and strategic behavior combine to yield prices and allocations of assets in real financial markets. There has been no quantum leap forward, just incremental learning at the margin. This should, we suppose, actually be gratifying, for it shows the robustness of our primitive and most cherished assumptions. Unfortunately, at times the models that are developed and that are pushing back the frontier only marginally are incredibly complicated, and one has to wonder if the complexity is worth it. One doubts that quantum leaps in knowledge are going to come from models that need a myriad of questionable assumptions.

Where next for auctions in corporate finance? We would suggest three areas for focus. First, data will be key for further empirical discovery, and this could in turn lead to new theoretical developments. We believe that auctions of private companies and auctions in bankruptcy are two areas that may yield significantly better data in the future and where the returns to clever empirical work would be large. Second, on the theoretical side, it is clear that some of the best work to date has been on what might appear as the second-order institutional practices, such as non-cash bids, toeholds, bidder discrimination, and reserve prices. Much progress has been made in understanding the role of these practices, while at the same time reinforcing the importance and validity of the overall auction-based framework. Third, we would like to see more work done with non-standard informational and valuation assumptions. The general symmetric model is extremely powerful, but does it really capture many of the real settings that we observe? We should expect that heterogeneity of bidders will be manifested in many ways and will turn out to affect the equilibria quite strongly, especially in regard to bidders' profits. Efficiency of the allocation will also become inherently more interesting of a question, and initial work suggests it will be harder to achieve.

We would confidently make the prediction, though, that auctions in corporate finance will be a much-studied topic for years to come. Our very strong recommendation would be for all PhD students to get a thorough grounding in auction theory.

## References


Aghion, P., Hart, O., Moore, J., 1992. The economics of bankruptcy reform. Journal of Law, Economics and Organization 8, 523–546.

Andrade, G., Mitchell, M., Stafford, E., 2001. New evidence and perspectives on mergers. Journal of Economic Perspectives 15, 103–120.

Back, K., Zender, J., 1993. Auctions of divisible goods: On the rationale for the treasury experiment. Review of Financial Studies 6, 733–764.

Bagwell, L.S., 1992. Dutch auction repurchase: An analysis of shareholder heterogeneity. Journal of Finance 47, 71–105.

Baird, D.G., 1986. The uneasy case of corporate reorganization. Journal of Legal Studies 15, 127–147.

Benveniste, L.M., Spindt, P.A., 1989. How investment bankers determine the offer price and allocation of new issues. Journal of Financial Economics 24, 343–361.

Betton, S., Eckbo, B.E., 2000. Toeholds, bid-jumps, and expected payoffs in takeovers. Review of Financial Studies 13, 841–882.

Betton, S., Eckbo, B.E., Thorburn, K.S., 2005. The toehold puzzle. Working Paper, Concordia University Dartmouth College.

Betton, S., Eckbo, B.E., Thorburn, K.S., 2006. Corporate takeovers. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 2. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam. Chapter 15, forthcoming.

Biais, B., Faugeron-Crouzet, A.M., 2002. IPO auction: English, Dutch, ... French and internet. Journal of Financial Intermediation 11, 9–36.

Boone, A., Mulherin, H., 2003. Corporate restructuring and corporate auctions. Working Paper, College of William and Marry, Claremont McKenna College.

Boone, A., Mulherin, H. (2006a). How are firms sold? Journal of Finance, in press.

Boone, A., Mulherin, H. (2006b). Do auctions induce a winner's curse? New Evidence from the corporate takeover market. Working Paper, University of Georgia.

Bradley, M., Desai, A., Kim, E.H., 1983. The rationale behind inter-firm tender offers: Information or synergy? Journal of Financial Economics 11, 141–153.

Bulow, J., Huang, M., Klemperer, P., 1999. Toeholds and takeovers. Journal of Political Economy 107, 427–454.

Bulow, J., Klemperer, P., 1996. Auction versus negotiations. American Economic Review 86, 180–194.

Bulow, J., Roberts, J., 1989. The simple economics of optimal auctions. Journal of Political Economy 97, 1060–1090.

Burkart, M., 1995. Initial shareholdings and overbidding in takeover contests. Journal of Finance 50, 1491–1515.

Cai, J., Vijh, A.M., 2006. Incentive effects of stock and option holdings for target and acquirer CEOs. Journal of Finance, in press.

Che, Y.-K., Gale, I., 1998. Standard auctions with financially constrained bidders. Review of Economic Studies 65, 1–22.

Clayton, M.J., Ravid, S.A., 2002. The effect of leverage of bidding behavior: theory and evidence from the FCC auction. Review of Financial Studies 15, 723–750.

Cramton, P., Schwartz, A., 1991. Using auction theory to inform takeover regulation. Journal of Law, Economics and Organization 7, 27–53.

Dasgupta, P.S., Hammond, P.J., Maskin, E.S., 1979. The implementation of social choice rules: some general results on incentive compatibility. Review of Economic Studies 46, 185–216.

Dasgupta, S., Sengupta, K., 1993. Sunk investment, bargaining, and choice of capital structure. International Economic Review 34, 203–220.

Dasgupta, S., Tao, Z., 1998. Incomplete contracts, ownership rights, and the optimality of equity joint ventures. Journal of Economic Behavior and Organizations 37, 391–413.

Dasgupta, S., Tao, Z., 2000. Bargaining, bonding, and partial ownership. International Economic Review 41, 609–635.

Dasgupta, S., Tsui, K., 2003. A “matching auction” for targets with heterogeneous bidders. Journal of Financial Intermediation 12, 331–364.
Dasgupta, S., Tsui, K., 2004. Auction with cross-shareholdings. Economic Theory 24, 163–194.
Degeorge, F., Derrien, F., Womack, K.L., 2004. Quid pro quo in IPOs: Why book-building is dominating auctions. Working Paper, University of Lugano; University of Toronto; Dartmouth College.
DeMarzo, P.M., Kremer, I., Skrzypacz, A., 2005. Bidding with securities: Auctions and security design. American Economic Review 95, 936–959.
Easterbrook, F.H., 1990. Is corporate bankruptcy efficient? Journal of Financial Economics 27, 411–417.
Eckbo, B.E., Giammarino, R., Heinkel, R., 1990. Asymmetric information and medium of exchange in takeovers: Theory and tests. Review of Financial Studies 3, 651–675.
Eckbo, B.E., Thorburn, K.S., 2000. Gains to bidder firms revisited: Domestic and foreign acquisitions in Canada. Journal of Financial and Quantitative Analysis 35, 1–25.
Eckbo, B.E., Thorburn, K.S., 2003. Control benefit and CEO discipline in automatic bankruptcy auctions. Journal of Financial Economics 69, 227–258.
Eckbo, B.E., Thorburn, K.S., 2005. Bidding in mandatory bankruptcy auctions: Theory and evidence. Working Paper.
Fishman, M.J., 1988. A theory of pre-emptive takeover bidding. RAND Journal of Economics 19, 88–101.
Fishman, M.J., 1989. Preemptive bidding and the role of the medium of exchange in acquisitions. Journal of Finance 44, 41–57.
French, K.R., McCormick, R.E., 1984. Sealed bids, sunk costs, and the process of competition. Journal of Business 57, 417–441.
Gal-Or, E., 1985. First mover and second mover advantages. International Economic Review, 649–653.
Gal-Or, E., 1987. First mover disadvantages with private information. Review of Economic Studies, 279–292.
Gilson, S.C., 1997. Transactions costs and capital structure choise: Evidence from financially distressed firms. Journal of Finance 52, 161–196.
Grossman, S., Hart, O., 1980. Disclosure laws and takeover bids. Journal of Finance 35, 323–334.
Hansen, R.G., 1985a. Empirical testing of auction theory. American Economic Review 75, 156–159.
Hansen, R.G., 1985b. Auction with contingent payments. American Economic Review 75, 862–865.
Hansen, R.G., 1986. Sealed bid versus open auctions: The evidence. Economic Inquiry 24, 125–142.
Hansen, R.G., 2001. Auction of companies. Economic Inquiry 39, 30–43.
Hansen, R.G., Thomas, R.S., 1998. Auctions in bankruptcy: Theoretical analysis and practical guidance. International Review of Law and Economics 18, 159–185.
Harris, M., Raviv, A., 1981. Allocation mechanisms and the design of auctions. Econometrica 49, 1477–1499.
Hart, O., LaPorta, R., Lopez-de-Silanes, F., Moore, J., 1997. A new bankruptcy procedure that uses multiple auctions. European Economic Review 41, 461–473.
Hartzell, J., Ofek, E., Yermack, D., 2004. What’s in it for me? CEOs whose firms are acquired. Review of Financial Studies 17, 37–61.
Hausch, D.B., Seward, J.K., 1993. Signalling with dividends and share repurchases: A choice between deterministic and stochastic cash disbursements. Review of Financial Studies 6, 121–154.
Hendricks, R., Porter, R., 1988. An empirical study of an auction with asymmetric information. American Economic Review 78, 865–883.
Hendricks, K., Porter, R., 1992. Joint bidding in the federal OCS auctions. American Economic Review 82, 506–511.
Herzel, L., Shepro, W., 1990. Bidders and Targets: Mergers and Acquisitions in the U.S. Blackwell, Cambridge, MA.
Hirshleifer, J., 1971. The private and social value of information and the reward to inventive activity. American Economic Review 61, 561–574.
Holmstrom, B., Kaplan, S.N., 2001. Corporate governance and merger activity in the United States: Making sense of the 1980s and 1990s. Journal of Economic Perspectives 15, 121–144.
Jagannathan, R., Sherman, A., 2006. Why do IPO auctions fail? NBER Working Paper.
Jarrell, G.A., Brikley, J.A., Netter, J.M., 1988. The market for corporate control: The empirical evidence since 1980. Journal of Economic Prospectives 2 (1), 49–68.

Jennings, R.H., Mazzeo, M.A., 1993. Competing bids, target management resistance, and the structure of takeover bids. Review of Financial Studies 6, 883–909.
Jensen, M.C., 2004. The agency costs of overvalued equity. Harvard NOM Research Papar No. 04-26 (May 2004) and ECGI-Finance Working Paper No. 39/2004. http://ssrn.com/abstract=480421.
Jovanovic, B., Braguinsky, S., 2004. Bidder discounts and target premia in takeovers. American Economic Review 94, 46–56.
Kahan, M., Klausner, M., 1996a. Lockups and the market for corporate control. Stanford Law Review 48, 1539–1571.
Kahan, M., Klausner, M., 1996b. Path dependence in corporate contracting: Increasing returns, herding behavior and cognitive biases. Washington University Law Quarterly 74, 317–325.
Kandel, S., Sarig, O., Wohl, A., 1999. The demand for stocks: An empirical analysis of IPO stock auctions. Review of Financial Studies 12, 227–247.
Kerins, F., Kutsuna, K., Smith, R., 2003. Why are IPO underpriced? Evidence from Japan's hybrid auction-method offerings. Working Paper, Claremont Graduate University.
Klemperer, P.D., 1998. Auctions with almost common values. European Economic Review 42, 757–769.
Klemperer, P.D., 2000. Why every economist should learn some auction theory. In: Dewatripont, M., Hansen, L., Turnovsky, S. (Eds.), Advances in Economics and Econometrics. Cambridge University Press, Cambridge.
Kremer, I., Nyborg, K.G., 2004a. Divisible good auctions: The role of allocation rules. RAND Journal of Economics 35, 147–159.
Kremer, I., Nyborg, K.G., 2004b. Underpricing and market power in uniform price auctions. Review of Financial Studies 17, 849–877.
Krishna, V., 2002. Auction Theory. Academic Press, San Diego, CA.
Lang, L., Stulz, R.M., Walkling, R.A., 1991. A test of the free cash flow hypothesis: The case bidder returns. Journal of Financial Economics 29, 315–335.
Li, T., Perrigne, I.M., Vuong, Q., 2002. Structural estimation of the affiliated private value model. RAND Journal of Economics 33, 171–193.
Loderer, C., Martin, K., 1990. Corporate acquisitions by listed firms: The experience of a comprehensive sample. Financial Management 19, 17–33.
Maskin, E., Riley, J., 2000a. Asymmetric auctions. Review of Economic Studies 67, 413–438.
Maskin, E., Riley, J., 2000b. Equilibrium in sealed high bid auctions. Review of Economic Studies 67, 439–454.
McAfee, R.P., McMillan, J., 1987. Auctions and biddings. Journal of Economic Literature 25, 699–738.
McAfee, R.P., McMillan, J., 1996. Analyzing the airwaves auction. Journal of Economic Perspectives 10, 159–175.
Milgrom, P.R., Weber, R.J., 1982a. A theory of auctions and competitive bidding. Econometrica 50, 1089–1122.
Milgrom, P.R., Weber, R.J., 1982b. The value of information in a sealed-bid auction. Journal of Mathematical Economics 10, 105–114.
Milgrom, P.R., 2000. Putting auction theory to work: The simultaneous ascending auction. Journal of Political Economy 108, 245–272.
Miller, M., 1977. Debt and taxes. Journal of Finance 32, 261–275.
Miller, M., Rock, K., 1985. Dividend policy under asymmetric information. Journal of Finance 40, 1031–1051.
Moeller, S., Schlingemann, F., Stulz, R., 2004. Firm size and the gains from acquisitions. Journal of Financial Economics 73, 201–228.
Morck, R., Shleifer, A., Vishny, R.W., 1990. Do managerial objectives drive bad acquisitions? Journal of Finance 45, 31–48.
Myerson, R., 1981. Optimal auction design. Mathematics of Operation Research 6, 58–73.
Persons, J., 1994. Signaling and takeover deterrence with stock repurchases: Dutch auctions versus fixed price tender offers. Journal of Finance 49, 1373–1402.

Povel, P., Singh, R., 2006. Takeover contests with asymmetric bidders. Review of Financial Studies, in press.
Reece, D.K., 1978. Competitive bidding for offshore petroleum lease. Bell Journal of Economics 9, 364–384.
Rhodes-Kropf, M., Viswanathan, S., 2000. Corporate re-organization and non-cash auctions. Journal of Finance 55, 1807–1849.
Rhodes-Kropf, M., Viswanathan, S., 2004. Market valuation and merger waves. Journal of Finance 59, 2685–2718.
Rhodes-Kropf, M., Viswanathan, S., 2005. Financing auction bids. RAND Journal of Economics, in press.
Riley, J., Samuelson, W., 1981. Optimal auctions. American Economic Review 71, 381–392.
Roll, R., 1986. The Hubris hypothesis of corporate takeovers. Journal of Business 59, 197–216.
Salant, D.J., 1997. Up in the air: GTE's experience in the MTA auction for personal communication services licenses. Journal of Economics and Management Strategy 6, 549–572.
Singh, R., 1998. Takeover bidding with toeholds: The case of owner's curse. Review of Financial Studies 11, 679–704.
Sherman, A., 2005. Global trends in IPO methods: Book building versus auctions with endogenous entry. Journal of Financial Economics 78, 615–649.
Shleifer, A., Vishny, R., 1986. Large shareholders and corporate control. Journal of Political Economy 94, 461–488.
Shleifer, A., Vishny, R., 2003. Stock market driven acquisitions. Journal of Financial Economics 70, 295–311.
Thorburn, K., 2000. Bankruptcy auctions: Costs, debt recovery, and firm survival. Journal of Financial Economics 58, 337–368.
Weiss, L.A., 1990. Bankruptcy resolution: Direct costs and violation of priority of claims. Journal of Financial Economics 27, 285–314.
Wilson, R., 1979. Auctions of shares. Quarterly Journal of Economics 94, 675–689.

*Chapter 4*

# BEHAVIORAL CORPORATE FINANCE*

MALCOLM BAKER

*Harvard Business School and NBER, Soldiers Field, Boston, MA 02163, USA*
*e-mail: mbaker@hbs.edu*

RICHARD S. RUBACK

*Harvard Business School, Soldiers Field, Boston, MA 02163, USA*
*e-mail: rruback@hbs.edu*

JEFFREY WURGLER

*Stern School of Business and NBER, 44 West Fourth Street, Suite 9-190, New York, NY 10012-1126, USA*
*e-mail: jwurgler@stern.nyu.edu*

## Contents



* The authors are grateful to Heitor Almeida, Nick Barberis, Zahi Ben-David, Espen Eckbo, Xavier Gabaix, Dirk Hackbarth, Dirk Jenter, Augustin Landier, Alexander Ljungqvist, Ulrike Malmendier, Jay Ritter, David Robinson, Hersh Shefrin, Andrei Shleifer, Meir Statman, Theo Vermaelen, Ivo Welch, and Jeffrey Zweibel for helpful comments. Baker and Ruback gratefully acknowledge financial support from the Division of Research of the Harvard Business School.

*Handbook of Corporate Finance, Volume 1*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01004-X*

**Abstract**

Research in behavioral corporate finance takes two distinct approaches. The first emphasizes that investors are less than fully rational. It views managerial financing and investment decisions as rational responses to securities market mispricing. The second approach emphasizes that managers are less than fully rational. It studies the effect of nonstandard preferences and judgmental biases on managerial decisions. This survey reviews the theory, empirical challenges, and current evidence pertaining to each approach. Overall, the behavioral approaches help to explain a number of important financing and investment patterns. The survey closes with a list of open questions.

**Keywords**

irrational investors, irrational managers, investment policy, financial policy, market timing, catering

## 1. Introduction

Corporate finance aims to explain the financial contracts and the real investment behavior that emerge from the interaction of managers and investors. Thus, a complete explanation of financing and investment patterns requires an understanding of the beliefs and preferences of these two sets of agents. The majority of research in corporate finance assumes a broad rationality. Agents are supposed to develop unbiased forecasts about future events and use these to make decisions that best serve their own interests. As a practical matter, this means that managers can take for granted that capital markets are efficient, with prices rationally reflecting public information about fundamental values. Likewise, investors can take for granted that managers will act in their self-interest, rationally responding to incentives shaped by compensation contracts, the market for corporate control, and other governance mechanisms.

This paper surveys research in behavioral corporate finance. This research replaces the traditional rationality assumptions with potentially more realistic behavioral assumptions. The literature is divided into two general approaches, and we organize the survey around them. Roughly speaking, the first approach emphasizes the effect of *investor* behavior that is less than fully rational, and the second considers *managerial* behavior that is less than fully rational. For each line of research, we review the basic theoretical frameworks, the main empirical challenges, and the empirical evidence. Of course, in practice, both channels of irrationality may operate at the same time; our taxonomy is meant to fit the existing literature, but it does suggest some structure for how one might, in the future, go about combining the two approaches.

The "irrational investors approach" assumes that securities market arbitrage is imperfect, and thus that prices can be too high or too low. Rational managers are assumed to perceive mispricings, and to make decisions that may encourage or respond to mispricing. While their decisions may maximize the short-run value of the firm, they may also result in lower long-run values as prices correct. In the simple theoretical framework we outline, managers balance three objectives: fundamental value, catering, and market timing. Maximizing fundamental value has the usual ingredients. Catering refers to any actions intended to boost share prices above fundamental value. Market timing refers specifically to financing decisions intended to capitalize on temporary mispricings, generally via the issuance of overvalued securities and the repurchase of undervalued ones.

Empirical tests of the irrational investors model face a significant challenge: measuring mispricing. We discuss how this issue has been tackled and the ambiguities that remain. Overall, despite some unresolved questions, the evidence suggests that the irrational investors approach has a considerable degree of descriptive power. We review studies on investment behavior, merger activity, the clustering and timing of corporate security offerings, capital structure, corporate name changes, dividend policy, earnings management, and other managerial decisions. We also identify some disparities between the theory and the evidence. For example, while catering to fads has potential to reduce long-run value, the literature has yet to clearly document significant long-term value losses.

The second approach to behavioral corporate finance, the "irrational managers approach", is less developed at this point. It assumes that managers have behavioral biases, but retains the rationality of investors, albeit limiting the governance mechanisms they can employ to constrain managers. Following the emphasis of the current literature, our discussion centers on the biases of optimism and overconfidence. A simple model shows how these biases, in leading managers to believe their firms are undervalued, encourage overinvestment from internal resources, and a preference for internal to external finance, especially internal equity. We note that the predictions of the optimism and overconfidence models typically look very much like those of agency and asymmetric information models.

In this approach, the main obstacles for empirical tests include distinguishing predictions from standard, non-behavioral models, as well as empirically measuring managerial biases. Again, however, creative solutions have been proposed. The effects of optimism and overconfidence have been empirically studied in the context of merger activity, corporate investment-cash flow relationships, entrepreneurial financing and investment decisions, and the structure of financial contracts. Separately, we discuss the potential of a few other behavioral patterns that have received some attention in corporate finance, including bounded rationality and reference-point preferences. As in the case of investor irrationality, the real economic losses associated with managerial irrationality have yet to be clearly quantified, but some evidence suggests that they are very significant.

Taking a step back, it is important to note that the two approaches take very different views about the role and quality of managers, and have very different normative implications as a result. That is, when the primary source of irrationality is on the investor side, long-term value maximization and economic efficiency requires insulating managers from short-term share price pressures. Managers need to be insulated to achieve the flexibility necessary to make decisions that may be unpopular in the marketplace. This may imply benefits from internal capital markets, barriers to takeovers, and so forth. On the other hand, if the main source of irrationality is on the managerial side, efficiency requires reducing discretion and obligating managers to respond to market price signals. The stark contrast between the normative implications of these two approaches to behavioral corporate finance is one reason why the area is fascinating, and why more work in the area is needed.

Overall, our survey suggests that the behavioral approaches can help to explain a range of financing and investment patterns, while at the same time depend on a relatively small set of realistic assumptions. Moreover, there is much room to grow before the field reaches maturity. In an effort to stimulate that growth, we close the survey with a short list of open questions.

## 2. The irrational investors approach

We start with one extreme, in which rational managers coexist with irrational investors. There are two key building blocks here. First, irrational investors must influence securities prices. This requires limits on arbitrage. Second, managers must be smart in the sense of being able to distinguish market prices and fundamental value.

The literature on market inefficiency is far too large to survey here. It includes such phenomena as the January effect; the effect of trading hours on price volatility; post-earnings-announcement drift; momentum; delayed reaction to news announcements; positive autocorrelation in earnings announcement effects; Siamese twin securities that have identical cash flows but trade at different prices, negative "stub" values; closed-end fund pricing patterns; bubbles and crashes in growth stocks; related evidence of mispricing in options, bond, and foreign exchange markets; and so on. These patterns, and the associated literature on arbitrage costs and risks, for instance short-sales constraints, that facilitate mispricings, are surveyed by Barberis and Thaler (2003) and Shleifer (2000). In the interest of space, we refer the reader to these excellent sources, and for the discussion of this section we simply take as given that mispricings can and do occur.

But even if capital markets are inefficient, why assume that corporate managers are "smart" in the sense of being able to identify mispricing? One can offer several justifications. First, corporate managers have superior information about their own firm. This is underscored by the evidence that managers earn abnormally high returns on their own trades, as in Muelbroek (1992), Seyhun (1992), or Jenter (2005). Managers can also create an information advantage by managing earnings, a topic to which we will return, or with the help of conflicted analysts, as for example in Bradshaw, Richardson, and Sloan (2003).

Second, corporate managers also have fewer constraints than equally "smart" money managers. Consider two well-known models of limited arbitrage: De Long et al. (1990) is built on short horizons and Miller (1977) on short-sales constraints. CFOs tend to be judged on longer horizon results than are money managers, allowing them to take a view on market valuations in a way that money managers cannot.[1] Also, short-sales constraints prevent money managers from mimicking CFOs. When a firm or a sector becomes overvalued, corporations are the natural candidates to expand the supply of shares. Money managers are not.

Third and finally, managers might just follow intuitive rules of thumb that allow them to identify mispricing even without a real information advantage. In Baker and Stein (2004), one such successful rule of thumb is to issue equity when the market is particularly liquid, in the sense of a small price impact upon the issue announcement. In the presence of short-sales constraints, unusually high liquidity is a symptom of the fact that the market is dominated by irrational investors, and hence is overvalued.

### *2.1. Theoretical framework*

We use the assumptions of inefficient markets and smart managers to develop a simple theoretical framework for the irrational investors approach. The framework has roots in

[1] For example, suppose a manager issues equity at \$50 per share. Now if those shares subsequently double, the manager might regret not delaying the issue, but he will surely not be fired, having presided over a rise in the stock price. In contrast, imagine a money manager sells (short) the same stock at \$50. This might lead to considerable losses, an outflow of funds, and, if the bet is large enough, perhaps the end of a career.

Fischer and Merton (1984), De Long et al. (1989), Morck, Shleifer, and Vishny (1990b), and Blanchard, Rhee, and Summers (1993), but our particular derivation borrows most from Stein (1996).

In the irrational investors approach, the manager balances three conflicting goals. The first is to maximize fundamental value. This means selecting and financing investment projects to increase the rationally risk-adjusted present value of future cash flows. To simplify the analysis, we do not explicitly model taxes, costs of financial distress, agency problems or asymmetric information. Instead, we specify fundamental value as

$$f(K, \cdot) - K,$$

where $f$ is increasing and concave in new investment $K$. To the extent that any of the usual market imperfections leads the Modigliani–Miller (1958) theorem to fail, financing may enter $f$ alongside investment.

The second goal is to maximize the current share price of the firm's securities. In perfect capital markets, the first two objectives are the same, since the definition of market efficiency is that prices equal fundamental values. But once one relaxes the assumption of investor rationality, this need not be true, and the second objective is distinct. In particular, the second goal is to "cater" to short-term investor demands via particular investment projects or otherwise packaging the firm and its securities in a way that maximizes appeal to investors. Through such catering activities, managers influence the temporary mispricing, which we represent by the function

$$\delta(\cdot),$$

where the arguments of $\delta$ depend on the nature of investor sentiment. The arguments might include investing in a particular technology, assuming a conglomerate or single-segment structure, changing the corporate name, managing earnings, initiating a dividend, and so on. In practice, the determinants of mispricing may well vary over time.

The third goal is to exploit the current mispricing for the benefit of existing, long-run investors. This is done by a "market timing" financing policy whereby managers supply securities that are temporarily overvalued and repurchase those that are undervalued. Such a policy transfers value from the new or the outgoing investors to the ongoing, long-run investors; the transfer is realized as prices correct in the long run.[2] For simplicity, we focus here on temporary mispricing in the equity markets, and so $\delta$ refers to the difference between the current price and the fundamental value of equity. More generally, each of the firm's securities may be mispriced to some degree. By selling a fraction of the firm $e$, long run shareholders gain

$$e\delta(\cdot).$$

[2] Of course, we are also using the market inefficiency assumption here in assuming that managerial efforts to capture a mispricing do not completely destroy it in the process, as they would in the rational expectations world of Myers and Majluf (1984). In other words, investors underreact to corporate decisions designed to exploit mispricing. This leads to some testable implications, as we discuss below.

We leave out the budget constraint, lumping together the sale of new and existing shares. Instead of explicitly modeling the flow of funds and any potential financial constraints, we will consider the reduced form impact of $e$ on fundamental value.

It is worth noting that other capital market imperfections can lead to a sort of catering behavior. For example, reputation models in the spirit of Holmstrom (1982) can lead to earnings management, inefficient investment, and excessive swings in corporate strategy even when the capital markets are not fooled in equilibrium.[3] Viewed in this light, the framework here is relaxing the assumptions of rational expectations in Holmstrom, in the case of catering, and Myers and Majluf (1984), in the case of market timing.

Putting the goals of fundamental value, catering, and market timing into one objective function, the irrational investors approach has the manager choosing investment and financing to

$$\max_{K,e} \lambda\big[f(K,\cdot) - K + e\delta(\cdot)\big] + (1-\lambda)\delta(\cdot),$$

where $\lambda$, between zero and one, specifies the manager's horizon. When $\lambda$ equals one, the manager cares only about creating value for existing, long-run shareholders, the last term drops out, and there is no distinct impact of catering. However, even an extreme long-horizon manager cares about short-term mispricing for the purposes of market timing, and thus may cater to short-term mispricing to further this objective. With a shorter horizon, maximizing the stock price becomes an objective in its own right, even without any concomitant equity issues.

We take the managerial horizon as given, exogenously set by personal characteristics, career concerns, and the compensation contract. If the manager plans to sell equity or exercise options in the near term, his portfolio considerations may lower $\lambda$. However, the managerial horizon may also be endogenous. For instance, consider a venture capitalist who recognizes a bubble. He might offer a startup manager a contract that loads heavily on options and short-term incentives, since he cares less about valuations that prevail beyond the IPO lock-up period. Career concerns and the market for corporate control can also combine to shorten horizons, since if the manager does not maximize short-run prices, the firm may be acquired and the manager fired.

Differentiating with respect to $K$ and $e$ gives the optimal investment and financial policy of a rational manager operating in inefficient capital markets:

$$f_K(K,\cdot) = 1 - \left(e + \frac{1-\lambda}{\lambda}\right)\delta_K(\cdot),$$

$$-f_e(K,\cdot) = \delta(\cdot) + \left(e + \frac{1-\lambda}{\lambda}\right)\delta_e(\cdot).$$

[3] For examples, see Stein (1989) and Scharfstein and Stein (1990). For a comparison of rational expectations and inefficient markets in this framework, see Aghion and Stein (2006).

In words, the first condition is about investment policy. The marginal value created from investment is weighed against the standard cost of capital, normalized to be one here, net of the impact that this incremental investment has on mispricing, and hence its effect through mispricing on catering and market timing gains. The second condition is about financing. The marginal value lost from shifting the firm's current capital structure toward equity is weighed against the direct market timing gains and the impact that this incremental equity issuance has on mispricing, and hence its effect on catering and market timing gains. This is a lot to swallow at once, so we consider some special cases.

*Investment policy*. Investment and financing are separable if both $\delta_K$ and $f_e$ are equal to zero. Then the investment decision reduces to the familiar perfect markets condition of $f_K$ equal to unity. Real consequences of mispricing for investment thus arise in two ways. In Stein (1996) and Baker, Stein, and Wurgler (2003), $f_e$ is not equal to zero. There *is* an optimal capital structure, or at least an upper bound on debt capacity. The benefits of issuing or repurchasing equity in response to mispricing are balanced against the reduction in fundamental value that arises from too much (or possibly too little) leverage. In Polk and Sapienza (2004) and Gilchrist, Himmelberg, and Huberman (2005), there is no optimal capital structure, but $\delta_K$ is not equal to zero: mispricing is itself a function of investment. Polk and Sapienza focus on catering effects and do not consider financing ($e$ equal to zero in this setup), while Gilchrist et al. model the market timing decisions of managers with long horizons ($\lambda$ equal to one).

*Financial policy*. The demand curve for a firm's equity slopes down under the natural assumption that $\delta_e$ is negative, e.g., issuing shares partly corrects mispricing.[4] When investment and financing are separable, managers act like monopolists. This is easiest to see when managers have long horizons, and they sell down the demand curve until marginal revenue $\delta$ is equal to marginal cost $-e\delta_e$. Note that price remains above fundamental value even after the issue: "corporate arbitrage" moves the market toward, but not all the way to, market efficiency.[5] Managers sell less equity when they care about short-run stock price ($\lambda$ less than one, here). For example, in Ljungqvist, Nanda, and Singh (2006), managers expect to sell their own shares soon after the IPO and so issue less as a result. Managers also sell less equity when there are costs of suboptimal leverage.

*Other corporate decisions*. Managers do more than simply invest and issue equity, and this framework can be expanded to accommodate other decisions. Consider dividend policy. Increasing or initiating a dividend may simultaneously affect both fundamental value, through taxes, and the degree of mispricing, if investors categorize stocks

[4] Gilchrist, Himmelberg, and Huberman (2005) model this explicitly with heterogeneous investor beliefs and short-sales constraints.

[5] Total market timing gains may be even higher in a dynamic model where managers can sell in small increments down the demand curve.

according to payout policy as they do in Baker and Wurgler (2004a). The tradeoff is

$$-f_d(K,\cdot) = \left(e + \frac{1-\lambda}{\lambda}\right)\delta_d(\cdot),$$

where the left-hand side is the tax cost of dividends, for example, and the right-hand side is the market timing gain, if the firm is simultaneously issuing equity, plus the catering gain, if the manager has short horizons. In principle, a similar tradeoff governs the earnings management decision or corporate name changes; however, in the latter case, the fundamental costs of catering would presumably be small.

### 2.2. Empirical challenges

The framework outlined above suggests a role for securities mispricing in investment, financing, and other corporate decisions. The main challenge for empirical tests in this area is measuring mispricing, which by its nature is hard to pin down. Researchers have found several ways to operationalize empirical tests, but none of them is perfect.

*Ex ante misvaluation*. One option is to take an *ex ante* measure of mispricing, for instance a scaled-price ratio in which a market value in the numerator is related to some measure of fundamental value in the denominator. Perhaps the most common choice is the market-to-book ratio: a high market-to-book suggests that the firm may be overvalued. Consistent with this idea, and the presumption that mispricing corrects in the long run, market-to-book is found to be inversely related to future stock returns in the cross-section by Fama and French (1992) and in the time-series by Kothari and Shanken (1997) and Pontiff and Schall (1998). Also, extreme values of market-to-book are connected to extreme investor expectations by Lakonishok, Shleifer, and Vishny (1994), La Porta (1996), and La Porta et al. (1997).

One difficulty that arises with this approach is that the market-to-book ratio or another *ex ante* measure of mispricing may be correlated with an array of firm characteristics. Book value is not a precise estimate of fundamental value, but rather a summary of past accounting performance. Thus, firms with excellent growth prospects tend to have high market-to-book ratios, and those with agency problems might have low ratios—and perhaps these considerations, rather than mispricing, drive investment and financing decisions. Dong et al. (2005) and Ang and Cheng (2006) discount analyst earnings forecasts to construct an arguably less problematic measure of fundamentals than book value.

Another factor that limits this approach is that a precise *ex ante* measure of mispricing would represent a profitable trading rule. There must be limits to arbitrage that prevent rational investors from fully exploiting such rules and trading away the information they contain about mispricing. But on a more positive note, the same intuition suggests that variables like market-to-book are likely to be a more reliable mispricing metric in regions of the data where short-sales constraints and other (measurable) arbitrage costs and risks are most severe. This observation has been exploited as an identification strategy.

*Ex post misvaluation*. A second option is to use the information in future returns. The idea is that if stock prices routinely decline after a corporate event, one might infer that they were inflated at the time of the event. However, as detailed in Fama (1998) and Mitchell and Stafford (2000), this approach is also subject to several critiques.

The most basic critique is the joint hypothesis problem: a predictable "abnormal" return might mean there was misvaluation *ex ante*, or simply that the definition of "normal" expected return (e.g., CAPM) is wrong. Perhaps the corporate event systematically coincides with changes in risk, and hence the return required in an efficient capital market. Another simple but important critique regards economic significance. Market value-weighting or focusing on NYSE/AMEX firms may reduce abnormal returns or cause them to disappear altogether.

There are also statistical issues. For instance, corporate events are often clustered in time and by industry—IPOs are an example considered in Brav (2000)—and thus abnormal returns may not be independent. Barber and Lyon (1997) and Lyon, Barber, and Tsai (1999) show that inference with buy-and-hold returns (for each event) is challenging. Calendar-time portfolios, which consist of an equal- or value-weighted average of all firms making a given decision, have fewer problems here, but the changing composition of these portfolios adds another complication to standard tests. Loughran and Ritter (2000) also argue that such an approach is a less powerful test of mispricing, since the clustered events have the worst subsequent performance. A final statistical problem is that many studies cover only a short sample period. Schultz (2003) shows that this can lead to a small sample bias if managers engage in "pseudo"-market timing, making decisions in response to past rather than future price changes.

Analyzing aggregate time series resolves some of these problems. Like the calendar time portfolios, time series returns are more independent. There are also established time-series techniques, e.g., Stambaugh (1999), to deal with small-sample biases. Nonetheless, the joint hypothesis problem remains, since rationally required returns may vary over time.

But even when these econometric issues can be solved, interpretational issues may remain. For instance, suppose investors have a tendency to overprice firms that have genuinely good growth opportunities. If so, even investment that is followed by low returns need not be *ex ante* inefficient. Investment may have been responding to omitted measures of investment opportunities, not to the misvaluation itself.

*Cross-sectional interactions*. Another identification strategy is to exploit the finer cross-sectional predictions of the theory. In this spirit, Baker, Stein, and Wurgler (2003) consider the prediction that if $f_e$ is positive, mispricing should be more relevant for financially constrained firms. More generally, managerial horizons or the fundamental costs of catering to sentiment may vary across firms in a measurable way. Of course, even in this approach, one still has to proxy for mispricing with an *ex ante* or *ex post* method. To the extent that the hypothesized cross-sectional pattern appears strongly in the data, however, objections about the measure of mispricing lose some steam.

### 2.3. Investment policy

Of paramount importance are the real consequences of market inefficiency. It is one thing to say that investor irrationality has an impact on capital market prices, or even financing policy, which lead to transfers of wealth among investors. It is another to say that mispricing leads to underinvestment, overinvestment, or the general misallocation of capital and deadweight losses for the economy as a whole. In this subsection we review research on how market inefficiency affects real investment, mergers and acquisitions, and diversification.

#### 2.3.1. Real investment

In the rational managers, irrational investors framework, mispricing influences real investment in two ways. First, investment may itself be a characteristic that is subject to mispricing ($\delta_K > 0$ above). Investors may overestimate the value of investment in particular technologies, for example. Second, a financially constrained firm ($f_e > 0$ above) may be forced to pass up fundamentally valuable investment opportunities if it is undervalued.

Most research has looked at the first type of effect. Of course, anecdotal evidence of this effect comes from bubble episodes; it was with the late 1920s bubble fresh in mind that Keynes (1936) argued that short-term investor sentiment is, at least in some eras, a major or dominant determinant of investment. More recent US stock market episodes generally viewed as bubbles include the electronics boom in 1959–1962, growth stocks in 1967–1968, the "nifty fifty" in the early 1970s, gambling stocks in 1977–1978, natural resources, high tech, and biotechnology stocks in the 1980s, and the Internet in the late 1990s; see Malkiel (1990) for an anecdotal review of some of these earlier bubbles, and Ofek and Richardson (2003) on the Internet. See Kindleberger (2000) for an attempt to draw general lessons from bubbles and crashes over several hundred years, and for anecdotal remarks on their sometimes dramatic real consequences.

The first modern empirical studies in this area asked whether investment is sensitive to stock prices over and above direct measures of the marginal product of capital, such as cash flow or profitability. If it is not, they reasoned, then the univariate link between investment and stock valuations likely just reflects the standard, efficient-markets $Q$ channel. This approach did not lead to a clear conclusion, however. For example, Barro (1990) argues for a strong independent effect of stock prices, while Morck, Shleifer, and Vishny (1990b) and Blanchard, Rhee, and Summers (1993) conclude that the incremental effect is weak.

The more recent wave of studies has taken a different tack. Rather than controlling for fundamentals and looking for a residual effect of stock prices, they try to proxy for the mispricing component of stock prices and examine whether it affects investment. In this spirit, Chirinko and Schaller (2001, 2006), Panageas (2004), Polk and Sapienza (2004), and Gilchrist, Himmelberg, and Huberman (2005) all find evidence that investment is sensitive to proxies for mispricing. Of course, the generic concern is that the mispricing proxies are still just picking up fundamentals. To refute this, Polk and Sapienza, for example, consider the finer prediction that investment should be more sensitive to short-term mispricing when managerial horizons are shorter. They find that investment is indeed more sensitive to mispricing proxies when share turnover is higher, i.e., where the average shareholder's horizon is shorter.

The second type of mispricing-driven investment is tested in Baker, Stein, and Wurgler (2003). Stein (1996) predicts that investment will be most sensitive to mispricing in equity-dependent firms, i.e., firms that have no option but to issue equity to finance their marginal investment, because long-horizon managers of undervalued firms would rather underinvest than issue undervalued shares. Using several proxies for equity dependence, Baker et al. confirm that investment is more sensitive to stock prices in equity-dependent firms.

Overall, the recent studies suggest that some portion of the effect of stock prices on investment is a response to mispricing, but key questions remain. The actual magnitude of the effect of mispricing has not been pinned down, even roughly. The efficiency implications are also unclear. Titman, Wei, and Xie (2004) and Polk and Sapienza (2004) find that high investment is associated with lower future stock returns in the cross section, and Lamont (2000) finds a similar result for planned investment in the time series. However, sentiment and fundamentals seem likely to be correlated, and so, as mentioned previously, even investment followed by low returns may not be *ex ante* inefficient. Finally, even granting an empirical link between overpricing and investment, it is hard to determine the extent to which managers are rationally fanning the flames of overvaluation, as in the catering piece of our simple theoretical framework, or are simply just as overoptimistic as their investors. We return to the effects of managerial optimism in the second part of the survey.

### *2.3.2. Mergers and acquisitions*

Shleifer and Vishny (2003) propose a market timing model of acquisitions. They assume that acquirers are overvalued, and the motive for acquisitions is not to gain synergies, but to preserve some of their temporary overvaluation for long-run shareholders. Specifically, by acquiring less-overvalued targets with overpriced stock (or, less interestingly, undervalued targets with cash), overvalued acquirers can cushion the fall for their shareholders by leaving them with more hard assets per share. Or, if the deal's value proposition caters to a perceived synergy that causes the combined entity to be overvalued, as might have happened in the late 1960s conglomerates wave (see below), then the acquirer can still gain a long-run cushion effect, while offering a larger premium to the target.

The market timing approach to mergers helps to unify a number of stylized facts. The defensive motive for the acquisition, and the idea that acquisitions are further facilitated when catering gains are available, help to explain the time-series link between merger volume and stock prices, e.g., Golbe and White (1988).[6] The model also predicts that cash acquirers earn positive long-run returns while stock acquirers earn negative long-run returns, consistent with the findings of Loughran and Vijh (1997) and Rau and Vermaelen (1998).

Recent papers have found further evidence for market timing mergers. Dong et al. (2005) and Ang and Cheng (2006) find that market-level mispricing proxies and merger volume are positively correlated, and (within this) that acquirers tend to be more overpriced than targets.[7] They also find evidence that offers for undervalued targets are more likely to be hostile, and that overpriced acquirers pay higher takeover premia. Rhodes-Kropf, Robinson, and Viswanathan (2005) also link valuation levels and merger activity. Bouwman, Fuller, and Nain (2006) find evidence suggestive of a short-term catering effect. In high-valuation periods, investors welcome acquisition announcements, yet the subsequent returns of mergers made in those periods are the worst. Baker, Foley, and Wurgler (2006) find that FDI outflows, which are often simply cross-border acquisitions, increase with the current aggregate market-to-book ratio of the acquirer's stock market and decrease with subsequent returns on that market. All of these patterns are consistent with overvaluation-driven merger activity.

An unresolved question in the Shleifer–Vishny framework is why managers would prefer a stock-for-stock merger to an equity issue if the market timing gains are similar. One explanation is that a merger more effectively hides the underlying market timing motive from investors. Baker, Coval, and Stein (2006) consider another mechanism that can also help explain a generic preference for equity issues via merger.[8] The first ingredient of the story is that the acquiring firm faces a downward sloping demand curve for its shares, as in Shleifer (1986) and Harris and Gurel (1986). The second ingredient is that some investors follow the path of least resistance, passively accepting the acquirer's shares as consideration even when they would not have actively participated in an equity issue. With these two assumptions, the price impact of a stock-financed merger can be much smaller than the price impact of an SEO. Empirically, inertia is a prominent feature in institutional and especially individual holdings data that is associated with smaller merger announcement effects.

[6] See Rhodes-Kropf and Viswanathan (2004) for a somewhat different misvaluation-based explanation of this link, and Jovanovic and Rousseau (2002) for an explanation based on technological change in efficient markets.

[7] A related prediction of the Shleifer–Vishny framework is that an overvalued acquirer creates value for long-term shareholders by acquiring a fairly valued or simply less overvalued target. Savor (2006) tests this proposition by comparing the returns of successful acquirers to those that fail for exogenous reasons, such as a regulatory intervention. Successful acquirers perform poorly, as in Loughran and Vijh (1997), but unsuccessful acquirers perform even worse.

[8] For example, in the case of S&P 100 firms over 1999–2001, Fama and French (2005) find that the amount of equity raised in mergers is roughly 40 times that raised in SEOs.

### 2.3.3. Diversification and focus

Standard explanations for entering unrelated lines of business include agency problems or synergies, e.g., internal capital markets and tax shields. Likewise, moves toward greater focus are often interpreted as a triumph of governance. While our main task is to survey the existing literature, the topics of diversification and focus have yet to be considered from a perspective where investors are less than fully rational. So, we take a short detour here. We ask whether the evidence at hand is consistent with the view that the late-1960s conglomerate wave, which led to conglomerates so complex they were still being divested or busted up decades later, was in part driven by efforts to cater to a temporary investor appetite for conglomerates.

Investor demand for conglomerates appears to have reached a peak in 1968. Ravenscraft and Scherer (1987, p. 40) find that the average return on 13 leading conglomerates was 385% from July 1965 to June 1968, while the S&P 425 gained only 34%. Diversifying acquisitions were being greeted with a positive announcement effect, while other acquisitions were penalized (Matsusaka, 1993). Klein (2001) finds a "diversification premium" of 36% from 1966–1968 in a sample of 36 conglomerates. Perhaps responding to these valuation incentives, conglomerate mergers accelerated in 1967 and peaked in 1968 (Ravenscraft and Scherer, 1987, pp. 24, 161, 218).

Conglomerate valuations started to fall in mid-1968. Between July 1968 and June 1970, the sample followed by Ravenscraft and Scherer lost 68%, three times more than the S&P 425. Announcement effects also suggest a switch in investor appetites: diversification announcements were greeted with a flat reaction in the mid- to late-1970s and a negative reaction by the 1980s (Morck, Shleifer, and Vishny, 1990a). Klein finds that the diversification premium turned into a discount of 1% in 1969–1971 and 17% by 1972–1974, and a discount seems to have remained through the 1980s (Lang and Stulz, 1994; Berger and Ofek, 1995). Again, possibly in response to this shift in catering incentives, unrelated segments began to be divested, starting a long trend toward focus (Porter, 1987; Kaplan and Weisbach, 1992).[9] Overall, while systematic evidence is lacking, the diversification and subsequent re-focus wave seems to fit the catering model well.

## 2.4. Financial policy

The simple theoretical framework suggests that long-horizon managers may reduce the overall cost of capital paid by their ongoing investors by issuing overpriced securities and repurchasing underpriced securities. Here, we survey the evidence on the extent to which market timing affects equity issues, repurchases, debt issues, cross-border issues, and capital structure.

[9] In a case study of the diversification and subsequent refocus of General Mills, Donaldson (1990) writes that the company spent some effort "to verify the dominant trends in investor perceptions of corporate efficiency, as seen in the company study of the impact of excessive diversification on the trend of price-earnings multiples in the 1970s" (p. 140).

### 2.4.1. Equity issues

Several lines of evidence suggest that overvaluation is a motive for equity issuance. Most simply, in the Graham and Harvey (2001) anonymous survey of CFOs of public corporations, two-thirds state that "the amount by which our stock is undervalued or overvalued was an important or very important consideration" in issuing equity (p. 216). Several other questions in the survey also ask about the role of stock prices. Overall, stock prices are viewed as more important than nine out of ten factors considered in the decision to issue common equity, and the most important of five factors in the decision to issue convertible debt.

Empirically, equity issuance is positively associated with plausible *ex ante* indicators of overvaluation. Pagano, Panetta, and Zingales (1998) examine the determinants of Italian private firms' decisions to undertake an IPO between 1982 and 1992, and find that the most important is the market-to-book ratio of seasoned firms in the same industry. Lerner (1994) finds that IPO volume in the biotech sector is highly correlated with biotech stock indexes. Loughran, Ritter, and Rydqvist (1994) find that aggregate IPO volume and stock market valuations are highly correlated in most major stock markets around the world. Similarly, Marsh (1982) examines the choice between (seasoned) equity and long-term debt by UK quoted firms between 1959 and 1974, and finds that recent stock price appreciation tilts firms toward equity issuance. In US data, Jung, Kim, and Stulz (1996) and Hovakimian, Opler, and Titman (2001) also find a strong relationship between stock prices and seasoned equity issuance.

Of course, there are many non-behavioral reasons why equity issuance and market valuations should be positively correlated. More specific evidence for equity market timing comes from the pattern that new issues earn low subsequent returns. In an early test, Stigler (1964) tried to measure the effectiveness of the S.E.C. by comparing the *ex post* returns of new equity issues (lumping together both initial and seasoned) from 1923–1928 with those from 1949–1955. If the S.E.C. improved the pool of issuers, he reasoned, then the returns to issuers in the latter period should be higher. But he found that issuers in both periods performed about equally poorly relative to a market index. Five years out, the average issuer in the pre-S.E.C. era lagged the market by 41%, while the average underperformance in the later period was 30%.

Other sample periods show similar results. Ritter (1991) examines a sample of IPOs, Speiss and Affleck-Graves (1995) examine SEOs, and Loughran and Ritter (1995) examine both. And, Ritter (2003) updates these and several other empirical studies of corporate financing activities. The last paper's sample includes 7,437 IPOs and 7,760 SEOs between 1970 and 1990. Five years out, the average IPO earns lower returns than a size-matched control firm by 30%, and the average SEO underperforms that benchmark by 29%. Gompers and Lerner (2003) fill in the gap between the samples of Stigler (1964) and Loughran and Ritter (1995). Their sample of 3,661 IPOs between 1935 and 1972 shows average five-year buy-and-hold returns that underperform the value-weighted market index by 21% to 35%.[10] Thus, a rough summary of non-overlapping samples is that, on average, US equity issues underperform the market somewhere in the ballpark of 20–40% over five years.

In a test that speaks closely to the question of opportunistic timing of new investors, Burch, Christie, and Nanda (2004) examine the subsequent performance of seasoned equity issued via rights offers, which are targeted to a firm's ongoing shareholders, and firm commitment offers, which are targeted to new shareholders. In their 1933 to 1949 sample, a period in which rights offers were more common, they find underperformance entirely concentrated in the latter group. This fits exactly with the framework sketched above, which emphasizes the opportunistic timing of *new* investors.

If equity issues cluster when the market as a whole is overvalued, the net gains to equity market timing may be even larger than the underperformance studies suggest. Baker and Wurgler (2000) examine whether equity issuance, relative to total equity and debt issuance, predicts aggregate market returns between 1927 and 1999. They find that when the equity share was in its top historical quartile, the average value-weighted market return over the next year was negative 6%, or 15% below the average market return. Henderson, Jegadeesh, and Weisbach (2006) find a similar relationship in several international markets over the period 1990 to 2001. In 12 out of the 13 markets they examine, average market returns are higher after a below-median equity share year than after an above-median equity share year.[11]

The equity market timing studies continue to be hotly debated. Some authors highlight the joint hypothesis problem, proposing that the reason why IPOs and SEOs deliver low returns is that they are actually less risky. For more on this perspective, see Eckbo, Masulis, and Norli (2000), Eckbo and Norli (2004), and Chapter 6 by Eckbo, Masulis and Norli in this volume. In a recent critique, Schultz (2003) points out that a small-sample bias he calls "pseudo market timing" can lead to exaggerated impressions of underperformance when abnormal performance is calculated in "event time". The empirical relevance of this bias has yet to be pinned down. Schultz (2003, 2004) argues that it may be significant, while Ang, Gu, and Hochberg (2005), Dahlquist and de Jong (2004), and Viswanathan and Wei (2004) argue that it is minor.[12] The key issue concerns

[10] Gompers and Lerner also confirm what Brav and Gompers (1997) found in a later sample: while IPOs have low absolute returns, and low returns relative to market indexes, they often do not do worse than stocks of similar size and book-to-market ratio. One interpretation is that securities with similar characteristics, whether or not they are IPOs, tend to be similarly priced (and mispriced) at a given point in time.

[11] Note that these aggregate predictability results should probably not be interpreted as evidence that "managers can time the aggregate market". A more plausible explanation is that broad waves of investor sentiment lead many firms to be mispriced in the same direction at the same time. Then, the *average* financing decision will contain information about the *average* (i.e., market-level) mispricing, even though individual managers are perceiving and responding only to their *own* firm's mispricing.

[12] Butler, Grullon, and Weston (2005) take Schultz's idea to the time-series and argue that the equity share's predictive power is due to an aggregate version of the pseudo market timing bias. Baker, Taliaferro, and Wurgler (2006) reply that the tests in Butler et al. actually have little relevance to the bias, and that simple simulation techniques show that small-sample bias can account for only one percent of the equity share's actual predictive coefficient.

the variance in the number of security issues over time. Schultz assumes a nonstationary process for this time series. This means that the number of security issues can explode or collapse to zero for prolonged periods of time, and the simulated variance of equity issuance exceeds the actual experience in the U.S.

We leave the resolution to future research, but we stress that the returns studies should not be considered in isolation. Survey evidence was mentioned above. Other relevant results include Teoh, Welch, and Wong (1998a, 1998b), who find that the equity issuers who manage earnings most aggressively have the worst post-issue returns (we return to earnings management below). Jain and Kini (1994), Mikkelson, Partch, and Shah (1997), and Pagano, Panetta, and Zingales (1998) find that profitability deteriorates rapidly following the initial offering, and Loughran and Ritter (1997) document a similar pattern with seasoned issues. Jenter (2005) finds that seasoned equity offerings coincide with insider selling. When viewed as a whole, the evidence indicates that market timing plays a nontrivial role in equity issues.

### *2.4.2. Repurchases*

Undervaluation is an important motive for repurchases. Brav et al. (2005) survey 384 CFOs regarding payout policy, and "the most popular response for all the repurchase questions on the entire survey is that firms repurchase when their stock is a good value, relative to its true value: 86.6% of all firms agree" (p. 26). Other work finds positive abnormal returns for firms that conduct repurchases, suggesting that managers are on average successful in timing them. Ikenberry, Lakonishok, and Vermaelen (1995) study 1,239 open market repurchases announced between 1980 and 1990. Over the next four years, the average repurchaser earned 12% more than firms of similar size and book-to-market ratios. Ikenberry, Lakonishok, and Vermaelen (2000) find similar results in a recent sample of Canadian firms.

The evidence shows that managers tend to issue equity before low returns, on average, and repurchase before higher returns. Is there a ballpark estimate of the reduction in the cost of equity, for the average firm, that these patterns imply? Without knowing just how the "rational" cost of equity varies over time, this question is hard to answer. However, suppose that rationally required returns are constant. By following aggregate capital inflows and outflows into corporate equities, and tracking the returns that follow these flows, Dichev (2004) reports that the average "dollar-weighted" return is lower than the average buy-and-hold return by 1.3% per year for the NYSE/Amex, 5.3% for Nasdaq, and 1.5% (on average) for 19 stock markets around the world. Put differently, if NYSE/Amex firms had issued and repurchased randomly across time, then, holding the time series of realized returns fixed, they would have paid 1.3% per year more for the equity capital they employed.

Of course, this reduction in the cost of equity capital is not evenly distributed in the cross section of firms. The difference between Nasdaq and NYSE/Amex gives a hint of this. For the many mature firms that rarely raise external equity, the gains may be negligible. For other firms that access the capital markets repeatedly through seasoned equity issues and stock-financed mergers, the gains may be much larger.

### *2.4.3. Debt issues*

A few papers have examined debt market timing, i.e., raising debt when its cost is unusually low. Survey evidence lends some initial plausibility to timing in this market as well. In particular, Graham and Harvey (2001) find that interest rates are the most cited factor in debt policy decisions: CFOs issue debt when they feel “rates are particularly low”. Expectations about the yield curve also appear to influence the *maturity* of new debt. Short-term debt is preferred “when short-term rates are low compared to long-term rates” and when “waiting for long-term market interest rates to decline”. Clearly, CFOs do not believe in the textbook version of the expectations hypothesis, under which the cost of debt is equal across maturities. At the same time, CFOs do not confess to exploiting their private information about credit quality, instead highlighting general debt market conditions.

On the empirical side, Marsh (1982), in his sample of UK firms, finds that the choice between debt and equity does appear to be swayed by the level of interest rates. And Guedes and Opler (1996) examine and largely confirm the survey responses regarding the effect of the yield curve. In a sample of 7,369 US debt issues between 1982 and 1993, they find that maturity is strongly negatively related to the term spread (the difference between long- and short-term bond yields), which was fluctuating considerably during this period.

Is debt market timing successful in any sense? In aggregate data, Baker, Greenwood, and Wurgler (2003) examine the effect of debt market conditions on the maturity of debt issues and, perhaps more interestingly, connect the maturity of new issues to subsequent bond market returns. Specifically, in US Flow of Funds data between 1953 and 2000, the aggregate share of long-term debt issues in total long- and short-term debt issues is negatively related to the term spread, just as Guedes and Opler find with firm-level data. Further, because the term spread is positively related to future excess bond returns—i.e., the difference in the returns of long-term and short-term bonds, or the realized relative cost of long- and short-term debt—so is the long-term share in debt issues. Perhaps simply by using a naïve rule of thumb, “issue short-term debt when short-term rates are low compared to long-term rates”, managers may have timed their debt maturity decisions so as to reduce their overall cost of debt. Of course, such a conclusion is subject to the usual risk-adjustment caveats.

Unfortunately, the data on individual debt issues and their subsequent returns does not approach the level of detail of the IPO and SEO data. But one intriguing pattern that has been uncovered is that debt issues are followed by low *equity* returns. Speiss and Affleck-Graves (1999) examine 392 straight debt issues and 400 convertible issues between 1975 and 1989. The shares of straight debt issuers underperform a size- and book-to-market benchmark by an insignificant 14% over five years (the median underperformance is significant), while convertible issuers underperform by a significant

37%. There is also a suggestion that the riskiest firms may be timing their idiosyncratic credit quality, despite the survey answers on this point: the shares of unrated issuers have a median five-year underperformance of 54%. If the equity did so poorly, the debt issues presumably also did poorly. In a much broader panel, Richardson and Sloan (2003) also find that net debt issuance is followed by low stock returns.

There are several potential explanations for this pattern. Certainly, equity overvaluation would be expected to lower the cost of debt directly—credit risk models routinely include stock market capitalization as an input—so the relationship with subsequent stock returns may reflect debt market timing *per se*. Or, managerial and investor sentiment is correlated; managers may tend to be most optimistic precisely when capital is cheap, and thus raise and invest as much as they can from any source. This story combines investor and managerial irrationality and so does not fit neatly within our taxonomy, but seems like a promising approach for future work. A third possibility, outlined in Baker, Stein, and Wurgler (2003), is that equity overvaluation relaxes a binding leverage constraint, creating debt capacity that subsequently gets used up. But debt is always correctly priced in this setting, so debt market timing *per se* is not possible.

### *2.4.4. Cross-border issues*

The evidence in Froot and Dabora (1999) suggests that relative mispricings across international securities markets are possible, even between particularly liquid markets such as the US and the UK. This raises the possibility of international market timing. Along these lines, Graham and Harvey (2001) find that among US CFOs who have considered raising debt abroad, 44% implicitly dismissed covered interest parity in replying that lower foreign interest rates were an important or very important consideration in their decision.[13]

In practice, most international stock and bond issues are made on the US and UK markets. Henderson, Jegadeesh, and Weisbach (2006) find that when total foreign issues in the US or the UK are high, relative to respective GDP, subsequent returns on those markets tend to be low, particularly in comparison to the returns on issuers' own markets. In a similar vein, and consistent with the survey evidence mentioned above, foreign firms tend to issue more debt in the US and the UK when rates there are low relative to domestic rates.

### *2.4.5. Capital structure*

As an accounting identity, every firm's capital structure is the cumulative outcome of a long series of incremental financing decisions, each driven by the need to fund some investment project, consummate a merger, or achieve some other purpose. To the extent that market timing is a determinant of any of these incremental financing decisions,

[13] Almost all equity raised by US corporations is placed in domestic markets, so Graham and Harvey do not ask about the determinants of international stock issues.

then, it may help to explain the cross-section of capital structure. In particular, if market timing-motivated financing decisions are not quickly rebalanced away, low-leverage firms will tend to be those that raised external finance when their stock prices were high, and hence those that tended to choose equity to finance past investments and mergers, and vice-versa for high leverage firms.[14]

This market timing theory of capital structure is developed and tested in Baker and Wurgler (2002). In an effort to capture the historical coincidence of market valuations and the demand for external finance in a single variable, they construct an "external finance weighted-average" of a firm's past market-to-book ratios. For example, a high value would mean that the firm raised the bulk of its external finance, equity *or* debt, when its market-to-book was high. If market timing has a persistent impact on capital structure, Baker and Wurgler argue, this variable will have a negative cross-sectional relationship to the debt-to-assets ratio, even in regressions that control for the current market-to-book ratio. In a broad Compustat sample from 1968 to 1999, a strong negative relationship is apparent.

This evidence has inspired debate. On one hand, Hovakimian (2006) argues that equity issues do not have persistent effects on capital structure, and that the explanatory power of the weighted average market-to-book arises because it contains information about growth opportunities, a likely determinant of target leverage, that is not captured in current market-to-book. Leary and Roberts (2005), Kayhan and Titman (2004), Flannery and Rangan (2006) also argue that firms rebalance toward a target. Alti (2005) looks specifically at the time series variation in IPO leverage, finding that an initial and statistically significant response to hot issues markets is short-lived.

On the other hand, Huang and Ritter (2005) show that the tendency to fund a financing deficit with equity decreases with proxies for the cost of equity capital. And, Welch (2004) and Huang and Ritter (2005), like Fama and French (2002), argue that firms rebalance their capital structures much more slowly, so that shocks to capital structure are long lived. Moreover, Chen and Zhao (2004b) point out that mean reversion in leverage is not definitive evidence for a tradeoff theory. Because leverage is a ratio, shocks tend to cause mean reversion mechanically. In an analysis of the choice between equity and debt issues, which avoids this problem, Chen and Zhao (2004a) find that deviation-from-target proxies have little explanatory power, while market-to-book and past stock returns are very important.

### *2.5. Other corporate decisions*

In this subsection, we consider what the irrational investors approach has to say about dividend policy, firm name changes, and earnings management.[15] We also discuss recent work that looks at executive compensation from this perspective.

[14] Similarly, one could articulate a simple theory of debt maturity structure as reflecting the historical coincidence of debt issuance and debt market conditions like the term spread.

[15] We put dividend policy in this section and repurchases in the financing section, because, unlike a repurchase, pro-rata dividends do not change the ownership structure of the firm, and there is no market timing

### *2.5.1. Dividends*

The catering idea has been applied to dividend policy. Long (1978) provides some early motivation for this application. He finds that shareholders of Citizens Utilities put different prices on its cash dividend share class than its stock dividend share class, even though the value of the shares' payouts are equal by charter. In addition, this relative price fluctuates. The unique experiment suggests that investors may view cash dividends *per se* as a salient characteristic, and in turn raises the possibility of a catering motive for paying them.

Baker and Wurgler (2004a) outline and test a catering theory of dividends in aggregate US data between 1963 and 2000. They find that firms initiate dividends when the shares of existing payers are trading at a premium to those of nonpayers, and dividends are omitted when payers are at a discount. To measure the relative price of payers and nonpayers, they use an *ex ante* measure of mispricing they call the "dividend premium". This is just the difference between the average market-to-book ratios of payers and nonpayers. They also use *ex post* returns, and find that when the rate of dividend initiation increases, the future stock returns of payers (as a portfolio) are lower than those of nonpayers. This is consistent with the idea that firms initiate dividends when existing payers are relatively overpriced. Li and Lie (2005) find similar results for dividend changes.

Time-varying catering incentives also appear to shed light on the "disappearance" of dividends. Fama and French (2001) document that the percentage of Compustat firms that pay dividends declines from 67% in 1978 to 21% in 1999, and that only a part of this is due to the compositional shift towards small, unprofitable, growth firms which are generally less likely to pay dividends. Baker and Wurgler (2004b) observe that the dividend premium switched sign from positive to negative in 1978 and has remained negative through 1999, suggesting that dividends may have been disappearing in part because of the consistently lower valuations put on payers over this period. An analysis of earlier 1963–1977 data also lends support to this idea. Dividends "appeared", "disappeared", and then "reappeared" in this period, and each shift roughly lines up with a flip in the sign of the dividend premium. In UK data, Ferris, Sen, and Yui (2006) find that dividends have been disappearing during the late 1990s, and that a dividend premium variable formed using UK stocks lines up with this pattern.

The evidence suggests that the dividend supply responds to catering incentives, but why does investor demand for payers vary over time? One possibility is that "dividend clienteles" vary over time, for example with tax code changes. However, in US data, the dividend premium is unrelated to the tax disadvantage of dividend income, as is the rate of dividend initiation. Shefrin and Statman (1984) develop explanations for why investors prefer dividends based on self-control problems, prospect theory, mental accounting, and regret aversion. Perhaps these elements vary over time. Baker and Wurgler (2004a) argue that the dividend premium reflects sentiment for "risky" nonpaying

benefit or cost. For this reason, it fits more naturally with the category of corporate decisions that might influence the level of mispricing, but do not otherwise transfer value among investors.

growth firms versus "safe" dividend payers, since it falls in growth stock bubbles and rises in crashes. Fuller and Goldstein (2003) show more explicitly that payers outperform in market downturns. Perhaps investors seek the perceived safety of cash dividends in these gloomy periods, and bid up the shares of payers.

There are clear limitations to a catering theory of dividends, however. For one, it is a descriptive theory of whether firms pay dividends at all, not how much—in US data, at least, the dividend premium does not explain aggregate fluctuations in the level of dividends. DeAngelo, DeAngelo, and Skinner (2004) report that the aggregate dollar value of dividends has increased in real terms, as dividends have become concentrated in a smaller faction of traded firms. Also, it works better for explaining initiations than omissions, and it has little to say about the strong persistence in dividend policy. Catering is probably best viewed as one building block in an overall descriptive theory of dividend policy.

### *2.5.2. Firm names*

Name changes provide some of the simplest and most colorful examples of catering. In frictionless and efficient markets, firm names should be about as irrelevant as dividends. But there is a low fundamental cost of changing names, and perhaps through a name change a firm can create a salient association with an overpriced category of stocks.

Evidence of a catering motive for corporate names is most prominent in bubbles. In the 1959–1962 era which Malkiel (1990) refers to as the "tronics boom", firms "often included some garbled version of the word 'electronics' in their title even if the companies had nothing to do with the electronics industry" (p. 54). Systematic evidence has been assembled for the Internet bubble. Cooper, Dimitrov, and Rau (2001) find that 147 (generally small) firms changed to "dotcom" names between June 1998 and July 1999, as Internet valuations were rapidly rising. Catering to Internet sentiment did seem to deliver a short-term price boost: the authors report an average announcement effect of 74% for their main sample, and an even larger effect for the subset that had little true involvement with the Internet. Interestingly, Cooper et al. (2005) find that names were also used to *dissociate* companies from the Internet sector, as prices started crashing. Between August 2000 and September 2001, firms that dropped their dotcom name saw a positive announcement effect of around 70%. The effect was almost as large for firms that dropped the dotcom name but kept an Internet business focus, and for the "double dippers" which dropped the name they had newly adopted just a few years earlier.

The names of mutual funds also seem to be sensitive to investor sentiment. Cooper, Gulen, and Rau (2005) find that fund names shift away from styles that experience low returns and toward those with high returns. The authors find that name changes do not predict fund performance, yet inflows increase dramatically, even for "cosmetic" name changers whose underlying investment style remains constant. Presumably, then, the name change decision is driven in part by the desire to attract fund inflows, which increase the fund's size and the fees its managers earn. Indeed, Cooper et al. find that the inflow effect is increased when money is spent to advertise the "new" styles. While we group this study with other name changes, it actually involves an investment policy decision, in the sense that the goal of the name change is to increase the fundamental value of the franchise.

### 2.5.3. Earnings management

The quarterly net income figure that managers report to shareholders does not equal actual economic cash flows, but instead includes various non-cash accruals, some of which are fairly discretionary. According to the survey by Graham, Harvey, and Rajgopal (2005), CFOs believe that investors care more about earnings per share than cash flows.[16]

As the irrational investors theory predicts, managers with "short horizons" are especially likely to manage earnings. Bergstresser and Philippon (2006) find that accruals management increases as the CEO's compensation, via stock and options holdings, becomes more sensitive to current share prices. Sloan (1996) finds that firms with high accruals earn low subsequent returns, which suggests that earnings management may be successful in boosting share price, or at least in maintaining overvaluation. Consistent with the view that managers use earnings management to fool investors and issue overvalued equity, Teoh, Welch, and Wong (1998a, 1998b) find that initial and seasoned equity issuer underperformance is greatest for firms that most aggressively manage pre-issue earnings.

An interesting and largely unexplored question is whether earnings management has serious consequences for investment. Graham, Harvey, and Rajgopal (2005) present CFOs with hypothetical scenarios and find that 41% of them would be willing to pass up a positive-NPV project just to meet the analyst consensus EPS estimate. Direct evidence of this type of value loss is difficult to document, but Jensen (2005) presents a range of anecdotes, and highly suggestive empirical studies include Teoh et al. (1998a, 1998b), Erickson and Wang (1999), Bergstresser, Desai, and Rauh (2006), and Pshisva and Suarez (2004). The last three papers report that earnings management activity increases prior to stock acquisitions.

### 2.5.4. Executive compensation

In the theoretical framework at the beginning of this section, we assumed that managers may have the incentive to cater to short-term mispricing. One question is why shareholders do not set up executive compensation contracts to force managers to take the long view.[17] Bolton, Scheinkman, and Xiong (2005) suggest that short horizons may be an equilibrium outcome. They study the optimal incentive compensation contract for the

[16] There is a large literature in financial accounting on corporate earnings management. Here, we offer a brief and incomplete review, focusing on the link between earnings management and corporate financing decisions.
[17] A separate but related question is how managers compensate lower level employees within the firm. Bergman and Jenter (2006) argue that rational managers may minimize costs by paying optimistic employees

dynamic speculative market of Scheinkman and Xiong (2003), in which two groups of overconfident investors trade shares back and forth as their relative optimism fluctuates. The share price in this market contains a speculative option component, reflecting the possibility that nonholders might suddenly become willing to buy at a high price. Bolton et al. find that the optimal contract may induce the CEO to take costly actions that exacerbate differences of opinion, thus increasing the value of the option component of stock prices, at the expense of long-run value.

## 3. The irrational managers approach

The second approach to behavioral corporate finance takes the opposite extreme, in which irrational managers operate in efficient capital markets. To be more precise, by irrational managerial behavior we mean behavior that departs from rational expectations and expected utility maximization of the manager. We are not interested in rational moral hazard behavior, such as empire building, stealing, and plain slacking off. Instead, we are concerned with situations where the manager believes that he is actually close to maximizing firm value—and, in the process, some compensation scheme—but is in fact deviating from this ideal.[18]

As in the irrational investors approach, an extra building block is required. In order for less-than-fully-rational managers to have an impact, corporate governance must be limited in its ability to constrain them into making rational decisions. In general, an assumption of limited governance seems like a reasonable one to maintain. Takeover battles and proxy fights are notoriously blunt tools. Boards may be more a part of the problem than the solution if they have their own biases or are pawns of management. And unlike in a traditional agency problem, which arises when there is a conflict of interest between managers and outside investors, standard incentive contracts have little effect: an irrational manager may well think that he is maximizing value. Finally, in the US, a significant element of managerial discretion is codified in the business judgment rule. See Adams, Almeida, and Ferreira (2005) and Bertrand and Schoar (2003) for direct evidence that managers have discretion, and Shleifer and Vishny (1997) for a broader review of corporate governance institutions.

The psychology and economics literatures relevant to managerial behavior are vast. For us, the main themes are that individuals do not always form beliefs logically, nor do these beliefs convert to decisions in a consistent and rational manner—see Gilovich, Griffin, and Kahneman (2002) and Kahneman and Tversky (2000) for collected works. Thus far, most research in corporate finance has focused on the positive illusions of optimism and overconfidence. Illustrating the pattern of optimism, Weinstein (1980) finds

in overvalued equity, in the form of options grants. Benartzi (2001) offers a foundation for this sort of optimism, showing that employees have a tendency to extrapolate past returns, and as a consequence hold too much company stock. See also Core and Guay (2001) and Oyer and Schaeffer (2005).

[18] Our focus is on corporate finance decisions. Camerer and Malmendier (2005) discuss the impact of less than fully rational behavior in other parts of organizations.

that subjects tend to believe themselves to be more likely than average to experience positive future life events (e.g., owning own home, living past 80) and less likely to experience negative events (being fired, getting cancer). Illustrating overconfidence in one's own skills, Svenson (1981) finds that 82% of a sample of students placed themselves in the top 30% in terms of driving safety.

There are good reasons to focus on these particular biases in a managerial setting. First, they are strong and robust, having been documented in many samples, in particular samples of managers (Larwood and Whittaker, 1977; March and Shapira, 1987; Ben-David, 2004). Second, they are often fairly easy to integrate into existing models, in that optimism can be modeled as an overestimate of a mean and overconfidence as an underestimate of a variance. Third, overconfidence leads naturally to more risk-taking. Even if there is no overconfidence on average in the population of *potential* managers, those that are overconfident are more likely to perform extremely well (and extremely badly), placing them disproportionately in the ranks of upper (and former) management. And fourth, even if managers start out without bias, an attribution bias—the tendency to take greater responsibility for success than failure (e.g., Langer and Roth, 1975)—may lead successful managers to *become* overconfident, as in Gervais and Odean (2001).

After reviewing the theory and evidence on optimism and overconfidence, we turn briefly to potential applications of bounded rationality and reference-point preferences. Given the state of the literature, our treatment there is necessarily more speculative. Further, we do not discuss at all the impact of several other judgmental biases, such as representativeness, availability, anchoring, and narrow framing—not because we believe them to be unimportant, but because no systematic studies of their impacts on corporate finance decisions have yet been conducted.

### *3.1. Theoretical framework*

The idea of managerial optimism and overconfidence in finance dates at least to Roll (1986). The derivation below is in the spirit of Heaton (2002) and Malmendier and Tate (2005), as modified to match our earlier notation as much as possible. We start by assuming the manager is optimistic about the value of the firm's assets and investment opportunities. He then balances two conflicting goals. The first is to maximize *perceived* fundamental value. To capture this, we augment fundamental value with an optimism parameter $\gamma$,

$$(1+\gamma)f(K,\cdot)-K,$$

where $f$ is increasing and concave in new investment $K$. Note that here, the manager is optimistic about both the assets in place ($f$ can include a constant term) and new opportunities. Once again, if traditional market imperfections cause the Modigliani and Miller (1958) theorem to fail, financing may enter $f$ alongside investment.

The manager's second concern is to minimize the *perceived* cost of capital. We assume here that the manager acts on behalf of existing investors, because of his own stake in the firm and fiduciary duty. This leads to a similar setup to the market timing objective in Section 2.1, except that an optimistic manager believes there is never a good time to issue equity. In particular, since the capital market is efficient and values the firm at its true fundamental value of $f - K$, the manager believes that the firm is undervalued by $\gamma f$, and thus in selling a fraction of the firm $e$ he perceives that existing, long-run shareholders will lose

$$e\gamma f(K, \cdot).$$

Putting the two concerns together, the optimistic manager chooses new investment and financing to solve

$$\max_{K,e}(1+\gamma)f(K,\cdot) - K - e\gamma f(K,\cdot).$$

We do not explicitly include a budget constraint. Instead, again to keep the notation simple, we consider its reduced-form impact on $f$.

Differentiating with respect to $K$ and $e$ gives the optimal investment and financial policy of an optimistic manager operating in efficient capital markets:

$$f_K(K,\cdot) = \frac{1}{1+(1-e)\gamma}, \quad \text{and}$$
$$(1+\gamma)f_e(K,\cdot) = \gamma\big(f(K,\cdot) + ef_e(K,\cdot)\big).$$

Put into words, the first condition is about investment policy. Instead of setting the marginal value created from investment equal to the true cost of capital, normalized to be one here, managers overinvest, to the point where the marginal value creation is less than one. The more optimistic ($\gamma$) is the manager and the less equity ($e$) he is forced to raise in financing investment, the greater the problem. The second is about financing. The marginal value lost from shifting the firm's current capital structure away from equity is weighed against the perceived market timing losses. As in the analysis of irrational investors, we consider some special cases.

*Investment policy*. If there is no optimal capital structure, so that $f_e$ is equal to zero, the manager will not issue equity, setting $e$ to zero, and there is no interaction among financing, internal funds, and investment. In this case, the optimistic manager will clearly overinvest: $f_K$ is less than unity. In Heaton (2002) and Malmendier and Tate (2005), there is an optimal capital structure, or more precisely an upper bound on debt. If the manager needs equity to invest ($f_e$ greater than zero, here), the degree of overinvestment falls.

Needing equity is akin to having little cash or cash flow available for investment. Thus in this setup, investment can be strongly related to current cash flow and profits, controlling for investment opportunities. This leads to a behavioral foundation for the Jensen (1986) agency costs of free cash flow. But instead of receiving private benefits of control, managers are simply overconfident and overinvest from current resources as a result. Leverage reduces the degree of overinvestment by increasing $f_e$, thereby increasing equity issues $e$ and reducing $K$.

In a more complex specification, these conclusions may change. One might have the manager optimistic only about assets in place, in which case there is no overinvestment, and there will typically be underinvestment as a firm approaches its debt capacity. Also, it is worth emphasizing that we are examining optimism in isolation here. Layering on other imperfections, such as risk aversion, may mean that optimism moves investment from an inefficiently *low* level toward the first best, as in Gervais, Heaton, and Odean (2003) and Goel and Thakor (2002). In a related vein, Hackbarth (2004) argues that managerial optimism and overconfidence can reduce the underinvestment associated with debt overhang, as in Myers (1977).

*Financial policy*. An optimistic manager never sells equity unless he has to. If there is an upper bound on leverage ($f_e$ greater than zero, here), optimism predicts a 'pecking order' of financing decisions: the manager relies on internal capital and debt and uses outside equity only as a last resort. Again, other imperfections may mitigate the aversion to equity. If the manager is risk averse with an undiversified position in the firm's equity, for example, he may wish to issue equity even though it is below what he thinks it to be worth.

*Other corporate decisions*. It is not as easy to incorporate other decisions into this framework. Consider dividend policy. If the manager is more optimistic about future cash flow and assets in place than outside investors, he might view a dividend payment as more sustainable. On the other hand, if he views future investment opportunities, and hence funding requirements, as greater, he might be reluctant to initiate or increase dividends and retain internal funds instead. This analysis requires a more dynamic model of investment and cash flow and a decomposition of firm value into assets in place and growth opportunities.

### *3.2. Empirical challenges*

If the main obstacle to testing the irrational investors approach is finding a proxy for misvaluation, the challenge here is to identify optimism, overconfidence, or the behavioral bias of interest. Without an empirical measure, the irrational managers approach is difficult to distinguish from traditional agency theory, in particular. That is, in Stein (2003), an empire-building manager will

$$\max_{K,e} (1+\gamma) f(K) - K - c(e),$$

where $\gamma$ reflects the preference for or the private benefits that come with presiding over a larger firm, as in Jensen and Meckling (1976) or Grossman and Hart (1988), rather than optimism. Rational investors recognize the agency problem up front, so $c$ reflects the cost of raising outside equity, and management and existing shareholders bear the agency costs.

This reduced form is almost identical to the objective function of an optimistic manager. Both can generate overinvestment, underinvestment, cash flow-investment sensitivities, pecking order financing, and so forth. Moreover, Stein points out that the agency model is itself hard to distinguish from models of costly external finance built on asymmetric information. Thus, to test the behavioral theories, one must separate the $\gamma$ related to overconfidence and optimism from the $\gamma$ that arises from agency or asymmetric information problems.

### *3.3. Investment policy*

Despite the obvious difficulty of obtaining direct, manager-level measures of optimism and overconfidence, evidence is accumulating that these biases do affect business investment.

#### *3.3.1. Real investment*

We begin with startup investments. The evidence indicates that entrepreneurial startups are generally made under a halo of overconfidence and optimism. Cooper, Woo, and Dunkelberg (1988) find that 68% of entrepreneurs think that their startup is more likely to succeed than comparable enterprises, while only 5% believe that their odds are worse, and a third of entrepreneurs view their success as essentially guaranteed. The survey responses of French entrepreneurs tabulated in Landier and Thesmar (2005) also seem consistent with an initial underestimation of the task of starting a firm: at startup, 56% expect "development" in the near future, and 6% expect "difficulties".

The actual performance of startup investments is more sobering. Landier and Thesmar find that when surveyed three years into their endeavor, only 38% of French entrepreneurs expect further "development" while 17% anticipate "difficulty". Leaving profitability aside entirely, only half of all startups survive more than three years (Scarpetta et al., 2002). Moskowitz and Vissing-Jorgensen (2002) argue more generally that the return on private equity in the US between 1952 and 1999 is lower than seems justified given the undiversified nature of entrepreneurial investment. As a whole, the evidence on startup investments seems consistent with the overconfidence that Camerer and Lovallo's (1999) experimental subjects display when making entry decisions.

Optimism also appears to influence investment in more mature firms. Merrow, Phillips, and Myers (1981) compare forecast and actual construction costs for pioneer process plants in the energy industry. There is a strong optimism bias in project cost forecasts, with actual costs typically more than double the initial estimates. Statman and Tyebjee (1985) survey several other studies of this sort, involving military hardware, drugs, chemicals, and other development projects, and conclude that optimistic biases in cost and sales forecasts are fairly widespread.

Malmendier and Tate (2005) provide cross-sectional tests of the effects of optimism in a broader sample. They form a clever manager-level proxy for optimism: the propensity for a manager to voluntarily hold in-the-money stock options in his own firm. The intuition is that since the CEO's human capital is already so exposed to firm-specific risk, voluntarily holding in-the-money options can be seen as a strong vote of optimism.[19] With this optimism proxy in hand for a large sample of US firms between 1980 and 1994, Malmendier and Tate find that the sensitivity of investment to cash flow is higher for the more optimistic CEOs. This sensitivity is especially high for optimistic CEOs in equity-dependent firms, that is, in situations where perceived financial constraints are most binding. Their results support the predictions of the basic optimism model.

While the empirical evidence that optimism affects investment may not seem extensive, keep in mind that optimism, as discussed earlier, shares many predictions with more established theories, and thus is a candidate to explain various earlier results. For example, the fact that managers invest rather than pay out cash windfalls (Blanchard, Lopez-de-Silanes, and Shleifer, 1994) looks like a moral hazard problem, but is also consistent with optimism. Likewise, some investment patterns that look like adverse-selection-driven costly external finance may actually reflect a mistaken managerial belief that external finance is costlier. A possible example is the higher investment-cash flow sensitivities among younger and entrepreneurial firms (Schaller, 1993), which as noted above appear to be run by especial optimists.

### *3.3.2. Mergers and acquisitions*

Roll (1986) pioneered the optimism and overconfidence approach to corporate finance with his "hubris" theory of acquisitions. He suggests that successful acquirers may be optimistic and overconfident in their own valuation of deal synergies, and fail to properly account for the winner's curse. Roll interprets the evidence on merger announcement effects, surveyed by Jensen and Ruback (1983) and more recently by Andrade, Mitchell, and Stafford (2001) and Moeller, Schlingemann, and Stulz (2005), as well as the lack of evidence of fundamental value creation through mergers, as consistent with this theory.

More recently, Malmendier and Tate (2006) develop this argument and use their proxy for CEO optimism, outlined above, to test it. They find a number of patterns consistent with the optimism and overconfidence theory. First, optimistic CEOs complete more mergers, especially diversifying mergers, which are perhaps of more dubious value. Second, optimism has its biggest effect among the least equity dependent firms, i.e., when managers do not have to weigh the merger against an equity issue that they, as optimists, would perceive as undervalued. Third, investors are more skeptical about bid announcements when they are made by optimistic CEOs. This last result is consistent with the theme of irrational managers operating in efficient markets.[20]

[19] Malmendier and Tate find that the propensity to voluntarily retain in-the-money options is not significantly related to future abnormal stock returns, supporting their assumption that such behavior indeed reflects optimism rather than genuine inside information.

[20] For additional, anecdotal evidence on the role of hubris in takeovers, see Hietala, Kaplan, and Robinson (2003) and Shefrin (2000, Chapter 16).

### *3.4. Financial policy*

Direct empirical tests of how optimism and overconfidence affects financing patterns is not extensive. Existing work addresses capital structure and financial contracting.

#### *3.4.1. Capital structure*

The basic optimism model predicts a pecking order financing policy, as pointed out by Heaton (2002). Thus, much of the existing evidence of pecking-order policies, from Donaldson (1961) to Fama and French (2002), is at face value equally consistent with pervasive managerial optimism. And the notion of pervasive managerial optimism does not seem farfetched. In Graham's (1999) survey, almost two-thirds of CFOs state their stock is undervalued while only three percent state it is overvalued. Such responses are all the more striking given the fact that the survey was taken shortly *before* the Internet crash.

To distinguish optimism from other explanations of pecking order behavior (for example, adverse selection as in Myers and Majluf, 1984), a natural test would use cross-sectional variation in measured optimism to see whether such behavior is more prevalent in firms run by optimists. To our knowledge, exactly this test has yet to be conducted, but certain results in Malmendier and Tate (2005, 2006) have a closely related flavor. First, and as noted above, firms run by optimists (as identified by the Malmendier and Tate options-based proxies for optimism) display a higher sensitivity of investment to internal cash flow. Second, managers classified as optimistic show a differentially higher propensity to make acquisitions when they are not dependent on external equity.

#### *3.4.2. Financial contracting*

Landier and Thesmar (2005) examine financial contracting between rational investors and optimistic entrepreneurs.[21] They highlight two aspects of contracting with optimists. First, because optimists tend to inefficiently persist in their initial business plan, the optimal contract transfers control when changes are necessary. (Kaplan and Stromberg, 2003, find that contingent transfers of control are common features of venture capital contracts.) Second, because optimists believe good states to be more likely, they are willing to trade some control and ownership rights in bad states for greater claims in good ones; in this sense, the optimal contract "pays the entrepreneur with dreams". Ultimately, optimists may self-select into short-term debt, as it transfers payments and control to the investor in states that seem unlikely to occur, while realistic entrepreneurs prefer less risky long-term debt.

Landier and Thesmar find some empirical evidence of this separation in a data set of French entrepreneurs. Among other results, they find that the use of short-term debt

[21] Manove and Padilla (1999) also consider how banks separate optimists and realists. They focus on the overall efficiency of the credit market.

is positively related to an *ex post* measure of optimistic expectations, the difference between realized growth and initial growth expectations. They also find that the use of short-term debt is positively related to psychologically-motivated instruments for expectations, such as regional sunlight exposure and rates of mental depression.

### *3.5. Other behavioral patterns*

In the remainder of the survey, we briefly explore patterns other than optimism and overconfidence, in particular bounded rationality and reference-point preferences.

#### *3.5.1. Bounded rationality*

Perhaps the simplest deviation from the benchmark of full rationality is bounded rationality, introduced by Simon (1955). Bounded rationality assumes that some type of cognitive or information-gathering cost prevents agents from making fully optimal decisions. Boundedly-rational managers cope with complexity by using rules of thumb that ensure an acceptable level of performance and, hopefully, avoid severe bias. Conlisk (1996) reviews the bounded rationality literature.

Rules of thumb are hardly uncommon in financial management. For example, the net present value criterion is the optimal capital budgeting rule (in efficient markets), yet in practice managers employ various simpler rules. Surveying practice in the 1970s, Gitman and Forrester (1977) find that less than 10% of 103 large firms use NPV as their primary technique, while over 50% use the IRR rule, which avoids a cost of capital calculation. The Graham and Harvey (2001) survey of CFOs also finds that the IRR rule is more widely used than NPV, and over 50% of CFOs use the payback period rule, an even less sophisticated rule that requires neither a cost of capital input nor forecasts of cash flows beyond a cutoff date. Graham and Harvey also find that among managers who do use a discounting procedure, it is common to apply a firm-wide discount rate rather than a project-specific rate, again in stark contrast to normative principles.[22]

Other instances of rule-based management include the use of simple targets for capital structures and payouts. Graham and Harvey (2001) find that 10% of the CFOs in their sample use a "very strict" target debt–equity ratio and 34% use a "somewhat tight" target or range. Such leverage targets are typically defined in terms of book value, and Welch (2004) confirms that market leverage is, to a large extent, allowed to float with stock prices. Likewise, the Lintner (1956) field interviews revealed a set of common rules of thumb in payout policy that led him to an empirically accurate specification for dividends.

[22] A good question is whether the use of such rules is better understood as an agency problem than as bounded rationality. That is, executives might use simple rules to shorten the workday and save time for golf. However, Graham and Harvey find that high-ownership managers are if anything *less* likely to use NPV and *more* likely to use the payback period rule.

### 3.5.2. Reference-point preferences

Psychological experiments and intuition suggest that people value *changes* in economic states, such as wealth or performance, not just levels. This is reflected in the value function in Kahneman and Tversky's (1979) prospect theory, which is defined in terms of gains and losses relative to a reference point.

In corporate finance, the most developed application of reference-point preferences has been to IPO underpricing, the pattern that the initial offering price is, on average, below the market price that prevails after a day of trading. (For more on this, see the chapter by Ljungqvist in this volume.) There are, of course, many non-behavioral explanations for this pattern. Loughran and Ritter (2002) develop an explanation that combines reference-point preferences and mental accounting (Thaler, 1980, 1985). They assume that issuing managers mentally account for two quantities in judging an offering's success: the (perceived) gain from the gap between the first day closing price and a natural reference point, the midpoint of the file price range; and the (real) loss from the dilutive effect of the underpricing. If the gain is judged to outweigh the loss, where each is evaluated with the prospect theory value function, the executives are net satisfied. Intuitively, they may be too overwhelmed by the "windfall" gain versus the reference point to complain much about underpricing.[23]

This setup is designed, in part, to explain the pattern that underpricing is greater when the offer price is above the initial file price range. Loughran and Ritter (2002) find that in issues where the offer price is below the minimum of the file price range, first-day returns are a relatively small 4%, on average, while those priced above the maximum have average first-day returns of 32%. This is consistent with issuers acquiescing in severe underpricing only when they are simultaneously getting good news in the form of upward revisions from the filing range.[24] Ljungqvist and Wilhelm (2005) test some of the behavioral underpinnings of the Loughran and Ritter view. Using data on the ownership stakes of executives in IPO firms, they crudely proxy for the proposed notion of issuer satisfaction by taking the dollar amount of executives' perceived "gain" from revisions from the midpoint of the file price range and subtracting the dollar amount of dilution due to underpricing. They find that executive teams that are more "satisfied" with their IPOs by this criterion are more likely to use the same underwriter for seasoned offerings, and to pay higher fees for those transactions.

A different application of reference-point thinking is the widely asserted, but less well documented, managerial propensity to "throw good money after bad". Such behavior is most relevant for us to the extent that it reflects something more than rational

[23] Loughran and Ritter assume that the underwriter prefers underpricing, perhaps because it generates profitable rent-seeking activities among investors, e.g., trading with the underwriter's brokerage arm, or because it reduces marketing costs.

[24] See Benveniste and Spindt (1989) for an alternative explanation for this asymmetry based on information gathering in the book-building process; and Edelen and Kadlec (2005) for an alternative explanation, based on sample truncation bias related to the withdrawl of IPOs whose prospects deteriorate during the waiting period.

career concerns, e.g., a situation where the manager tries to distort the updating process to maintain high compensation. Shefrin (2001) offers several anecdotes concerning major corporate investments that have the flavor of good money after bad, and Statman and Sepe (1989) find that the market reaction to the termination of historically unprofitable investment projects is positive, suggesting that investors recognize that executives have a tendency to continue poor projects. Related evidence comes from the Guedj and Scharfstein (2004) study of drug development decisions. Those authors find that single-product early stage firms appear highly reluctant to abandon their only viable drug candidates, even when the results of clinical trials are less than promising. Some combination of agency, managerial optimism, and a gambling-to-get-back-to-even attitude seems like a plausible explanation for these results.

## 4. Conclusion

The behavioral corporate finance literature has matured to the point where one can now sketch out a handful of canonical theoretical frameworks and use them to organize the accumulated evidence of dozens of empirical studies. This survey suggests that the behavioral approaches to corporate finance offer a useful complement to the other paradigms in the field. They deliver intuitive and sometimes quite compelling explanations for important financing and investing patterns, including some that are difficult to reconcile with existing theory.

In its current state of flux, the field offers a number of exciting research questions. We close by highlighting just a few. In no particular order, we wonder:

- Are behavioral factors at the root of why managers do not more aggressively pursue the tax benefits of debt, as in Graham (2000)? Hackbarth (2004) develops a theoretical argument along these lines.
- While the existing literature has generally considered the two approaches separately, the irrational manager and irrational investor stories can certainly coexist. Would a model featuring a correlation between investor and managerial sentiment, for example, lead to new insights?
- What are the determinants of managerial "horizons", and how can they be measured and appropriately governed? Polk and Sapienza (2004) and Gaspar, Massa, and Matos (2005) use share turnover by investors to proxy for shareholder horizons.
- To what extent should the venture capital industry be viewed as an institution that identifies and caters to emerging pockets of investor sentiment?
- What determines investor sentiment, and how is it managed through corporate investor relations? Potential avenues to consider are interactions with past stock market returns, technological change and the valuation of new industries, media coverage, financial analysts and financial reporting, or investment banking. Brennan and Tamarowski (2000) offer an overview of investor relations.
- Do equity and debt market timing reduce the overall cost of capital by a large amount, or just a little? Dichev (2004) offers an approach here.

- To what extent can features of financial contracts be understood as a response to assorted behavioral biases? Williamson took first steps here. Regarding consumer contracts, Della Vigna and Malmendier (2004) suggest that credit cards and health club contracts, among others, are shaped by naïve expectations and time inconsistent preferences.
- What is the impact of investor inertia and limited attention on corporate finance? Recent papers by Baker, Coval, and Stein (2006) and Della Vigna and Pollet (2006) consider stock swaps and the timing of corporate disclosure. Hirshleifer and Welch (2002) develop implications for organizations.
- How should one approach the proper regulation of inefficient markets and financial reporting?
- What are the limits of corporate arbitrage, including detecting and generating mispricing, maintaining reputation, and avoiding fraud?
- Can a catering approach help to explain the diversification and subsequent re-focus wave that has taken place in the US since the late-1960s? We speculated in Section 2.3.2, but are aware of no systematic studies.
- How significant is the economy-wide misallocation of capital caused by collected behavioral distortions, and in particular how do these distortions interact with traditional capital market imperfections? For example, if there is underinvestment due to agency or asymmetric information, bubbles may bring investment closer to the efficient level.
- What are the behavioral underpinnings of Lintner's (1956) dividend model?
- If bounded rationality or investor pressures lead managers to rely on specific performance metrics, will third parties exploit this? The marketing of takeovers and financing vehicles as EPS-improving transactions by investment banks is a potential example. More generally, what profit opportunities are created by behavioral biases of investors and managers?
- To what extent are corporate "hedging" policies actually directional bets? The evidence in Brown, Crabb, and Haushalter (2005) and Faulkender (2005) suggests that in many companies, interest rate risk management and the use of derivatives has little to do with textbook hedging.
- In the Introduction, we pointed out that the normative implication of the irrational investors approach is to insulate managers from short-term market pressures, while the implication of the irrational managers approach is to obligate them to follow market prices. What, in the end, is the right balance?

## References

Adams, R., Almeida, H., Ferreira, D., 2005. Powerful CEOs and their impact on corporate performance. Review of Financial Studies 18, 1403–1432.

Aghion, P., Stein, J., 2006. Growth vs. margins: Business-cycle implications of giving the stock market what it wants. Working Paper, Harvard University.

Alti, A., 2005. How persistent is the impact of market timing on capital structure? Journal of Finance, in press.

Andrade, G., Mitchell, M.L., Stafford, E., 2001. New evidence and perspectives on mergers. Journal of Economic Perspectives 15, 103–120.

Ang, J.S., Cheng, Y., 2006. Direct evidence on the market-driven acquisitions theory. Journal of Financial Research 29, 199–216.

Ang, A., Gu, L., Hochberg, Y.V., 2005. Is IPO underperformance a peso problem? Journal of Financial and Quantitative Analysis, in press.

Baker, M., Coval, J., Stein, J., 2006. Corporate financing decisions when investors take the path of least resistance. Journal of Financial Economics, in press.

Baker, M., Foley, C.F., Wurgler, J., 2006. Multinationals as arbitrageurs? The effect of valuations on foreign direct investment. Working Paper, Harvard University.

Baker, M., Greenwood, R., Wurgler, J., 2003. The maturity of debt issues and predictable variation in bond returns. Journal of Financial Economics 70, 261–291.

Baker, M., Stein, J., 2004. Market liquidity as a sentiment indicator. Journal of Financial Markets 7, 271–299.

Baker, M., Stein, J., Wurgler, J., 2003. When does the market matter? Stock prices and the investment of equity-dependent firms. Quarterly Journal of Economics 118, 969–1006.

Baker, M., Taliaferro, R., Wurgler, J., 2006. Predicting returns with managerial decision variables: Is there a small sample bias? Journal of Finance 61, 1711–1730.

Baker, M., Wurgler, J., 2000. The equity share in new issues and aggregate stock returns. Journal of Finance 55, 2219–2257.

Baker, M., Wurgler, J., 2002. Market timing and capital structure. Journal of Finance 57, 1–32.

Baker, M., Wurgler, J., 2004a. A catering theory of dividends. Journal of Finance 59, 271–288.

Baker, M., Wurgler, J., 2004b. Appearing and disappearing dividends: The link to catering incentives. Journal of Financial Economics 73, 271–288.

Barber, B.M., Lyon, R., 1997. Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. Journal of Financial Economics 43, 341–372.

Barberis, N., Thaler, R.H., 2003. A survey of behavioral finance. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. North-Holland, Amsterdam.

Barro, R.J., 1990. The stock market and investment. Review of Financial Studies 3, 115–132.

Ben-David, I., 2004. Optimism in managerial forecasts and plans. Working Paper, University of Chicago.

Benartzi, S., 2001. Excessive extrapolation and the allocation of 401(k) accounts to company stock. Journal of Finance 56, 1747–1764.

Benveniste, L., Spindt, P., 1989. How investment bankers determine the offer price and allocation of new issues. Journal of Financial Economics 24, 343–362.

Berger, P., Ofek, E., 1995. Diversification's effect on firm value. Journal of Financial Economics 37, 39–65.

Bergman, N.K., Jenter, D., 2006. Employee sentiment and stock option compensation. MIT Working Paper.

Bergstresser, D., Desai, M., Rauh, J.D., 2006. Earnings manipulation and managerial investment decisions: Evidence from sponsored pension plans. Quarterly Journal of Economics, in press.

Bergstresser, D., Philippon, T., 2006. CEO incentives and earnings management. Journal of Financial Economics 80, 511–529.

Bertrand, M., Schoar, A., 2003. Managing with style: The effect of managers on firm policies. Quarterly Journal of Economics 118, 1169–1208.

Blanchard, O., Lopez-de-Silanes, F., Shleifer, A., 1994. What do firms do with cash windfalls? Journal of Financial Economics 36, 337–360.

Blanchard, O., Rhee, C., Summers, L., 1993. The stock market, profit, and investment. Quarterly Journal of Economics 108, 115–136.

Bolton, P., Scheinkman, J., Xiong, W., 2005. Pay for short-term performance: Executive compensation in speculative market. Working Paper, Princeton University.

Bouwman, C., Fuller, K., Nain, A., 2006. The performance of stock-price driven acquisitions. Review of Financial Studies, in press.

Bradshaw, M.T., Richardson, S.A., Sloan, R.G., 2003. Pump and dump: An empirical analysis of the relation between corporate financing activities and sell-side analyst research. Working Paper, University of Pennsylvania.

Brav, A., 2000. Inference in long-horizon event studies: A Bayesian perspective. Journal of Finance 55, 1979–2016.

Brav, A., Graham, J., Harvey, C.R., Michaely, R., 2005. Payout policy in the 21st century. Journal of Financial Economics 77, 483–528.

Brav, A., Gompers, P.A., 1997. Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture capital and nonventure capital-backed companies. Journal of Finance 52, 1791–1822.

Brennan, M., Tamarowski, C., 2000. Investor relations, liquidity and stock prices. Journal of Applied Corporate Finance 12, 26–37.

Brown, G., Crabb, P., Haushalter, D., 2005. Are firms successful at selective hedging? UNC Working Paper.

Burch, T.R., Christie, W.G., Nanda, V., 2004. Do firms time equity offerings? Evidence from the 1930s and 1940s. Financial Management 33, 5–23.

Butler, A.W., Grullon, G., Weston, J.P., 2005. Can managers forecast aggregate market returns? Journal of Finance 60, 963–986.

Camerer, C., Lovallo, D., 1999. Optimism and excess entry: An experimental approach. American Economic Review 89, 306–318.

Camerer, C., Malmendier, U., 2005. Behavioral economics of organizations. In: Diamond, P., Vartiainen, H. (Eds.), Economic Institutions and Behavioral Economics. Yrjö Jahnsson Foundation, Princeton University Press, Princeton, in press.

Chen, L., Zhao, X., 2004a. Firm financing decisions. Working Paper, Michigan State University.

Chen, L., Zhao, X., 2004b. Profitability, mean reversion of leverage ratios, and capital structure choices. Working Paper, Michigan State University.

Chirinko, R.S., Schaller, H., 2001. Business fixed investment and 'bubbles': The Japanese case. American Economic Review 91, 663–680.

Chirinko, R.S., Schaller, H., 2006. Glamour vs. value: The real story. Working Paper, Emory University.

Conlisk, J., 1996. Why bounded rationality? Journal of Economic Literature 34, 669–700.

Cooper, A.C., Woo, C.Y., Dunkelberg, W.C., 1988. Entrepreneurs' perceived chances for success. Journal of Business Venturing 3, 97–108.

Cooper, M.J., Dimitrov, O., Rau, P.R., 2001. A rose.com by any other name. Journal of Finance 56, 2371–2388.

Cooper, M.J., Gulen, H., Rau, P.R., 2005. Changing names with style: Mutual fund name changes and their effects on fund flows. Journal of Finance 60, 2825–2858.

Cooper, M.J., Korana, A., Osobov, I., Patel, A., Rau, P.R., 2005. Managerial actions in response to a market downturn: valuation effects of name changes in the dot.com decline. Journal of Corporate Finance 11, 319–335.

Core, J., Guay, W., 2001. Stock option plans for non-executive employees. Journal of Financial Economics 61, 253–287.

Dahlquist, M., de Jong, F., 2004. Pseudo market timing: Fact or fiction? Working Paper, University of Amsterdam.

De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R., 1989. The size and incidence of losses from noise trading. Journal of Finance 44, 681–696.

De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R., 1990. Noise trader risk in financial markets. Journal of Political Economy 98, 703–738.

DeAngelo, H., DeAngelo, L., Skinner, D.J., 2004. Are dividends disappearing? Dividend concentration and the consolidation of earnings. Journal of Financial Economics 72, 425–456.

Della Vigna, S., Malmendier, U., 2004. Contract design and self-control: Theory and evidence. Quarterly Journal of Economics 119, 353–402.

Della Vigna, S., Pollet, J., 2006. Investor inattention, firm reaction, and Friday earnings announcements. Working Paper, University of California, Berkeley, CA.

Dichev, I.D., 2004. What are stock investors' actual historical returns? Working Paper, University of Michigan.

Donaldson, G., 1961. Corporate Debt Capacity: A Study of Corporate Debt Policy and the Determination of Corporate Debt Capacity. Division of Research, Graduate School of Business Administration, Harvard University, Boston, MA.

Donaldson, G., 1990. Voluntary restructuring: The case of General Mills. Journal of Financial Economics 27, 114–141.

Dong, M., Hirshleifer, D., Richardson, S., Teoh, S.H., 2005. Does investor misvaluation drive the takeover market? Working Paper, Ohio State University.

Eckbo, B.E., Masulis, R.A., Norli, O., 2000. Seasoned public offerings: Resolution of the 'new issues puzzle'. Journal of Financial Economics 56, 251–292.

Eckbo, B.E., Norli, O., 2004. Liquidity risk, leverage, and long-run IPO returns. Journal of Corporate Finance 11, 1–35.

Edelen, R.M., Kadlec, G.B., 2005. Issuer surplus and the partial adjustment of IPO prices to public information. Journal of Financial Economics 77, 347–373.

Erickson, M., Wang, S.-W., 1999. Earnings management by acquiring firms in stock for stock mergers. Journal of Accounting and Economics 27, 149–176.

Fama, E.F., 1998. Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics 49, 283–306.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E.F., French, K.R., 2001. Disappearing dividends: Changing firm characteristics or lower propensity to pay? Journal of Financial Economics 60, 3–44.

Fama, E.F., French, K.R., 2002. Testing tradeoff and pecking order predictions about dividends and debt. Review of Financial Studies 15, 1–37.

Fama, E.F., French, K.R., 2005. Financing decisions: Who issues stock? Journal of Financial Economics 76, 549–582.

Faulkender, M., 2005. Hedging or market timing? Selecting the interest rate exposure of corporate debt. Journal of Finance 60, 931–962.

Ferris, S.P., Sen, N., Yui, H.-P., 2006. God save the Queen and her dividends: Corporate payouts in the UK. Journal of Business, in press.

Fischer, S., Merton, R.C., 1984. Macroeconomics and finance: The role of the stock market. Carnegie–Rochester Conference Series on Public Policy 21, 57–108.

Flannery, M., Rangan, K., 2006. Partial adjustment toward target capital structures. Journal of Financial Economics 79, 469–506.

Froot, K.A., Dabora, E.M., 1999. How are stock prices affected by the location of trade? Journal of Financial Economics 53, 189–216.

Fuller, K.P., Goldstein, M.A., 2003. Dividend policy and market movements. Working Paper, University of Georgia.

Gaspar, J.-M., Massa, M., Matos, P., 2005. Shareholder investment horizons and the market for corporate control. Journal of Financial Economics 76, 135–165.

Gervais, S., Heaton, J.B., Odean, T., 2003. Overconfidence, investment policy, and executive stock options. Working Paper, Duke University.

Gervais, S., Odean, T., 2001. Learning to be overconfident. Review of Financial Studies 14, 1–27.

Gilchrist, S., Himmelberg, C., Huberman, G., 2005. Do stock price bubbles influence corporate investment? Journal of Monetary Economics 52, 805–827.

Gilovich, T., Griffin, D., Kahneman, D. (Eds.), 2002. Heuristics and Biases: The Psychology of Intuitive Judgment. Cambridge University Press, Cambridge.

Gitman, L.J., Forrester Jr., J.R., 1977. A survey of capital budgeting techniques used by major US firms. Financial Management 6, 66–71.

Goel, A.M., Thakor, A.V., 2002. Do overconfident managers make better leaders? Working Paper, DePaul University.

Golbe, D.L., White, L.J., 1988. A time series analysis of mergers and acquisitions in the US economy. In: Auerbach, A.J. (Ed.), Corporate Takeovers: Causes and Consequences. University of Chicago Press, NBER.

Gompers, P.A., Lerner, J., 2003. The really long-run performance of initial public offerings: The pre-Nasdaq evidence. Journal of Finance 58, 1355–1392.
Graham, J.R., 1999. Duke University CFO outlook survey 1999Q2, http://www.duke.edu/~jgraham/99q2/q299ind.htm.
Graham, J.R., 2000. How big are the tax benefits of debt? Journal of Finance 55, 1901–1941.
Graham, J.R., Harvey, C.R., 2001. The theory and practice of corporate finance: Evidence from the field. Journal of Financial Economics 60, 187–243.
Graham, J.R., Harvey, C.R., Rajgopal, S., 2005. The economic implications of corporate financial reporting. Journal of Accounting and Economics 40, 3–73.
Grossman, S.J., Hart, O.D., 1988. One share-one vote and the market for corporate control. Journal of Financial Economics 20, 175–202.
Guedes, J., Opler, T., 1996. The determinants of the maturity of corporate debt issues. Journal of Finance 51, 1809–1833.
Guedj, I., Scharfstein, D., 2004. Organizational scope and investment: Evidence from the drug development strategies of biopharmaceutical firms. MIT Sloan Working Paper.
Hackbarth, D., 2004. Determinants of corporate borrowing: A behavioral perspective. Working Paper, Indiana University.
Harris, L., Gurel, E., 1986. Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures. Journal of Finance 41, 815–829.
Heaton, J.B., 2002. Managerial optimism and corporate finance. Financial Management 31, 33–45.
Henderson, B.J., Jegadeesh, N., Weisbach, M.S., 2006. World markets for raising new capital. Journal of Financial Economics 82, 63–101.
Hietala, P., Kaplan, S.N., Robinson, D.T., 2003. What is the price of hubris? Using takeover battles to infer overpayments and synergies. Financial Management 32, 1–32.
Hirshleifer, D., Welch, I., 2002. An economic approach to the psychology of change: amnesia, inertia, and impulsiveness. Journal of Economics and Management Strategy 11, 379–421.
Holmstrom, B., 1982. Managerial incentive problems—a dynamic perspective. Review of Economic Studies 66, 169–182. Republished in 1999.
Hovakimian, A., 2006. Are observed capital structures determined by equity market timing? Journal of Financial and Quantitative Analysis 41, 221–243.
Hovakimian, A., Opler, T., Titman, S., 2001. The debt-equity choice. Journal of Financial and Quantitative Analysis 36, 1–24.
Huang, R., Ritter, J., 2005. Testing the market timing theory of capital structure. Working Paper, University of Florida.
Ikenberry, D., Lakonishok, J., Vermaelen, T., 1995. Market underreaction to open market share repurchases. Journal of Financial Economics 39, 181–208.
Ikenberry, D., Lakonishok, J., Vermaelen, T., 2000. Stock repurchases in Canada: Performance and strategic trading. Journal of Finance 55, 2373–2398.
Jain, B.A., Kini, O., 1994. The post-issue operating performance of IPO firms. Journal of Finance 49, 1699–1726.
Jensen, M.C., 1986. Agency costs of free-cash-flow, corporate finance, and takeovers. American Economic Review 76, 323–329.
Jensen, M.C., 2005. Agency costs of overvalued equity. Financial Management 34.
Jensen, M., Meckling, W., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics 3, 305–360.
Jensen, M.C., Ruback, R.S., 1983. The market for corporate control: The scientific evidence. Journal of Financial Economics 11, 5–50.
Jenter, D., 2005. Market timing and managerial portfolio decisions. Journal of Finance 60, 1903–1949.
Jovanovic, B., Rousseau, P., 2002. The Q-theory of mergers. American Economic Review 92, 198–204.
Jung, K., Kim, Y.C., Stulz, R.M., 1996. Timing, investment opportunities, managerial discretion, and the security issue decision. Journal of Financial Economics 42, 159–185.

Kahneman, D., Tversky, A. (Eds.), 2000. Choices, Values, and Frames. Cambridge University Press, Cambridge.
Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 263–291.
Kaplan, S.N., Stromberg, P., 2003. Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. Review of Economic Studies 70, 281–315.
Kaplan, S.N., Weisbach, M., 1992. The success of acquisitions: Evidence from divestitures. Journal of Finance 47, 107–138.
Kayhan, A., Titman, S., 2004. Firms' histories and their capital structures. NBER Working Paper No. 10526.
Keynes, J.M., 1936. The General Theory of Employment, Interest, and Money. Macmillan, London.
Kindleberger, C., 2000. Manias, Panics, and Crashes. Wiley, New York.
Klein, P.G., 2001. Were the acquisitive conglomerates inefficient? RAND Journal of Economics 32, 745–761.
Kothari, S.P., Shanken, J., 1997. Book-to-market, dividend yield, and expected market returns: A time-series analysis. Journal of Financial Economics 44, 169–203.
La Porta, R., 1996. Expectations and the cross-section of stock returns. Journal of Finance 51, 1715–1742.
La Porta, R., Lakonishok, J., Shleifer, A., Vishny, R., 1997. Good news for value stocks: Further evidence on market efficiency. Journal of Finance 52, 859–874.
Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541–1578.
Lamont, O., 2000. Investment plans and stock returns. Journal of Finance 55, 2719–2745.
Landier, A., Thesmar, D., 2005. Financial contracting with optimistic entrepreneurs: Theory and evidence. Working Paper, University of Chicago.
Lang, L., Stulz, R., 1994. Tobin's q, corporate diversification and firm performance. Journal of Political Economy 102, 1248–1280.
Langer, E.J., Roth, J., 1975. Heads I win tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. Journal of Personality and Social Psychology 32, 951–955.
Larwood, L., Whittaker, W., 1977. Managerial myopia: Self-serving biases in organizational planning. Journal of Applied Psychology 62, 194–198.
Leary, M., Roberts, M.R., 2005. Do firms rebalance their capital structures? Journal of Finance, in press.
Lerner, J., 1994. Venture capitalists and the decision to go public. Journal of Financial Economics 35, 293–316.
Li, W., Lie, E., 2005. Dividend changes and catering incentives. Journal of Financial Economics 80, 293–308.
Lintner, J., 1956. Distribution of incomes of corporations among dividends, retained earnings and taxes. American Economic Review 46, 97–113.
Ljungqvist, A., Nanda, V., Singh, R., 2006. Hot markets, investor sentiment, and IPO pricing. Journal of Business, in press.
Ljungqvist, A., Wilhelm, W., 2005. Does prospect theory explain IPO market behavior? Journal of Finance 60, 1759–1790.
Long, J., 1978. The market valuation of cash dividends: A case to consider. Journal of Financial Economics 6, 235–264.
Loughran, T., Ritter, J., 1995. The new issues puzzle. Journal of Finance 50, 23–51.
Loughran, T., Ritter, J., 1997. The operating performance of firms conducting seasoned equity offerings. Journal of Finance 52, 1823–1850.
Loughran, T., Ritter, J., 2000. Uniformly least powerful tests of market efficiency. Journal of Financial Economics 55, 361–389.
Loughran, T., Ritter, J., 2002. Why don't issuers get upset about leaving money on the table in IPOs? Review of Financial Studies 15, 413–443.
Loughran, T., Ritter, J., Rydqvist, K., 1994. Initial public offerings: International insights. Pacific-Basin Finance Journal 2, 165–199.
Loughran, T., Vijh, A.M., 1997. Do long-term shareholders benefit from corporate acquisitions? Journal of Finance 52, 1765–1790.

Lyon, R., Barber, B.M., Tsai, C.-L., 1999. Improved methods for tests of long-run abnormal stock returns. Journal of Finance 54, 165–201.
Malkiel, B., 1990. A Random Walk Down Wall Street. W.W. Norton, New York.
Malmendier, U., Tate, G., 2005. CEO overconfidence and corporate investment. Journal of Finance, in press.
Malmendier, U., Tate, G., 2006. Who makes acquisitions? CEO overconfidence and the market's reaction. Working Paper, University of California, Berkeley, CA.
Manove, M., Padilla, A.J., 1999. Banking (conservatively) with optimists. RAND Journal of Economics 30, 324–350.
March, J.G., Shapira, Z., 1987. Managerial perspectives on risk and risk taking. Management Science 33, 1404–1418.
Marsh, P., 1982. The choice between equity and debt: An empirical study. Journal of Finance 37, 121–144.
Matsusaka, J.G., 1993. Takeover motives during the conglomerate merger wave. RAND Journal of Economics 24, 357–379.
Merrow, E.W., Phillips, K.E., Myers, C.W., 1981. Understanding cost growth and performance shortfalls in pioneer process plants. Rand, Santa Monica, CA.
Mikkelson, W.H., Partch, M., Shah, K., 1997. Ownership and operating performance of companies that go public. Journal of Financial Economics 44, 281–307.
Miller, E.M., 1977. Risk, uncertainty, and divergence of opinion. Journal of Finance 32, 1151–1168.
Mitchell, M.L., Stafford, E., 2000. Managerial decisions and long-term stock price performance. Journal of Business 73, 287–330.
Modigliani, F., Miller, M.H., 1958. The cost of capital, corporation finance, and the theory of investment. American Economic Review 48, 655–669.
Moeller, S., Schlingemann, F., Stulz, R., 2005. Wealth destruction on a massive scale? A study of acquiring-firm returns in the recent merger wave. Journal of Finance 60, 757–782.
Morck, R., Shleifer, A., Vishny, R.W., 1990a. Do managerial objectives drive bad acquisitions? Journal of Finance 45, 31–48.
Morck, R., Shleifer, A., Vishny, R.W., 1990b. The stock market and investment: Is the market a sideshow? Brookings Papers on Economic Activity 2, 157–215.
Moskowitz, T.J., Vissing-Jorgensen, A., 2002. The returns to entrepreneurial investment: A private equity premium puzzle? American Economic Review 92, 745–778.
Muelbroek, L., 1992. An empirical analysis of illegal insider trading. Journal of Finance 47, 1661–1699.
Myers, S.C., 1977. Determinants of corporate borrowing. Journal of Financial Economics 5, 147–175.
Myers, S.C., Majluf, N.S., 1984. Corporate financing and investment decisions when firms have information that investors do not have. Journal of Financial Economics 13, 187–221.
Ofek, E., Richardson, M., 2003. DotCom mania: The rise and fall of internet stocks. Journal of Finance 58, 1113–1138.
Oyer, P., Schaeffer, S., 2005. Why do some firms give stock to all employees? An empirical examination of alternative theories. Journal of Financial Economics 76, 99–133.
Pagano, M., Panetta, F., Zingales, L., 1998. Why do companies go public? An empirical analysis. Journal of Finance 53, 27–64.
Panageas, S., 2004. Speculation, overpricing, and investment: Theory and empirical evidence. Working Paper, University of Pennsylvania.
Polk, C., Sapienza, P., 2004. The real effects of investor sentiment. NBER Working Paper No. 10563.
Pontiff, J., Schall, L.D., 1998. Book-to-market ratios as predictors of market returns. Journal of Financial Economics 49, 141–160.
Porter, M.E., 1987. From competitive advantage to corporate strategy. Harvard Business Review 65, 43–59.
Pshisva, R., Suarez, G.A., 2004. Earnings management and stock market-driven acquisitions: Evidence from the 1990s. Working Paper, Harvard University.
Ravenscraft, D.J., Scherer, F.M., 1987. Mergers, Sell-Offs, and Economic Efficiency. Brookings Institution, Washington, DC.
Rau, P.R., Vermaelen, T., 1998. Glamour, value and the post-acquisition performance of acquiring firms. Journal of Financial Economics 49, 223–253.

Rhodes-Kropf, M., Viswanathan, S., 2004. Market valuation and merger waves. Journal of Finance 59, 2685–2718.

Rhodes-Kropf, M., Robinson, D., Viswanathan, S., 2005. Valuation waves and merger activity: The empirical evidence. Journal of Financial Economics 77, 561–604.

Richardson, S.A., Sloan R.G., 2003. External financing and future stock returns. Working Paper, University of Pennsylvania.

Ritter, J., 1991. The long-run performance of initial public offerings. Journal of Finance 42, 365–394.

Ritter, J., 2003. Investment banking and securities issuance. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. North-Holland, Amsterdam.

Roll, R., 1986. The Hubris hypothesis of corporate takeovers. Journal of Business 59, 197–216.

Savor, P., 2006. Do stock mergers create value for acquirers? Working Paper, Harvard University.

Scarpetta, S., Hemmings, P., Tressel, T., Woo, J., 2002. The role of policy and institutions for productivity and firm dynamics: Evidence from micro and industry data. OECD Working Paper.

Schaller, H., 1993. Asymmetric information, liquidity constraints and Canadian investment. Canadian Journal of Economics 26, 552–574.

Scharfstein, D., Stein, J., 1990. Herd behavior and investment. American Economic Review 80, 465–479.

Scheinkman, J., Xiong, W., 2003. Overconfidence and speculative bubbles. Journal of Political Economy 111, 1183–1219.

Schultz, P., 2003. Pseudo market timing and the long-run underperformance of IPOs. Journal of Finance 58, 483–517.

Schultz, P., 2004. Pseudo market timing and the stationarity of the event-generating process. Working Paper, University of Notre Dame.

Seyhun, H.N., 1992. Why does aggregate insider trading predict future stock returns? Quarterly Journal of Economics 107, 1303–1331.

Shefrin, H., 2000. Beyond Greed and Fear. Harvard Business School Press, Boston, MA.

Shefrin, H., 2001. Behavioral corporate finance. Journal of Applied Corporate Finance 14, 113–124.

Shefrin, H., Statman, M., 1984. Explaining investor preference for cash dividends. Journal of Financial Economics 13, 253–282.

Shleifer, A., 1986. Do demand curves for stocks slope down? Journal of Finance 41, 579–590.

Shleifer, A., 2000. Inefficient Markets: An Introduction to Behavioral Finance. Oxford University Press, Oxford.

Shleifer, A., Vishny, R., 1997. A survey of corporate governance. Journal of Finance 52, 737–783.

Shleifer, A., Vishny, R., 2003. Stock market driven acquisitions. Journal of Financial Economics 70, 295–312.

Simon, H.A., 1955. A behavioral model of rational choice. Quarterly Journal of Economics 69, 99–118.

Sloan, R., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? The Accounting Review 71, 289–316.

Speiss, K., Affleck-Graves, J., 1995. Underperformance in long-run stock returns following seasoned equity offerings. Journal of Financial Economics 38, 243–267.

Speiss, K., Affleck-Graves, J., 1999. The long-run performance of stock returns following debt offerings. Journal of Financial Economics 54, 45–73.

Stambaugh, R.F., 1999. Predictive regressions. Journal of Financial Economics 54, 375–421.

Statman, M., Sepe, J.F., 1989. Project termination announcements and the market value of the firm. Financial Management 18, 74–81.

Statman, M., Tyebjee, T.T., 1985. Optimistic capital budgeting forecasts: An experiment. Financial Management 14, 27–33.

Stein, J.C., 1989. Efficient capital markets, inefficient firms: a model of myopic corporate behavior. Quarterly Journal of Economics 104, 655–669.

Stein, J.C., 1996. Rational capital budgeting in an irrational world. Journal of Business 69, 429–455.

Stein, J.C., 2003. Agency, information and corporate. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. North-Holland, Amsterdam.

Stigler, G.J., 1964. Public regulation of the securities markets. Journal of Business 37, 117–142.

Svenson, O., 1981. Are we all less risky and more skillful than our fellow drivers? Acta Psychologica 47, 143–148.
Teoh, S.H., Welch, I., Wong, T.J., 1998a. Earnings management and the long-run market performance of initial public offerings. Journal of Finance 53, 1935–1974.
Teoh, S.H., Welch, I., Wong, T.J., 1998b. Earnings management and the underperformance of seasoned equity offerings. Journal of Financial Economics 50, 63–99.
Thaler, R.H., 1980. Towards a positive theory of consumer choice. Journal of Economic Behavior and Organization 1, 39–60.
Thaler, R.H., 1985. Mental accounting and consumer choice. Marketing Science 4, 199–214.
Titman, S., Wei, K.C.J., Xie, F., 2004. Capital investments and stock returns. Journal of Financial and Quantitative Analysis 39, 677–700.
Viswanathan, S., Wei, B., 2004. Endogenous events and long run returns. Working Paper, Duke University.
Weinstein, N.D., 1980. Unrealistic optimism about future life events. Journal of Personality and Social Psychology 39, 806–820.
Welch, I., 2004. Stock returns and capital structure. Journal of Political Economy 112, 106–131.

*Chapter 5*

# BANKS IN CAPITAL MARKETS*

STEVEN DRUCKER

*Columbia Business School, Uris Hall, 3022 Broadway, New York, NY 10027, USA*
*e-mail: sd2281@columbia.edu*

MANJU PURI

*Fuqua School of Business, 1 Towerview Drive, Duke University and NBER, Durham, NC 27708, USA*
*e-mail: mpuri@duke.edu*

## Contents



* The focus of this survey is by its very nature limited in its scope and, inevitably, we have left out many important papers. We apologize to those who feel that their research has been ignored or misrepresented.

*Handbook of Corporate Finance, Volume 1*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01005-1*

## Abstract

Banks are an important source of funding in economies all around the world, making it vital to understand how banks directly and indirectly affect funding through capital markets. Few issues have perhaps been as controversial as the appropriate scope of bank activities and whether banks should participate directly in capital market activities, providing both lending and other services, such as underwriting. We review the arguments and theoretical models that consider the consequences of commercial banks engaging in investment banking activities, and we examine the empirical evidence on the potential for conflicts of interest, which focuses on the pricing and long run performance of debt and equity underwritten securities, both in the United States and internationally. A related topic is whether investment banks and commercial banks can co-exist as underwriters. We summarize the theoretical and empirical literature, focusing on the effect that bank lending has had on underwriter fees and the ability of banks to win underwriting mandates, as well as how investment banks have adapted to commercial bank entry into investment banking. We also consider the indirect role of commercial banks in capital markets, providing a summary of banks' ability to signal the quality of borrowers through their decisions to originate and sell loans. Finally we examine related topics, such as the effects of banks holding equity and engaging in venture capital activities, and we suggest research directions.

## Keywords

commercial banks, investment banks, underwriters, certification, conflicts of interest, Glass–Steagall

## 1. Introduction

Banks are an important source of funding in economies around the world. Through syndicated loans arranged by commercial banks, industrial firms borrowed 1.4 trillion dollars in 2003 and 13.2 trillion dollars between 1993 and 2003.[1] The public capital markets have also proved to be a very important funding source. Between 1993 and 2003, industrial firms issued 10.2 trillion dollars of public debt and 2.3 trillion dollars of common stock.[2] Nearly 40 percent of equity issuance and 20 percent of debt issuance occurred in the United States. These facts raise an important question—how do banks directly and indirectly affect funding through capital markets?

Few regulatory issues have been as controversial as the appropriate scope of bank activities. Should banks participate directly in capital markets, providing not just lending services but also other services for the firm, such as public security underwriting? Both academics and regulators have debated this issue for decades. In the United States, commercial banks were permitted to underwrite public securities prior to 1933. However, the stock market crash of 1929 raised concerns over the potential for conflicts of interest and the fear that commingling of investment and commercial banking increased the riskiness of the financial system. In response, Congress passed the Glass–Steagall Act, which effectively prohibited banks from underwriting securities and set the basis for the following sixty year separation of commercial and investment banking. While there has been much rhetoric on potential conflicts of interest when banks combine lending with underwriting, the academic literature on this subject burgeoned only recently.

We begin by reviewing some of the arguments and theoretical models that analyze the implications of banks combining lending with underwriting. Much of the focus of these studies is on the potential for conflicts of interest that can occur when banks use their private information from lending relationships in underwriting their borrowers' public securities. These conflicts of interest are weighed against potential benefits, such as the bank being able to credibly certify the quality of its borrowers to outside investors and generate cost savings from informational economies of scope. This survey deals with these issues and its scope is defined by our perception of this literature.

The theoretical analyses provide a framework for empirical tests of conflicts of interest. These papers analyze the pricing and long run performance of commercial bank-underwritten securities. The first papers use data on public security offerings from before the enactment of the 1933 Glass–Steagall Act while more recent research uses data from the 1980s and beyond, after the relaxation and eventual repeal of the Glass–Steagall Act. Additional studies test whether banks can use organizational means to reduce the potential for conflicts of interest, and other papers examine the effects of a financial intermediary holding equity claims in firms. We also summarize the international evidence on the interaction between commercial banks and capital markets.

[1] Estimates are from Loan Pricing Corporation, which gathers its loan data from SEC filings, large loan syndicators, and a staff of reporters.

[2] Global issuance. Estimates from Thomson Financial.

There are considerable differences in regulatory environments and quality of the financial markets across countries, so additional insight into the causes and consequences of potential conflicts of interest can be ascertained from these studies. As the literature has grown, researchers have become increasingly sophisticated in using many different empirical methodologies to test for the presence of conflicts of interest. We highlight these methods throughout this survey.

Can investment banks, which generally do not provide lending services, co-exist with commercial banks? Some theoretical models suggest that this is a realistic possibility. For example, issuers may choose commercial banks when economies of scope are large and choose investment banks when the costs from conflicts of interest are sizable. Other models point out the possibility that investment banks and commercial banks can co-exist by charging different underwriter fees that reflect their relative benefits and costs. Another possibility, not generally addressed in the theoretical literature, is that investment banks will compete with commercial banks by expanding their lending activities. We survey the empirical literature on this topic by highlighting a number of papers that examine the effects of commercial bank re-entry into underwriting on the costs of intermediation, the impact of lending relationships on underwriting fees, and whether lending influences the likelihood of winning underwriting mandates.

In addition to the direct interaction between commercial banks and the capital markets, there is an indirect role of commercial banks on capital markets. Through screening and monitoring, banks gather private information about their borrowers. Even if banks do not directly participate in underwriting, banks' lending decisions can still signal the quality of firms to investors. Generally, researchers have examined this possibility by quantifying the firm's stock price reaction to loan initiations, renewals, and sales. Other studies examine if lending relationships provide positive information to outsiders by documenting the effects of bank loans from non-underwriting banks on the pricing of public security offerings. We provide a detailed summary of banks' ability to convey quality to outsiders through signaling.

Finally, we explore a number of areas where more research is needed. One such topic concerns the ability of banks to hold equity in firms, which is currently limited in the United States but is allowed in other countries, such as Germany and Japan. There is some evidence from the United States on the effects of banks and other financial intermediaries holding equity through venture capital subsidiaries. However, the consequences of banks holding equity remain unclear and highlight the need for additional research so that we can more fully understand the interaction between banks and capital markets.

## 2. Commercial banks as underwriters: Theoretical literature

When a commercial bank underwrites a firm's public securities, a number of benefits may arise. First, the private information that a bank gathers in the lending process may be valuable in public security underwriting. Insider banks know more about a firm's prospects than outsiders due to their screening and monitoring of loans. Compared with investment banks, which do not generally acquire private information through lending activities, commercial banks have lower costs of information production. This advantage can allow commercial banks to gather more information about their clients and be better certifiers of firm value than investment banks. Second, banks may achieve informational economies of scope by jointly delivering lending and underwriting services and re-using the same client specific information for several purposes (see e.g. Benston, 1990; Saunders and Walter, 1994). As a result, informational economies of scope can lower transaction costs and reduce the costs of intermediation.

However, the potential benefits of commercial banks as underwriters of public securities can be limited by costs that can occur due to conflicts of interest from banks' incentives to misuse their private information. A bank may privately know that a firm has poor prospects but attempt to protect its own interests by certifying that the firm is of high quality and underwriting public securities, with the hope that investors will subscribe to the issue. The firm then can use the proceeds to pay down its bank loans at the expense of outside investors. This activity benefits the bank in two ways—in addition to earning a fee on the security underwriting, the bank reduces its overall portfolio exposure to default risk. A commercial bank that lends and underwrites may face other conflicts of interest that it may attempt to exploit. For example, banks may issue loans to third-party investors on the condition that these funds are used to support the price of a new issuance of public securities. In this case, the supporting of the security price through bank loans could send incorrect signals to investors and other new issuers regarding the true performance of the underwriter, making the bank appear to be a better underwriter than in truth. As another example, the bank may attempt to "tie" the provision or pricing of credit to the firm's use of the bank's investment banking services. By threatening to reduce the availability of credit or increase the cost of borrowing, the client may then face costs from higher-priced or lower-quality services, with the bank reaping the rewards.[3] Of course, conflicts of interest may be mitigated by the bank's concern for harming ongoing client relationships and its own reputation. It is likely that short-term gains from exploiting these conflicts are offset in the long run by these concerns, which can affect the ability of the bank to generate future business and profits.

Theoretical papers by Kanatas and Qi (1998, 2003), Puri (1999), and Rajan (2002) contrast the benefits that can arise from certification and informational economies of scope with the costs from conflicts of interest. These papers provide formal analyses of allowing banks to extend their business beyond traditional lending activities, and these studies produce some implications for the pricing of public securities, the firm's choice of underwriter, and the costs of financial intermediation.[4]

[3] See Walter (2004) for a thorough analysis of potential conflicts of interest in financial services firms.

[4] In a slightly different vein, Boot and Thakor (1997a, 1997b) examine the impact of the choice between universal and functionally separate banking, and argue that a financial system in its infancy will be bank dominated.

Rajan (2002) and Kanatas and Qi (1998, 2003) examine the implications of some costs and benefits of universal banking. Rajan (2002) examines if, with unrestricted competition, commercial banks with expanded powers will naturally evolve as efficient institutions. He shows that unrestricted competition does not necessarily lead to efficient institutions if the markets in which institutions compete are not naturally competitive. The intuition is that in producing one service (say lending), the integrated producer obtains the possibility of an ex-post rent in producing the second service (say underwriting). This rent can arise because the private information that the bank attains through lending may allow the bank to "capture" the firm.[5] So long as the ex-post rent is greater than the inefficiency that the integrated producer brings to underwriting, the bank can secure the customer's underwriting business. In this setting, universal banks can deter the emergence of other specialized organizational forms. Rajan (2002) argues that this is one plausible scenario for financial institutions not to evolve in the socially optimal way. Of course, whether these conditions apply is an empirical question. Rajan (2002) points out that if underwriting markets are competitive, then commercial banks will be forced to internalize the costs of the structure that they choose. In such a case, regulators can rely on commercial banks to make the right decision about whether to enter into the security underwriting.

Kanatas and Qi (1998) focus on the trade-offs between informational economies of scope and conflicts of interest. The authors assume the existence of the incentive conflict where the bank underwrites low quality firms' securities in order to pay down its bank loans. This incentive conflict limits the ability of the bank to credibly certify the quality of firms that use its underwriting services. Therefore, outside investors pool high-quality security issues with low-quality issues, which increases the financing costs of high-quality firms. High-quality firms can avoid being pooled with low-quality issuers by either using an independent underwriter or borrowing from a lending-only bank. However, by doing so, the firm forgoes any benefits that could arise due to informational economies of scope from using the same bank for both lending and underwriting services. Therefore, universal banks underwrite securities for firms when the benefits of scope economies outweigh the costs from conflicts of interest. In a related study, Kanatas and Qi (2003) develop a model in which economies of scope are a double-edged sword for the universal bank. On the positive side, informational economies of scope provide a cost advantage to universal banks (which is shared with clients) that enables universal banks to lock-in their clients' future business. However, on the negative side, the fact that relationships are more durable reduces the incentive for the universal bank to place effort into underwriting the clients' securities. In this model, firms trade-off the benefit of lower costs of dealing with a universal bank with the greater likelihood

[5] The firm can be captured for two reasons. First, the bank has lower costs of information production in security underwriting, which deters competition from other underwriters. Second, the bank's information creates a lemons problem for the firm in that other underwriters will be skeptical of the quality of firms that do not use their universal bank as underwriter.

of successful capital market financing from choosing an investment bank. Two key implications of the model are that universal banks will likely be selected as the underwriter when economies of scope are large, and underwriting allows banks to form and maintain strong relationships in multiple product lines, with firms that use universal banks for underwriting more likely to use the same universal bank for bank lending services.

Puri (1999) models the trade-off between commercial banks' potential to be better certifiers of firm value and the conflict of interest that can arise from the bank misrepresenting the value of a firm's securities in order to use the proceeds to repay bank loans. The formal model is a repeated game where investors are rational and update their beliefs about banks given the last period action, which allows reputation concerns to be captured. Commercial banks know if the firm is good or bad due to previous loan monitoring activities. Commercial banks can underwrite bad firms in order to pay down pre-existing claims, but if investors observe this action, then this observation will reduce the reputation of the bank and its future profits. Investment banks, which do not know the quality of the firm, can incur an investigative cost to determine the true quality of the firm. If they choose not to investigate and subsequently underwrite a low quality firm, the investment bank will suffer through reputation loss. The trade-offs that each underwriter faces determine equilibrium strategies and the prices that the market assigns to underwritten securities. A key result is that commercial banks are likely to obtain better prices for underwritten securities than investment banks when the costs of information production are high, as might be seen in junior and informational sensitive securities.[6] These results help provide theoretical underpinnings for many of the results in the empirical literature. This paper also examines if it is possible for commercial banks and investment banks to coexist in equilibrium and derives sufficient conditions for coexistence in which the level of rent extraction and the relative underwriting fees adjust so that firms are indifferent between going to commercial banks and investment banks. In Section 4, we survey the empirical evidence on competition between commercial banks and investment banks, including the literature on underwriting fees.

## 3. Empirical evidence on conflicts of interest

Like the theoretical literature, the empirical studies also examine the benefits and costs of universal banking. Much of the focus of the empirical literature is on the effect of bank's lending, and the private information contained therein, on commercial banks' ability to certify firm value in the presence of potential conflicts of interest. In this section, we provide an overview of this literature. The literature examines these issues over multiple periods of time in the United States. In Section 3.1, we provide a summary

[6] Of course, there are other scenarios where an investment bank can achieve higher prices than a commercial bank, such as when the costs of investigation for the investment bank are sufficiently small, or when both types of underwriters are perceived to have low reputations.

of the evidence from before the enactment of the 1933 Glass–Steagall Act, which prohibited commercial banks from underwriting public securities for nearly sixty years. Section 3.2 provides a review of studies that use data from the late 1980s and beyond, after the relaxation and repeal of the Glass–Steagall Act. In Section 3.3, we explore another strand of the literature that examines if commercial banks can use organizational means to mitigate the potential for conflicts of interest. Also, there are papers that explore the consequences of a financial intermediary holding equity claims in firms. We summarize these studies in Section 3.4. Throughout, we highlight the many different methodologies that have been used to test for the presence of conflicts of interest.

There are two primary ways that researchers examine whether banks are net certifiers of firm value or if commercial banks are subject to conflicts of interest. The first method is to examine the ex ante pricing of public securities. The foundation of these studies is that rational investors should anticipate whether commercial banks or investment banks have a higher net certification effect, and price the securities accordingly. If investors perceive that conflicts of interest are large, then commercial bank-underwritten securities will be priced lower than similar investment bank-underwritten securities, while if conflicts of interest are small, then commercial bank issues will achieve higher prices. The second method is to examine the ex post performance of underwritten securities. If commercial bank-underwritten securities perform worse than *ex ante similar* securities that are underwritten by investment banks, then this would be consistent with commercial banks underwriting securities that they privately know to be of lower quality, which is indicative of conflicts of interest. In general, there is little support for banks' exploiting conflicts of interest. In fact, many studies find commercial banks to be net certifiers of firm value.

### *3.1. Before the 1933 Glass–Steagall Act*

Prior to 1933, commercial banks were permitted to underwrite public securities. However, after the stock market crash of 1929, concerns over the potential for conflicts of interest and the fear that the commingling of investment and commercial banking increases the riskiness of the financial system prompted Congress to enact the Glass–Steagall Act of 1933, which prohibited commercial banks from engaging in public security underwriting. Popular support for the Act came from investigations by the Pecora Committee (U.S. Senate Committee on Banking and Currency, 1933–1934), which examined alleged abuses at the security affiliates of commercial banks, in particular, National City Company and Chase Securities Corporation.[7] However, many scholars have argued that evidence of these concerns was anecdotal and little verification was provided that any abuses were systematic in nature (see e.g. Carosso, 1985; Benston,

[7] See Kelly (1985) for details on the legal history of the Glass–Steagall Act.

1990).[8] Below, we highlight the formal empirical analyses that explore if the committee's concerns were justified.

Puri (1996) examines the ex ante pricing of industrial bonds and preferred stock during the period January 1927 through September 1929. She regresses the yield of the securities on a dummy variable that indicates if the issue is commercial bank-underwritten, and she includes control variables for bond characteristics and issuer characteristics that could also affect the yield.[9] She finds that, relative to investment bank issues, commercial bank-underwritten issues have a significantly lower yield, which is consistent with commercial banks having a net certification effect. Of course, there are other explanations (other than net certification) that could account for this yield difference. Hence the author conducts a number of tests to determine if yield differences are higher in junior and more information sensitive securities as suggested in Puri (1999). She finds that having a commercial bank underwriter has a significantly larger effect on yield in samples where private information is likely to be more important. For example, the strongest effects are in the preferred stock sample, which is junior and more sensitive to information than bonds. There are also stronger effects for new issues than seasoned issues and non investment-grade issues than investment-grade securities. Further, there is little effect of underwriter type on foreign government bond issues, which are not information intensive.

The baseline tests in Puri (1996) use OLS regressions of yield on control variables and a bank underwriting dummy, and use the coefficient on the dummy to infer whether a bank underwriting lowers yields. This is a standard approach prevalent in empirical banking and corporate finance research. Puri (1996) also conducts additional tests to examine whether the lower yield of bank underwritings can be attributed to the private information held by banks. Her approach is to estimate private information as a residual and use its correlation with the next-stage dependent variable as a basis for testing whether private information matters. In the specific implementation of this approach in Puri (1996), a probit model is used to determine the probability of being bank underwritten. The estimates are used to compute the inverse Mills ratio, which is a proxy for private information because it is the expectation of the residuals not explained by public information. The coefficient for the inverse Mills ratio is negative, consistent with a net certification effect for commercial bank underwritten offerings. Interestingly, Puri's technique parallels a similar approach used subsequently in the insurance literature, where Chiappori and Salanie (2000) test whether customers buying more comprehensive automobile insurance coverage have private information that they have higher accident probabilities. Like Puri (1996), Chiappori and Salanie (2000) also

[8] In recent times, regulators have raised questions on the firm-level and competitive effects of the relaxation and repeal of the Glass–Steagall Act (see e.g. Mester, 1996; Berger, Demsetz, and Strahan, 1999; Santomero and Eckles, 2000).

[9] The yield is defined as the premium of the ex ante yield of the security over the ex ante yield of a government bond of nearest maturity issued in the same month.

estimate private information as a residual and use its correlation with the second stage dependent variable as a basis for testing whether private information matters.

Clearly, many different approaches can be used to assess the benefits of bank underwriting, and a number of these techniques are utilized in studies of the post-1990 period and in examinations of underwriter fees, which we will discuss later. For example, an alternate approach is to use endogenous switching models (see e.g. Fang, 2005; Song, 2004 for applications and Maddala (1983) for details on the model). These models generalize the two-stage approach used in Puri (1996) by allowing commercial banks and investment banks to have separate yield equations. This relaxes the assumption that the variables that affect yield have the same effect for investment bank and commercial bank issues. Estimating the model involves two steps. First, the researcher runs a probit model to determine the probability that the issuer chooses a commercial bank or investment bank underwriter. In the second step, the researcher estimates two yield equations separately for investment bank and commercial bank issues, including independent variables that affect yield as well as the inverse Mills ratio. Interestingly, when we apply endogenous switching methodology to industrial bonds in the pre-1933 data, similar effects are found.

The evidence that investors paid more for bank underwritten-securities pre-Glass–Steagall suggests that commercial banks are net certifiers of firm value. However, this raises an important question of interpretation. Namely, did investors pay more for bank-underwritten securities because they rationally believed them to be of better quality, or were investors naïve and banks took advantage of them so that investors paid higher prices for worse securities? This question can be addressed by examining the ex post performance of bank underwritten securities. Ang and Richardson (1994), Kroszner and Rajan (1994), and Puri (1994) examine the ex post performance of securities using data from the pre-Glass–Steagall period. As noted, if commercial bank-underwritten securities perform worse than ex ante similar securities that are underwritten by investment banks, then this would be consistent with commercial banks underwriting securities that they privately know to be of lower quality. All three studies find no evidence to support the existence of conflicts of interest.

Ang and Richardson (1994) examine the long-run performance of bonds, using a comprehensive sample of 647 bond issuances over the years 1926 through 1930. The authors compare the default rate of commercial bank and investment bank-underwritten bonds based on the default status of the bonds at two points in time (1934 and 1939) and find that the default rates are similar for investment bank- and commercial bank-underwritten securities. While this analysis is limited because the authors do not control for differences in the characteristics of issuers across the two types of underwriter, the results suggest that conflicts of interest did not override the certification ability of commercial banks.

Kroszner and Rajan (1994), using data from the first quarters of the years 1921 through 1929, examine the relative performance of industrial bonds that are underwritten by commercial banks with those that are investment bank-underwritten. The main measure of bond performance is the default rate because reliable price data is scarce for this time period. To compare default performance between ex ante similar bonds, the authors use two methods: (i) matched-security tests, where bonds originated by commercial bank affiliates are matched to similar investment bank-underwritten bonds based on observable characteristics; and, (ii) logit analysis. For the matched-security tests, the authors create a sample of ex ante similar commercial bank and investment bank-underwritten securities, using the credit rating as the primary measure of bond quality. In total, Kroszner and Rajan (1994) find 121 industrial bond matches, where the bonds have the same initial credit rating, are issued within six months of each other, have similar maturity and size, and have the same conversion provision.[10] Using this sample, the authors find that at the end of each year after 1924, there are fewer cumulative defaults among commercial bank-underwritten issues, and by the end of the sample period in 1940, 32 percent of investment bank-underwritten bonds defaulted relative to 23 percent of bonds that were underwritten by commercial banks. By dollar volume, approximately 28 percent of investment bank-underwritten issues default by 1940, compared with only 11 percent of commercial bank-underwritten issues. Further, not only do investment bank-underwritten issues default more frequently, but they also default earlier in their lives. All of these findings suggest that commercial bank-underwritten issues performed better than similar, investment bank-underwritten issues, which is inconsistent with commercial banks succumbing to conflicts of interest.

In addition, Kroszner and Rajan (1994) perform a log-rank test using the sample of matched securities. This test takes into account both the number of defaults and the timing of defaults by comparing the mortality rates of the two groups of bonds. Consistent with their initial findings, the main result of this test is that the survival rate of commercial bank-underwritten bonds is significantly higher than investment bank-underwritten bonds. Importantly, these differences are strong in the non investment-grade sample, but insignificant in the investment-grade sample. Since incentive conflicts created by information asymmetries between underwriters and investors are larger in low quality issues, this result supports the view that conflicts of interest were not large during the pre-Glass–Steagall period.

Kroszner and Rajan (1994) confirm the lower default probability of commercial bank-underwritten issues using logit analysis in which they estimate whether the type of underwriter affects the probability of default, after controlling for security and firm characteristics. The logit analysis complements the matched-security tests by allowing the authors to use data on all of the investment bank-underwritten issues (instead of just the smaller sample that is matched to commercial bank-underwritten issues) and providing means to control for other factors that may be correlated with default. According to the estimates from logit models, underwriting by a commercial bank reduces the probability of default by 11 percent, with large and significant reductions in default probabilities seen among the lowest quality issues. In economic terms, an 11 percent

[10] If there are multiple matches, the authors use other criteria, such as collateralization status, to select the best match.

difference in the probability of default is approximately the same as the difference in default probability between investment-grade bonds and unrated bonds.

Puri (1994) also examines the long run default performance of bank-underwritten issues. The author uses both the cumulative mortality rate and probit models to examine the default performance of bonds. The cumulative mortality rate allows for an accurate comparison of default probability by measuring default rates on bonds that have been outstanding for equal periods of time, adjusted for calls, maturities, and previous defaults.[11] Using a sample of industrial bond issues during the period January 1927 through September 1929, Puri (1994) finds that the cumulative mortality rate is significantly higher for non-bank underwritten issues than bank underwritten issues for 3, 5, and 7 years from the issue date.[12] These results are particularly strong in the non-investment grade sample for all time periods. The results support the view that banks were not exploiting conflicts of interest. While the mortality rate analysis is better than an unconditional comparison of default rates, the probit model allows the researcher to control for other important factors that might influence the probability of default. Consistent with the mortality analysis, the results of the probit model strongly indicate that commercial bank underwritings of industrial bonds and preferred stock defaulted less often, and foreign government bonds defaulted with similar probability. Interestingly, Puri (1994) finds that there was a selection bias in the Senate hearings that lead to the Glass–Steagall Act. The two banks that bore the brunt of the investigation underwrote securities that had a significantly higher default rates than that of other banks and were not representative of bank underwriters in general.

Together, the ex ante pricing results and the long run performance studies paint a convincing picture. Commercial bank-underwritten securities received higher prices. Investors rationally paid higher prices because in the long run these securities performed better than ex ante similar offerings. This suggests that conflicts of interest were not dominant in bank-underwritings during the pre Glass–Steagall period.

### *3.2. The late 1980s and beyond*

During the late 1980s and throughout the 1990s, commercial banks were gradually allowed to re-enter underwriting markets. In 1987, the Federal Reserve permitted individual bank holding companies to establish Section 20 subsidiaries that could to a limited extent engage in "bank ineligible" activities.[13] However, the subsidiaries had to be separately capitalized and separated from the lending parent by information, finance,

[11] See Altman (1989) for a formal definition and discussion.

[12] The sample period for this study provides for a more uniform regulatory and economic environment, as it starts after the passage of the McFadden Act, which legally allowed national banks to underwrite debt securities, and ends before the stock market crash of October 1929.

[13] Section 20 of the Glass–Steagall Act prevented commercial banks from affiliating with a company "engaged principally" in the "issue, flotation, underwriting, public sale, or distribution at wholesale or retail or through syndicate participation of stocks, bonds, debentures, notes or other securities".

and resource firewalls. In 1989, Section 20 affiliates were permitted to underwrite corporate debt, and in 1990, the Federal Reserve granted equity underwriting powers. The Federal Reserve set an initial revenue cap on bank ineligible activities at 5 percent of the gross revenue of the Section 20 subsidiary, and the cap was raised to 10 percent in 1989 and then to 25 percent in December 1996.[14] In 1997, the Federal Reserve removed the majority of firewalls between Section 20 subsidiaries and their bank holding company parents, and on November 12, 1999, the Gramm–Leach–Bliley Act (Financial Modernization Act) effectively repealed the Glass–Steagall Act.

There are a number of papers that use more recent data to examine the pricing of securities underwritten by commercial banks. As in the pre Glass–Steagall period, most of the evidence points to a net certification effect for commercial banks. Gande et al. (1997) examine the pricing of debt securities from January 1, 1993 to March 31, 1995, a period when commercial banks' underwriting affiliates were constrained by regulation that limited their ability to generate revenues and faced significant firewalls that could reduce information flow between the underwriting affiliate and the parent commercial bank. The authors are able to measure the amount of lending exposure between the issuer and the underwriter, which, as per the theory, should be important in determining security prices. The authors find that commercial banks primarily underwrite small issues, which is consistent with a positive role of banks in bringing smaller issuers to the market. Importantly, after controlling for bond characteristics, issuer characteristics, and underwriter attributes, the authors find that underwritings where the bank has existing lending exposure have significantly lower yields for lower credit rated (Caa-Ba3) issues, but no difference on the less informationally sensitive, higher rated issues.[15] Again, these results are consistent with bank underwriting being valuable for lower credit rated issues due to a net certification effect. Further, if conflicts of interest are present, they are likely to be highest when the purpose of the debt issuance is to refinance existing bank debt because in these issues, the bank may misrepresent the quality of the firm so that the issuer can raise more money to pay down its existing bank loans. Among this sample of issues, the effect of lending exposure on yields is economically and statistically insignificant, indicating a lack of conflicts of interest. As an additional robustness check, the authors create a proxy for private information by estimating the residuals in a probit model where the dependent variable is one if lending bank is the underwriter and independent variables are observable factors that affect underwriter choice. These residuals are found to be correlated with reduced yields for lower credit rated issues, after controlling for publicly available bond characteristics, consistent with a net certification effect.

[14] Note that the other revenue of the Section 20 subsidiary comes from "eligible" activities, such as swaps origination and government bond underwriting.

[15] For lower credit rated issues, a one-unit increase in LN(Amount of lending exposure) reduces yields by 27 basis points for lower-credit rated issues. An alternative measure, PROP(STAKE), which is the lending exposure over the amount of the debt issue size, produces similar results.

A follow-up paper by Roten and Mullineaux (2002) uses similar methods as Gande et al. (1997), but examines a later time period: January 1, 1995 to December 31, 1998. During this period, many of the restrictions on commercial bank underwriting were relaxed. Roten and Mullineaux (2002) find the benefits of bank underwriting in this later period show up in reduced underwriting fees rather than in net yields. There is more work on the underwriting fee differentials between commercial and investment banks that we will discuss in more detail in Section 4.1.

Thus far, we have focused on the effect of commercial bank underwriting on public debt issues. A few recent papers examine equity issues. In equity markets, an indirect cost of initial public offerings (IPOs) is underpricing, where the price of the security at offering is, on average, below the price prevailing in the market shortly after the IPO.[16] It is well documented that IPOs are underpriced, and many theoretical papers indicate that IPO underpricing arises from asymmetric information problems regarding the issuing firm's value (see e.g. Rock, 1986; Benveniste and Spindt, 1989; Benveniste and Wilhelm, 1990; Allen and Faulhaber, 1989; Grinblatt and Hwang, 1989; Welch, 1989, 1992). The benefits of bank lending relationships are likely to be especially important when a firm goes public due to the substantial uncertainty about a firm's value. However, the consequences of conflicts of interest can be more severe in IPOs due to equity being junior to debt and the pronounced asymmetric problems with private firms.

Schenone (2004) examines the effect of having a banking relationship with the underwriter of the IPO on the firm's IPO underpricing. If conflicts of interest are high, then investors may perceive stocks underwritten by relationship banks to be riskier than other IPOs. Using a sample of 306 IPOs from 1998 through 2000, the author finds that IPOs underwritten by a firm's relationship bank are less underpriced than IPOs where the firm does not have lending relationships with any potential underwriter. In addition, there is no significant difference in underpricing relative to firms that could have, but do not, use their relationship bank as underwriter. These results indicate that IPOs with relationship banks are, at a minimum, not perceived to be riskier than other IPOs, supporting that conflicts of interest do not override the certification ability of the bank.

Benzoni and Schenone (2004) examine the long run performance of equity issues that are underwritten by the firms' relationship banks relative to those issues that are underwritten by other commercial bank and investment bank underwriters. The focus on ex post performance is similar to Ang and Richardson (1994), Kroszner and Rajan (1994), and Puri (1994). The main differences are that Benzoni and Schenone (2004) use modern data from 1998 through 2000 and examine equity issues as opposed to debt issues. The authors examine the impact of lending relationships on the firm's long run equity performance in two ways. First, for each of the 306 IPO firms, the authors construct 2-year buy-and-hold returns for the firm's stock as well as the buy-and-hold returns for two benchmark portfolios, one of which is specific to each firm and is comprised of

[16] Underpricing is defined as the differences between the first closing pricing and the offer price, divided by the offer price. See Ritter (1998) for a survey of empirical evidence.

the returns on six portfolios of stock ranked by size and book-to-market, and the other is the CRSP value-weighted market portfolio. Using feasible generalized least squares to account for cross-sectional correlation in the stock returns of firms, the authors estimate the impact of having an existing lending relationship with the underwriter on the long-run returns of the IPO firm relative to the benchmark portfolios, controlling for firm characteristics, IPO characteristics, and other factors. The authors find that IPOs underwritten by relationship banks perform no better or worse than issues underwritten by outside commercial or investment banks. This result is inconsistent with relationship banks misrepresenting the quality of the firm's that they underwrite. Second, for each IPO underwritten by the firm's relationship bank, the authors find a similar matched IPO from the sample of non-relationship bank IPOs based on the dates of the IPOs and the book-to-market ratios of the firms. Benzoni and Schenone (2004) form a portfolio of long positions in the relationship bank IPO firms and short positions in the matched sample of non-relationship bank IPO firms. The authors regress the weekly portfolio returns on Fama and French's (1993) market, size, and book-to-market returns and examine if there are abnormal returns associated with this portfolio. The authors do not find significant abnormal returns, indicating that relationship bank-underwritten firms perform similarly to the matched sample, which is again inconsistent with relationship banks misrepresenting the quality of the firm's that they underwrite.

As in the pre Glass–Steagall period, the evidence from the late 1980s and beyond suggests that conflicts of interest are not dominant in bank underwritings. The ex ante pricing results indicate that when the firm and underwriting commercial bank have a lending relationship, the public security prices are no worse and sometimes better than similar issues underwritten by investment banks or non-relationship commercial banks. These results are robust to different methodologies, time periods, and types of security. Further, the long run performance of relationship bank-underwritten IPOs are no worse than similar IPOs that are underwritten by non-relationship banks, which is inconsistent with the existence of conflicts of interest.

### *3.3. Mitigating conflicts of interest: Organizational structure and syndicates*

The aforementioned papers examine the trade-offs between certification and conflicts of interest by analyzing the ex-ante yield of debt, the underpricing of equity securities, and the ex post performance of securities. The evidence suggests that bank certification at least cancels out and may outweigh potential conflicts of interest. While these studies take as given, and attempt to quantify, the relative magnitude of these offsetting effects, a number of papers examine if there are ways for commercial banks to take action to reduce the potential for conflicts of interest. In other words, can commercial banks credibly commit to certifying firm value in order to mitigate any perception that they will exploit conflicts of interest?

Puri (1996) and Kroszner and Rajan (1997) examine if the organizational structure of the financial institution can mitigate potential conflicts of interest. During the pre-Glass–Steagall period, commercial banks organized their investment banks as either internal securities departments within the bank or as separately incorporated affiliates with their own boards of directors. By forming independent entities, banks may be able to credibly commit to not exploit potential conflicts of interest that could be pronounced due to the likelihood of increased information flows in an internal structure. Puri (1996) finds that affiliate underwritings do not have significantly lower yields than similar investment bank issues, while in-house underwritings have significantly lower yields when compared with investment bank issues. Her results do not support the view that independent entities were beneficial. However, these results contrast with Kroszner and Rajan (1997). Using a sample of 422 industrial bonds from 1925 through 1929, the authors compare the initial yields on issues underwritten by internal departments with issues that are underwritten by separate affiliates. The results of their multivariate regression suggest that independent affiliate-underwritten issues have yields that are significantly lower than internal department-underwritten bonds, by 12 to 23 basis points. These results are consistent with the independent structure allowing for credible commitment. There are some differences in the samples and approaches of these two studies. Clearly, more research is needed to sort out this question.

Narayanan, Rangan, and Rangan (2004) and Song (2004) explore another way for commercial banks to credibly commit to certify firm value and avoid conflicts of interest. These authors examine the role of syndicate structure in underwriting. Narayanan, Rangan, and Rangan (2004) focus on the possibility that a relationship bank may co-manage an issuance with a reputable, non-lending underwriter in order to commit against opportunistic behavior. Using 1,640 seasoned equity issuances from the years 1994 through 1997, Narayanan, Rangan, and Rangan (2004) find that the proportion of syndicate co-manger roles to lead manager roles for relationship banks is about three times higher than for non-relationship banks. Also, relationship banks are significantly more likely than non-lending banks to co-manage an issue with an independent, high reputation lead manager. Further, an examination of the pricing of these issues reveals that issues where a relationship bank is a co-manager exhibit similar levels of underpricing as issues where only investment banks are underwriters. Taken together, these results are consistent with the view that relationship banks use the syndicate structure to credibly commit against exploiting potential conflicts of interest. Importantly, similar to the results in Roten and Mullineaux (2002) for debt issuances, Narayanan, Rangan, and Rangan (2004) show that while issuers do not receive better pricing on their equity issuance, the issuer benefits from reduced underwriting fees. We discuss this further in Section 4.1.

Another interpretation of the underpricing results in Narayanan, Rangan, and Rangan (2004) is that relationship banks do not improve the certification ability of the syndicate. Otherwise, we would observe lower underpricing on issues where relationship banks are co-managers. However, to really examine if co-managing allows relationship banks to improve their net certification of issues relative to lead managing, one would have to contrast underpricing between issues that are co-managed by relationship banks with similar issues that are lead managed by relationship banks. This comparison is not provided in Narayanan, Rangan, and Rangan (2004) due to a lack of commercial bank lead managed issues during the time period. However, Song (2004) is able to make a related comparison in the bond underwriting market. Using a sample of 2,345 corporate bond issues from 1991 to 1996, Song (2004) examines the clienteles and bond pricing associated with three different syndicate structures: (i) commercial bank-lead syndicates; (ii) syndicates with only investment banks; and, (iii) hybrid syndicates where an investment bank leads the issue and commercial banks are co-managers. Song uses an endogenous switching model with six equations: three selection equations, which capture the likelihood of choosing a given syndicate structure over the other options, and three yield equations, one for each of the three syndicates.[17] The results of her model indicate that commercial banks are more likely to co-manage an issue rather than serve as lead manager when the purpose of the issue is to refinance bank debt and the issuer has more loans from the commercial bank underwriters. Since these issues are more likely to be prone to conflicts of interest, the results are consistent with the view that acting as a co-manager allows commercial banks to mitigate perceptions that they will exploit conflicts of interest. However, bond yields are similar when commercial banks are lead managers as opposed to co-managers. This suggests that co-managing does not improve the certification ability of commercial banks.

### *3.4. Conflicts of interest from equity holdings: Evidence from venture capital*

Much of the focus so far has been on the trade-off between the private information from lending allowing banks to be better certifiers of firm value with the potential for conflicts of interest from misusing the information. In this section, we again explore this trade-off, but examine some of the different effects that can occur when underwriters are equity holders in the firm. The evidence from venture capital can provide insight into the potential consequences of allowing banks to hold equity in firms.

Some authors maintain that allowing banks to hold equity claims helps increase a financial intermediary's credibility in certifying the firm's value (see e.g. Leland and Pyle, 1977), which provides a formal analysis of how equity holdings in the firm can provide a signal of firm value). However, as Puri's (1999) model points out, the horizon for which equity is held is critical to this certification. If the bank can retire its financial claim through the proceeds of the equity issuance, then holding equity can hurt the credibility of the bank more than holding debt. There is some empirical evidence on the impact of equity holdings on the certification ability of the underwriter, derived from comparing IPOs where the underwriter has gained an equity stake through venture capital investments and other IPOs where the underwriter does not have an equity claim. Both papers that we survey do not find evidence of conflicts of interest.

[17] In another application of this method, Fang (2005) studies the relation between investment bank reputation and the prices of underwritten bonds and uses separate pricing equations for high and low reputation underwriters. Fang (2005) finds that more reputable underwriters obtain lower yields and charge higher fees.

Using a sample of 885 venture-backed IPOs from December 1972 to December 1992, Gompers and Lerner (1999) compare long run performance, liquidation probability, and underpricing based on if an underwriter in the IPO holds a venture stake. To examine long-run performance, Gompers and Lerner (1999) calculate the 5-year buy-and-hold excess return, which is the firm's buy-and-hold return minus the 5-year buy-and-hold return of the portfolio of firms with the same size and book-to-market ratio. This comparison reveals that issues where the investment bank held an equity stake perform just as well, and by some measures significantly better than, non-affiliated offerings. These results are inconsistent with the existence of conflicts of interest. Further, the authors explore if excess returns are influenced by the percentage of venture investors' equity sold at the time of the IPO. If venture investors are attempting to take advantage of outsider investors, then higher fractions of equity sold should result in lower excess returns. They find no evidence to support this hypothesis, and in fact, find the opposite to be true. In addition, Gompers and Lerner (1999) examine the probability that a firm is liquidated within 5 years of the IPO and find no significant relationship between liquidation probability and using an underwriter that has a venture claim in the firm. Again, these results suggest that conflicts of interest are not a concern. Further supporting this view, when examining underpricing, the authors do not find a significant difference between IPOs that are underwritten by affiliated underwriters and independent underwriters.

Li and Masulis (2004) also examine the impact of venture capital investments by IPO underwriters on the net certification ability of the underwriter. However, as opposed to Gompers and Lerner (1999) who treat all existing venture relationships as equally important, Li and Masulis (2004) examine if the size of the equity ownership by the underwriter affects IPO underpricing and the probability of being delisted in the future. This approach is similar to that in Gande et al. (1997). Using a sample of 1,480 venture-backed IPOs from 1993 to 2000, the authors find that IPO underpricing decreases as the share of the underwriter's equity ownership increases, even after controlling for other factors that can influence underpricing. The underpricing results are consistent with certification effects overriding any conflicts of interest. In support, the authors find that among issues that are more uncertain, which is proxied for by the firm's stock volatility during the year following IPO, venture investments reduce underpricing more than for less risky issues. This suggests that the prior information from the venture investment allows the underwriter to reduce informational asymmetries. Similar to Gompers and Lerner's (1999) evaluation of liquidation, Li and Masulis (2004) find no significant relationship between underwriter shareholdings and the likelihood of subsequent stock delisting, which is consistent with conflicts of interest not being pronounced.

The evidence in these two studies highlight that conflicts of interest are not a concern when an underwriter holds an equity stake in the firm. Further, the analysis in Li and Masulis (2004) suggests there are benefits from the underwriter holding an equity stake, with affiliated underwriters being net certifiers of firm value and allowing for firms to reduce their direct costs of going public.

## 4. Empirical evidence on competition between commercial and investment banks

The empirical evidence in Section 3 on conflicts of interest raises an important point: If commercial banks are net certifiers of firm value and conflicts of interest can be mitigated, then commercial banks may be superior underwriters compared to investment banks. If so, can both types of underwriter co-exist? How can competition between them affect the services they offer to firms?

By combining lending with underwriting, banks may achieve informational economies of scope by jointly delivering lending and underwriting services and re-using the same client specific information for several purposes (see e.g. Benston, 1990; Saunders and Walter, 1994). As emphasized in Kanatas and Qi (2003), informational economies of scope can lower transaction costs and can theoretically reduce underwriting fees if banks pass along costs savings to firms. Puri (1999) derives sufficient conditions for commercial banks and investment banks to coexist. One implication of this analysis drawn out in the paper is that, under some circumstances, commercial banks may charge higher underwriting fees than investment banks. We survey the literature on underwriting fees in Section 4.1. The analyses in Kanatas and Qi (1998, 2003) emphasize that large scope economies from combining lending and underwriting will be important in determining if an issuer selects its commercial bank lender as public security underwriter. Rajan (2002) points out that the bank's information advantage from lending may allow it to secure the underwriting mandates of its borrowers. In Section 4.2, we summarize the studies that examine the effect of bank lending on underwriter selection. In Section 4.3, we provide some additional evidence on how investment banks are adapting to competition from commercial banks.

### *4.1. Underwriting fees*

We begin by providing evidence that commercial bank entry after the relaxation of the Glass–Steagall Act caused lower overall underwriting fees, consistent with a pro-competitive effect on corporate securities underwriting markets. Gande, Puri, and Saunders (1999), using a sample of 2,992 debt issues from 1985 through 1996, document that following bank entry into debt underwriting in 1989, the gross spread, or underwriting fee, declined significantly.[18] Further, this decline is more pronounced in samples where commercial banks gained a larger market share (non-investment-grade and smaller issues). This result stands in contrast to equity markets where commercial banks had not yet gained much market share, and where similar declines in gross spreads are not observed in this time period.

Gande, Puri, and Saunders (1999) also find that among the sample of 1,180 debt issues between 1989 and 1996, commercial banks and investment banks charge similar

[18] The authors capture the impact of bank competition on gross spreads in two ways. First, they use a dummy variable that is one after 1989, when banks were first allowed to underwrite corporate debt. Second, they use the logarithm of commercial banks market share in debt underwriting.

fees. A number of studies build on this comparison of fees by focusing directly on the impact of lending relationships on gross spreads. In public debt markets, two studies find that prior lending relationships reduce the gross spread. Using a regression framework that controls for bond and issuer characteristics, Roten and Mullineaux (2002) find that prior lending relationships lower debt underwriting fees by 10 basis points during the 1995 to 1998 period. Yasuda (2005) gathers a sample of 1,535 bond issues from 1993 to 1997 and uses a more advanced methodology to examine the impact of lending relationships on fees. The author points out that studies that examine the effect of lending relationships on fees use the equilibrium pricing outcomes that are observed. Yasuda (2005) argues that the gross spread that is observed is likely to be lower on average than the unconditional distribution of the gross spread. Therefore, the author imputes the implied gross spread for each of the other underwriters that the firm could have selected to underwrite the issue. Using the Expectation-Maximization Algorithm that accounts for this downward bias in observed gross spread, Yasuda (2005) estimates a joint model of the gross spread and the firm's selection of underwriter. The gross spread is modeled as a function of bond and issuer characteristics, as well as if the potential underwriter was an arranger on any of the firm's prior loans before 1993. The choice of underwriter is a function of the implied gross spread, bond and issuer characteristics, and existing lending relationships. Yasuda (2005) also finds that lending relationships significantly decrease the gross spread by approximately nine basis points.

As in debt underwriting, the evidence suggests that lending reduces the gross spreads of equity offerings. Three papers use the framework developed by Altinkilic and Hansen (2000), who find that gross spreads for seasoned equity offerings are U-shaped with respect to the size of the offering. Theoretically, U-shaped curves can arise because scale economies cause gross spreads to decline initially, but as issue size increases, higher placement costs can override the benefits of scale economies, causing gross spreads to increase. Narayanan, Rangan, and Rangan (2004), using seasoned equity offerings from 1994 to 1997, include a variable that captures if a commercial bank in the underwriting syndicate has a lending relationship with the issuing firm. They find that the existence of a lending relationship reduces gross spreads by 46 basis points, which is significant at the one percent level. This result is consistent with informational economies of scope from combining lending and underwriting. Drucker and Puri (2005) study "concurrent lending" and underwriting, which occurs when the underwriter of a seasoned equity offering provides a loan to the issuer between six months before and six months after the issuance. As part of their study, the authors examine the impact of concurrent lending and prior lending on seasoned equity offering gross spreads. The authors argue that informational economies of scope are likely to be large when issuers receive a loan concurrently because the information from the lending transaction is directly re-usable in the equity offering. The authors extend the Altinkilic and Hansen (2000) model to include variables that control for firm characteristics and prior underwriting relationships as well as variables that indicate if the lead underwriter provided concurrent loans or had a prior lending relationship with the issuer. For a sample of 2,301 seasoned equity offerings from 1996 through 2001, concurrent lending without a prior lending relationship significantly reduces gross spreads by 18 basis points and concurrent lending where a prior lending relationship exists results in gross spreads that are 36 basis points lower. Prior lending relationships without a concurrent loan also cause gross spreads to be reduced significantly, by 36 basis points. These discounts are consistent with the existence of informational economies of scope. Further, the discounts for concurrent and prior lending relationships are significant in the sample of non-investment-grade issuers, where economies of scope from combining lending and underwriting are likely to be larger. Bharath et al. (2004) use a sample of 283 initial public offerings and estimate U-shaped models that separately include three different measures of lending relationship strength. These measures capture if the firm and underwriter have a prior lending relationship, the proportion of the firm's loans over the five years prior to the IPO where the underwriter had a lead role, and the dollar-based percentage of the firm's loans where the underwriter had a lead role. In all three cases, gross spreads are significantly lower by 19 to 26 basis points.

### *4.2. Underwriter selection*

In general, the evidence points to better pricing and lower underwriting fees from using a relationship bank as underwriter. Presumably, the benefits of using a relationship lender as security underwriter will influence the firm's choice of underwriter. Is there evidence that lending relationships allow underwriters to increase their likelihood of winning underwriting mandates? Four recent papers examine the effect of lending on a firm's choice of underwriter, and all find that lending increases the likelihood of winning underwriting business.

Drucker and Puri (2005) examine if lending around the time of a securities offering (concurrent lending) and prior lending impact the choice of seasoned equity underwriter. The authors use McFadden's (1973) choice model to examine the choice of underwriter. The authors allow the choice of underwriter to depend upon concurrent and prior lending, firm characteristics, and attributes that are specific to the relationships between each firm and potential underwriter, such as the analyst coverage and the quality of the coverage that potential underwriters provide for the firm, the reputation of potential underwriters, and any existing underwriting relationships. The results of this model reveal that both concurrent lending and prior lending increase the likelihood of the bank being selected as the lead underwriter. Further, the authors examine if concurrent lending increases the likelihood that underwriters are selected for future equity underwriting business. Using a nested logit model in which the issuer first chooses if it will re-issue in the equity market and then chooses if it will keep the same underwriter or switch to a new underwriter, the authors find that concurrent lending increases the likelihood that investment banks keep future underwriting business, which is consistent with lending fostering an ongoing relationship between underwriters and firms.

In a related paper, Bharath et al. (2004) examines the impact of prior lending on capturing debt and equity underwriting business. The authors use a logit model, allowing each issuer to choose among the top-20 investment banks and any commercial bank that could underwrite its securities. They allow the choice of underwriter to depend upon the size of the issue, existing lending and underwriting relationships, and the reputation of the underwriters in both the underwriting and lending market. Bharath et al. (2004) find that prior lending relationships significantly increase the likelihood of winning debt underwriting mandates and being selected as lead manager on IPOs.

Ljungqvist, Marston, and Wilhelm (2006) also provide evidence related to the influence of bank lending relationships on underwriter selection. Using a sample of 16,625 debt and equity deals over the period December 1993 to December 2002, the authors estimate a probit model, providing each firm with the potential to choose any of the 16 most active underwriters. In their model, the choice of underwriter depends upon bank-firm underwriting and lending relationships, as well as bank reputation and analyst characteristics and behavior. Ljungqvist, Marston, and Wilhelm (2006) find that the probability of winning both equity and debt underwriting business is increasing in the bank's share of the issuer's prior loans.

Yasuda (2005) provides an examination of the impact of existing lending relationships on the choice of debt underwriter during the period 1993 to 1997. As previously explained, Yasuda (2005) estimates a joint model of the gross spread and the firm's selection of underwriter, allowing the firm to choose between sixteen underwriters. The joint framework allows the author to include in the underwriter selection equation the estimated fee that each underwriter would have charged the issuer to underwrite the offering. Therefore, Yasuda (2005) can examine if the lending relationship influences underwriter selection above and beyond any effect that charging lower gross spreads has on underwriter selection. The author finds that prior lending relationships significantly increase the likelihood that the lending bank wins the bond underwriting business over and above the effect of the gross spread discount. Further, lending relationships have a stronger impact on underwriter selection among junk rated issuers and new issuers, where a bank's private information is likely to be most valuable. These results are consistent with Kanatas and Qi's (1998, 2003) theoretical models, which indicate that lenders will select their bank as underwriter when there are likely to be large informational economies of scope. The estimates of Yasuda's (2005) model indicate that firms are willing to pay a higher underwriter fee to banks with which they have a prior relationship. One of the major benefits of this framework is that the author can explicitly calculate how much more an issuer is willing to pay. For the sample mean issue size of \$180 million, an issuer is willing to pay \$2.23 million more to use a relationship commercial bank and \$2.62 million to have a relationship investment bank as underwriter. Junk rated issuers and first time issuers, where the value of a bank's private information is likely to be largest, are willing to pay even more. These results are consistent with a certification effect for relationship banks.

### *4.3. Can investment banks survive?*

Overall, the empirical evidence shows that using relationship banks as underwriters improves the pricing of issues and lowers fees, and both prior lending relationships and lending around the time of a security issuance increase the probability that an underwriter will be selected as underwriter. Further, as Yasuda (2005) points out, lending relationships increase the likelihood of selection above and beyond any cost reductions. Given these facts, is it possible for investment banks to remain viable underwriters?

One possibility is that investment banks can remain viable competitors by expanding their lending activities. Some evidence of investment bank lending and its effects on financing costs and the choice of underwriter is provided in Drucker and Puri (2005). They document that investment banks are now making loans and are competing aggressively with commercial banks by providing loans around the time of seasoned equity offerings.[19] While the authors show that, in concurrent deals, investment banks cannot compete with commercial banks on the yield spreads that they can charge for the loan, concurrent lending and prior lending allow investment banks to provide lower gross spreads on the equity offering, with investment bank lending significantly reducing gross spreads by between 26 and 44 basis points. These results are consistent with lending by investment banks creating sufficient economies of scope to allow investment banks to be viable competitors with commercial banks. This inference is bolstered by an examination of issuers' selection of their underwriter. As with commercial banks, when investment banks have prior lending relationships or provide concurrent loans, their likelihood of being selected as equity underwriter increases. Further, investment banks are able to leverage concurrent deals into extended relationships by capturing future underwriting business.

The evidence in Yasuda (2005) also hints that investment banks may use lending to compete with commercial banks. Her model shows that for the mean debt issue, issuers that have a lending relationship with an investment bank underwriter are willing to pay more to use the relationship investment bank as underwriter. Further, investment bank lending increases the likelihood of winning the underwriting mandate. These results suggest that firms value lending relationships with investment banks, and as a result, lending may allow investment banks to remain competitive.

## 5. International evidence

Thus far, we have examined the impact of commercial banks as underwriters in the United States. However, there is some international evidence on differences between the ex ante pricing and ex post performance of commercial bank and investment bank-underwritten issues, additional analyses on the extent to which commercial banking relationships affect the choice of underwriter, and evidence on other potential conflicts of interest. Below, we survey the literature related to commercial bank underwriting in Japan, Canada, and Israel.

[19] In fact, of the 201 issues where the underwriter provides a loan concurrently with the seasoned equity offering, investment banks are the underwriter in 110 cases.

### *5.1. Japan*

Japan and the United States have similar regulatory histories regarding the ability of commercial banks to underwrite securities. In Japan, commercial banks were allowed to underwrite securities until 1948. However, much like the Glass–Steagall Act in the United States, Article 65 of the Securities and Exchange Act of 1948 effectively prohibited commercial banks from running securities businesses. The Financial System Reform Act of 1992, which came into effect in 1993, allowed commercial banks to again underwrite securities through subsidiaries, and in a short period of time, commercial banks gained significant market share in corporate bond underwriting.[20] Despite the similarities in their regulatory histories, as opposed to banks in the United States, Japanese commercial banks have historically operated in a main bank system where banking relationships are strong and long-term.[21] Therefore, the trade-offs between conflicts of interest and certification should be pronounced in Japan, and there is likely to be a strong impact of existing banking relationships on competition for underwriting mandates.

Konishi (2002) examines the pricing and long-term default performance of industrial bonds underwritten by commercial banks as compared with investment banks during the pre-war period in Japan (January 1919–December 1927). Using the same framework to examine ex-ante pricing as in Puri (1996), the author finds no difference in the yields of commercial bank-underwritten and investment bank-underwritten bonds, consistent with conflicts of interest not dominating the certification effect. To examine long-term default performance, Konishi (2002) follows Puri (1994) and calculates cumulative mortality rates as well as uses a probit model to estimate if the probability of default is influenced by bank underwriting, after controlling for other important factors. The results of the mortality analysis indicate that commercial bank-underwritten issues default significantly less often than investment bank issues at time horizons from three to seven years after issuance. Further, the probit analysis of default probability also shows that commercial bank issues are significantly less likely to default. Together, these results suggest that conflicts of interest were not a problem when banks underwrote public securities in pre-war Japan, which is consistent with the evidence from the pre-Glass–Steagall period in the United States.

There are three papers that study the pricing of industrial bonds in Japan after the Financial Systems Reform Act, comparing commercial bank and investment bank issues. In each of these papers, the authors identify if the underwriter has an outstanding loan to the issuer and also if the underwriter owns shares in the firm. As in Gande et al. (1997), Roten and Mullineaux (2002), and Schenone (2004), the identification of these prior relationships allows for a richer testing ground. These papers are by Hamao and Hoshi (2002), Takaoka and McKenzie (2006) and Liu and Kang (2004). In examining

[20] See Hoshi and Kashyap (1999) for more details on financial deregulation in Japan and its consequences.
[21] See Hoshi (1996) for a discussion of main bank relationships and universal banking in Japan.

yields, Hamao and Hoshi (2002) find weak evidence for conflicts of interest, Takaoka and McKenzie (2006) find weak evidence of certification, and Liu and Kang (2004) find no evidence of conflicts of interest in terms of differential effects between commercial banks and investment banks. Takaoka and McKenzie (2006) additionally examine if bank entry post-1993 lowered commissions and yields, and they find some supporting evidence. Each of these papers has slightly different samples and methodology. Clearly more research is needed here to answer this question.

Yasuda (2006) also examines the Japanese bond market after the Financial Systems Reform Act but focuses on the effect of bank relationships on competition in the underwriting market. Using a framework similar to Yasuda (2005), the author estimates a joint model of the gross spread and the firm's selection of underwriter, allowing the firm to choose between fifteen underwriters. The estimates indicate that having a prior lending relationship significantly increases the probability of being selected as bond underwriter, and stronger lending relationships increase the likelihood of selection by more than weaker relationships. This is consistent with the evidence from the United States. In addition, issuers are willing to pay a higher fee (+0.513%) for underwriting services from banks with which they have pre-existing lending relationships, all else equal. This suggests that banking relationships provide additional value to the firm. Interestingly, the results in Yasuda (2005) indicate that issuers in the United States are willing to pay 1.238% more to use a relationship commercial bank as underwriter, which provides some indication that there could be more benefits in the United States from issuers using a relationship bank as underwriter.

### *5.2. Canada*

In Canada, universal banking began after deregulation occurred in 1987. Within thirteen months after the June 1987 change in law permitting bank entry, all six of Canada's chartered banks had an underwriting division. The commercial banks gained significant market share during the following years. Ursel and Ljucovic (1998) examine the relationship between commercial bank underwriting (as modeled by a dummy variable that indicates if the underwriter is bank-owned) and underpricing using a data set of 111 Canadian IPOs between July 1987 and December 1994. The authors are limited by data constraints that prevent them from tracking existing lending relationships between banks and firms, which would enable stronger conclusions. Using a parsimonious specification, the authors find that commercial bank-underwritten issues have lower underpricing, but after controlling for other important factors, such as reputation, Ursel and Ljucovic (1998) no longer find a significant difference.

Hebb and Fraser (2002) examine the relationship between commercial bank underwriting and bond yields using 356 non-convertible bond issues from 1987 to 1997. The authors find that commercial bank underwritten issues have a yield that is lower by 20 basis points, consistent with commercial banks being net certifiers. For the issues where the authors are able to identify the issuer's primary lender they find that the existence of a lending relationship does not affect bond yield spreads. The Canadian data from both the equity and debt markets suggest that conflicts of interest are limited by commercial banks certification ability.

### *5.3. Israel*

In Israel, banks are highly universal in nature, managing investment funds and controlling underwriting affiliates. The close links between the investment fund and the bank allow for researchers to examine this potential source of conflicts of interest. Ber, Yafeh, and Yosha (2001) perform such an analysis, examining if conflicts of interest are present when banks underwrite Israeli IPOs. In this study, the authors gather data on 128 IPOs of manufacturing firms from 1991 to 1994. For each of the issues, the authors identify if the firm has a prior lending relationship with the underwriter and if the bank's investment fund purchases the firm's stock at the time of the offering and in the aftermarket. Ber, Yafeh, and Yosha (2001) examine the effect of these relationships on the accounting and stock performance for one year following the issue as well as the underpricing of the issue. In terms of accounting profitability, the authors find that firms that are underwritten by bank lenders significantly outperform other issuers. Further, they find that these better performing firms were similar ex ante to other IPO firms based on publicly available information, indicating that banks underwrite superior firms. This is inconsistent with the existence of conflicts of interest. However, when the authors examine the long run stock performance, their results indicate that the stocks of firms with a bank underwriter-lender exhibit significantly negative excess returns during the first year that are significantly different than the excess of returns of firms that do not have a bank lender-underwriter.[22] Also, an examination of first day returns reveals that issues involving a bank lender-underwriter are significantly overpriced. How can the strong accounting performance and poor stock performance be reconciled? The authors find that much of the poor stock performance comes from issues where the bank's fund management division made significant purchases. Ber, Yafeh, and Yosha (2001) conclude that the results indicate a conflict of interest, as banks overpriced these IPOs, favoring the IPO client firms at the expense of investors in the bank's investment fund.

## 6. The indirect role of commercial banks on capital markets

Throughout this article, we have documented that the empirical literature has generally found commercial banks to be certifiers of firm value when they combine lending and underwriting activities. However, even if banks cannot directly participate in capital markets through underwriting, banks' actions and lending decisions might still affect outside stakeholders in firms. Can banks, which have private information about

[22] The authors use a market model approach and compare the excess returns for each firm for the first year after IPO.

a firm's prospects, signal the quality of firms to outside investors through their lending decisions? Fama (1985) and Diamond (1991) provide theoretical analyses of these questions. Fama (1985) claims that banks are "special" with respect to other financial intermediaries in their ability to gather and process private information and their ability to certify firm value to outsiders. His argument relies on two important observations. First, bank borrowers are usually depositors at the bank, which creates an information advantage for banks relative to other financial intermediaries because they have access to private information provided by the ongoing history of bank deposits. The deposit record makes it cheaper for banks to monitor and screen potential borrowers. Second, bank loans are generally low priority claims, so the granting and renewal of bank loans provides positive signals to higher priority lenders, allowing these higher priority lenders to avoid monitoring the firm.[23] Therefore, bank loans reduce the need for outsiders to generate duplicate information, allowing bank loans to reduce overall information costs. Since outsiders use the bank loans as positive signals of firm value, according to this analysis, bank loans are important conveyers of private information to the capital markets.

Building on insights in Fama (1985), Diamond (1991) develops a model in which banks have a comparative advantage relative to capital markets in funding younger, smaller and less well-known firms due to their ability to screen and monitor borrowers. Through ongoing lending relationships in which the bank monitors the firm, young firms can develop a credit record to obtain a sound reputation. The acquisition of reputation allows the firm to access the public debt markets later in the "life-cycle." In this model, the banks' superior access to private information from screening and monitoring activities allows the bank to convey information about borrower quality and signal creditworthiness to the capital markets.

The analyses in Fama (1985) and Diamond (1991) highlight banks' role as information producers. One implication of these studies is that if the private information gathered in the lending process provides banks with a comparative advantage over other intermediaries and allows firms to build a reputation, then the granting and renewal of bank loans will provide a positive signal to outside investors of the bank's private information, particularly when borrowers are young and informationally-opaque. Conversely, the selling of loans may be a negative signal. In Section 6.1, we survey the empirical studies that test this claim by examining the borrowing firm's stock price response to bank loan announcements, renewals, and sales. Another implication of the analyses in Fama (1985) and Diamond (1991) is that by conveying private information to the market through lending decisions, bank loans reduce the need for outsiders to

[23] In contrast to Fama's (1985) theory, Carey (1995) shows that in a sample of 18,000 syndicated loans made between 1986 and 1993, 99% of the loans are senior. Welch (1997) aroves that bank loans are senior to reduce deadweight costs from Organized banks contesting priority in financial distress. Stih, Fama's (1985) suggestion that banks are compactively advantaged over capital market participants in screening and monitoring borrowers is well supported in the literature (see e.g. Diamond, 1984; Ramakrishnan and Thakor, 1984; Boyd and Prescott, 1986). See Mayer and Vives (1993) for a comprehensive survey.

generate duplicate information and reduce information asymmetries between firms and capital markets. Therefore, even if the bank cannot underwrite the firm's public securities, the existence of a bank loan may result in higher security prices. In Section 6.2, we summarize the empirical evidence on the effects of bank loans from non-underwriting banks on the pricing of public security offerings.

### 6.1. *Market reaction to loan announcements, renewals, and sales*

If the announcement of bank loans conveys positive private information to investors, then the borrowing firm should realize an abnormal return around the event date. James (1987) provides the first in-depth analysis of the impact of bank loan announcements on a firm's equity returns, as he compares the abnormal returns associated with bank loan announcements with the returns generated by announcements of other financings.[24] James (1987) selects 300 companies at random from the Center for Research on Security Prices (CRSP) daily return files and searches the *Wall Street Journal Index* for announcements of public straight debt offerings, private placements of debt, and bank borrowing agreements over the period 1974–1983. The bank loan agreements consist of new credit agreements and the expansion of existing agreements. In total, James (1987) finds 207 financing announcements, which are comprised of 80 bank loan agreements, 37 private placements (which are primarily arranged by insurance companies), and 90 public straight debt offerings.[25] He uses a market model to obtain estimates of abnormal stock returns around the announcement of the financing events. Using two-day announcement period abnormal returns, James (1987) finds that bank loan agreements produce an abnormal return of $+1.93\%$, which is significant at the one percent level. In contrast, the author finds that announcements of public debt offerings produce a statistically insignificant abnormal return of $-0.11\%$, and private placements produce an average abnormal return of $-0.91\%$, which is significant at the ten percent level. The positive reaction to bank loan agreements and the negative reaction to the other financings, which are not arranged by commercial banks, suggest that there is some benefit to the intermediation process provided by commercial banks and bank loans. However, since the abnormal returns may be driven by differences in the characteristics of the issues rather than the special nature of bank lending, James (1987) further refines his tests by grouping the types of announcements based on the purpose of the financing, the maturity of the issuances, the debt rating of the issuer, and the size of the borrower. His analysis indicates that differences in the abnormal performance are not driven by these characteristics, strengthening the view that bank loan agreements signal the bank's positive private information about a firm's prospects to the capital markets.

[24] Mikkelson and Partch (1986) first discovered that bank credit line announcements cause abnormal returns, but this analysis was a small aspect of their study.

[25] In James (1987) and future studies that examine announcement effects, the authors take great care to remove any announcements that are potentially "contaminated" by other news information, such as dividend declarations, earnings announcements, or other financings.

Building on James (1987), Lummer and McConnell (1989) make an important distinction between new bank credit agreements and revisions to already existing credit agreements. If the announcement effects are significant for new bank credit agreements, then this suggests that banks can transmit private information to the capital markets at the outset of a loan agreement due to the initial screening of the client. However, if announcement effects are pronounced among loan renewals and revisions, then this suggests that banks are able to convey private information from their ongoing monitoring activities to capital markets. To construct their sample, Lummer and McConnell (1989) search the *Wall Street Journal Index* for announcements of credit agreements involving commercial banks and U.S. corporations covered by CRSP for the period 1976 to 1986, and they find 728 announcements that meet their criteria. Of the 728 announcements, 371 are new credit agreements and 357 concern existing agreements. Using the same methodology employed by James (1987), the authors employ an event-time study of stock returns over the two-day period encompassing the announcement day in the *Wall Street Journal* and the previous day. Consistent with James (1987), the authors find an announcement-period excess return of $+0.61\%$, which is significant at the one percent level. Importantly, the authors find that the positive abnormal return is driven by revised credit agreements, which produce a highly significant positive abnormal return of $+1.24\%$. In contrast, the sample of new credit agreements produces a statistically insignificant announcement-period excess return of $-0.01\%$. Further, favorable revisions produce positive abnormal returns while negative revisions and cancellations that are initiated by the lender produce strongly negative announcement-period excess returns. Importantly, the results hold up when the authors use multivariate regression models that control for other characteristics of loan initiations and renewals that could be driving the results, such as the sizes, maturity, secured status, and structure of the contract. Overall, these results support the view that the private information that banks transmit to the capital markets arises from the monitoring activities of the bank that take place over the course of an ongoing relationship rather than from the screening of the borrower during a loan initiation.

An empirical study by Slovin, Johnson, and Glascock (1992) examines if announcement effects differ by the size of the firm. Based on insights in Diamond's (1991) model, the authors claim that since there is more public information available for larger firms rather than smaller firms, banks do not have to provide as intense screening and monitoring services for larger borrowers. Therefore, if announcement day abnormal returns reflect a bank's private information that is gathered through screening and monitoring, then bank loan announcement effects should decrease in firm size. To test the hypothesis, Slovin, Johnson, and Glascock (1992) classify the sample firms into small and large based on the median market value of equity of all listed CRSP firms in the year of a given announcement. A search of the *Wall Street Journal Index* over 1980 to 1986 produces 273 favorable loan announcements, of which 156 are for small firms and 117 are for large firms. The results indicate a statistically insignificant reaction to large firm loan announcements ($+0.48\%$, $z$-statistic $= 1.58$), but a large positive abnormal return for small firm loan announcements ($+1.92\%$, $z = 5.35$). Furthermore, for small firms, both initiations and renewals produce positive abnormal returns, while for large firms, neither initiations nor renewals have significant announcement period excess returns. These results, which are robust to a multivariate specification, support the view that banks gather more private information when screening and monitoring small firms rather than large firms.

Best and Zhang (1993) also examine the information content of bank loan agreements. The authors claim that if there are reliably accurate public signals of firm value, then bank loan announcements should convey little additional information to the market. Alternatively, in cases where public signals are noisy, then the information content of bank loans should be pronounced. To examine this possibility, the authors split their sample according to whether financial analysts' percentage earnings prediction errors during the year prior to the announcement are high or low. The results, which are based on 491 usable loan announcements over the period 1977 to 1989, indicate that announcement day excess returns are significant for the high prediction error sample (+0.6031%, $z = 2.99$) but not for the low prediction error sample (+0.0444%, $z = 0.28$). These findings indicate that bank's private information is valuable when public information is unclear. The authors also examine if there is evidence consistent with the view that banks expend more effort to monitor a firm when public signals indicate that a firm's prospects have changed. To do so, Best and Zhang (1993) test if abnormal returns differ if the most recent earnings forecast revisions are positive, unchanged, or negative. The authors find that for firms that receive positive earnings forecast revisions, loan announcement abnormal returns are insignificant, but for firms who receive negative earnings forecast revisions and have noisy forecasts of earnings, loan announcements produce significant abnormal returns. All of the results withstand a multivariate specification that allows the authors to control for other factors that could influence abnormal returns. One interpretation of the results is that banks do little further monitoring and screening when public signals are reliable and positive, but when public signals are noisy and firms prospects change for the worse, banks expend additional effort on monitoring.

In another study, Billett, Flannery, and Garfinkel (1995) examine if the lender's identity affects the announcement day abnormal returns. The key motivation for this breakdown is that loan announcements from higher-quality lenders, who could have better monitoring abilities, may be more informative to outsiders than loan announcements from lower-quality lenders. To examine if lender identity matters, the authors examine if a bank's credit rating causes differences in the announcement day abnormal returns. Billett, Flannery, and Garfinkel (1995) search Dow Jones News Retrieval Service for the time period 1980 to 1989 and find 626 usable loan announcements. Using the same basic methodology as in previous studies, the authors find that loan announcements where the lender is a high-quality lender (rated AAA) produce a significantly positive abnormal return of +0.320%, while loans from low-quality lenders (rated BAA or below) are negative (−0.233%) and statistically insignificant.[26] Further, mean abnormal returns for

[26] One difference between Billett, Flannery, and Garfinkel (1995) and the previous papers is that this study uses one-day event windows because the authors are able to identify if the announcement occurred during the

loans from AAA lenders significantly exceed the excess returns from lenders rated BAA or lower. While the univariate results indicate that the announcement effects are concentrated among loans from high-quality lenders, as in previous studies, the authors refine their test through regressions that control for other characteristics that could be driving the result (such as differences in borrower characteristics). Even after controlling for these other factors, higher quality lenders continue to be associated with significantly higher abnormal announcement returns.

Overall, these studies find that the announcement day abnormal returns are significantly positive for loan announcements, stronger for loan renewals and changes rather than initiations, larger for smaller firms rather than larger firms, stronger when public information about the firm is noisy rather than clear, and larger for higher quality lenders. It is important to note that the positive abnormal returns around loan announcements contrasts with the strongly negative announcement effects of equity, the moderately negative abnormal returns around convertible debt issuance, and the insignificant abnormal returns surrounding straight public debt announcements (see Asquith and Mullins (1986), Mikkelson and Partch (1986), and Masulis and Korwar (1986) for equity issuances; Eckbo (1986) and Mikkelson and Partch (1986) for debt issuances). These latter results are consistent with a firm's public security issuance decision revealing its private information on its prospects (Myers and Majluf, 1984) and strengthen the view that bank loan announcements convey positive private information to the capital markets.[27]

To further examine the information content of loans, Dahiya, Puri, and Saunders (2003) take a different approach and study the announcement of a *sale* of a borrower's loans by its lending bank. The termination of a banking relationship through a loan sale may convey a negative signal to the market about a firm's prospects. To test this hypothesis, the authors employ the event-study methodology, using a sample of loan sale announcements by the originating bank for sub-par loans. The data is collected by cross-matching loan sale information in two market newsletters (*Gold Sheets* and *Bank Letter*) with CRSP.[28] Using 3-day, 5-day, and 7-day event windows, the authors find a highly significant negative abnormal return of between $-1.61\%$ and $-8.11\%$. These results are consistent with a negative information effect arising from loan sales. The authors also perform two additional tests to examine if loan sales are valid signals of a bank's negative private information about the firm. First, the authors find that firms that have their loans sold are more likely to file for bankruptcy than other comparable firms

trading day. The data used in previous studies does not allow for the authors to distinguish announcements that occur during or after the trading day, which forces the use of two-day event windows.

[27] See however, some contrary evidence in Billett, Flannery, and Garfinkel (2003) who find in the long run, bank loan announcements appear no different from seasoned equity or public debt issuance. Thus the positive abnormal return is subsequently followed by a negative return.

[28] Dahiya, Puri, and Saunders (2003) focus on the sales of seasoned sub-par loans, where the information effects of bank sales are likely to be highest and where they have a more representative sample of loan sales.

and firms that are performing poorly.[29] Second, the authors find that firms that have their loans sold are not the worst performers in their respective industries during the year before their loan was sold, indicating that ex-ante, publicly available information alone may not have allowed outsiders to identify the true weakness of these firms. These results support the view that loan sales by the original lender provide negative private information to outside investors.

### *6.2. Non underwriter-bank loans and public security pricing*

In addition to conveying private information to the market through announcements and sales, the analysis in Fama (1985) and Diamond (1991) indicates that bank loans reduce the need for outsiders to generate duplicate information. This may allow bank lending to reduce overall information costs. Further, bank loans can help resolve information asymmetries between management and outside investors that could induce managers to refrain from issuing equity and foregoing positive net present value investments (Myers and Majluf, 1984). One way to test if bank loans reduce information-related costs is to empirically examine the impact of existing loans on the pricing of a firm's public security issuance. If bank loans reduce information costs, then the existence of bank loans should result in higher security pricing.

James and Weir (1990) investigate how an established relationship with a bank affects IPO underpricing. The authors develop a theoretical model that predicts that due to the information benefits of having an existing lending relationship, firms with an established lending relationship will experience less severe underpricing when they go public. To test the model, the authors collect a sample of 549 IPOs for non-financial firms that occurred between 1980 and 1983 and identify 417 firms with existing bank borrowing relationships. After controlling for factors that have been identified to affect underpricing (i.e. the reputation of the underwriter, the age of the firm, the offering size, and the shares offered by insiders), having bank loans outstanding or a bank-credit agreement significantly reduces underpricing by 8.5 percent. This result suggests that bank loans reduce the information costs associated with issuing public securities.[30]

In addition to having a positive effect on equity issuances, existing lending relationships may also reduce the costs of public debt financings. Datta, Iskandar-Datta, and Patel (1999) examine this possibility. The authors argue that if banks have superior monitoring ability, then the presence of a bank lending relationship should lower information costs associated with raising public debt, which will be reflected through a lower

[29] For this test, the authors are able to expand their sample to 53 firms that have a sub-par loan sold. They were forced to use the smaller sample of 15 loan sales in the event-study due to missing information on the precise date of sale.

[30] However, the authors do not find a statistically significant difference between the effect of bank loans and long term debt on IPO underpricing, which does not support the hypothesis that bank loans play a special role in reducing information costs for IPOs.

at-issue yield spread (higher price) of a firm's first public debt issuance. The authors focus on first debt issuances because firms issuing seasoned debt are already monitored by public debtholders, which can make it difficult to distinguish if the private banking relationship drives any results. Further, first public debt offers are undertaken by younger and smaller firms, where asymmetric information is likely to be high. Datta, Iskandar-Datta, and Patel (1999) collect a sample of 98 initial public offerings of straight debt that occurred between 1971 and 1994 and determine that 64 firms have bank debt at the time of public issuance. In their main specification, after controlling for firm and bond characteristics, as well as differences in risk, the authors find that the existence of a bank lending relationship during the year prior to the public debt issuance reduces the at-issue yield spread of the first public bond offering by 84 basis points.[31]

The findings in these two papers support the view that the existence of a bank relationship reduces the information costs of accessing the public equity and debt markets. The results are consistent with bank lending agreements being valuable because the existence of a banking relationship increases a firm's public security prices. These findings complement the evidence on the market's reaction to loan announcements, renewals, and sales in that they emphasize the importance of the information content of bank loans.

## 7. Extensions

Thus far, we have confined ourselves to the interaction of banks and capital markets. There are, of course, many interesting and important areas in banking and financial intermediation that are not directly related to the main focus of this survey but where additional research is needed.[32] Some of these areas are mentioned here.

### *7.1. Banks as equity holders*

An area where banks might expand their activities but traditionally have not been allowed to, at least in the U.S., is in holding equity stakes. As opposed to Japan and Germany, where banks are allowed to hold equity, banks in the U.S. are allowed to hold equity only though restructuring bad loans (see e.g. James, 1995), or through some provisions in venture capital (see e.g. Hellmann, Lindsey, and Puri (2006)).[33] As previously noted, there are efficiencies from underwriters holding equity in firms; venture

[31] The results in Datta, Iskandar-Datta, and Patel (1999) are somewhat mixed. In another specification, they find a significantly *positive* relationship between a banking relationship and the at-issue yield spread when they only control for firm characteristics and if the bond has a call provision. The negative relationship arises once the authors control for subordination and covenants.

[32] For an excellent survey of many of the other areas in financial intermediation not covered here, see Gorton and Winton (2002).

[33] The advantages and disadvantages of allowing banks to hold equity have been analyzed in a number of models (see e.g. Berlin, John, and Saunders, 1996; Stiglitz, 1985; Winton, 2003). Empirical work on bank

capitalists as equity holders reduce IPO underpricing when they underwrite and, further, gross spreads on IPOs decrease in the underwriters' shareholdings of the firm (Li and Masulis, 2004).

Why and when do banks choose to invest in equity, and what are the implications for the firm? There is surprisingly little research on this issue. Hellmann, Lindsey, and Puri (2006) explore this topic by focusing on the impact of bank venture capital relationships on the bank's core lending division. Venture relationships may allow the bank to foster an ongoing lending relationship. Also, the private information from the venture relationship may reduce the bank's cost of lending due to informational economies of scope, allowing firms to benefit from lower loan yield spreads. To examine these issues, the authors collect detailed information on all venture capital investments for the period 1980 through 2000 and gather lending data for the 10,583 venture backed customers. To examine if venture relationships increase the likelihood that the bank investor will forge a lending relationship with the firm, the authors estimate a conditional logit model in which each firm can choose among banks. The results reveal that the venture relationship does indeed increase the likelihood of being selected as lender, even after controlling for the bank's share of the lending market and the firm's public status. These results are confirmed through another test at the aggregate level in which the authors find that banks are more likely, on average, to lend to companies with whom they have a prior venture relationship.

To examine loan pricing, Hellmann, Lindsey, and Puri (2006) match loans where the lender has a prior venture relationship with the firm ("relationship loans") with similar loans where no venture relationship exists ("non-relationship loans") and compare the yield spreads of the matched loans. Since it is difficult to match loans directly based on multiple relevant characteristics, the authors use propensity score matching, which reduces the multiple-dimension matching problem to a single-dimension, called the propensity score.[34] These methods take into account the fact that the characteristics of relationship loans may differ significantly from non-relationship loans and ensure that such observed characteristics are not driving the results. Using various estimators, the authors find that relationship loans have significantly lower yield spreads, by 18 to 26 basis points. In sum, the results suggest that as venture capitalists, banks tend to be strategic investors in equity and use venture capital relationships to foster a lending relationship that results in efficiencies that benefit both banks and firms. The results

control rights through board seats and equity holding is found in Germany by Gorton and Schmid (2000), who find that banks use their equity holding and board seats to improve firm performance. For Japan, Kaplan and Minton (1994) find banks are more likely to get board seats following poor firm performance, and Weinstein and Yafeh (1998) and Morck, Nakamura, and Shivdasani (2000) find that Japanese firms with a main bank have lower growth and profitability than others.

[34] To employ the methodology, the authors first run a probit model, where the dependent variable is one if the loan is a relationship loan and zero otherwise, and the independent variables are the matching dimensions, which include loan and borrower characteristics. Each loan is assigned a propensity score, which is the predicted probability from the probit model. See Heckman, Ichimura, and Todd (1997, 1998) for more details. This method of matching loans is also used in Drucker and Puri (2005).

highlight the impact of organizational form on the incentives and behavior of investors. This is an area worthy of more study.

### *7.2. Beyond screening and monitoring*

The central idea behind much of the banking literature is that banks have access to private information about the firm. The bank's ability to generate information has implications on the firm's financing decision. Can banks play other roles for firms that go beyond screening and monitoring? In studies of venture capitalists, there is some evidence that venture capitalists do not simply screen and monitor but also help provide costly effort in the form of support activities for the firm. When financing the firm, venture capitalists expect to help the founder professionalize the company (Kaplan and Stromberg, 2001, 2004). Also, firms financed by venture capitalists are more likely to professionalize early and are more likely to get their product to market (Hellmann and Puri, 2000, 2002). Lerner (1995), Baker and Gompers (2003), and Hochberg (2004) find that venture capitalists play an important role in determining the composition of the board of directors. Lindsey (2004) finds some evidence that venture capitalists facilitate strategic alliances of firms within their portfolio. Of course, one could argue that the main difference between banks and venture capitalists is that banks typically provide only debt financing while venture capitalists have equity-based contracts. However, in many countries around the world, banks are not prohibited from taking equity stakes. Yet, other than some evidence on banks' role on boards of directors (see e.g. Kroszner and Strahan, 2001), we have little evidence that banks play a support role for their borrowers. Is there limited evidence because banks do not provide these services or simply because this possibility has not been explored by researchers? Again, this is an area where more research is warranted.

### *7.3. Loan sales*

The loan sales market is rapidly growing, and loans sales are a major source of funding for banks and a way for banks to manage risk.[35] While a number of studies have formed and tested theories of the loan sales market, a consensus has not been reached on the functioning of this very important market. We summarize the literature on the two prevailing information-based theories of loan sales – the "monitoring technology hypothesis", and the "comparative advantage hypothesis".[36] We also provide some recent evidence on the effects of loan sales on corporate borrowers.

[35] See Gorton and Haubrich (1990) for early evidence on loan sales and Yago and McCarthy (2004) and Thomas and Wang (2004) for more recent developments in the loan sales market. See Gorton and Haubrich (1990) for empirical evidence on loan sales market size and trends.

[36] Berger and Udell (1993) provide details on nine competing theories. Non-information based theories include the "diversification hypothesis", which claims that loan sales provide a way for banks with limited opportunities to diversify their loan portfolio (Demsetz, 2000; Haubrich and Thomson, 1996; Pavel and Phillis, 1987), the "regulatory tax hypothesis", which suggests that regulatory taxes on on-balance sheet

Berger and Udell (1993) develop the "monitoring technology hypothesis", which attempts to explain loan sales as a reaction to improvements in monitoring and information technology.[37] As information technology improves, banks can sell loans to direct lenders because these loan buyers increase their ability to monitor loans. For high quality borrowers, the monitoring cost advantage of banks falls below the signaling costs of intermediation, which enables the sale of loans. An important implication of the theory is that banks keep risky, essentially illiquid loans for which their monitoring advantage is important. Berger and Udell (1993) find empirical support.

Drucker and Puri (2006) also find evidence consistent with the monitoring technology hypothesis. Using a sample of loans that are originated between 1999 and 2004, the authors identify individual loans that are traded in the secondary market. They find that banks sell the loans of more informationally transparent borrowers—larger firms who have long-term debt credit ratings. The monitoring cost advantage of banks is presumably smaller for these types of loans. Interestingly, sold loans have additional, tighter financial covenants as compared with loans to similar firms which are not traded in the secondary market. This is consistent with loan buyers directly monitoring borrowers through covenants.

The "comparative advantage hypothesis" argues that loan sales arise out of exogenous differences in the comparative advantages of financial intermediaries. Researchers have explored a number of different comparative advantages that could motivate loan sales.[38] Hess and Smith (1994) claim that banks may have a comparative advantage in originating and servicing loans but not in funding or interest risk management. Pavel and Phillis (1987) provide empirical support, showing that banks with origination and servicing advantages have a higher probability of selling loans and also sell more loans. Carlstrom and Samolyk (1995) assume that banks have an advantage in finding and screening profitable local projects and loan sales arise because, otherwise, financially constrained banks would have to pass up positive investments when there were many good opportunities in the local market. Some empirical studies have supported this theory, as a typical bank with a binding capital constraint is more likely to sell a higher

activities result in banks using off-balance sheet activities, such as loan sales (Pennacchi, 1988; Pavel and Phillis, 1987), the "collateralization hypothesis", in which loan sales provide a mechanism to shift risk from risk-averse to risk-neutral investors (Benveniste and Berger, 1986, 1987) or to help avoid debt overhang for banks (James, 1988), and the "moral hazard hypothesis", which suggests that banks use loan sales to book income immediately and increase leverage to take advantage of deposit insurance (Benveniste and Berger, 1986; James, 1988).

[37] This theory is an extension of Bhattacharya and Thakor (1993), who find that intermediary monitoring dominates direct monitoring when the benefits from scale economies in monitoring exceed the costs of signaling the value of assets to investors.

[38] In addition to the information-based comparative advantages that are discussed here, Pennacchi (1988) discusses another comparative advantage that is based on funding differences between banks. Loan sales provide a means by which the inexpensive funds that are raised by some banks can be used to finance the loans at other, higher cost banks. The empirical evidence on this non-information-based view is mixed (see e.g. Berger and Udell, 1993; Haubrich and Thomson, 1996).

proportion of loans than an unconstrained bank (Haubrich and Thomson, 1996; Pavel and Phillis, 1987), unconstrained banks are more likely to buy loans (Demsetz, 2000), and strong local origination opportunities are positively related to loan selling (Demsetz, 1994, 2000).

Recent papers have explored the effect of loan sales on corporate borrowers. It may be costly for borrowers to have their loans sold, particularly if they need to renegotiate their loans in the future, as they will have to deal with additional lenders that may not take a long-term view of the company's prospects. Further, there is a concern that loan sales harm lending relationships. Guner (2006) examines if borrowers receive an offsetting benefit through lower loan interest rates. Guner (2006) identifies banks that were active loan sellers during the 1987 through 1993 period and shows that borrowers of these banks indeed received lower loan interest rates. Importantly, the interest rate reductions are concentrated among borrowers that are more likely to have their loans sold based on ex ante characteristics. Drucker and Puri (2006), using data that covers the time period 1999 through 2004, show that borrowers whose loans are sold receive additional bank loans, both in the year of the loan sale and in the future. These results are consistent with loan selling increasing borrowers' access to bank loans. Contrary to concerns that lending relationships are harmed by loan selling, Drucker and Puri (2006) show that borrowers whose loans are sold are more likely to retain their lending relationships. One explanation is that loan sales let banks manage their lending risks up front, which permits banks to extend loans to their relationship borrowers in the future.

### *7.4. Bank organizational form*

There is a growing amount of work on the nature of information collected by banks from their clients, and on how the organizational form of the bank may be more conducive to collecting some kinds of information as opposed to others. A key empirical finding is that large banks tend to lend to large companies and small banks tend to lend to small companies (see Berger, Kashyap, and Scalise, 1995; Berger et al., 1998; Berger and Udell, 1996; Nakamura, 1994; Peek and Rosengren, 1996; Strahan and Weston, 1996, 1998; Sapienza, 2002). Stein (2002) argues that the key difference between small and large business lending is that small business lending relies on "soft" information, which is information that cannot be directly verified by anyone other than the agent who produces the information. Small banks are better at processing soft information while large banks are better at processing verifiable "hard" information, such as financial statements, public credit ratings, and formalized records.

Since research shows that relationships are important for small companies (see e.g. Petersen and Rajan, 1994), it is vital to understand the effects on small firms of the growth in the size of banks and the increased reliance on hard information. There are a few empirical papers that examine the role of hard and soft information in the credit decisions of banks. Liberti (2002) examines a hierarchical structure change in a corporate commercial lending division of a foreign bank in Argentina. He finds that managers with more independence base their pricing decision more heavily on soft information than managers with more limited decision authority. Berger and Udell (1996) show that large banks do not reduce credit to small firms whose credit worthiness can be judged by examining hard information, such as their financial ratios. Another study by Cole, Goldberg, and White (1999) uses the National Survey of Small Business Finances to examine the decision by banks to accept or reject credit applications by small firms. They find that larger banks make credit allocations based on standard hard information criteria, such as figures that can be obtained in financial statements. Mian (2004) finds evidence consistent with foreign banks with larger distance between their head offices and local branches avoiding informationally difficult credit, where soft information is likely to be more important. Also, Berger et al. (2005) use a sample of small business loans and find that firms with financial records borrow from banks that are larger, on average.

### *7.5. Bank-based vs. market-based economies*

Many economies are largely bank-dependent and capital markets are not well developed. Are banks and stock markets substitutes? This remains an important question. There has been some theoretical work on this subject. Allen (1993) and Allen and Gale (1999) argue that banks and stock markets fundamentally differ in the way that they process information, in that banks are inherently more conservative. Thus, stock market based economies are more likely to embrace new technologies. In contrast, Dow and Gorton (1997) argue that banks and stock markets are alternative institutions for the savings/investment process. There is now growing empirical research on bank based and stock market based systems.[39] The bulk of evidence seems to suggest that both financial intermediaries and markets matter for growth. However, the results are far from conclusive and more research is needed.

## 8. Concluding remarks

There has been a large amount of research on the implications of allowing banks to expand their activities beyond traditional lending into underwriting. There is convincing evidence that, at least in the United States, commercial banks do not suffer from conflicts of interest and can be net certifiers of firm value when underwriting public securities. This is seen through the ex ante pricing and ex post performance of commercial bank-underwritten securities. The results are robust across different time periods, different securities, and the use of different empirical methodology. The international evidence on conflicts of interest from commercial bank underwriting is mixed. However, the discrepancies may be partially explained by the varying regulatory environments and quality of the financial markets in these countries. Future research will benefit from empirical tests that explicitly account for these differences.

[39] See Levine (2004) for an excellent review.

Many empirical studies document that, in both debt and equity offerings, borrowers receive lower underwriting fees when they use their lending bank as underwriter. Both prior lending and concurrent lending increases the likelihood that the bank will win underwriting mandates. These results seem to imply that commercial banks will crowd out specialized investment banks. However, recent evidence suggests that investment banks are competing with commercial banks by developing lending units. Investment bank lending raises serious issues for regulators, yet there is limited evidence on the consequences of investment bank lending. More research is needed in this area.

Banks also play an indirect role in capital markets. Empirical studies of stock market reaction to loan initiations, renewals, and sales confirm that banks can signal the quality of firms to outside investors through their lending decisions. Additional evidence suggests that the existence of a bank lending relationship reduces the costs of information acquisition for capital market participants.

Overall, there are positive effects from the interaction between commercial banks' lending activities and the capital markets. However, banks in the U.S. are allowed to hold equity only though restructuring bad loans or through venture capital investments. Should commercial banks in the U.S. be allowed to expand their ability to hold equity holdings of firms? In general, we observe that there is some evidence of efficiencies from financial intermediaries being able to hold equity stakes. For example, when commercial banks hold equity in firms through venture capital subsidiaries, they foster an ongoing lending relationship that results in efficiencies that benefit firms through lower loan pricing. Would there be positive effects from the interaction between commercial banks' equity holdings and capital markets? When banks hold equity in firms do they provide value added services similar to those provided by venture capitalists, such as professionalization or support in the human resources area? What can we learn from examining the effects of commercial bank equity holding in other countries, such as Germany and Japan? These issues are yet unresolved and promise to be at the forefront of continued regulatory debate on the scope of bank activities.

## References


Allen, F., 1993. Stock markets and resource allocation. In: Vives, C., Vives, X. (Eds.), Capital Markets and Financial Intermediation. Cambridge University Press, Cambridge, pp. 81–104.

Allen, F., Faulhaber, G., 1989. Signaling by underpricing in the IPO market. Journal of Financial Economics 23, 303–323.

Allen, F., Gale, D., 1999. Diversity of opinion and financing new technologies. Journal of Financial Intermediation 8, 68–89.

Altinkilic, O., Hansen, R., 2000. Are there economies of scale in underwriting fees? Evidence of rising external financing costs. Review of Financial Studies 13, 191–218.

Altman, E., 1989. Measuring corporate bond mortality and performance. Journal of Finance 44, 909–922.

Ang, J., Richardson, R., 1994. The underwriting experience of commercial bank affiliates prior to the Glass–Steagall Act: A re-examination of evidence for passage of the act. Journal of Banking and Finance 18, 351–395.

Asquith, P., Mullins, D., 1986. Equity issues and offering dilution. Journal of Financial Economics 15, 73–85.

Baker, M., Gompers, P., 2003. The determinants of board structure at the time of the IPO. Journal of Law and Economics 46, 569–598.
Benston, G.J., 1990. The Separation of Commercial and Investment Banking. Oxford University Press, New York.
Benveniste, L., Berger, A., 1986. An empirical analysis of standby letters of credit. In: Proceedings of the 16th Conference on Bank Structure and Competition, pp. 387–412.
Benveniste, L., Berger, A., 1987. Securitization with recourse: An instrument that offers uninsured bank depositors sequential claims. Journal of Banking and Finance 11, 403–424.
Benveniste, L., Spindt, P., 1989. How investment bankers determine the offer price and allocation of new issues. Journal of Financial Economics 24, 343–346.
Benveniste, L., Wilhelm, W., 1990. A comparative analysis of IPO proceeds under alternative regulatory environments. Journal of Financial Economics 28, 173–207.
Benzoni, L., Schenone, C., 2004. Conflict of interest or certification? Evidence from IPOs underwritten by the firm's relationship bank. Working Paper, University of Minnesota.
Ber, H., Yafeh, Y., Yosha, O., 2001. Conflict of interest in universal banking: Bank lending, stock underwriting, and fund management. Journal of Monetary Economics 47, 189–218.
Berger, A., Demsetz, R., Strahan, P., 1999. The consolidation of the financial services industry: Causes, consequences, and implications for the future. Journal of Banking and Finance 23, 135–194.
Berger, A., Kashyap, A., Scalise, J., 1995. The transformation of the U.S. banking industry: What a long strange trip it's been. Brookings Papers on Economic Activity 2, 55–201.
Berger, A., Miller, N., Petersen, M., Rajan, R., Stein, J., 2005. Does function follow organizational form? Evidence from the lending practices of large and small banks. Journal of Financial Economics 76, 237–269.
Berger, A., Saunders, A., Scalise, J., Udell, G., 1998. The effects of bank mergers and acquisitions on small business lending. Journal of Financial Economics 50, 187–229.
Berger, A., Udell, G., 1993. Securitization, risk, and the liquidity problem in banking. In: Klausner, M., White, L. (Eds.), Structural Change in Banking. Irwin, Homewood, IL.
Berger, A., Udell, G., 1996. Universal banking and the future of small business lending. In: Saunders, A., Walter, I. (Eds.), Financial System Design: The Case for Universal Banking. Irwin, Burr Ridge, IL, pp. 559–627.
Berlin, M., John, K., Saunders, A., 1996. Bank equity stakes in borrowing firms and financial distress. Review of Financial Studies 9, 889–920.
Best, R., Zhang, H., 1993. Alternative information sources and the information content of bank loans. Journal of Finance 48, 1507–1523.
Bharath, S., Dahiya, S., Saunders, A., Srinivasan, A., 2004. So what do I get? The bank's view of lending relationships. Working Paper, New York University.
Bhattacharya, S., Thakor, A., 1993. Contemporary banking theory. Journal of Financial Intermediation 3, 2–50.
Billett, M., Flannery, M., Garfinkel, J., 1995. The effect of lender identity on a borrowing firm's equity return. Journal of Finance 50, 699–718.
Billett, M., Flannery, M., Garfinkel, J., 2003. Are bank loans special? Evidence on the post-announcement performance of bank borrowers. Working Paper, University of Florida.
Boot, A., Thakor, A., 1997a. Banking scope and financial innovation. Review of Financial Studies 10, 1099–1131.
Boot, A., Thakor, A., 1997b. Financial system architecture. Review of Financial Studies 10, 693–733.
Boyd, J., Prescott, E., 1986. Financial intermediary-coalitions. Journal of Economic Theory 38, 211–232.
Carey, M., 1995. Are bank loans mispriced? Working Paper, Board of Governors of the Federal Reserve System.
Carlstrom, C., Samolyk, K., 1995. Loan sales as a response to market-based capital constraints. Journal of Banking and Finance 19, 627–646.
Carosso, V., 1985. Comment. In: Walter, I. (Ed.), Deregulating Wall Street. Wiley, New York, pp. 89–91.

Chiappori, P., Salanie, B., 2000. Testing for asymmetric information in insurance markets. Journal of Political Economy 108, 56–78.

Cole, R., Goldberg, L., White, L., 1999. Cookie-cutter vs. character: The microstructure of small business lending by large and small banks. In: Blanton, J., Williams, A., Rhine, S. (Eds.), Federal Reserve Bank of Chicago Conference on Business Access to Capital and Credit, pp. 362–389.

Dahiya, S., Puri, M., Saunders, A., 2003. Bank borrowers and loan sales: New evidence on the uniqueness of bank loans. Journal of Business 76, 563–582.

Datta, S., Iskandar-Datta, M., Patel, A., 1999. Bank monitoring and the pricing of corporate public debt. Journal of Financial Economics 51, 435–449.

Demsetz, R., 1994. Evidence on the relationship between regional economic conditions and loan sales activity. In: Proceedings of the 30th Annual Conference on Bank Structure and Competition, pp. 370–380.

Demsetz, R., 2000. Bank loan sales: A new look at the motivations for secondary market activity. Journal of Financial Research 23, 192–222.

Diamond, D., 1984. Financial intermediation and delegated monitoring. Review of Economic Studies 51, 393–414.

Diamond, D., 1991. Monitoring and reputation: The choice between bank loans and directly placed debt. Journal of Political Economy 99, 689–721.

Dow, J., Gorton, G., 1997. Stock market efficiency and economic efficiency: Is there a connection? Journal of Finance 52, 1087–1129.

Drucker, S., Puri, M., 2005. On the benefits of concurrent lending and underwriting. Journal of Finance 60, 2763–2799.

Drucker, S., Puri, M., 2006. On loan sales, loan contracting, and leading relationships. Working paper, Duke University.

Eckbo, B., 1986. Valuation effects of corporate debt offerings. Journal of Financial Economics 15, 119–152.

Fama, E., 1985. What's different about banks? Journal of Monetary Economics 10, 10–19.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fang, L., 2005. Investment bank reputation and the price and quality of underwriting services. Journal of Finance 60, 2729–2761.

Gande, A., Puri, M., Saunders, A., 1999. Bank entry, competition, and the market for corporate securities underwriting. Journal of Financial Economics 54, 165–195.

Gande, A., Puri, M., Saunders, A., Walter, I., 1997. Bank underwriting of debt securities: modern evidence. Review of Financial Studies 10, 1175–1202.

Gompers, P., Lerner, J., 1999. Conflict of interest in the issuance of public securities: Evidence from venture capital. Journal of Law and Economics 42, 1–28.

Gorton, G., Haubrich, J., 1990. The loan sales market. In: Kaufman, G. (Ed.), Research in Financial Services: Private and Public Policy, vol. 2. JAI, Greenwich.

Gorton, G., Schmid, F., 2000. Universal banking and the performance of German firms. Journal of Financial Economics 58, 29–80.

Gorton, G., Winton, A., 2002. Financial Intermediation. NBER Working Paper 8928. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. North Holland, Amsterdam.

Grinblatt, M., Hwang, C., 1989. Signaling and the pricing of new issues. Journal of Finance 44, 393–420.

Güner, A.B., 2006. Loan sales and the cost of corporate borrowing. Review of Financial Studies 19, 687–716.

Hamao, Y., Hoshi, T., 2002. Bank underwriting of corporate bonds: Evidence from Japan after the financial system reform of 1993. Working Paper, University of Southern California.

Haubrich, J., Thomson, B., 1996. Loan sales, implicit contracts, and bank structure. Review of Quantitative and Financial Accounting 7, 137–162.

Hebb, G., Fraser, D., 2002. Conflict of interest in commercial bank security underwritings: Canadian evidence. Journal of Banking and Finance 26, 1935–1949.

Heckman, J., Ichimura, H., Todd, P., 1997. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. Review of Economic Studies 64, 605–654.

Heckman, J., Ichimura, H., Todd, P., 1998. Matching as an econometric evaluation estimator. Review of Economic Studies 65, 261–294.
Hellmann, T., Lindsey, L., Puri, M., 2006. Building relationships early: Banks in venture capital. Working Paper, Duke University.
Hellmann, T., Puri, M., 2000. The interaction between product market and financing strategy: The role of venture capital. Review of Financial Studies 13, 959–984.
Hellmann, T., Puri, M., 2002. Venture capital and the professionalization of start-up firms: Empirical Evidence. Journal of Finance 57, 169–197.
Hess, A., Smith, C., 1994. Elements of mortgage securitization. In: James, C. (Ed.), Studies in Financial Institutions: Commercial Banks. McGraw-Hill, New York.
Hochberg, Y., 2004. Venture capital and corporate governance in the newly public firm. Working Paper, Cornell University.
Hoshi, T., 1996. Bank to the future: Universal banking in Japan. In: Saunders, A., Walter, I. (Eds.), Universal Banking: Financial System Design Reconsidered. Irwin, Chicago, pp. 205–244.
Hoshi, T., Kashyap, A., 1999. The Japanese banking crisis: Where did it come from and how will it end? WBER Working Paper 7250.
James, C., 1987. Some evidence on the uniqueness of bank loans. Journal of Financial Economics 19, 217–235.
James, C., 1988. The use of loan sales and standby letters of credit by commercial banks. Journal of Monetary Economics 22, 395–422.
James, C., 1995. When do banks take equity? An analysis of bank loan restructurings and the role of public debt. Review of Financial Studies 8, 1209–1234.
James, C., Weir, P., 1990. Borrowing relationships, intermediation, and the cost of issuing public securities. Journal of Financial Economics 28, 149–171.
Kanatas, G., Qi, J., 1998. Underwriting by commercial banks: Incentive conflicts, scope economies, and project quality. Journal of Money, Credit, and Banking 30, 119–133.
Kanatas, G., Qi, J., 2003. Integration of lending and underwriting: Implications of scope economies. Journal of Finance 58, 1167–1191.
Kaplan, S., Minton, B., 1994. Appointments of outsiders to Japanese boards: Determinants and implications for managers. Journal of Financial Economics 36, 225–258.
Kaplan, S., Strömberg, P., 2001. Venture capitalists as principals: Contracting, screening, and monitoring. American Economic Review, Papers and Proceedings, 426–430.
Kaplan, S., Strömberg, P., 2004. Contracts, characteristics, and actions: Evidence from venture capital analyses. Journal of Finance 59, 2177–2210.
Kelly III, E., 1985. Legislative history of the Glass–Steagall Act, I. In: Walter, I. (Ed.), Deregulating Wall Street. Wiley, New York, pp. 231–254.
Konishi, M., 2002. Bond underwriting by banks and conflicts of interest: Evidence from Japan during the pre-war period. Journal of Banking and Finance 26, 767–793.
Kroszner, R., Rajan, R., 1994. Is the Glass–Steagall act justified? A study of the U.S. experience with universal banking before 1933. American Economic Review 84, 810–832.
Kroszner, R., Rajan, R., 1997. Organization structure and credibility: Evidence from commercial bank securities before the Glass–Steagall Act. Journal of Monetary Economics 39, 475–516.
Kroszner, R., Strahan, P., 2001. Bankers on boards: Monitoring, conflicts of interest, and lender liability. Journal of Financial Economics 62, 415–452.
Leland, H., Pyle, D., 1977. Informational asymmetries, financial structure, and financial intermediation. Journal of Finance 32, 371–387.
Lerner, J., 1995. Venture capitalists and the oversight of private firms. Journal of Finance 50, 301–318.
Levine, R., 2004. Finance and growth: Theory and evidence. NBER Working Paper, 10766.
Li, X., Masulis, R., 2004. Venture capital investments by IPO underwriters: Certification or conflict of interest? Working Paper, Vanderbilt University.
Liberti, J., 2002. Initiative, incentives and soft information: How does delegation impact the role of bank relationship managers? Working Paper, London Business School.

Lindsey, L., 2004. The venture capital Keiretsu effect: An Empirical analysis of strategic alliances among portfolio firms. Working Paper, Arizona State University.
Liu, W., Kang, J., 2004. Is universal banking justified? Evidence from bank underwriting of corporate bonds in Japan. Working Paper, Michigan State University.
Ljungqvist, A., Marston, F., Wilhelm, W., 2006. Competing for securities underwriting mandates: Banking relationships and analyst recommendations. Journal of Finance 61, 301–340.
Lummer, S., McConnell, J., 1989. Further evidence on the bank lending process and the capital market response to bank loan agreements. Journal of Financial Economics 25, 99–122.
Maddala, G., 1983. Limited Dependent and Qualitative Variables in Econometrics. Econometrics Society Monographs, vol. 3. Cambridge University, Cambridge.
Masulis, R., Korwar, A., 1986. Seasoned equity offerings: An empirical investigation. Journal of Financial Economics 15, 85–119.
Mayer, C., Vives, X. (Eds.), 1993. Capital Markets and Financial Intermediation. Cambridge University Press, Cambridge.
McFadden, D., 1973. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York.
Mester, L., 1996. Repealing Glass–Steagall: The past points the way to the future. Federal Reserve Bank of Philadelphia Business Review (July), 3–18.
Mian, A., 2004. Distance constraints: The limits of foreign lending in poor economies. Working Paper, University of Chicago.
Mikkelson, W., Partch, M., 1986. Valuation effects of securities offerings and the issuance process. Journal of Financial Economics 15, 31–60.
Morck, R., Nakamura, M., Shivdasani, A., 2000. Banks, ownership structure, and firm value in Japan. Journal of Business 73, 539–567.
Myers, S., Majluf, N., 1984. Corporate financing and investment decisions when firms have information that investors do not have. Journal of Financial Economics 13, 157–187.
Nakamura, L., 1994. Small borrowers and the survival of the small bank: Is mouse bank mighty or Mickey? Federal Reserve Bank of Philadelphia Business Review (November/December), 3–15.
Narayanan, R., Rangan, K., Rangan, N., 2004. The role of syndicate structure in bank underwriting. Journal of Financial Economics 72, 555–580.
Pavel, C., Phillis, D., 1987. Why commercial banks sell loans: An empirical analysis. Economic Perspectives of the Federal Reserve Bank of Chicago (July/August), 3–14.
Peek, J., Rosengren, E., 1996. Small business credit availability: How important is size of lender? In: Saunders, A., Walter, I. (Eds.), Financial System Design: The Case for Universal Banking. Irwin, Burr Ridge, IL, pp. 628–655.
Pennacchi, G., 1988. Loan sales and the cost of bank capital. Journal of Finance 43, 375–395.
Petersen, M., Rajan, R., 1994. The benefits of lending relationships: Evidence from small business data. Journal of Finance 49, 3–37.
Puri, M., 1994. The long-term default performance of bank underwritten security issues. Journal of Banking and Finance 18, 397–418.
Puri, M., 1996. Commercial banks in investment banking: Conflict of interest or certification role? Journal of Financial Economics 40, 373–401.
Puri, M., 1999. Commercial banks as underwriters: Implications for the going public process. Journal of Financial Economics 54, 133–163.
Rajan, R., 2002. An investigation into the economics of extending bank powers. Journal of Emerging Market Finance 1, 125–156.
Ramakrishnan, R., Thakor, A., 1984. Information reliability and a theory of financial intermediation. Review of Economic Studies 22, 415–432.
Ritter, J., 1998. Initial Public Offerings. In: Logue, D., Seward, J. (Eds.), Warren, Gorham, and Lamont Handbook of Modern Finance. WGL/RIA, Boston and New York, pp. 5–30.
Rock, K., 1986. Why new issues are underpriced. Journal of Financial Economics 15, 187–212.

Roten, I., Mullineaux, D., 2002. Debt underwriting by commercial bank-affiliated firms and investment banks: more evidence. Journal of Banking and Finance 26, 689–718.
Santomero, A., Eckles, D., 2000. The determinants of success in the new financial services environment. FRBNY Economic Policy Review (October), 11–23.
Sapienza, P., 2002. The effects of banking mergers on loan contracts. Journal of Finance 57, 329–368.
Saunders, A., Walter, I., 1994. Universal banking in the United States. What could we gain? What could we lose? Oxford University, New York.
Schenone, C., 2004. The effect of banking relationships on the firm's IPO underpricing. Journal of Finance 59, 2903–2958.
Slovin, M., Johnson, S., Glascock, J., 1992. Firm size and the information content of bank loan announcements. Journal of Banking and Finance 16, 1057–1071.
Song, W., 2004. Competition and coalition among underwriters: The decision to join a syndicate. Journal of Finance 59, 2421–2444.
Stein, J., 2002. Information production and capital allocation: Decentralized versus hierarchical firms. Journal of Finance 57, 1891–1922.
Stiglitz, J., 1985. Credit markets and the control of capital. Journal of Money, Credit and Banking 17, 133–152.
Strahan, P., Weston, J., 1996. Small business lending and bank consolidation: Is there cause for concern? Federal Reserve Bank of New York Current Issues in Economics and Finance 2, 1–6.
Strahan, P., Weston, J., 1998. Small business lending and the changing structure of the banking industry. Journal of Banking and Finance 22, 821–845.
Takaoka, S., McKenzie, C., 2006. The impact of bank entry in the Japanese corporate bond underwriting market. Journal of Banking and Finance 30, 59–83.
Thomas, H., Wang, Z., 2004. The integration of bank syndicated loan and junk bond markets. Journal of Banking and Finance 28, 299–329.
Ursel, N., Ljucovic, P., 1998. The impact of bank ownership of underwriters on the underpricing of initial public offerings. Canadian Journal of Administrative Science 15, 17–27.
Walter, I., 2004. Conflicts of interest and market discipline in financial services firms. In: Borio, C., et al. (Eds.), Market Discipline Across Countries and Industries. MIT Press, Cambridge, MA, pp. 175–186.
Welch, I., 1989. Seasoned offerings, imitation costs and the underpricing of initial public offerings. Journal of Finance 44, 421–449.
Welch, I., 1992. Sequential sales, learning and cascades. Journal of Finance 47, 695–732.
Welch, I., 1997. Why is bank debt senior? A thoery of asymmetry and claim priority based on influence costs. Review of Financial Studies 10, 1203–1236.
Weinstein, D., Yafeh, Y., 1998. On the costs of a bank-centered financial system: Evidence from the changing main bank relations in Japan. Journal of Finance 53, 635–672.
Winton, A., 2003. Institutional liquidity needs and the structure of monitored finance. Review of Financial Studies 16, 1273–1313.
Yago, G., McCarthy, D., 2004. The U.S. Leveraged Loan Market: A Primer. Milken Institute.
Yasuda, A., 2006. Bank relationships and competition in the underwriting market: Evidence from Japan. Journal of Financial Economics, in press.
Yasuda, A., 2005. Do bank relationships affect the firm's underwriter choice in the corporate-bond underwriting market? Journal of Finance 60, 1259–1292.

*Chapter 6*

# SECURITY OFFERINGS*

B. ESPEN ECKBO

*Tuck School of Business, Dartmouth College, Hanover, NH 03755, USA*
*e-mail: b.espen.eckbo@dartmouth.edu*

RONALD W. MASULIS

*Owen Graduate School of Management, Vanderbilt University, 401 21st Avenue South, Nashville, TN 37203, USA*
*e-mail: ronald.masulis@owen.vanderbilt.edu*

ØYVIND NORLI

*Norwegian School of Management BI, Nydalsveien 37, N-0442 Oslo, Norway*
*e-mail: oyvind.norli@bi.no*

## Contents



* We thank Robert S. Hansen, Jay Ritter and Xueping Wu for suggestions.

*Handbook of Corporate Finance, Volume 1*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01006-3*

**Abstract**

This essay surveys the extant literature and adds to the empirical evidence on issuance activity, flotation costs, and valuation effects of security offerings. We focus primarily on public offerings of equity for cash, although we also review and present new evidence on debt offerings and private placements. The essay has four major parts: (1) We review aggregate issue activity in exchange listed securities from 1980 through 2004. Following the IPO, only about one-half of the publicly traded firms undertake a public security offering of *any* type, and only about one-quarter undertake a SEO. Thus, SEOs are relatively rare, which is consistent with adverse selection costs being an important consideration when raising cash externally. (2) We review the evidence on direct issue costs across security types and flotation methods, including the more recent SEO underpricing phenomenon. A large number of studies provide evidence on the determinants of underwriter compensation, and confirm the importance of variables capturing information asymmetries and underwriter competition. (3) We survey and interpret the valuation effects of security issue announcements. In the period since the Eckbo and Masulis (1995) survey, many studies examining announcement-period stock returns have focused on the effects of flotation method choice and foreign offerings. The well-known negative average announcement effect observed for U.S. SEOs appears to be a somewhat U.S.-specific phenomenon. (4) We review and extend evidence on the performance of issuing firms in the five year post-issue period. The literature proposes either a risk based-explanation or a behavioral explanation for the phenomenon of low average realized returns following IPOs and SEOs. Standard factor model regressions fail to reject the null that the low average returns are commensurate with issuers' risk exposures. Recent theoretical developments suggest that lower risk levels following equity issues may be linked to issuers' investment activity, a promising direction for future research.

**Keywords**

security offering, IPO, SEO, debt offer, flotation method, underwriting, rights offer, private placement, shelf registration, adverse selection, announcement returns, long run performance

## 1. Introduction

Security offerings are a very visible and important activity in the life of a firm. Their visibility arises in part because of the typically large amount of new capital raised relative to an issuer's existing capital base or asset size. The motives for security offerings are quite varied. The most common reason given for these actions is to raise capital for capital expenditures and new investment projects. Other reasons explored in the literature include the need to refinance or replace existing or maturing securities, to modify a firms capital structure, to exploit private information about securities intrinsic value, to exploit periods when financing costs are historically low, to finance mergers and acquisitions, to facilitate asset restructuring such as spin-offs and carve-outs, to shift wealth and risk bearing among classes of securities, to improve the liquidity of existing securities, to create more diffuse voting rights and ownership, to strengthen takeover defenses and to facilitate blockholder sales, privatizations, demutualizations and reorganizations.

This survey focuses exclusively on security offerings *for cash*, and then primarily to the *public*—although we also track private placements to some extent. Non-cash offerings, such as securities issued as employee compensation, and the many variants of security swaps, are covered elsewhere in this Handbook. For example, stocks issued as part of employee compensation plans are covered extensively in Aggarwal (2007, Chapter 17). Equity-for-equity swaps associated with mergers and takeovers are evidenced in Betton, Eckbo, and Thorburn (2007, Chapter 15). Security swaps associated with financial restructurings of non-distressed firms are covered in Eckbo and Thorburn (2007, Chapter 16), and senior-for-junior security swaps by firms in financial distress are examined in Hotchkiss et al. (2007, Chapter 14).

The decision to issue securities draws on all of the core areas in financial economics: asset pricing theory, capital structure theory, managerial investment incentives, financial institutions, contracting, and corporate governance. Moreover, there is a wealth of available data, particularly with the emergence in the 1990s of the comprehensive, machine-readable, transactions-oriented data base provided by the Security Data Corporation (SDC), with data back to 1980. Yet, there is surprisingly little consensus on key determinants of the security issuance decision and its economic effects on the firm.

The very existence of elaborate schemes for marketing security offerings to the public—including book building and road shows by underwriters—speaks to the importance of information asymmetries in the market for public issues. Moreover, judging from the recent regulatory focus on investor protection (e.g., the Sarbanes–Oxley Act of 2002), public security offerings for cash are relatively vulnerable to potential conflicts of interests. As such, these security issues are also the prime empirical laboratory for exploring models of capital structure choice—including the "pecking order" of (Myers, 1984)—as well as selling-mechanism designs that presume the public is substantially less informed than the issuer about the true value of the security issued.[1] While the survey provides information on the number of initial public offerings (IPOs) and private

[1] Time series evidence on the pecking order theory is surveyed in Frank and Goyal (2007, Chapter 12).

placements, the main focus is on issuances by exchange-listed firms—both seasoned equity offerings (SEOs) and debt issues.

We have four main objectives: (1) To survey the level of aggregate security issue activity and some of the characteristics of issuing firms; (2) to review direct issue costs across security types and selling mechanisms; (3) to survey and interpret the valuation effect of security issue announcements; and (4) to review and extend evidence on the performance of issuing firms in the five year post-issue period.

Mapping out the SDC data base, we start by providing an overview of aggregate issue activity in the U.S. over the period 1980–2003. We separate industrial firms from public utilities, and financial issuers from non-financial companies. We track primarily the largest security classes, such as common stock (IPOs, SEOs, and private placements) and debt (both straight and convertible), but provide some information on unit offerings, dual offerings, and foreign offerings (ADR and GDR) as well. We review potential determinants of the wave-like pattern of aggregate security offerings. At the firm level, we review evidence that links the security offering frequency through time. This includes the time period between the IPO and the first follow-on SEO, between two successive SEOs, and between debt and equity issues. Overall, this evidence confirms and generalize the early finding of Mikkelson and Partch (1986) that equity issues for cash are rare—both on an absolute level and relative to public debt issues.

Our second objective is to survey the nature and magnitude of direct issue costs, including the more recent phenomenon of SEO underpricing.[2] At the most basic level of economic analysis, firms minimize direct costs of raising capital. Yet, surprisingly few papers try to estimate the direct issue cost function. Following the adverse selection model of Myers and Majluf (1984), the literature has been preoccupied with the potential for wealth transfer caused by security offerings. We confirm the conclusion of Eckbo and Masulis (1995) that the adverse selection framework is the leading theoretical explanation for the announcement-induced abnormal stock returns for seasoned public offerings of debt and equity. However, the current evidence does not rule out the influence of direct transaction costs on a firm's issue decision, but is less supportive of wealth transfer concerns.

Understanding issue costs and the issue decision requires a thorough understanding of alternative selling mechanisms. We review how different selling mechanisms are designed to deal with different forms of information asymmetry, and the associated total issue costs. The literature here is sparse, leaving the link between contracting theory and optimal selling mechanisms design a fertile area for future research. One area in which this has immediate practical importance is in the choice between auctions and firm-commitment underwriting (fixed price) offerings, as witnessed in the recent Google IPO. Establishing the efficiency of the auction mechanism is also essential to the literal interpretation of an offering-price discount (underpricing) as "money left on the table" for shareholders of the issuing firm (Loughran and Ritter, 2002).

[2] We touch only briefly on IPO underpricing, which is the topic of Ljungqvist (2007, Chapter 7).

A third major objective of the survey is to both review and provide additional evidence on short- and long-term performance of issuing firms. In the period after the review of Eckbo and Masulis (1995), studies reporting short-term, announcement-period abnormal stock returns have focused in particular on the effect of the flotation method choice and of foreign offerings. Interestingly, the well-known negative announcement effect of the average SEO in the U.S. appears to be somewhat of a U.S.-specific phenomenon. While Eckbo and Masulis (1995) did not cover long-run performance studies, in this survey we provide our own large-scale analysis in addition to surveying the evidence in existing studies.

As in Loughran and Ritter (1995) and Eckbo and Norli (2004), we find that *total* returns are relatively low following security offerings, and in particular following IPOs. The low post-issue total return is most noticeable after IPO clusters ("hot" IPO periods). These clusters raise issues concerning selection bias and what Shultz (2003) terms "pseudo-timing" evidence. Overall, consistent with the conclusions of Eckbo, Masulis, and Norli (2000), Brav, Geczy, and Gompers (2000) and Eckbo and Norli (2005), but contrary to the inference Ritter (2003) draws from his survey, we conclude that the preponderance of the evidence fails to reject the hypothesis of zero *abnormal* returns in the post-issue period. This conclusion is robust to alternative definitions of expected returns, and it holds whether the issue is an IPO, a SEO, a private placement, or a (straight or convertible) debt offering.

The survey is organized as follows. Section 2 provides an overview of major regulatory rules and restrictions guiding security issues in the U.S. The section covers both regulations by the Securities and Exchange Commission (SEC), and self-regulatory authority rules issued by stock exchanges and the National Association of Security Dealers (NASD). Section 2 also summarizes the overall issue activity in the SDC population of U.S. issuers, 1980–2004. Section 3 reviews direct issue costs across major flotation methods, with a major emphasis on underwriting costs and understanding the underwriting process. Section 4 examines the flotation method choice and summarizes the evidence on the valuation effects of security offering announcements (both U.S. and internationally). Section 5 examines various theories for post-issue stock price performance, and presents the results of an original long-term return analysis performed on our SDC sample. Section 6 provide concluding remarks.

## 2. The security offering process

Equity offerings come in many colors and flavors, from IPOs to SEOs, public offers to private placements, classes of stock with differing cash flow and voting rights, from domestic issues to global issues and from warrants to employee/management stock options to convertible debt. They are also sold using many different mechanisms, from a firm commitment underwriting contract to a rights offering to a discriminatory or non-discriminatory auction, to more exotic methods such as privatizations, carve-outs, employee stock ownership plans (ESOPs), equity bonus plans, mutual-to-stock conversions, forced conversions of convertible securities (including conversions of venture capital held securities at the IPO), equity financed acquisitions, dividend reinvestment plans and funding pension plans with your own stock.

Legal systems, tax codes, securities regulations and the treatment of investors of a country are likely to have a significant bearing on the level of security offering activity as well as the choice of flotation methods. Over the last 25 years, there have been major changes in securities regulations in the U.S. and other major capital markets. We review some of these major changes and the trends in the evolution of security regulation in the next section.

### *2.1. U.S. securities regulations*

The U.S. regulatory environment is anchored on two major laws. The first major law is the Securities Act of 1933, which requires issuers of securities to sell the entire issue at a single offer price to all investors, to meet filing rules and extensive disclosure requirements prior to the offering date. Under the regulations implementing this law, prospective issuers must file an S-1 statement with SEC prior to the offering. Within approximately 30 days, the SEC will send the issuer a letter of comment asking for additional disclosures and request amendments to the registration statement. The issuer sends a response and after several exchanges of letters, the SEC will typically declare the registration effective. Once the filing statement is approved, the issuer can proceed with the offering. The second major act is the Securities Exchange Act of 1934 which mandates that issuers of publicly held securities make periodic disclosures through public filings of annual 10-K, quarterly 10-Q and occasional 8-K statements, when material changes occur.

There are several exemptions from the registration requirements under the Securities Act for small issues, private placements, mergers and reorganizations. While privately placed securities are exempt from registration requirements, these securities can not be resold for a year without being publicly registered with the SEC.

In recent years U.S. securities regulations have moved toward more rapid disclosure of material changes in company conditions, less delay in securities issuance and an easing of restrictions on private placements and foreign security issuance in the U.S. and the use of U.S. accounting standards under "generally accepted accounting standards" (GAAP). However, these changes appear to be more than offset for foreign issuers and small U.S. issuers by the passage of the Sarbanes–Oxley Act of 2002 which requires major changes in Board of Directors committee structure, auditor independence and certification of company financial disclosures.

As of March 1982, the SEC adopted Rule 415 Shelf Registration, which enabled public companies to sell securities more rapidly. Under the Rule, issuers register securities that can be sold from time to time over a two year period, with offer terms at each sale set in light of current market conditions and other factors. The Rule permits an issuer to avoid the delays involved in filing a new registration statement at each sale date. This flotation method was only available to larger, financially sound issuers meeting the following requirements: common stock (with or without voting rights) having a market value of at least $75 million, no defaults on any debt, preferred stock or rental payments for 3 years, all SEC disclosure requirements have been met for the last 3 years and the firm's debt is investment grade.

Under U.S. securities regulations, a foreign issuer has a choice of issuing either publicly or privately held equity or debt in the U.S. Typically, a foreign issuer of equity in the U.S. employs an American Depository Receipt (ADR) or Global Depository Receipt (GDR) mechanism which eliminates the domestic investors need to undertake foreign exchange transactions to acquire and dispose of these securities and convert cash dividend payments to dollars. An ADR is a financial instrument backed by a depository bank owning the underlying foreign shares, to which the ADR has a fractional claim, but which pays cash distributions and trades in dollars and settles trades in the U.S. market. Arbitrage keeps the prices of the underlying shares and the ADR in close alignment after adjusting for foreign exchange movements. GDRs are similar financial instruments which pay cash distributions and trade in a specific foreign currency and settle trades on a particular foreign stock exchange.

In April 1990 the SEC approved Rule 144A, which allows immediate sale and resale of private placements to "qualified institutional buyers" (QIBs) without having to register these securities or hold them for a year, as previously required.[3] This rule was particularly aimed at reducing regulatory costs and improving the liquidity of privately placed securities issued by privately held companies and foreign issuers. It gives privately held U.S. firms the ability to either privately place securities with accredited and sophisticated investors pursuant to Section 4.2 of the 1933 Securities Act or Rule 506 of Regulation D or to sell them to QIBs as a Rule 144A issue. The approval of Rule 144A also has the effect of allowing international firms to gain access to U.S. institutional investors without having to meet the strict disclosure and GAAP accounting requirements of U.S. public companies.

Under U.S. regulation, there are several ways a foreign company can tap the U.S. capital market. A firm can first make a small Rule 144A private placement and trade over-the-counter, which is called a Level I program. If it chooses to list on a U.S. exchange, it moves to a Level II program. Alternatively, it may undertake a Level III public offer of stock in the U.S. with listing on a U.S. stock exchange. An issuer can simply undertake a large 144A private placement or a firm can begin by seeking Level I or II market listing in the U.S., followed by a public offering. One key benefit of a 144A private placement is that a foreign issuer can raise capital in the U.S. sooner, since the issuer does not have to meet U.S. accounting and disclosure standards to tap this market. However, the stock's issue price is likely to be significantly discounted for its lower liquidity in the private placement market. In addition, issuers often need to obtain

[3] QIB typically refers to an institution (e.g., insurance companies, investment companies and pension funds) that own or invest $100 million in securities of non-affiliated companies.

home market regulatory approval before initiating any foreign trading in its securities. There can also be home country restrictions on foreign sales of domestic securities and purchases of foreign securities by domestic investors.

Under Regulation T of the Securities and Exchange Act of 1934, the Federal Reserve Board of Governors establishes rules to limit the portion of a security's market value that can be loaned to the investor by a broker. These margin requirements are established for the purpose of reducing selling pressure on investors who financed their security purchases with loans. Thus, in market downturns, investors borrowing on margin are required to put up additional collateral when their securities fall in value. This can force many liquidity impaired investors to sell securities to raise collateral or if they fail to meet the call for added collateral, the broker can sell their securities and close out their margin loans. Either event can create a cascading pattern of sell orders, which has been alleged to destabilize the stock market.

The SEC regulates the financial condition of brokerage firms and the short selling of securities by investors and underwriters. In the normal case of investor short selling, brokerage houses and institutional investors lend securities to short sellers, who immediately sell these securities in the stock market, knowing that at a future date they will be obligated to purchase these same securities in the stock market to close out their short positions with their lenders.

SEC regulations concerning public offerings of securities underwent sweeping changes as of December 1, 2005. One major innovation is the creation of a new category of issuers called "well known seasoned issuers" (WKSI) with special filing exemptions. WKSIs are publicly listed firms (involuntary filers) that are eligible to issue shelf offerings, which are current and timely in their reporting obligations over the past year. They must also meet one of two conditions; (1) have outstanding a minimum of $700 million of common equity market capitalization world-wide that is held by non-affiliates, or (2) if they are only registering non-convertible securities other than common equity, that during the past three years they have issued non-convertible securities other than common equity in registered primary offerings with an aggregate value of $1 billion.[4]

Under the new rules, a WKSI can have oral or written communication with investors before during and after the offering process. WKSIs are also given automatic shelf registration status. They are permitted to register unspecified amounts of different specified types of securities on Form S-3 or F-3 (only non-convertible securities excluding common equity if only condition (2) above is met) without allocating between primary and secondary offerings. These registration statements are automatically effective on filing without SEC review. Issuers can also add further classes of securities and eligible majority owned subsidiary securities after the registration statement is effective, provided they make a post-effective amendment to the offering's registration statement.

[4] Majority owned subsidiaries of these firms also may be considered to be "well-known seasoned issuers" if the securities issued are non-convertible securities other than common equity, are fully and unconditionally guaranteed by the parent and are of investment grade.

A second major change in SEC regulations is increased disclosure requirements in registration statements and 10-K statements concerning risk factors. Third, Rule 415 will no longer limit the amount of securities registered on a shelf registration statement to an amount intended to be offered and sold within two years of the effective date of the registration statement. In practice the SEC has allowed shelf registration statements to remain effective for many years. Under the new rules, the shelf registration can only be used for three years. The new rules allow seasoned issuers to conduct primary offerings immediately after the effectiveness of a shelf registration statement. Shelf issuers may also conduct "at-the-market" equity offerings (sales at varying prices rather than a conventional fixed price offer) without existing volume limitations and without needing to identify the potential underwriters.

WKSIs are permitted to omit the plan of distribution, the names of any selling security holders, the description of securities to be offered, and the allocation between primary and secondary shares. This information can be incorporated in prospectus supplements and post-effective date amendments to the shelf registration statement.

Foreign private issues are able to take advantage of the relaxation of the gun-jumping rules (communications occurring prior to the effective date of the registration statement) and the revised shelf registration rules to the same extent as domestic issuers. Moreover, automatic shelf registration will make it much easier for foreign private issuers that are WKSIs to conduct rights offerings in the U.S.

Other changes in SEC regulations include giving issuers a safe harbor from being in violation of security regulations for written communications of regularly released factual information made before or during an offering and commonly released forward-looking information (e.g., earnings forecasts) made before or during an offering, allowing issuers a wider range of oral and written communications while the offering is in registration, allowing electronic delivery of filing materials to shareholders, and allowing analysts reports of new issues under a wide range of situations, even for analysts affiliated with an underwriter.

Parallel to U.S. securities regulation, there are similar national regulatory authorities around the globe. The International Organization of Securities Commissions (IOSC) is a global organization of national security regulators created to foster cooperation in promoting high standards of regulation in order to maintain efficient and sound capital markets; to establish standards and effective surveillance of international securities transactions and to promote effective enforcement of these standards. Among its recent achievements, the IOSC in 1998 adopted a comprehensive set of objectives and principles of securities regulation, which today are recognized by the world financial community as international benchmarks for all markets. In 2002 the IOSC endorsed a memorandum of understanding among securities regulators around the world, designed to facilitate the enforcement of security regulation and the exchange of information. Looking internationally, there has been an increase in disclosure regulation and increased regulation and enforcement of insider trading activity.

In addition to securities regulation, several other recent laws and rules of self regulatory organizations also have impacted the security offering process. In 1999, the

Glass–Steagall Act which prohibited commercial banks and their subsidiaries from affiliating with securities firms or underwriting corporate securities was effectively repealed by the Gramm–Leach–Bliley Financial Modernization Act. The passage of this law had a direct effect on the securities market by increasing competition for corporate underwriting assignments by allowing entry by commercial banks who could have prior lending relationships with issuers (see also Drucker and Puri, 2007, Chapter 5, this volume).

Self-Regulatory Authorities (NYSE, NASD) impose various listing requirements on firms trading securities on their exchanges. In addition, the NASD has responsibility for regulating many of the activities of broker-dealers and underwriters. In recent years, both the NYSE and the Nasdaq have imposed new corporate governance requirements on firms listing in their markets. The NYSE also prohibits listed firms from inducing dual shares with unequal voting rights since 1994.[5]

The passage of the Sarbanes–Oxley Act of 2002 has enhanced shareholder voting rights by encouraging more independent boards and requiring outside directors take on major governance roles within the board of directors. This Act has increased the credibility of firm disclosure requirements by requiring greater auditor independence and the CEO and CFO to personally certify the company's annual financial statements.

### *2.2. Alternative flotation methods*

Table 1 summarizes the major flotation method choices observed for IPOs, SEOs and debt offerings. The table starts with "firm commitment" underwriting, which is the primary choice of publicly traded U.S. firms. Here, an underwriter syndicate guarantees the proceeds of the issue (net of fees) and organizes the sale of the shares. Given the prominence of this flotation method, we discuss key aspects of the underwriting process before commenting on the other flotation methods listed in Table 1.

#### *2.2.1. The firm commitment underwriting process*

The time line in a firm commitment offering is roughly as follows: The issuer contacts an investment bank to form a syndicate guaranteeing the offering. The lead underwriter performs due diligence (examining the financial status of the issuer), registers the issue with the SEC, and presents a preliminary prospectus ("red herring") to key investors and clients in a "road show". The preliminary prospectus specifies only a possible price range for the offering as the firm is not permitted to sell shares prior to SEC registration. When the SEC approves the issue, the firm meets with the underwriter syndicate and sets the final offer price ("pricing meeting") and the offer typically starts the following day. The underwriter guarantee requires a firm offer price, so the guarantee period starts

[5] Exceptions are firms with dual class shares prior to listing such as Ford Motor Co., Berkshire Hathaway, which was grandfathered when these requirements were first implemented.

Table 1
Flotation methods

*Firm commitment*. An underwriter contractually commits to purchase an entire security issue at a fixed price discount from the public offering price. All shares are sold to the public at the same price and the underwriter generally has the power to allocate the issue if there is excess demand. This process may involve book building or a fixed price placing

*Rights*. Short lived in-the-money warrants to buy a fixed number of new shares at fixed price, which are distributed to existing shareholders on a pro rata basis. These rights can often be resold to other investors. On the warrant expiration date, unexercised warrants are sometimes redistributed to shareholders who do exercise their rights

*Standby rights*. These contracts represent rights offers combined with a standby underwriting contract. The underwriter guarantees to exercise all unexercised warrants delivered to them at the warrant expiration date. Underwriter will often short-sell the stock and buy rights in the secondary market ("layoff") during the offering period to the lessen uncertainty about the number of unexercised warrants they will need to exercise and to receive higher compensation. Compensation is in form of a fixed pre-commitment fee and a variable take-up fee that is proportional to the number of rights exercised by the underwriter

*Private placement*. An issuer privately negotiates a sale of stock to qualified investors. There are registered private placements and restricted private placements. Resale of the stock is generally restricted to other qualified investors for one year, unless the issue has an effective registration statement covering the resale of these securities. Restricted private placements are unregistered offers (no prospectus is required) that fall under Regulation D or Regulation S. Regulation S private placements are sold outside the U.S., while Regulation D allows private placements within the U.S. Regulation D prohibits an issuer from soliciting the general public under Rules 505 and 506. Under Regulation D, issuers of private placements are exempt from SEC disclosure requirements such as having a prospectus. Issuers must target mostly accredited investors (wealthy or sophisticated investors). Issuers may distribute an offering memorandum, but cannot advertise or solicit investors. If unaccredited investors participate in the offer, then the offering size is limited to $5 million under Rule 505, though the number of accredited investors also involved is unlimited. Under Rule 506, the offer size is unlimited, but the number of accredited investors is limited to at most 35

*PIPE (Private investments in public equity)*. Private investment in public equity. A public company sells equity through a privately negotiated sale. These offering may or may not include some form of issuer price guarantee against a subsequent share price drop, but they generally include a large discount from the security's market price

*Shelf issue*. Financially strong public companies can register to sell up to a certain number of shares over the next two years using a list of possible underwriters. The registration allows the sale of one or more equity issues or alternatively the sale of one or more debt issues, the choice of debt or equity must be made at the filing date

*Universal shelf issue*. Similar to shelf issues except that the issuer can choose to sell either debt or equity

*Direct public offering*. Issuer sells equity directly to investors without the use of bank as a financial intermediary. If the sale involves interstate distribution of the securities, then a brief filing statement with the SEC is required. A short form registration of an offering under $5 million in a 12 month period is allowed under Regulation A. Under Regulation D, Rule 504 provides for offerings up to $1 million in a 12 month period by filing a Form D (Form D registration or small corporate offering registration (SCOR).

(*Continued on next page*)

Table 1
(*Continued*)

*Best effort*. Investment banks do not underwrite these security issues, instead they only guarantee to do their best to sell/market the issue. If less than a fixed percentage of an issue is sold, the entire issue is usually cancelled

*DRIPS*. Dividend reinvestment plans allow shareholders to buy more shares in lieu of receiving cash dividends. The shares may be sold at a small discount.

*Sealed bid auction*. This is a traditional method of selling IPOs. Typically a fixed number of shares are sold on a specific date, where the rules of the auction are publicly announced considerably in advance of the auction date. Sealed bids can generally be submitted over a specified period of time for a specific number of shares. The auction can be fixed price so that all accepted bids are paid the same purchase price (Dutch auction), or it can be a discriminatory auction where each accepted bid pays the bid price (Boston auction). In a nondiscriminatory auction, investors bid for parts of an issue at their bid price. Bids are ordered and a stop out price is determined where demand equals shares offered. All shares are sold at that price to those investors bidding at the stop-out price or higher. In a discriminatory auction, all offers at or above the stop-out price are accepted, but each investor pays the price they bid. Prior to the auction rules are announced concerning the bidding process, determination of the bidder purchase price and share allocation process. There are also often minimum bid price requirements. Other more complicated rules are also possible and are typically used in privatizations

with the pricing meeting and expires at the end of the offer period. Since the typical (successful) offering is fully sold out over a couple of days, the effective firm commitment guarantee period is also typically short.

The following summarizes key aspects and terminology associated with the firm commitment underwriting process.

*Board of directors approval*. Approval is necessary before an offering can occur and it is also necessary to get prior shareholder authorization of any shares that will be issued, though most companies typically have shareholders authorize large numbers of shares far in advance of their possible use.

*Choice of lead underwriters*. Competing underwriters make presentations to the issuer, though many publicly listed issuers have long standing investment banking and commercial banking relationships with one or more potential underwriters.

*Advisory role of underwriters*. Lead underwriters advise the issuer on the security's price, the timing of the offering, the size of the offering, desirable and undesirable offering characteristics, road show mechanics and meeting various regulatory requirements.

*Syndicate formation*. Lead (and co-lead) underwriters often line up other banks to help underwrite and distribute shares. Syndicate members sign legal contracts to underwrite or distribute a certain number of shares in return for underwriting and distribution fees. Lead underwriters tend to take the largest portion of the underwriting risk. In most underwriting contracts, all banks share in any loses associated with unsold shares that are later resold in the secondary market.

*Syndicate roles and compensation*. Lead underwriters form and coordinate syndicates and receive the management fees. Some banks share underwriting risk and underwriting fees while other banks may help distribute shares and receive distribution fees. Lee et al. (1996) discuss the typical breakdown of underwriting syndicate compensation for IPOs.

*Due diligence investigation*. Underwriters must investigate the issuer and certify that the issue price is fair.

*Prospectus*. An issuer must produce a document describing the security offering and its financial condition with the help of its underwriter. The due diligence investigation helps assemble the information needed to meet SEC filing requirements.

*Registration process*. An issue must be registered in advance with the SEC. This must include a preliminary prospectus or red herring and later a final prospectus. In the U.S. and many other countries this will include an initial price range for the proposed offering.

*Effective date*. Security registration statements that must be filed prior to a security offering are said to be effective after they are reviewed by the SEC staff and any concerns are resolved. The date of SEC approval is termed the effective date of the security offering's registration statement, after which selling of the issue can occur.

*A seasoned issuer*. A reporting company that is eligible to use SEC Form S-3 or F-3 to register primary offerings of securities.

*A well-known seasoned issuer*. Publicly listed firms (involuntary filers) eligible to issue shelf offerings, which are current and timely in their reporting obligations over the past year. They must also (1) have outstanding a minimum of $700 million of common equity market capitalization world-wide that is held by non-affiliates or (2) if they are only registering non-convertible securities other than common equity, they have issued non-convertible securities other than common equity in registered primary offerings for cash $1 billion aggregate amount of during the past three years.

*Exchange listing process*. An issuer may seek a preliminary assessment of whether subsequent to a successful offering its stock is likely to meet an exchange's listing requirements. Plans to list on an exchange will be reported in the registration document.

*Quiet period*. U.S. regulation which prohibits firms going public and their underwriters from disclosing sales and earnings forecasts not in the prospectus starting before the firm announces its IPO and ending 40 calendar days after the offer.[6] This also precludes stock analysts affiliated with an underwriter from covering the stock of an IPO for the same period.

[6] Prior to July 2002, the quiet period only lasted until 25 calendar days after the IPO.

*Road show*. To market a security offering, senior management and the lead underwriters travel to major cities to meet with potential investors to discuss the planned offering. An exemption to the "quiet period" regulations allows managers and underwriters to make limited oral disclosures during road show presentations, where attendance is restricted to institutional investors. However, in practice most managers and underwriters try to avoid releasing new information. Thus, this process may be more an information gathering and marketing effort by an underwriter than an information session that offers investors new information about the issuer.

*Book building process*. Underwriters solicit tentative offers from a select group of institutional investors and other potential investors to buy shares. Bids can be in several forms: strike bids to buy a specific number of shares at almost any market clearing price, limit bids where an investor submits a bid for a specific number of shares at a specific offer price and step-bids where an investor submits a number of limit bids for specific numbers of shares at different offer prices. The underwriter can use its allocation ability to reward investors for revealing information on demand in the book building process. Generally, investors can submit bids until the book closes and can revise or cancel their bids. This process may cause the issuer to revise the price range, which will necessitate filing an amendment with the SEC. At the end of this process the underwriters will have reasonably good estimate of institutional investor demand for the issue. Of course small retail investors may have a very different demand for the issue.[7]

*Signing underwriting contract and setting the offer price*. The Underwriter accepts security issue price risk when it signs the Underwriting Agreement to purchase the entire security issue at an agreed upon fixed price, usually within 24 hours of the start of the public offering. It is at this point that the final prospectus is printed. On the morning of the chosen offer date, the underwriter files a "price amendment" with the SEC on behalf of the issuer specifying the security's offer price. As Smith (1977) notes, this is similar to the underwriter selling a put option on the security issue to the issuer for a fee. Underwriters reject some potential issuers and vice versa when they disagree on the level of risk and the appropriate fee or when the underwriters are unable to meet all the potential demand for their services. Underwriters can also back out of tentative commitments to underwrite issues up until the day before the public offering date.

*Allocation of offering and overselling of offering*. The syndicate generally oversells the issue since the orders are not legally binding and can be withdrawn, though withdrawals are likely to trigger future loss of allocations in offerings. The lead underwriter generally determines who is allowed to buy shares in a hot offer and how much of their order is filled. These investors tend to be good (large) customers of the underwriter.

[7] For further analysis of the book building process in IPOs, see the studies by Benveniste and Spindt (1989), Benveniste and Wilhelm (1990), Cornelli and Goldreich (2001), Cornelli and Goldreich (2003) and Sherman and Titman (2002).

Some issues are also allocated to friends and family of the issuer's management and to CEOs of companies the underwriter is cultivating for future business.[8]

*Public offer date activities*. Underwriters confirm investor orders, allocate hot issues, and may buy shares in the secondary market to meet some of their commitments as a result of overselling the issue when the after-market price isn't rising relative to the offering price.

*Analyst coverage commitment*. Lead underwriters, co-managers and other syndicate members often commit to produce analyst coverage for the stock for a period after the offering. This is likely to enhance investor interest in the stock and improve the stock's liquidity. A survey of issuer managers finds that underwriter selection is strongly influenced by whether an underwriter has reputable industry analysts.[9]

*Market making commitment*. Lead underwriters generally commit to be active market makers in the stock for a period of time after the offering. Existing evidence shows that this market making is very important in the early seasoning of an issue, but typically declines in importance over the first year following listing. This market making activity is typically profitable for the lead underwriter.[10]

*Price support*. Lead underwriters often place limit orders to buy shares immediately after an offering without being subject to price manipulation restrictions. If an underwriter oversells an offering, which afterwards drops in price, then the underwriter can buy additional shares in the secondary market at a price at or below the offering price, rather than exercise its over-allotment option to buy additional securities from the issuer. This has the effect of supporting the secondary market price and avoids adding more shares into the secondary market. If the secondary market price rises relative to the offering price, then no price support activity is necessary. Instead, the underwriter can meet its commitments to customers of oversold issues by exercising its over-allotment options to buy shares at the offer price net of the underwriter discount.[11]

*Lock-up agreements*. Insiders and other large holders such as venture capitalists commit not to sell their shares for a period of time after the offering. The typical lock-up period is 180 days for IPOs. If the secondary market reception for the issue is very strong, the agreements may be terminated early.[12]

*Insider trading regulation*. U.S. SEC Rule 10b-5 prohibits a person in possession of material non-public information from using it to buy or sell company securities or to tip

[8] See Cornelli and Goldreich (2001, 2003) and Jenkinson and Jones (2004) for evidence on the book building and share allocation process and Loughran and Ritter (2002) for evidence of spinning.

[9] See Krigman, Shaw, and Womack (2001) and Brau and Fawcett (2006).

[10] For an analysis of post-IPO market making by underwriters see (Ellis, Michaely, and O'Hara, 2004).

[11] Price support or stabilization activity for IPOs is studied by Aggarwal (2000), Boehmer and Fishe (2003) and Cotter, Chen, and Kao (2004) and Lewellen (2006).

[12] The lock-up process and its expiration effects are studied by Brav and Gompers (2003), Field and Hanka (2001), Field, Cao, and Hanka (2004) and Brau, Lambson, and McQueen (2005).

others who do so. There is also a filing requirement after the sale or purchase by insiders of the firm's securities.

### 2.2.2. Other major flotation methods

Table 1 gives a summary of the various flotation methods available for security offerings. A more detailed description of these flotation methods follows.

In a "rights offer" current shareholders are given the right to purchase a (pro rata) portion of a new equity issue at a fixed price. A rights offer in the U.S. typically expires after a period of typically one month. The rights offer price is initially set at a discount from the current market price, but if the market price falls, the rights offer can end up being at a premium, which is likely to result in offer undersubscription or offer failure. Thus, a rights offer is like a short-lived in-the-money warrant distributed to current shareholders in the same manner as a stock dividend. It is also similar to a stock dividend in that the sale of new shares at a discount has the effect of diluting the current share price. Rights may or may not be transferable and unsubscribed rights may be reallocated among subscribing shareholders. In these non-underwritten offers, the issuer bears a risk of offering failure, but this risk can be reduced by increasing the size of the offering price discount.

In a "standby rights offer" the firm making the rights offer hires an underwriter to "stand by" and guarantee to take up whatever portion of the rights offer shareholders leave unsubscribed. The standby underwriter as a consequence bears price risk, and carries out a due diligence investigation and may pursue a book building process described above for firm commitment offerings. For these services, the underwriter charges a fixed "standby" fee. In addition, the underwriter typically charges a "takeup" fee on each share taken up under the guarantee. If there is a secondary market in the rights, it is common for the underwriter to be the primary purchaser of these rights.

In a private placement, the firm places the entire issue with a single investor or consortium of investors, bypassing current shareholders. As listed in Table 1 and discussed above, such issues are subject to a number of regulations primarily designed to protect investors.

A "shelf" offering refers to an issue that has been pre-registered with the SEC. With the introduction of SEC Rule 415 in 1983, financially strong companies are allowed to sell up to a certain number of shares over the next two years using a list of possible underwriters. Thus, shelf registration increases the flexibility and speed of issue over a two-year period.

Auctions present another mechanism for selling equity. This method is only rarely used in the U.S. (it was used recently by Google), but has been an important method in certain international markets including France. The auction design is flexible, but the most common is a sealed bid auction where all accepted bids pay the same price. There are often minimum bid (reserve) price requirements (see Dasgupta and Hansen (2007) and Jagannathan and Sherman (2006) for details on IPO auction procedures).

Table 2
Flotation methods used to sell various types of securities

| Security type | Flotation method |
|---|---|
| IPOs | Firm commitments, Auctions, Direct offerings, Private placements, Best efforts, Privatization methods, Mutual to stock conversions |
| SEOs | Firm commitments, Shelf issues, Universal shelf issues, Private placements, Direct offerings, Rights, Standbys, Auctions, Best efforts, Equity financed acquisitions, PIPES, DRIPS, ESOPs, Equity based bonus plans, Equity for debt exchange offers and swaps, Privatization methods |
| Convertible offers | Firm commitments, Private placements, Auctions, Direct offerings, Shelf issues, Universal shelf issues, Convertible debt for equity exchange offers and swaps, Convertible debt financed acquisitions |
| Debt offers | Firm commitments, Private placements, Auctions, Shelf issues, Universal shelf issues, Debt for equity exchange offers and swaps, Debt financed acquisitions |
| Private debt | Direct offerings, Private placements, Venture capital |
| Private equity | Direct offerings, Private placements, Venture capital |

A detailed economic analysis of the flotation method choice is given in Section 4, below. As indicated there, the importance of the various flotation methods listed in Table 1 varies across countries, with issuers in larger capital markets exhibiting different preferences than those in smaller capital markets. In the U.S. nearly all IPOs are sold through a book building mechanism. Internationally, a firm commitment contract with book building is the dominant IPO issuance method in most large capital markets, while auction methods are dominant in smaller capital markets with more concentrated share ownership. For evidence that IPO flotation methods vary across countries, see the survey of international IPOs by Loughran, Ritter, and Rydqvist (1994), and Ritter (2003).

Table 2 describes the flotation methods used to sell various types of securities. As the table highlights, seasoned equity issues and debt issues use a wider array of offering methods. Debt offerings tend to rely on the same flotation methods as seasoned equity issues. In the U.S., the primary SEO flotation methods are: firm commitment underwritten offers (either syndicated or not, U.S. or global), shelf registered offers (either equity or universal), standby underwritten rights offers, rights offers, best efforts, direct issues and private placements. Outside the U.S., the primary flotation methods used are rights and standby offers, however, auctions, bought deals, installment sales and other methods are also important. Some capital markets have their own particular flotation methods including the U.K., France and Singapore. Privatization methods tend to be very idiosyncratic across countries as is highlighted in a survey by Megginson and Netter (2001).

IPO flotation methods vary across capital markets of differing size as discussed in Loughran, Ritter, and Rydqvist (1994), and Ritter (2003). In the U.S. nearly all IPOs are sold through a book building mechanism. Internationally, the firm commitment book building method is dominant in most large capital markets, while auction methods are dominant in smaller capital markets with more concentrated share ownership, though there is some question as to whether auctions are successful more because book building is unavailable due to regulation or minimum offer size. Jagannathan and Sherman (2006) examine why IPO auctions are unsuccessful in the U.S. market.

### *2.3. Aggregate issuance activity, U.S. 1980–2003*

#### *2.3.1. Offering frequencies and cash proceeds*

In order to understand the patterns in security issuance activity by U.S. firms, we start with the grand population of 91,455 issues from the SDC over the period 1980–2003. We then eliminate 8,173 issues for which we are unable to match the issuing firm's name and Cusip number in Thomson Financial's SDC database with a corresponding exchange-listed firm name on the University of Chicago CRSP daily stock master file for the issue year. This leaves a total of 83,282 issues for analysis. We then restrict our focus to the following seven major security classes:

(1) Public offerings of straight debt ($N = 37{,}398$, of which 18,662 are shelf offerings),
(2) Private placements of straight debt ($N = 17{,}948$, of which 5,983 are reg-144A offerings),
(3) SEOs ($N = 11{,}151$, of which 1,645 are shelf offerings),
(4) Equity IPOs ($N = 9{,}987$, of which 1,063 are "unit" offerings–with warrants),
(5) Private placements of equity ($N = 2{,}145$, of which 83 are SEC regulation 144A offerings),
(6) Convertible debt offerings ($N = 1{,}545$), and
(7) ADRs (American depository receipt stock offerings, $N = 453$).

After excluding 2,655 "other" security issues, we are left with a sample of 80,627 security offerings.

Table 3 shows the annual frequency of offerings across the seven major security offering categories. A number of regularities emerge from this table:

- For both IPOs and SEOs, the number of issues exceed 600 in years 1983, 1993, 1996 and 1997 (particularly "hot" issue markets).
- The total number of straight debt offerings outnumber the total number of SEOs by approximately three to one (37,298 vs. 11,151).
- Firms use the shelf registration procedure for approximately half of the debt issues (18,662 of 37,398), while fifteen percent of the SEOs are shelf issues (1,645 of 11,151).
- Straight debt is issued through private placements in one-third of the offerings (17,948 of 55,346 straight debt offerings), while one in six equity issues are sold in private placements (2,145 of 13,296 seasoned equity issues).
- In approximately ten percent of the IPOs, the stock is sold with stock warrants, which is termed a unit offering.

Table 3
Annual distribution of the population of 80,627 security issues in the U.S., 1980–2003

| Year | Equity IPOs | | | Seasoned equity offerings | | | Public straight debt offerings | | |
|---|---|---|---|---|---|---|---|---|---|
| | All | Regular | Unit | All | Regular | Shelf | All | Regular | Shelf |
| 1980 | 127 | 108 | 19 | 382 | 382 | 0 | 314 | 314 | 0 |
| 1981 | 303 | 267 | 36 | 416 | 416 | 0 | 254 | 254 | 0 |
| 1982 | 117 | 98 | 19 | 444 | 417 | 27 | 375 | 226 | 149 |
| 1983 | 659 | 574 | 85 | 813 | 672 | 141 | 324 | 135 | 189 |
| 1984 | 332 | 261 | 71 | 251 | 242 | 9 | 376 | 164 | 212 |
| 1985 | 339 | 287 | 52 | 400 | 391 | 9 | 554 | 251 | 303 |
| 1986 | 675 | 584 | 91 | 507 | 500 | 7 | 860 | 322 | 538 |
| 1987 | 529 | 447 | 82 | 314 | 311 | 3 | 626 | 232 | 394 |
| 1988 | 276 | 234 | 42 | 140 | 139 | 1 | 529 | 197 | 332 |
| 1989 | 237 | 183 | 54 | 230 | 229 | 1 | 555 | 202 | 353 |
| 1990 | 205 | 175 | 30 | 188 | 184 | 4 | 468 | 146 | 322 |
| 1991 | 398 | 349 | 49 | 508 | 498 | 10 | 1327 | 627 | 700 |
| 1993 | 805 | 729 | 76 | 736 | 706 | 30 | 1789 | 953 | 836 |
| 1994 | 631 | 539 | 92 | 474 | 438 | 36 | 1597 | 676 | 921 |
| 1995 | 572 | 508 | 64 | 619 | 535 | 84 | 2253 | 945 | 1308 |
| 1996 | 857 | 785 | 72 | 767 | 674 | 93 | 2626 | 1084 | 1542 |
| 1997 | 606 | 570 | 36 | 736 | 518 | 218 | 3440 | 1493 | 1947 |
| 1998 | 380 | 371 | 9 | 562 | 360 | 202 | 3704 | 1634 | 2070 |
| 1999 | 531 | 523 | 8 | 438 | 354 | 84 | 3488 | 2051 | 1437 |
| 2000 | 382 | 378 | 4 | 397 | 304 | 93 | 3172 | 1953 | 1219 |
| 2001 | 126 | 119 | 7 | 427 | 244 | 183 | 2873 | 1680 | 1193 |
| 2002 | 169 | 166 | 3 | 422 | 255 | 167 | 2596 | 1410 | 1186 |
| 2003 | 128 | 128 | 0 | 502 | 267 | 235 | 2160 | 1197 | 963 |
| All | 9987 | 8924 | 1063 | 11151 | 9506 | 1645 | 37398 | 18736 | 18662 |

| Year | PP straight debt | | | PP common stock | | | Conv. debt | ADR |
|---|---|---|---|---|---|---|---|---|
| | All | Regular | Reg-144a | All | Regular | Reg-144a | | |
| 1980 | 2 | 2 | 0 | 0 | 0 | 0 | 93 | 2 |
| 1981 | 365 | 365 | 0 | 29 | 29 | 0 | 88 | 1 |
| 1982 | 429 | 429 | 0 | 34 | 34 | 0 | 66 | 3 |
| 1983 | 462 | 462 | 0 | 51 | 51 | 0 | 113 | 10 |
| 1984 | 408 | 408 | 0 | 37 | 37 | 0 | 66 | 9 |
| 1985 | 553 | 553 | 0 | 69 | 69 | 0 | 138 | 2 |
| 1986 | 735 | 735 | 0 | 67 | 67 | 0 | 203 | 6 |
| 1987 | 864 | 864 | 0 | 53 | 53 | 0 | 146 | 16 |
| 1988 | 1160 | 1159 | 1 | 80 | 80 | 0 | 35 | 8 |
| 1989 | 971 | 971 | 0 | 99 | 99 | 0 | 61 | 8 |
| 1990 | 907 | 892 | 15 | 69 | 66 | 3 | 33 | 2 |
| 1991 | 987 | 870 | 117 | 90 | 77 | 13 | 49 | 12 |
| 1992 | 956 | 764 | 192 | 89 | 84 | 5 | 63 | 19 |
| 1993 | 1234 | 806 | 428 | 96 | 88 | 8 | 89 | 39 |

(*Continued on next page*)

Table 3
(*Continued*)

| Year | PP straight debt | | | PP common stock | | | Conv. debt | ADR |
|---|---|---|---|---|---|---|---|---|
| | All | Regular | Reg-144a | All | Regular | Reg-144a | | |
| 1994 | 906 | 639 | 267 | 93 | 75 | 18 | 33 | 49 |
| 1995 | 623 | 413 | 210 | 64 | 60 | 4 | 30 | 34 |
| 1996 | 623 | 294 | 329 | 50 | 49 | 1 | 44 | 69 |
| 1997 | 865 | 290 | 575 | 58 | 58 | 0 | 43 | 56 |
| 1998 | 1006 | 309 | 697 | 49 | 40 | 9 | 23 | 19 |
| 1999 | 812 | 252 | 560 | 66 | 64 | 2 | 27 | 23 |
| 2000 | 543 | 126 | 417 | 89 | 87 | 2 | 31 | 19 |
| 2001 | 833 | 142 | 691 | 265 | 258 | 7 | 43 | 13 |
| 2002 | 705 | 103 | 602 | 256 | 249 | 7 | 11 | 15 |
| 2003 | 999 | 117 | 882 | 292 | 288 | 4 | 17 | 19 |
| All | 17948 | 11965 | 5983 | 2145 | 2062 | 83 | 1545 | 453 |

The SDC source contains a total of 91,455 issues over the 24-year sample period. Of these, 8,173 are excluded as the issuing firm could not be identified on the University of Chicago CRSP file using the SDC name and Cusip number and the CRSP Permno. Moreover, another 2,659 offerings are excluded as they do not belong to any of the issue categories shown below. "PP" denotes private placement; "Unit" offerings are equity offerings with warrants; "Shelf" offerings are pre-registered under SEC Rule 415; "ADR" denotes American depository receipt; and "Reg-144a" denotes private placement to a qualifying investor under SEC regulation 144a.

- Convertible debt issues represent only three percent of all debt issues (1,545 of 56,891) and has remained relatively stable in annual terms since 1990.
- ADRs represent 4% of all SEOs and have remained relatively stable in annual terms since 1991.

Table 4 provides the annual distribution of offering proceeds (in $billion) from the offerings in Table 3. Over the 24-year period, the proceeds from all offerings are in excess of $12 trillion. Dividing through by the total number of issues reveals the following interesting regularities concerning average issue sizes:

- The average IPO is 21% smaller than the average SEO: $68 vs. $86 million.
- The typical public debt issue is about three times the average SEO: $230 vs. $86 million.
- Private placement issues are roughly half the size of public issues: $46 vs. $86 million for SEOs, and $122 vs. $230 million for public debt issues.
- For SEOs, shelf offerings are on average twice as large as traditional registered offerings: $149 vs. $75 million.
- For public offerings of straight debt, shelf issues are on average slightly smaller than traditional registered offerings, $211 vs. $250 million.
- Convertible debt issues are of the same average size as the privately placed straight debt issues: $119 vs. $122 million.

Table 4
Annual distribution the total of $12,820 billion issue proceeds from the population of 80,627 U.S. security issues, 1980–2003 (all numbers in $billion)

| Year | Equity IPOs | | | Seasoned equity offerings | | | Public straight debt offerings | | |
|---|---|---|---|---|---|---|---|---|---|
| | All | Regular | Unit | All | Regular | Shelf | All | Regular | Shelf |
| 1980 | 1.23 | 1.15 | 0.08 | 11.57 | 11.57 | 0.00 | 32.26 | 32.26 | 0.00 |
| 1981 | 2.96 | 2.76 | 0.19 | 12.17 | 12.17 | 0.00 | 26.69 | 26.69 | 0.00 |
| 1982 | 1.33 | 1.27 | 0.06 | 15.33 | 13.48 | 1.84 | 32.47 | 18.59 | 13.87 |
| 1983 | 12.37 | 12.02 | 0.35 | 25.80 | 19.06 | 6.73 | 30.69 | 11.36 | 19.33 |
| 1984 | 3.83 | 3.53 | 0.30 | 6.14 | 5.63 | 0.50 | 46.12 | 20.05 | 26.08 |
| 1985 | 8.44 | 8.10 | 0.34 | 16.40 | 15.58 | 0.83 | 66.56 | 25.89 | 40.67 |
| 1986 | 21.57 | 21.21 | 0.36 | 21.04 | 20.52 | 0.52 | 130.61 | 45.60 | 85.00 |
| 1987 | 23.88 | 23.44 | 0.45 | 17.34 | 17.23 | 0.10 | 97.65 | 37.04 | 60.61 |
| 1988 | 23.75 | 23.44 | 0.31 | 6.13 | 6.08 | 0.04 | 88.26 | 36.24 | 52.02 |
| 1989 | 13.39 | 13.16 | 0.23 | 9.35 | 9.28 | 0.07 | 94.56 | 36.01 | 58.55 |
| 1990 | 10.11 | 9.92 | 0.19 | 9.04 | 8.93 | 0.11 | 79.40 | 21.36 | 58.04 |
| 1991 | 25.71 | 25.37 | 0.34 | 33.38 | 32.09 | 1.28 | 164.12 | 62.90 | 101.22 |
| 1992 | 40.30 | 39.68 | 0.62 | 34.29 | 33.41 | 0.88 | 235.56 | 99.52 | 136.04 |
| 1993 | 56.45 | 55.89 | 0.56 | 49.75 | 45.80 | 3.94 | 316.84 | 147.13 | 169.71 |
| 1994 | 33.32 | 32.65 | 0.67 | 31.83 | 27.76 | 4.08 | 243.10 | 130.09 | 113.02 |
| 1995 | 30.14 | 29.36 | 0.77 | 52.23 | 44.47 | 7.76 | 350.95 | 189.06 | 161.90 |
| 1996 | 49.50 | 48.74 | 0.75 | 66.36 | 56.44 | 9.92 | 410.91 | 228.62 | 182.28 |
| 1997 | 41.04 | 40.61 | 0.43 | 75.05 | 48.97 | 26.08 | 542.47 | 297.81 | 244.66 |
| 1998 | 42.60 | 42.50 | 0.10 | 62.06 | 41.89 | 20.16 | 799.35 | 431.16 | 368.19 |
| 1999 | 59.37 | 59.29 | 0.07 | 86.95 | 67.79 | 19.16 | 845.74 | 491.40 | 354.34 |
| 2000 | 54.43 | 54.40 | 0.03 | 99.19 | 70.20 | 28.99 | 835.86 | 444.07 | 391.79 |
| 2001 | 38.39 | 38.07 | 0.32 | 78.05 | 39.00 | 39.06 | 1075.83 | 594.64 | 481.19 |
| 2002 | 42.08 | 41.96 | 0.12 | 68.59 | 32.76 | 35.83 | 974.32 | 560.79 | 413.54 |
| 2003 | 40.87 | 40.87 | 0.00 | 70.99 | 35.51 | 35.48 | 1107.35 | 701.98 | 405.37 |
| All | 677.05 | 669.40 | 7.65 | 959.00 | 715.64 | 243.36 | 8,627.70 | 4,690.26 | 3,937.43 |

| Year | PP straight debt | | | PP common stock | | | Conv. debt | ADR |
|---|---|---|---|---|---|---|---|---|
| | All | Regular | Reg-144a | All | Regular | Reg-144a | | |
| 1980 | 0.08 | 0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 4.23 | 0.09 |
| 1981 | 10.89 | 10.89 | 0.00 | 0.16 | 0.16 | 0.00 | 4.57 | 0.06 |
| 1982 | 15.63 | 15.63 | 0.00 | 0.27 | 0.27 | 0.00 | 3.18 | 0.19 |
| 1983 | 21.24 | 21.24 | 0.00 | 0.34 | 0.34 | 0.00 | 6.11 | 0.62 |
| 1984 | 25.31 | 25.31 | 0.00 | 0.60 | 0.60 | 0.00 | 4.09 | 0.61 |
| 1985 | 43.63 | 43.63 | 0.00 | 1.31 | 1.31 | 0.00 | 7.10 | 0.03 |
| 1986 | 65.52 | 65.52 | 0.00 | 1.78 | 1.78 | 0.00 | 9.71 | 0.37 |
| 1987 | 65.67 | 65.67 | 0.00 | 2.12 | 2.12 | 0.00 | 9.68 | 4.40 |
| 1988 | 102.10 | 101.61 | 0.48 | 2.84 | 2.84 | 0.00 | 3.14 | 1.10 |
| 1989 | 99.64 | 99.64 | 0.00 | 7.27 | 7.27 | 0.00 | 5.42 | 0.91 |
| 1990 | 66.54 | 64.51 | 2.03 | 3.83 | 3.79 | 0.03 | 4.76 | 0.14 |
| 1991 | 56.46 | 48.64 | 7.83 | 4.03 | 2.62 | 1.41 | 7.83 | 3.13 |
| 1992 | 52.82 | 35.98 | 16.84 | 3.86 | 3.48 | 0.39 | 6.71 | 4.23 |

(*Continued on next page*)

Table 4
(*Continued*)

| Year | PP straight debt | | | PP common stock | | | Conv. debt | ADR |
|---|---|---|---|---|---|---|---|---|
| | All | Regular | Reg-144a | All | Regular | Reg-144a | | |
| 1993 | 81.97 | 41.58 | 40.39 | 3.28 | 1.86 | 1.42 | 9.41 | 7.26 |
| 1994 | 56.82 | 29.67 | 27.15 | 2.62 | 0.88 | 1.75 | 4.42 | 8.63 |
| 1995 | 50.67 | 23.10 | 27.57 | 2.20 | 1.80 | 0.39 | 6.31 | 5.18 |
| 1996 | 65.01 | 19.10 | 45.91 | 4.99 | 4.95 | 0.04 | 6.69 | 10.48 |
| 1997 | 134.73 | 18.36 | 116.37 | 5.55 | 5.55 | 0.00 | 8.97 | 9.85 |
| 1998 | 189.95 | 30.07 | 159.88 | 5.89 | 5.46 | 0.43 | 14.22 | 8.09 |
| 2000 | 170.50 | 11.46 | 159.04 | 9.70 | 9.48 | 0.22 | 15.58 | 7.92 |
| 2001 | 277.67 | 15.16 | 262.51 | 13.74 | 10.94 | 2.79 | 18.54 | 4.20 |
| 2002 | 135.79 | 10.88 | 124.91 | 7.52 | 6.73 | 0.78 | 7.74 | 4.85 |
| 2003 | 206.00 | 15.54 | 190.46 | 7.12 | 6.54 | 0.58 | 9.67 | 5.46 |
| All | 2,181.79 | 841.98 | 1,339.81 | 98.36 | 84.66 | 13.70 | 183.60 | 92.08 |

"PP" denotes private placement; "Unit" offerings are equity offerings with warrants; "Shelf" offerings are pre-registered under SEC Rule 415; "ADR" denotes American depository receipt; and "Reg-144a" denotes private placement to a qualifying investor under SEC regulation 144a.

- ADRs have a relatively larger average size of $203 million, compared to SEO average proceeds of $86 million.

Figure 1 and Figure 2 show the distribution of total issue proceeds across three categories of issuers: industrial firms, banks and financial institutions, and public utilities.[13] Industrial firms are by far the dominant issuers of SEOs throughout the entire 24-year period (Figure 1). Banks and financial institutions are a distant second, with utilities are a very distant third. Both industrial firms and banks/financial institutions have substantially greater total issue proceeds in the second half of the sample period.

On the debt side, banks and other financial institutions greatly dominate the amount raised from public offerings of straight debt (part (a) of Figure 2). Here industrial firms and utilities are a distant second and third. For private placements of debt, however, industrial issuers dominate, with banks and financial institutions a close second. As with equity issues, the proceeds from both public and private debt issues are substantially greater in the second half of the sample period.

### 2.3.2. *Time from IPO to follow-on offerings*

The need for new capital is undoubtedly a key motivation to go public for many private companies. The immediate need for capital is covered by the proceeds from the IPO—but, equally important, a public company subsequently has better access to the capital markets. This section reviews evidence on how rapidly new public companies in fact do come back to the market with a follow-on offering.

[13] Notice the different scales across the vertical axis of the three figures.

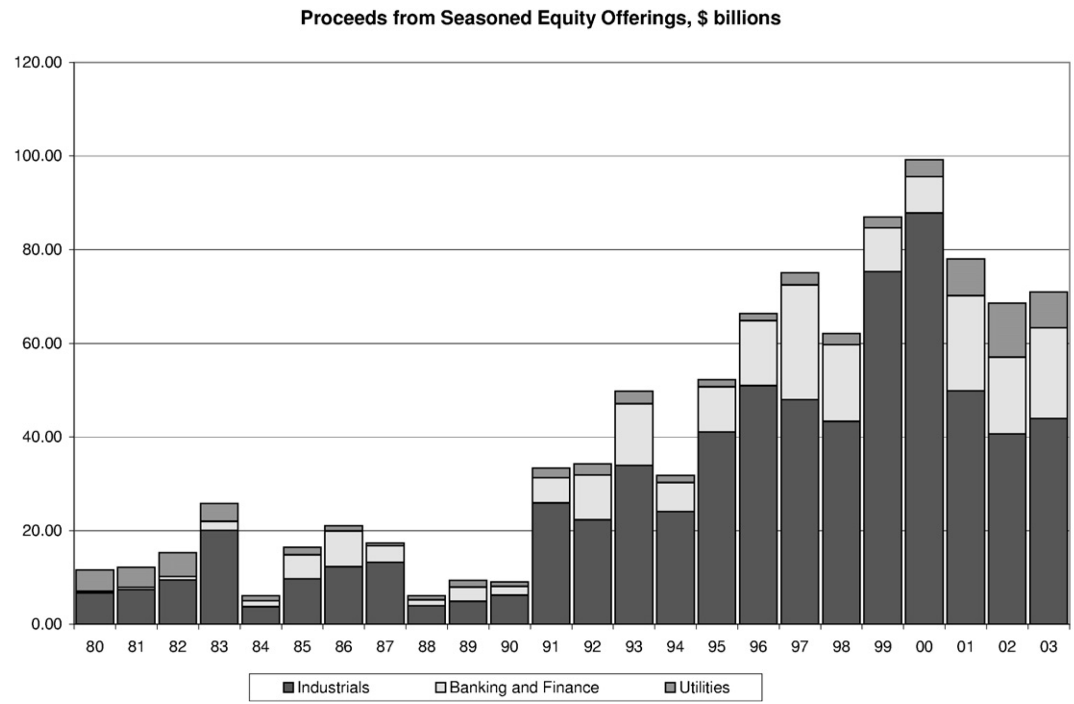
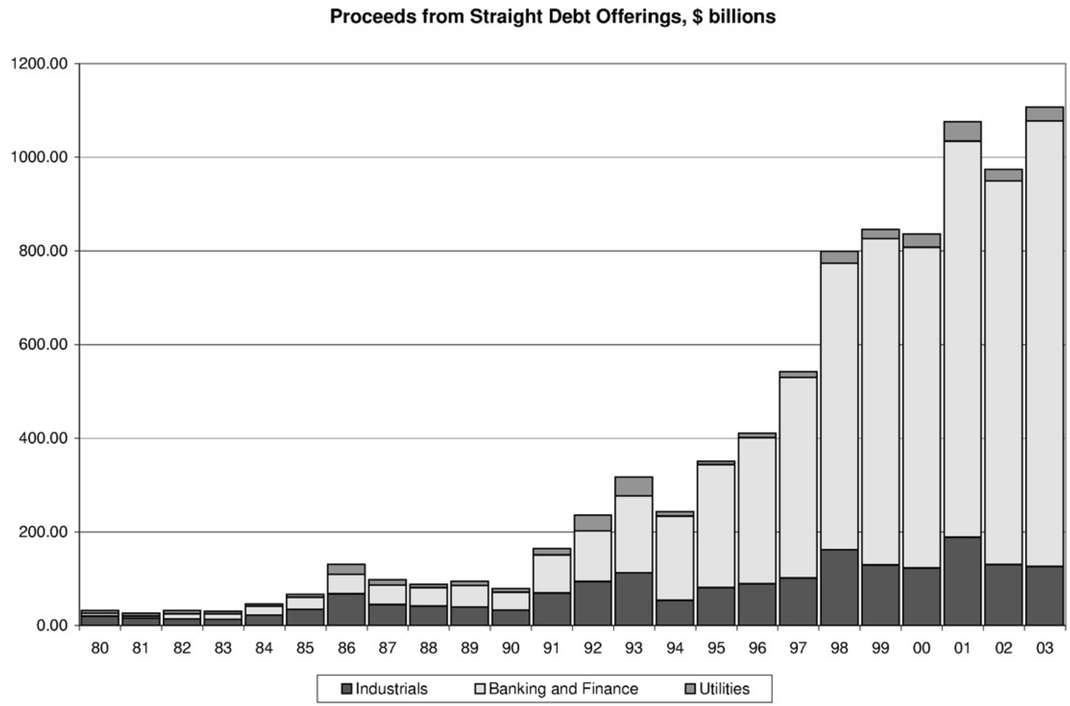
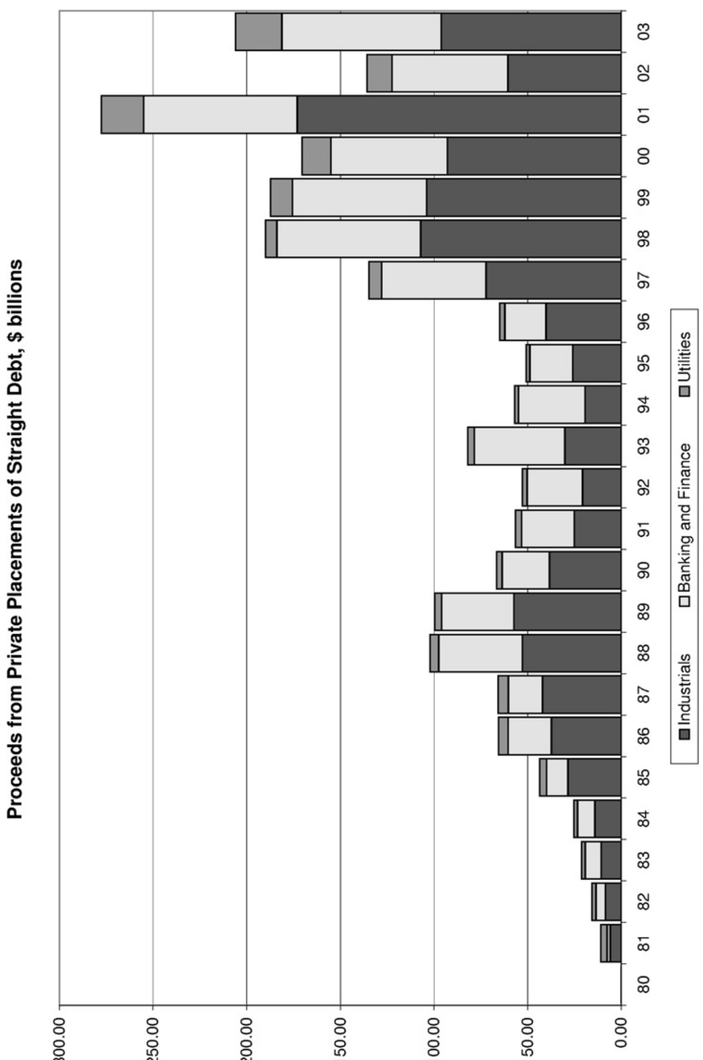



Fig. 1. Annual distribution of total issue proceeds in 11,151 SEOs by U.S. issuers, classified by whether the issuer is an Industrial Company, a Bank or Financial Institution, or a Public Utility, 1980–2003.




(a)

Fig. 2. Annual distribution of total issue proceeds in 37,398 public and 17,948 private issues of straight debt by U.S. companies, classified by whether the issuer is an Industrial Company, a Bank or Financial Institution, or a Public Utility, 1980–2003.




(b)

Fig. 2. *(Continued)*

Table 5
Time between an IPO and follow-on offerings, conditional on observing at least one follow-on offering, classified by security type, 1980–2000

| | *N* | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| A. First post-IPO issue regardless of security type | | | | | |
| Seasoned equity offerings | 1724 | 2.31 | 2.50 | 0.18 | 15.26 |
| Private placement of equity | 119 | 2.81 | 2.30 | 0.00 | 11.60 |
| Preferred equity | 61 | 2.64 | 2.61 | 0.07 | 10.98 |
| Convertible debt | 129 | 1.95 | 1.98 | 0.21 | 12.21 |
| Private placement of debt | 353 | 2.52 | 2.62 | 0.01 | 13.86 |
| Straight debt | 124 | 2.27 | 2.41 | 0.00 | 12.47 |
| Overall | 2531 | 2.35 | 2.51 | 0.00 | 16.44 |
| B. First post-IPO issue conditional on security type | | | | | |
| Seasoned equity offerings | 2665 | 3.30 | 3.26 | 0.13 | 19.70 |
| Private placement of equity | 214 | 3.80 | 3.11 | 0.00 | 18.70 |
| Preferred equity | 142 | 4.22 | 3.33 | 0.07 | 13.40 |
| Convertible debt | 315 | 3.62 | 3.15 | 0.21 | 15.58 |
| Private placement of debt | 1230 | 4.49 | 3.60 | 0.01 | 18.26 |
| Straight debt | 514 | 5.28 | 4.05 | 0.00 | 18.02 |

Source: Eckbo and Norli (2006). The table reports the number of calendar days between a firm's IPO date and the date of subsequent security offerings. The restriction that there must be at least one follow-on offering before 12/2000 (regardless of security type) restricts the sample from 6,092 to 2,531 IPOs. Panel A lists the time between the IPO and the first follow-on issue regardless of the type of security issued. Panel B shows the time between the IPO date and the date of the follow-on issue given that the security is of the type listed in the panel. $N$ is the number of security offerings after the IPO. For example, Panel B shows that there are a total of $N = 2{,}665$ (or 28%) SEOs following the 6,092 IPOs over the sample period.

Table 5, which appears in Eckbo and Norli (2006), shows descriptive statistics for follow-on security offerings made by 6,092 firms that went public during the period 1980–2000. A total of 3,579 firms (approximately 59%) do *no* follow-on offering during the sample period. Since firms going public in the last part of the sample period would have little time to do a follow-on offering, this number overstates the true fraction of non-follow-on firms. However, restricting the sample to the 3,750 IPOs that were completed in the period 1980–1993, which insures a minimum seven-year post-IPO period, a total of 1,977 firms (53%) did no follow-on offering during the interval 1980–2000. Overall, it appears that only one of two firms undertaking an IPO comes back to raise capital externally through a public security offering.[14]

[14] Firms that delist in the first few years after their IPO are even less likely to have any follow-on offerings. See Fama and French (2004) for information on survival frequencies in the population of listed firms. Eckbo and Norli (2005) show that delistings of IPO firms due to either acquisitions or bankruptcies in the first five years after the IPO is indistinguishable from the delisting frequency of seasoned firms matched on size and book-to-market ratio.

Panel A of Table 5 reports the average number of years between the IPO offer date and the first post-IPO security offering. In the sample of 6,092 IPOs, there are 1,724 firms that follow the IPO with a SEO as the first post-IPO security offering. The average number of years between the IPO and the SEO is 2.31 years. Panel A also shows that the SEO is the most common type of security offering to be made after the IPO. The second most common "first post-IPO offering" is a private placement of debt: 353 firms follow the IPO with this type of security.

The time from the IPO to the first security offering varies little across security types. The average time between the IPO and the follow-on security offering ranges from 1.95 years for convertible debt to 2.81 years for private placement of equity. Excluding convertible debt, the remaining five securities are offered on average between 2.27 and 2.81 years after the IPO. As suggested by Eckbo and Norli (2006), it appears that it takes on average 2.35 years to burn through the IPO proceeds, after which time companies may be selecting the security offering that minimizes issue costs.

Panel B of Table 5 reports the average number of years between the IPO and the first offering of security type $j$—regardless of whether or not security offering $j$ is the first to follow the IPO. Again, conditional on observing an IPO during the sample period, the most frequent security offering in our sample is SEOs. However, it is clear that if one does not condition on observing an IPO, the most common security offered is debt. As expected, the average number of years from the IPO to a specific security offering is longer than in Panel A of Table 5. The reason is that in Panel A each offering is required to be the first offering after the IPO. Panel B shows that following an equity IPO a convertible debt offering typically occurs sooner than a straight debt offering.

The finding that only one in two firms undertake a follow-on offerings is interesting. Although private firms almost certainly go public partly to get access to public security markets, external security issues (for cash) may be costly relative to internal financing. As discussed in Myers and Majluf (1984) and in Section 4 below, information asymmetries between the issuer and investors purchasing the issue may give rise to issue costs. These issue costs are found to be roughly proportional to the ex ante risk that an issue is overpriced, which leads Myers and Majluf (1984) to propose a financing pecking order. Internal equity (retained earnings) tops the pecking order, followed by debt securities and, finally, by external equity issues.

As surveyed by Frank and Goyal (2007), one prediction of the pecking order model is that debt ratios should be driven by the need for external funds. For example, the debt ratio should increase when firms experience a "financing deficit" (when retained earnings are insufficient to cover investment outlays). Shyam-Sunder and Myers (1999) find evidence consistent with this prediction. However, Frank and Goyal (2003) and Fama and French (2005) reach a different conclusion. Using a different sample than Shyam-Sunder and Myers (1999), Frank and Goyal (2003) find instead that net equity issues track financing deficits more closely than do net debt issues. Fama and French (2005) construct a measure of equity issues that includes any transaction that increases the split-adjusted number of shares outstanding. In addition to public equity offers for cash, such transactions include stock issues to employees, stock financed mergers, and rights offerings and direct purchase plans. Fama and French (2005) document that under their measure of equity issues, equity offerings are commonplace. For the three ten-year periods between 1973–2002 the authors find that 54%, 62%, and 72% of sample firms make net equity issues *every year*. They interpret this finding as a violation of the pecking order theory.

However, it is not clear that the evidence in Fama and French (2005), or studies of the Shyam-Sunder and Myers (1999) type of financing deficit, have the requisite power to reject the (basic) pecking order theory. Recall that this theory requires asymmetric information between the issuer and the investor purchasing the issue. A large proportion of the equity issues identified by Fama and French (2005) are stock swaps in mergers and acquisitions as well as stocks issued as part of employee compensation plans. It is difficult to imagine that stocks issued to CEOs give rise to adverse selection costs. Moreover, the ample opportunities for information exchange during merger negotiations also reduce adverse selection costs driven by information asymmetries. Also, given the *two-sided* information asymmetry associated with a stock exchange merger (the true value of the target shares is unknown to the bidder and vice versa), there is theoretical support for the proposition that the bidder prefers equity over cash or debt as the form of payment (see Eckbo, Giammarino, and Heinkel (1990) and the survey by Betton, Eckbo, and Thorburn (2007)). In sum, absent the requisite one-sided information asymmetry depicted in the original paper of Myers and Majluf (1984), evidence on the frequency of equity issues per se may have little power to test the pecking order. Of course, an equity issue for cash *does* satisfy this particular information asymmetry requirement since the value of cash is known to both sides of the transaction. As shown by Eckbo and Norli (2006) (Table 5 above), external equity issues for cash are indeed rare. This is consistent with the presence of external financing costs emanating from asymmetric information—as emphasized under the pecking order theory.

## 3. Flotation costs

To the extent that corporations choose among alternative financing methods so as to maximize the expected net proceeds of security offerings, flotation costs can have a large bearing on the choices an issuer makes. Broadly speaking, expected flotation costs includes components such as the expected issue announcement effect, expected underpricing, underwriter spread, expected out of pocket expenses, the probability of offer cancellation multiplied by the expected cost of cancellation,[15] and any short term incremental costs or benefits (if any) of moving away or towards a firm's target leverage ratio.

There is some disagreement on whether a security announcement is an expected flotation cost. Some researchers argue that a security offering announcement effect simply

[15] The expected cost of offer cancellation includes the loss of out of pocket expenses, management time and the expected opportunity costs of forgoing profitable investment projects if the offering isn't resurrected later.

conveys negative information about the issuing firm that managers alway knew, which would become public at some future date anyway, so why should it represent an issue cost? In contrast, other researchers view this announcement effect as capitalizing the direct and indirect effects of raising new equity capital, including empire building. At this point, we don't have resolution on this question. However, what we do know is that the typical negative announcement effect represents an expected permanent drop in the issue price. Furthermore, we view the early revelation of negative information about the issuer as an expected issue cost as well, as would any shareholder selling in the secondary market thereafter and as would any blockholder selling shares in a secondary offering. Evidence about security offering announcement effects is discussed extensively in Section 4, below.

While expected flotation costs tend not to change much over short time periods, market conditions and the firm's financial condition as well as the quality of publicly available information about the firm are all likely to vary substantially over longer periods of time (several years). For example, there are distinct differences in the level of underpricing needed to float a security issue and sizable differences in the likelihood of offer cancellation, both of which depend on current market conditions. Furthermore, our sample period has witnessed significant changes in securities regulations (such as shelf registration) and the competitive structure of the underwriting market—with the entry of commercial banks, investment banking industry consolidation and the increased internationalization of the security offering process—which can alter the level of underwriter competition and the pricing of their services.

Expected flotation costs also vary across firms at any point in time, depending of the characteristics of the issuers and the security offering. Thus, knowing these characteristics allows us to better forecast the expected flotation costs an issuer will bear from making a particular security offering. In the discussion to follow, we examine the existing evidence on the determinants of several of the flotation cost components.

### *3.1. Total flotation costs*

Flotation costs are made up of direct costs and indirect cost of selling a security through a public offering, where the direct costs include underwriter compensation, registration and listing fees, legal, accounting and printing expenses, etc. Underwriter compensation is made up of several components, the most important being the underwriter's gross spread or the difference between the public offering price and the underwriter purchase price. The other components of underwriter compensation include: an over-allotment option (typically this is a one month warrant to purchase an additional 15 percent of shares at the same price as the offering itself), plus long term warrants exercisable at the offer price, and extra reimbursements of underwriter expenses by the issuer.

Security sales also involve indirect flotation costs. The most important indirect cost is the typical underpricing costs associated with selling a security at a discount relative to both its prior trading day's closing price and its closing market price immediately following the public offering. Since an underwriter can allocate the issue, it is possible for an underwriter to capture much of the value associated with security offer underpricing. The effect of a security issue announcement on its offering price and the expected cost of a security issue delay or withdrawal are also potentially important indirect costs, which are discussed below. Finally, management time and energy devoted to the offering process is yet another significant, but hard to quantify indirect cost.

To summarize, expected flotation costs can be separated into direct and indirect cost components. Direct flotation costs are composed of:

- Fees to underwriters (including warrants and over-allotment options).
- Other out of pocket expenses, which include fees to accountants, law firms, listing fees, registration fees, printing, advertising as well as *road show* expenses and the cost of management time.

Indirect flotation costs include

- Issue underpricing, which can potentially be captured by underwriters through their power to allocate the issue to preferred customers and affiliates.
- Stock price reactions to initial offering announcements, which on average are negative, and any follow up announcements concerning changes in offer size and other characteristics.
- Costs of offering delays/cancellations.

Most of the extant literature focuses on the size and determinants of underwriting discounts (or fees) and security offering underpricing of equity offerings.

Early research on SEO flotation costs was conducted by the SEC staff in a series of studies and later by Smith (1977), who examined mean underwriter fees and other expenses. These two direct flotation cost components were examined across issue size categories and three major flotation methods. Flotation costs as a percent of gross proceeds were observed to fall with a rise in issue size. In addition, these costs were found to vary with flotation method for comparable size offers; more specifically, underwriter fees and other expenses were largest for firm commitments and cheapest for rights offers. Smith raised the question of why most U.S. firms appear to choose the highest cost flotation method and explored a number of possible added costs and benefits associated with each of these flotation methods. He was unable to explain away the puzzle. The question of whether there is a comparative advantage for alternative flotation methods was first raised by Hansen and Pinkerton (1982). A complication in undertaking this analysis is that the flotation method is an endogenous issuer decision, which could produce selection biases across the samples. These issues were more extensively studied by Eckbo and Masulis (1992) who re-examine the question of whether issuing firms fail to choose the flotation method that maximizes the net proceeds from their security sales. They uncover evidence consistent with the hypothesis that firms' choices do maximize net proceeds (further details on this issue are given in Section 4 below).

The issuer type and the flotation method choice generally affect both direct and indirect flotation costs of a particular security offering. As summarized in Table 6, in their sample of 1,249 SEOs over the period 1963–1981, Eckbo and Masulis (1992) report that the average direct cost of uninsured rights as a percent of total issue

Table 6
Total direct issue costs for U.S. issuers of seasoned equity, classified by issuer type and flotation method

| | Firm commitments | | Standby rights | | Uninsured rights | |
|---|---|---|---|---|---|---|
| Flotation costs | Ind | Utl | Ind | Utl | Ind | Utl |
| Number of observations | 351 | 639 | 42 | 89 | 26 | 23 |
| Underwriter compensation | 47 | 1.78 | 1.20 | 0.56 | | |
| ($ millions) | (1.03) | (1.32) | (0.47) | (0.34) | – | – |
| Other expenses | 0.16 | 0.14 | 0.36 | 0.38 | 0.11 | 0.45 |
| ($ millions) | (0.15) | (0.12) | (0.19) | (0.29) | (0.09) | (0.19) |
| Total costs | 1.72 | 1.92 | 1.59 | 0.94 | 0.11 | 0.45 |
| ($ millions) | (1.28) | (1.45) | (0.68) | (0.72) | (0.09) | (0.19) |
| Total costs/ | 6.09 | 4.23 | 4.03 | 2.44 | 1.82 | 0.51 |
| gross proceeds (%) | (5.53) | (3.82) | (3.32) | (2.07) | (0.94) | (0.22) |
| Total costs/ | 1.05 | 0.49 | 0.93 | 0.22 | 0.80 | 0.05 |
| market value common (%) | (0.68) | (0.41) | (0.57) | (0.18) | (0.30) | (0.02) |

Source: Eckbo and Masulis (1992). The sample size is 1,249 SEOs and the sample period 1963–1981. “Ind” denotes industrial issues and “Utl” denotes public utility. Data sources in the original study are the SEC Registered Offerings Statistics data tape and issue prospectuses. The cost of the offer price discount in firm commitment offers is not included, nor is the value of any “Green Shoe” options. In the standby rights category, the underwriter’s compensation is computed using the actual takeup fee based on subscription information.

proceeds is 1.82% for industrial issuer and 0.51% utility issuers. Despite a subscription rate that typically exceeds 70% (Hansen and Pinkerton, 1982; Eckbo and Masulis, 1992; Singh, 1997), the cost of standbys average as much as 4.03% of gross proceeds for industrials and 2.44% for utilities. Firm commitment offerings are the most expensive with average direct costs of 6.09% and 4.23% for industrial and utility issuers, respectively. Smith (1977), Hansen (1988) and Singh (1997) also presents costs of standby rights offerings consistent with those in Table 6. Furthermore, the low-cost status of uninsured rights holds internationally as well (e.g., Bøhren, Eckbo, and Michalsen, 1997; Slovin, Sushka, and Lai, 2000; Gajewski and Ginglinger, 2002).

Eckbo and Masulis (1992) also report that the average underpricing of SEOs in their firm commitment sample is very close to zero over their sample period (typically, the issue was offered at the previous closing price). As discussed below, this has since changed: it is now common to underprice a firm commitment SEO. Since current shareholders in a rights offer capture the value of underpricing through the value of the right, the development of underpricing in firm commitment SEOs further exacerbatesthe direct-cost disadvantage of this flotation method. It is clear that rights have lowest direct costs, while commitments is a firm's most expensive method, with standby rights in between.

Keep in mind that when comparing the costs of alternative flotation methods, one must control for firms' self-selection of the issue method. For example, as Hansen and Pinkerton (1982) point out, it is possible that observed flotation costs of uninsured rights are particularly low because this method is selected when a large blockholder is willing to guarantee subscription (which is typically the case). It is also possible that firms tend to select uninsured rights more generally when shareholder concentration is high, and when stock return variance is low. The point is that these and other characteristics can reduce direct flotation costs regardless of the chosen flotation method. To control for this effect, Eckbo and Masulis (1992) pool all flotation methods and use indicator variables for standbys and firm commitment issues in their cross-sectional regressions with direct issue costs as dependent variable. Conditional on various firm- and issue-specific factors, they conclude that the choice of an underwritten offer (standby or firm commitment) increases the flotation costs over and above uninsured rights, and that the choice of a firm commitment offer increases these costs further.

Lee et al. (1996) study direct flotation costs (underwriting spreads and other direct expenses as a percentage of offer gross proceeds) of IPOs, SEOs and issues of convertible and straight corporate debt over the 1990–1994 sample period. They find that the total direct issue costs are 11 percent for IPOs, 7.1 percent for SEOs, 3.8 percent for convertible debt and 2.2 percent for straight debt. They also document the frequency of issues with global tranches and over-allotment options. While debt offering flotation costs are low, it is important to keep in mind that debt issues, have a finite life of generally less than 10 years duration, especially taking into account sinking funds and callability. Thus, for a firm to have long term access to this debt capital, it is necessary to periodically refinance these debt issues, which involves repeated rounds of future flotation costs.

Public offering of debt can at times precede an IPO of stock, a phenomenon studied by Datta, Iskandar-Datta, and Patel (1997) and Cai, Ramchand, and Warga (2004). Firms issuing public debt are required to meet the SEC mandated financial disclosure requirements of public companies. Cai, Ramchand, and Warga (2004) report that subsequent IPOs by these firms are associated with significantly lower underpricing and lower price revisions from the midpoint of the filing range to the offer price. However, the lower underpricing is restricted to subsequent IPOs that have rated public debt, which tend to be financially stronger issuers. Also, public debt issues can be simultaneously offered with public equity issues, which is a financing decision studied by Hovakimian, Hovakimian, and Tehranian (2004).

### *3.2. Underwriter compensation*

Underwriter compensation is made up of three parts: management fees paid to the syndicate's lead underwriter or book runner, underwriting fees paid to the underwriters, and selling concessions to the syndicate members selling the shares to institutional and retail customers. In this literature, spreads are almost always measured as a percentage of offering size or gross proceeds. Most studies focus on either underwriter gross spread or underpricing costs, while very few studies estimate both the direct and indirect flotation costs of security offerings. Most studies also limit themselves to studying one security class, with SEOs being the most intensively examined offering type.

Moreover, most existing research on flotation costs focuses on the experiences of U.S. companies, primarily issuing common stock listed on major U.S. stock exchanges. Over the last 20 years, nearly all security offerings sold in the U.S. have relied on a firm commitment underwriting contract and a large majority of existing studies restrict their investigations to this sample. Most of these studies also limit their analysis to unregulated industrial firms. Since many of these studies also require the availability of machine readable accounting data, typically extracted from Compustat, the samples are further reduced by excluding many smaller firms not covered in this financial accounting database.

Kim, Palia, and Saunders (2005a) report underwriting spreads of industrial issues for SEOs, IPOs and straight corporate debt issues over the 1970–2000 period. They find that for the last three decades (i.e., 1970s, 1980s and 1990s) average underwriting spreads have fallen from 5.6 percent to 4.7 percent for SEOs, and from 7.7 percent to 6.7 percent for IPOs, with increased clustering of SEO spreads at 5 percent and IPO spreads at 7 percent. Similarly, average underwriting spreads have dropped in half from 1.6 percent to 0.8 percent for debt issues.

A consistent result found in the security offering literature is that underwriting spread rises with a security's total risk measured by return standard deviation over a pre-offering (SEOs) or post-offering (IPOs) estimation period. First, underwriting spreads are substantially larger for IPOs than SEOs, larger for SEOs than convertibles debt offers and smallest for straight debt offers. The average total risk (stock return standard deviation) of these classes of securities can likewise be ranked from highest to lowest. The rankings of total risk across security classes mirror those for security underwriter spreads: Total risk is on average highest for IPOs, followed by SEOs, then convertible debt and finally is smallest for straight debt. Within each of these security classes, there is also evidence that underwriter spreads are directly related to a security's return standard deviation.

The second major characteristic of security offerings found to reduce spreads is the offering size and this has been interpreted as an underwriting economy of scale effect due to the presence of large fixed costs, which exhibits increasing returns to scale. However, Altinkilic and Hansen (2000) takes issue with this interpretation. They point out that the observed fees do not fall steeply enough if they consist mostly of fixed costs. Thus, they argue that most of the fee is a variable cost, rather than a fixed cost. Offering size is also often measured as a percent of equity capitalization where it is interpreted as capturing an adverse selection effect.

A third very common characteristic used as a control variable is a measure of firm size, usually measured by firm book value of assets, market value of assets, equity market value (measured by book market value or debt plus equity market value), or firm annual sales. Firm size is generally interpreted as capturing asset diversification and the quality of publicly available information about the firm. These three characteristics are frequently used as control variables in this stream of literature examining underwriter spreads.

Two well cited studies of IPO underwriting spreads Chen and Ritter (2000) and Hansen (2001) document that these spreads strongly cluster at 7 percent, especially in the 1990s. However, in selecting their sample, Chen and Ritter exclude very large and very small issues where other levels of underwriting fees would most likely be observed. They interpret this as evidence that the market for underwriting services is oligopolistic. Hansen (2001) re-examines IPO underpricing without excluding relatively large and small issues and finds much greater variability in underwriting spreads. He also presents other evidence supporting the existence of a competitive underwriting market. More recently, Mullineaux and Roten (2005) compare IPO underwriting spreads by commercial banks and investment banks and find that commercial bank underwriters tend to be more concentrated at 7% than investment bank underwriters. Kim, Palia, and Saunders (2005b) examine trends in IPO and SEO underwriter spreads over the 1970–2004 period. They find evidence of a fall in IPO spreads over the 1990–2004 period, but no evidence of a change in SEO spreads, which is weak support for an increase in competition in the underwriting market.

In most studies of underwriter spreads, researchers take a particular focus, usually investigating an economic determinant of spreads that is not well documented in the literature, while controlling for other offering characteristics previously shown to affect spreads. For example, Kim, Palia, and Saunders (2005a) jointly study IPO underwriter spreads and underpricing, with particular focus on the interrelationship of underwriter spreads and underpricing. They argue that underpricing can be viewed as an additional form of compensation, which underwriters can capture through their power to allocate offers to favored customers. They find that IPO underwriter spread is positively related to IPO underpricing, a missing financial statement indicator and the inverse of the log of offer size and negatively related to the underwriter having a star analyst and issuer return volatility.

Turning to SEOs, Smith (1977) reports on direct flotation cost components classified by flotation method and offer size and scaled by gross proceeds. He calculates the mean values of both underwriter fees and other expenses across three major flotation methods; namely firm commitments, rights offers and standby offers. Smith finds that underwriter spreads average 5 percent of the offer price for firm commitments and that they range from over 10 percent for small issues to under 4 percent for very large issues.

Eckbo and Masulis (1992) study SEO underwriter spreads and flotation methods for industrials and utility issuers listed on NYSE and AMEX for nearly a 20 year period. They report underwriter fees and other flotation costs by flotation method and confirm

Smith (1977)'s findings that rights and standby offerings are less costly. Estimating determinants of direct flotation costs separately for industrial and utility issuers, they find for industrial issuers that flotation costs are negatively related to gross proceeds and average shareholding value and positively related to gross proceeds squared, return standard deviation, and percent change in shares. They emphasize the importance to flotation method choice of expected shareholder take-up in both rights and standby offers. Their evidence is consistent with the Myers and Majluf (1984) interpretation of the market's negative average announcement price reaction to an SEO as an upward revision in the market's expectation that the security is overvalued. They also find evidence that firms choose the flotation method that maximizes the net proceeds of their security offerings.

Altinkilic and Hansen (2000) study the determinants of underwriter spreads in industrial SEOs. They calculate mean underwriter spreads across offer size ranges and find that average spreads vary from 4.4 percent to 6.3 percent. They estimate the determinants of underwriter spreads as a function of the log of offer size, percent change in shares, return standard deviation and value of all underwritten industrial SEOs in the prior 3 months. They find that spread is significantly negatively related to log of offer size and positively related to percent change in shares, return standard deviation, the value of underwritten industrial SEOs in prior 3 months and the inverse of offer size when it is substituted for the log of offer size. Alternatively, Altinkilic and Hansen replace the log of offer size by the inverse of offer size and use it to estimate the slope of marginal spread. They find that the slope rises with offer size. This supports a rising variable cost of underwriting as offer size expands. Their perspective is that underwriter spreads are U shaped and that larger, less risky issuers have spreads that reach their minimum value at high offer sizes. Hansen (2001) examines whether this U shape spread phenomenon is present in IPO spreads prior to the rise of the 7% contract. He shows that IPO spreads are also consistent with rising variable costs, and are U-shaped. Corroborating evidence from German IPOs and SEOs is reported by Buhner and Kaserer (2002) and Kaserer and Kraft (2003) that marginal spreads are not decreasing in offer size. The Kaserer and Kraft analysis uses an principal components analysis within a generalized weighted least squares framework.

Kim, Palia, and Saunders (2005a) jointly study underwriter spreads and underpricing in SEOs as well as IPOs. They find that underwriter spreads are positively correlated with underpricing costs in SEOs and IPOs. They also investigate whether underwriter spreads are affected by market conditions, underwriter competition and issue characteristics using three stage least squares. They find SEO underwriter spread is positively related to underpricing, issuer leverage, missing financial statements, the inverse of the log of offer size and negatively related to market share of the top 25 underwriters, a top 25 underwriter indicator, indicator for bank entry into the underwriting market, and issuer profitability.

Butler, Grullon, and Weston (2005b) study the importance of SEO liquidity as a determinant of SEO underwriting spreads over the 1993–2000 period. They examine a broad range of liquidity measures including: quoted spread, effective spread, relative effective spread, quoted depth, trading volume, turnover, trade size, and a liquidity index of the above measures. They report that all the liquidity measures they examine are significant, with bid–ask spreads being positively related and the depth and activity levels being negatively related to underwriter spreads. They control for a broad range of other SEO characteristics and find that underwriting spreads are also negatively related to offer size, equity capitalization, share price, a multiple book manager indicator and positively related to return volatility, and Amex and Nasdaq indicators. In contrast, Altinkilic (2006) examines the role of underwriter market making immediately following SEOs to determine whether market making activities are partially paid by the underwriting spread. She argues that paying for market making in the underwriting spread takes pressure off the bid–ask spread, thus improving secondary market liquidity after the offer. Using abnormal share trading volume in the four weeks following the SEO as a proxy for market making costs, She finds that compensation for market making can explain 20% of the lead underwriter's total compensation, after controlling for other known determinants and that this underwriting fee component rises as the cost of market making rises.

More recently, Lee and Masulis (2006) examines the effect on SEO underwriting fees of financial accounting information quality, using a recent measure of accruals quality developed in the accounting literature by Dechow and Dichev (2002). They report that as the quality of issuer's financial accounting deteriorates, both SEO underwriting spreads, the negative announcement return, and frequency of offer withdrawals rise. They also find that a large number of other control variables are significant including log of net offer proceeds, secondary scale percentage, underwriter rank, log of total assets, stock return standard deviation and indicators for credit rated bonds and shelf offerings.

In another recent study, Drucker and Puri (1989) explore the effects of concurrent and prior lending and prior equity underwriting on the gross spreads of SEO. They find that a concurrent lending relationship, a prior lending relationship, or both, all reduce gross spreads. However, the effect of a concurrent lending relationship is stronger than a past relationship and a combined relationship is greater than a simple concurrent relationship. They also find that a past equity underwriting relationship reduces gross spreads, where they allow for a U-shaped spread following Altinkilic and Hansen (2000). This last result is consistent with several earlier studies of SEO underwriter competition that will be discussed later.

Table 7 summarizes the existing studies of underwriting spreads. The extant evidence shows that SEO underwriter spreads (1) exhibit a scale economy effect with diminishing marginal returns and (2) are negatively related to a firm's size and the offer's size relative to the issuer's equity capitalization. Finally, there is recent evidence that these underwriting spreads are negatively related to a security's liquidity and positively related to the quality of accounting information and existing and prior banking relationships. The evidence summarized in Table 7 is that SEO underwriting spreads are positively related to a firm commitment underwriting contract, percentage change in shares, inverse of offer size, log of offer size squared, underpricing, a missing financial statement indicator, bid–ask spread, prior SEO activity (prior 3 months), Amex and Nasdaq indicators.

Table 7
Evidence on underwriter spreads in IPOs and SEOs

| Study | Sample period | Explanatory variable | Sign |
|---|---|---|---|
| A. IPO studies | | | |
| Megginson and Weiss (1991) | 1983–1987 | Venture backing | − |
| | | Log(offer size) | − |
| | | Underwriter market share | − |
| | | Firm age | − |
| Hansen (2001) | 1980–1999 | Relative offer size | + |
| | | Stock return standard deviation | + |
| | | Log(offer size) | − |
| | | Secondary offering proportion | − |
| Kim, Palia, and Saunders (2005a) | 1970–2000 | Underpricing estimate | + |
| | | Missing financial statement (2 years) | + |
| | | Inverse log(issue size) | + |
| | | Stock return standard deviation | + |
| | | Lead underwriter's market share | − |
| | | Underwriter with All-Star analyst | − |
| B. SEO studies | | | |
| Smith (1977) | 1971–1975 | Firm commitment indicator | + |
| | | Offer size classes | − |
| Hansen and Torregrosa (1992) | 1978–1986 | Stock return residual st. error | + |
| | | Offer size | + |
| | | Log(offer size) | − |
| | | Log(equity market value) | − |
| | | Log(% manager shareholdings) | − |
| | | Syndicate manager portion of offer | − |
| Bhagat, Marr, and Thompson (1985) | 1982–1983 | Stock return residual variance | + |
| | | Log(offer size) | − |
| | | Stock beta | − |
| | | Utility firms | − |
| | | Shelf registrations | − |
| Eckbo and Masulis (1992) | 1963–1981 | Increased shares outstanding (%) | + |
| | | Stock return standard deviation | + |
| | | Log(offer size) squared | + |
| | | Log(offer size) | − |
| | | Ave shareholdings value | − |
| Denis (1993) | 1982–1985 | Log(offer size) ∗ shelf offer | + |
| | | Shelf issuer ∗ log(offer size) | + |
| | | Stock return variance (adjusted) | + |
| | | Log(offer size) | − |
| | | Shelf offer | − |
| | | Shelf issuer | − |
| Altinkilic and Hansen (2000) | 1990–1997 | Increase in shares outstanding (%) | + |
| | | Stock return variance | + |

(*Continued on next page*)

Table 7
(*Continued*)

| Study | Sample period | Explanatory variable | Sign |
|---|---|---|---|
| | | Recent SEO activity (prior 3 months) | + |
| | | Inverse of offer size | + |
| | | (offer size) | |
| Kim, Palia, and Saunders (2005a) | 1970–2000 | Underpricing estimate | + |
| | | Missed financial statement | + |
| | | Lead underwriter not in top 25 | + |
| | | Issuer leverage | + |
| | | Stock return standard deviation | + |
| | | Over-allotment option used | + |
| | | Inverse log(offer size) | + |
| | | Market cap ∗ Inverse log(offer size) | − |
| | | Herfindahl index in I-banking | − |
| | | Lead underwriters market share | − |
| | | Commercial bank market entry | − |
| | | Issuer profitability | − |
| Butler, Grullon, and Weston (2005a) | 1993–2000 | Bid–ask spread (%) | + |
| | | Log(stock return standard deviation) | + |
| | | Amex indicator | + |
| | | Nasdax indicator | + |
| | | Quoted depth | − |
| | | Trading volume | − |
| | | Issuer share turnover | − |
| | | Log(offer size) | − |
| | | Log(equity capitalization) | − |
| | | Log(share offer price) | − |
| | | Multiple syndicate book managers | − |
| Lee and Masulis (2006) | 1991–2002 | Poor accrual quality | + |
| | | Stock return standard deviation | + |
| | | Relative offer size | + |
| | | Shelf offering | − |
| | | Log(offer size) | − |
| | | Underwriter market share | − |
| | | Log(total assets) | − |
| Drucker and Puri (1989) | 1996–2001 | Inverse of offer size | + |
| | | Stock return standard deviation | + |
| | | Concurrent lending | − |
| | | Concurrent and prior lending | − |
| | | Prior lending | − |
| | | Prior equity underwriter | − |
| | | SEO market activity | − |

At the same time, SEO underwriting spreads are negatively related to offer size, issuer profitability, market depth, equity capitalization, share price, average shareholding value, market share of top 25 underwriter, commercial bank entry in the underwriting market, and multiple book managers.

### 3.3. Underpricing of SEOs

Underpricing is typically the most important indirect flotation costs in a security offering. There are several ways to measure underpricing of security issues. The offer price can be compared to the closing price, bid, ask or midpoint on the prior trading day or the first trade day following SEO completion. The offer price relative to the closing price on the offer date is generally termed the underpricing level. Researchers have also examined the offer price relative to the prior day's high and low prices.

We will focus most of our attention on recent empirical developments in IPO and SEO underpricing. Ljungqvist (2007) provides an excellent review of the theory and evidence on IPO underpricing elsewhere in this book. He concludes that much of the underpricing effect can be explained by information frictions including the Benveniste and Spindt (1989) theory that underwriters reward investors for information on issue demand through underpricing, as well as underwriter certification and various agency theory models which explore the conflict between IPO investors and issuer/management.

In a recent IPO study, Li and Masulis (2006) explore the effects of pre-IPO equity investments by major financial institutions including commercial banks, investment banks, venture capitalists and insurance companies, controlling for whether these financial institutions are also lenders to the firm or underwriters in its IPO. They examine these venture investment effects on IPO underpricing, offer price revisions from the filing range, post-IPO long run performance. Li and Masulis also employ a large number of other control variables used in earlier studies. They find evidence consistent with financial institution certification through venture investment, that is associated with lower IPO underpricing and offer price revisions and better long run performance. They also find that there are incremental certification effects as additional classes of financial institutions invest in these issuers. These results are robust to controlling for several forms of endogeneity. They also report that the coverage of pre-IPO loans is more completely reported in offering prospectuses than in the Dealscan loan database.

In another recent IPO study, Edelen and Kadlec (2005) develop a model of underpricing based on the probability of offer withdrawal and the importance of a successful offering. In essence, when the firm's stock price is rising before the offer day, managers are more willing to increase IPO underpricing to enhance the likelihood of a successful offering. Their model can explain why there is partial adjustment to public information released between the filing date and the offering date and it takes into account public information spillovers from the issuers industry. They report that their model can explain a large portion of the cross sectional dispersion in IPO underpricing and can explain hot issues markets. In their analysis, they use Heckman (1979)'s two step procedure where in the first step they estimate the probability of offer withdrawal and then in the second step they estimate the determinants of underpricing. They find that the estimated probability of an offer withdrawal has a significant negative effect on IPO underpricing. Their model also predicts an inverse relation between withdrawal frequency and industry stock returns between the filing and withdrawal dates. They argue that the asymmetric partial adjustment effect to industry information spillover effects found in earlier studies is due to a truncation regression bias and that once the withdrawal probability is taken into account this information spillover effect becomes symmetric.

Turning to SEO underpricing, Eckbo and Masulis (1992) examine mean and median underpricing by flotation methods for utility and industrial issues of NYSE and AMEX listed firms over the 1963–1981 period. They find that offer prices for firm commitments of industrial and utility issuers were on average underpriced by less than a half percent (i.e., 0.44 percent). Altinkilic and Hansen (2003) and Corwin (2003) investigate SEO underpricing of NYSE and Nasdaq listed stocks in more recent periods. Looking at mean underpricing by year, they find that it increases substantially in the 1990s relative to the 1980s. For example, Corwin (2003) reports that in the 1980s, it averaged 1.30 percent, while in the 1990s it averaged 2.92 percent. He observes that the rise in average underpricing of SEOs could be due in part to the large increase in the proportion of Nasdaq issuers, which in the 1990s were very young and with their asset values comprised mainly of risky intellectual property and growth options. However, a full explanation for SEO underpricing as well as its recent rise is still lacking.

Safieddine and Wilhelm (1996) analyzes the relationship of SEO underpricing to short selling. They examine offer date returns for industrials and utilities issuers with and without option trading and relate it to short interests in their stocks. They examine this activity before and after the enactment of Rule 10b-21, which prohibited using shares purchased at the offering price to close out short positions opened after the offering registration statement is filled. In this study, offering day returns are measured relative to the high and low prices on the day prior to the offer and the offer day. They report that underpricing is significantly negatively related to underwriter rank and a utility indicator and significantly positively related to abnormal short interest pre-Rule 10b-21 and an option trading indicator in the Rule 10b-21 period. They conclude that SEO offer dates exhibit abnormally high levels of short interest and option open interests and that SEO price discounts are positively related to these higher levels of short interest and option open interest. They also conclude that Rule 10b-21 appears to have curbed short selling activities and reduced underpricing, though Rule 10b-21 was implemented only three years earlier.

Kim and Shin (2004) re-examines the effects of short selling on underpricing using a longer and more recent sample period. They find that offer discounts are negatively related to underwriter rank and positively related to the Rule 10b-21 indicator, underwriter spread, and return volatility. Kim and Shin conclude that the SEC Rule was a partial cause for the temporal increased underpricing of NYSE listed stocks between the 1980s and 1990s, which runs counter to the conclusions of Safieddine and Wilhelm (1996). One serious concern with their study is that both underwriter rank and underwriter spreads are endogenously determined. Whether or not these results will hold up to taking this endogeneity in account is an open question.

Corwin (2003) reexamine the effect of Rule 10b-21 on underpricing using a model that excludes both underwriter rank and spread as regressors and draws a similar conclusion to Kim and Shin (2004). In his study, SEO underpricing is investigated for NYSE and Nasdaq listed stocks, with special emphasis on the differing market microstructure characteristics in the two marketplaces. He reports that underpricing is positively related to return standard deviation, average IPO underpricing in the month of SEO, relative offer size interacted with quartile indicators for the lowest stock prices and the highest stock return volatility and bid–ask spreads, and indicators for a negative 5 day pre-offer CAR, a tick size less than 0.25, and the Rule 10b-21 period. He also finds underpricing is negatively related to the closing price on day $-1$ and its interaction with offer price tick size less than 0.25, an NYSE indicator and the interaction of the negative 5 day pre-offer CAR indicator with the Rule 10b-21 period indicator. The negative NYSE indicator is consistent with the findings of Altinkilic and Hansen (2003) of greater underpricing for Nasdaq issues.

When Corwin estimates this model with Nasdaq quote data and adds several market microstructure variables, he finds similar findings, except that Nasdaq underpricing is also positively related to underwriter spread. He concludes that these changes can be explained by a variety of hypotheses related to asymmetric information (return standard deviation), temporary price pressure combined with inelastic demand relative offer size), short selling and manipulative trading (negative pre-offer CAR and Rule 10b-21 indicator), the informativeness of closing prices on the two exchanges (NYSE indicator), differences in underwriter pricing practices on these two exchanges (pre-offer price) and changes in the economics of the underwriting business (average IPO underpricing).

Kim, Palia, and Saunders (2005a) empirically examine the relationship between IPO and SEO underpricing and underwriter spreads. They find that underpricing is positively related to estimated underwriter spread. They also find that underpricing is positively related to the inverse log of issue size (consistent with Altinkilic and Hansen, 2003), the period with commercial bank underwriting and a prior 15 day momentum measure. They find SEO underpricing is negatively related to the market share of the top 25 underwriters, an indicator of a non top 25 lead underwriter and issuer equity market capitalization interacted with the inverse log of issue size. Their empirical analysis is based on a three-stage least squares model of underpricing and underwriter spread.

Evidence in several studies raises questions about the accuracy of the two benchmark prices used to measure underpricing, i.e., the offering day closing price and previous day's closing price. First, Altinkilic and Hansen (2006) report abnormal negative returns over the week prior to the SEO and abnormally high returns over the week following the SEO. Third, we know that underwriters can short sell shares of SEOs prior to the offering date and hedge them against their over-allotment options. Second, following an offering, stabilization activities can bias closing prices, cushioning price drops below the offer price for up to a month thereafter, though a couple of weeks or less is more common. Cotter, Chen, and Kao (2004) report price stabilization for SEOs is negatively related to offer price, trading volume, return variance and positively related to the interval between the filing and offer date. In addition, by looking at only completed SEOs, there can be some added selection bias where less favorably received offers are cancelled or delayed.

Of course accurate determination of the timing of an offering is critical to measure its price reactions, and Brown and Warner (1985) estimate the attenuation effect on measured market price reactions from inaccurate announcement dates. Another problem is that Lease, Masulis, and Page (1991) found a substantial proportion of SEOs are sold after the close of trading, rather than before the open, which is the more common occurrence. They used the Dow Jones time stamps to determine the actual time of day when the SEO is sold. Safieddine and Wilhelm (1996) use abnormal trading volume to determine the time of day when the SEO is sold and argue that this is more accurate approach. They also report a significant number of offers occurring after the market close.

A number of studies have investigated whether SEO underpricing is evidence of price pressure or a downward sloping demand curve. These studies include: Kadlec, Loderer, and Sheehan (1994), Corwin (2003), Meidan (2004) and Altinkilic and Hansen (2006). They report mixed results as to whether there is a downward sloping demand curve effect, short lived price pressure effect or adverse information effect similar to the observed effect of block trades. Kadlec, Loderer and Sheehan reports that in the months immediately surrounding an SEO there is evidence of a temporary stock price decline. Corwin (2003) finds SEO underpricing is positively related to relative offer size and interprets this as support for a price pressure effect. Meidan (2004) reports significant negative returns immediately before an SEO and significant positive returns immediately afterwards, which supports a price pressure effect. Altinkilic and Hansen (2003) report an unusually large negative mean return of $-2.6$ percent over the week prior to an SEO, followed by a small positive return in the week following the SEO, which is inconsistent with simple price pressure effect.

Table 8 provides a detailed summary of the empirical evidence from prior empirical studies on the determinants of underpricing of IPOs and SEOs. In light of the large number of explanatory variables studied, Table 9 provides a summary of these for easy reference. For the most part, the studies in this area report qualitatively consistent results for their effects on underpricing. Underpricing is found to be significantly related to (1) firm characteristics such firm size, financial condition, industry and share ownership structure, (2) security characteristics such as exchange listing, listed stock options, security volatility and market microstructure properties, and (3) offering characteristics such as offer size, offer price, underwriting syndicate, capital market conditions, other flotation costs and the likelihood of offer withdrawal.

Interestingly, venture capital backing, underwriter rank, and lead underwriter not in the top 25 are all often found to be significant, but with differing signs across the studies. This could reflect the endogeneity associated with the later two variables and underpricing. The varying sign of venture capital backing on underpricing is consistent with Habib and Ljungqvist (2001) who argue that the incentive to avoid underpricing an IPO will vary with the relative size of the primary and secondary shares that are offered. Thus, from this perspective it is important to model not only an indicator for venture backing, but also the size of venture shareholdings and whether these shares are being sold.

Table 8
Evidence on underpricing in IPOs and SEOs

| Study | Sample period | Explanatory variable | Sign |
|---|---|---|---|
| A. IPO studies | | | |
| Megginson and Weiss (1991) | 1983–1987 | Venture capital backing | − |
| | | Underwriter rank (market share) | − |
| | | Firm age | − |
| Booth and Chua (1996) | 1977–1988 | Log(offer price) | + |
| | | Underwriter rank ∗ firm commitment | − |
| | | Log(offer price) ∗ best effort | − |
| | | Prior IPO activity (past 3 month) | − |
| | | Prior IPO activity ∗ best efforts | − |
| | | Industry IPOs (12 month) ∗ best effort | − |
| Beatty and Welch (1996) | 1992–1994 | Underwriter compensation (%) | + |
| | | Log(1 + listed risks) | + |
| | | Underwriter rank (market share) | + |
| | | Inverse of offer price | − |
| | | Auditor market share (residual ) | − |
| | | Lawyer compensation (residual) | − |
| Carter, Dark, and Singh (1998) | 1979–1991 | Secondary offering (%) | + |
| | | Stock return standard deviation | + |
| | | Log(offer size) | − |
| | | Log(1 + firm age) | − |
| | | Underwriter rank (Carter–Manaster) | − |
| Gompers and Lerner (1999) | 1972–1992 | Log(equity capitalization) | + |
| | | Filing midpoint − offer price (%) | + |
| | | Log(book to market) | − |
| | | Underwriter rank | − |
| | | Underwriter and venture investor | − |
| Lowry and Shu (2002) | 1988–1995 | Lawsuit likelihood estimate | + |
| | | Market capitalization | + |
| | | Technology firm | + |
| | | Offer price − filing midpoint (%) | + |
| | | Cum. market return (prior 15 days) | + |
| | | Underwriter rank | − |
| | | Venture backed | − |
| | | NYSE/Amex listed | − |
| Lowry and Schwert (2002) | 1985–1997 | Technology firm | + |
| | | Filing midpoint − offer price | + |
| | | Filing midpoint − offer price > 0 | + |
| | | Underwriter rank | − |
| | | Log(real total assets) | − |
| | | National market system listed | − |
| | | Amex listed | − |
| Hansen (2001) | 1980–1999 | Underwriter rank | + |
| | | Stock return standard deviation | + |
| | | Secondary offering proportion | + |

(*Continued on next page*)

Table 8
(*Continued*)

| Study | Sample period | Explanatory variable | Sign |
|---|---|---|---|
| | | Log(offer size) | – |
| | | Leverage (Debt/Assets) | – |
| | | EBIT/Offer Proceeds | – |
| Habib and Ljungqvist (2001) | 1991–1995 | Underwriter spread estimate | + |
| | | Filing midpoint − offer price | + |
| | | Firm age | – |
| | | Log(sales) | – |
| | | Leverage | – |
| | | Increase in shares outstanding | – |
| | | Secondary shares (%) | – |
| | | Other expenses | – |
| Ljungqvist and Wilhelm (2003) | 1996–2000 | Targeted direct share programs | + |
| | | Proceeds for operating expenses | + |
| | | Estimated price revision | + |
| | | Estimated + price revision | + |
| | | High tech industry | + |
| | | Internet firm | + |
| | | 1999–2000 period | + |
| | | Venture capital shareholdings | – |
| | | Investment bank shareholdings | – |
| | | Corporate shareholdings | – |
| | | CEO shareholdings × Internet firm | – |
| | | Ownership concentration | – |
| | | Insider share sales | – |
| | | Venture capital share sales | – |
| | | Log(1 + firm age) | – |
| | | Secondary offer (%) | – |
| | | Increase in shares outstanding (%) | – |
| Edelen and Kadlec (2005) | 1985–2000 | Venture backing | + |
| | | Underwriter rank | + |
| | | Industry stk retns (filing to offer) | + |
| | | IPO underpricing (prior 30 days)[a] | + |
| | | Ave IPO offer price revision (30 day)[a] | + |
| | | IPO offer price-filing midpoint[a] | + |
| | | Log(offer size at filing) | – |
| | | IPO offer price revision < 0[a] | – |
| | | Estimated probability of withdrawal | – |
| Kim, Palia, and Saunders (2005a) | 1970–2000 | Underwriter spread estimate | + |
| | | Underwriter rank (market share) | + |
| | | Lead underwriter not in top 25 | + |
| | | Commercial banks enter market | + |
| | | Herfindahl index in I-banking | + |
| | | Cumulative mkt. ret. (prior 15 days) | + |
| | | Issuer profitability | – |

(*Continued on next page*)

Table 8
(*Continued*)

| Study | Sample period | Explanatory variable | Sign |
|---|---|---|---|
| | | Issuer leverage | − |
| | | Over-allotment option | − |
| Li and Masulis (2006) | 1993–2000 | Prior market return | + |
| | | Underwriter reputation | + |
| | | Venture capitalist share sale | + |
| | | Internet issuer | + |
| | | Global offering | + |
| | | Prior market return | + |
| | | New shares issued (%) | − |
| | | Bank shareholdings | − |
| | | Insurance co. shareholdings | − |
| | | Issuer bank loans | − |
| | | Venture capitalist shareholdings | − |
| | | CEO shareholdings | − |
| | | Fin'l institution shareholdings | − |
| | | Log(total assets) | − |
| | | IPO registration period | − |
| | | Big 6 auditor | − |
| B. SEO studies | | | |
| Bhagat, Marr, and Thompson (1985) | 1982–1983 | Stock return residual variance | + |
| | | Market return variance | + |
| | | Stock beta | − |
| | | Utility firms | − |
| Safieddine and Wilhelm (1996) | 1980–1988 | Abnormal short interest | + |
| | | Lead underwriter rank | − |
| | | Utility issuers | − |
| | 1989–1991 | Stock with listed options | + |
| | (Rule 10b-21) | Lead underwriter rank | − |
| | | Utility issuers | − |
| Corwin (2003) | 1980–1998 | Stk ret stnd deviation (prior 30 days) | + |
| | | Increase in shares outstanding (%) | + |
| | | Lowest market cap quartile | + |
| | | Highest standard deviation quartile | + |
| | | Lowest stock price quartile | + |
| | | Prior CAR $< 0$ (week prior to offer) | + |
| | | Offer price tick size $< 1/4$ | + |
| | | Rule 10b-21 in force | + |
| | | IPO underpricing in same month | + |
| | | Close − Bid on day $-1$ (%) $*$ Nasdaq | + |
| | | Log(stock price on day $-1$) | − |
| | | Log(stock price) $*$ Tick size $< 1/4$ | − |
| | | Prior CAR $< 0 *$ Rule 10b-21 | − |
| | | NYSE listed | − |
| Altinkilic and Hansen (2003) | 1990–1997 | Nasdaq listing | + |
| | | Relative offer size | + |

(*Continued on next page*)

Table 8
(*Continued*)

| Study | Sample period | Explanatory variable | Sign |
|---|---|---|---|
| | | Inverse of stock price | + |
| | | Stock return standard deviation | + |
| | | Offer size | − |
| | | Cumulative market return (from filing) | − |
| | | Cumulative abnormal stock return | − |
| | | Underwriter rank | − |
| Kim and Shin (2004) | 1983–1998 | Stk ret standard deviation (prior year) | + |
| | | Rule 10b-21 in force | + |
| | | Underwriter spread | + |
| | | Underwriter rank | − |
| Mola and Loughran (2004) | 1986–1999 | Nasdaq listing | + |
| | | Technology firm | + |
| | | Underwriter spread | + |
| | | Underwriter has top tier analyst | + |
| | | Offer price is an integer | + |
| | | Utility industry | − |
| | | Log(closing price on day −1) | − |
| | | Prior SEO | − |
| | | Underwriter rank | − |
| Kim, Palia, and Saunders (2005a) | 1970–2000 | Underwriter spread estimate | + |
| | | Commercial bank undrwrtrs allowed | + |
| | | Cumulative mkt. ret. (prior 15 days) | + |
| | | Inverse log(issue size) | + |
| | | Market cap ∗ Inverse log(issue size) | − |
| | | Underwriter rank (market share) | − |
| | | Lead underwriter not in top 25 | − |

### *3.4. Dependence between underpricing and underwriter spreads*

Mola and Loughran (2004) finds a significantly positive relationship for SEOs between underpricing and underwriter spreads. However, they do not fully control for the potential joint determination of these two costs. Kim, Palia, and Saunders (2005a) examines the relationship between underpricing and underwriter spreads. They find that in both SEOs and IPOs there is a positive relation between underwriter spreads and underpricing, though in the case of IPOs the relationship is driven by low quality issuers. They argue that these two flotation cost components can both be viewed as forms of underwriter compensation, which can be one explanation for their positive correlation. This evidence is consistent with Smith (1986), Hansen (1986) and Chen and Ritter (2000) who argue that underwriters and issuers jointly determining the direct and indirect costs of issuance.[16]

[16] Yeoman (2001) develops a model of net proceeds maximization where underwriter spreads and underpricing are interrelated. However, the predicted relationship is negative in his model.

Table 9
Summary of determinants of underpricing in IPOs and SEOs

| Variables with significantly positive effects | Variables with significantly negative effects |
|---|---|
| A. Issuer characteristics | |
| Firm size | Log of prior stock price |
| Technology issuer | Log of total sales |
| Internet issuer | Log of book to market |
| Prior cumulative stock return | Issuer profitability |
| Stock return's (or residual) standard dev. or variance | Percentage of tangible assets |
| Nasdaq listing | Firm age or Log(1 + firm age) |
| Stock with listed options | NYSE/Amex listed; Stock beta; Leverage; Prior SEO indicator; Utility issuer |
| B. Offer characteristics | |
| Log of offer price | Log of offer size |
| Offer price is an integer | Log of offer price * best effort |
| Offer price tick size less than 0.25 | Inverse of offer price |
| Offer price revision from midpoint of filing range | Underwriter rank (market share) |
| Proceeds used for operating expenses | Underwriter rank * firm commitment |
| Targeted direct share purchase programs | Lead underwriter not in the top 25 |
| Log(1 + listed risks in SEC filing) | Qualified independent underwriter employed |
| Abnormal short interest in stock | Over-allotment option used |
| Estimated likelihood of a lawsuit | Auditor market share |
| Underwriter rank | Big 6 auditor |
| Underwriter with top tier analyst | Legal compensation |
| Herfindahl index for investment banking | Prior week cumulative stock return * Rule 10b-21 |
| Lead underwriter not in top 25 | Log of prior stock price * indicator of offer price tick size less than 0.25; Filing to offer date interval; Estimated probability of offer withdrawal |

(*Continued on next page*)

To analyze the potential interdependence of spread and underpricing, Kim, Palia, and Saunders (2005a) employ three stage least squares to estimate the jointly determined underwriter spread and underpricing, which they note gives consistent estimates. They find three instruments that are significantly related to spreads, but not to underpricing (existence of a star analyst, issuers lacking two years of financial statements at the IPO date, standard deviation of daily stock returns for one year), and one instrument related to underpricing, but unrelated to spreads (market run-up over the prior 15 trading days). They point out that this interdependence raises some serious questions about the reliability of many earlier studies, which focus exclusively on underpricing or underwriter spreads, and generally do not control for the potential interdependence of these two flotation cost components. While the study makes a strong case for interdependence of underpricing and spreads, it is less convincing in its claims about the appropriate instruments needed to identify their three equation system.

Table 9
(*Continued*)

| Variables with significantly positive effects | Variables with significantly negative effects |
|---|---|
| C. Market conditions | |
| Prior cumulative market return | Prior IPO activity |
| 1999–2000 "Bubble" period | Prior IPO activity ∗ best effort |
| Average IPO underpricing in the prior month | Prior industry IPOs ∗ best efforts |
| Commercial banks allowed to underwrite securities | Industry stock returns (filing to offer date) |
| Rule 10b-21 in force | Negative industry stock returns (filing to offer date) |
| Global offering | Out of pocket expenses |
| Estimated or actual underwriter spread | Percentage secondary offer |
| Percent increase in shares outstanding and its interactions with: (1) lowest market capitalization quartile, (2) lowest stock price quartile, (3) highest stock return standard deviation quartile | Average offer price revisions in prior 30 days |
| D. Share ownership | |
| Venture capital backing | CEO shareholdings ∗ Internet firm |
| Venture capitalist selling shares | Issuer share ownership concentration; Investment bank shareholdings; Investment bank non-underwriter shareholdings; Commercial bank shareholdings; Commercial bank underwriter shareholdings; Venture capital backing; Venture capital shareholdings; Corporate shareholdings; Insurance company shareholder; Insider share sales; CEO share sale; Venture capital share sales; Commercial bank lender |

Another serious methodological issue is the extent to which various explanatory variables found to be correlated with underpricing and underwriter spreads are themselves endogenously determined. In this category, underwriter ranking has been most extensively studied and Ljungqvist and Wilhelm (2003) and Habib and Ljungqvist (2001) conclude that it is endogenously determined. Habib and Ljungqvist also find evidence that the some of underwriting fees and out of pocket expenses, which they call promotion costs are significantly related to underpricing and endogenously determined as well. Habib and Ljungqvist also test whether number of shares sold is endogenously determined and conclude that is not.

Another explanatory variable that is often used in explaining underpricing is the price revision from the filing range midpoint, measured by the offer price minus the midpoint, divided by the midpoint. Since underpricing is also a non-linear function of offer price, there is a danger that this strong empirical association is being driven mechanically by the common component in the two measures.

### 3.5. Offering delays and withdrawals

Another component of expected flotation costs is the costs of bearing most of the out of pocket expenses associated with preparing a security offering without realizing the benefits of actually raising capital due to an offering cancellation. In addition, this capital short fall can have adverse implications for a firm's ability to pursue the positive net present value projects that it has available to it and may have a negative effect on the timing, size and pricing of a subsequent security offering. Interestingly, several early studies by Mikkelson and Partch (1986) and Officer and Smith (1986) reported that announcements of SEO withdrawals are greeted by a positive market reaction. Examining both SEO and convertible debt withdrawals, Jensen and Pugh (1995) report similar positive stock reactions. Altinkilic and Hansen (2006) report that SEO withdrawals are preceded on average by a precipitous stock price drop of 17 percent. To the extent that offer cancellation has negative implications for the firm's financial condition and the size of flotation costs and its ability to pursue investment projects, this positive price reaction suggests that the market was skeptical about the profitability of the firm's planned investment projects or else was concerned that the reason for the stock offer was that the stock was seriously overvalued, following the logic of Myers and Majluf (1984).

Edelen and Kadlec (2005) explore the implications of the risk of offer cancellation on the pricing of the offering. They observe that as offer price discount rises the risk of offer cancellation falls. This can explain why issuers are willing to go forward with offerings that they know are underpriced and why positive information released between the filing and offering dates is only partially incorporated into the final offer price as documented by Hanley (1993). Taking into account that some firms will have greater need for funds than others, and that new public information about the stock's value will vary across offerings, they are also able to develop a model to predict which offers will be more underpriced. In estimating the probability of offer withdrawal using a probit model, they find that it is significantly *positively* related to industry returns between the filing and offering date, prior IPO initial returns (30 days), log of the offer size, and withdrawals of earlier IPOs and significantly negatively related to prior IPO offer price revisions between the filing and offering dates (prior 30 days), and underwriter rank.

### 3.6. Underwriter competition

There is conflicting evidence on whether the market for underwriter services is highly competitive or oligopolistic. Chen and Ritter (2000) argue that the high frequency of 7 percent underwriter spreads in IPOs is evidence that this market is far from perfectly competitive. Hansen (2001) reports a number of pieces of evidence about the IPO process that supports the contention that this market is highly competitive, such as an IPO with 7 percent underwriter spread does not contain abnormal profits relative to other IPOs, that there is no evidence of monopoly profits in underpricing or unusual charges in subsequent SEOs, and that the 7 percent contract has persisted despite the

Department of Justice investigation of collusion allegations following the release of the Chen and Ritter (2000) findings. Hansen (2001) also reports that measures of concentration in the IPO market are well below the level considered by the Department of Justice to be anticompetitive. He notes that underwriters compete in many dimensions in addition to underwriter spreads, so that convergence to a common spread like 7 percent is not strong evidence of anticompetitive behavior.

Dunbar (2000) studies market share changes of book managers of IPOs and finds that they are negatively related to IPO first day returns and underwriter compensation (fees) and positively with analyst reputation.[17] This suggests that underwriters are competing implicitly, if not explicitly, on the level of IPO underpricing and underwriter spreads, contrary to the popular notion that banks do not cut fees to attract business. Corwin (2003) finds that seasoned offers were underpriced by an average of 2.2 percent during the 1980s and 1990s, with the discount increasing substantially over time, and that underpricing is significantly related to underwriter pricing conventions such as price rounding and pricing relative to the bid quote. Mola and Loughran (2004) also documents the increased usage of price rounding in setting SEO offer prices. These results appear to suggest a weakening in underwriter competition.

Adding to this debate, Burch, Nanda, and Warther (2005) examine underwriting fees of repeat security issuers to determine the relation between loyalty to a bank underwriter and the fees charged. They find that loyalty is associated with lower fees for common stock offers, but higher fees for debt offers. For both offer types, firms that graduate to higher ranked banks face lower fees. They also show that firms, which tend to switch banks to improve analyst coverage, pay higher fees in common stock offers, but do not pay higher fees in debt offers.

In contrast to this evidence, Ellis, Michaely, and O'Hara (2004) report that while many firms "graduate" to better underwriters, most firms move laterally or are downgraded in terms of lead underwriter ranking. They show that firms that graduate to a higher ranked underwriter must pay a premium for the privilege (i.e., above the fee charged by the same underwriter to an existing client for a similar deal), and, similarly, firms that use a lower ranked underwriter for their equity offering must also pay a premium.

Krigman, Shaw, and Womack (2001) studies underwriter selection in IPOs and finds that the quality of the analyst team is a key factor in underwriter selection. They also find that better performing IPO firms often switch to higher ranked underwriters for their SEOs. In addition, they conducted a field-based survey of chief financial officers (CFOs) and chief executive officers (CEOs) of IPO firms, who later switched underwriters, as to which factors were most important to their underwriter selections. Their survey reveals that the most important factors for issuers' senior management in selecting a lead underwriter are underwriters' and analysts' reputations, with issue pricing

[17] Interestingly, Dunbar (2000) also finds that banks lose market share if they are associated with overpriced IPOs, consistent with Booth and Smith (1986)'s certification theory.

and market making ability being moderately important and underwriting fees being the least important attribute. This ranking suggests that competition over underwriting fees is unlikely to have much explanatory power empirically. Mola and Loughran (2004) estimates the determinants of SEO underwriter market share and finds that a highly regarded analyst team increases the underwriter market share by 1.5 percent, adjusting for other factors (see their Table 5).

Ellis, Michaely, and O'Hara (2000) report that lead underwriters are initially the most active market maker in IPO stocks. Ellis, Michaely, and O'Hara (2004) find that the economic significance of lead underwriter market making declines as IPO stocks become seasoned over the following year. Corwin and Schultz (2005) show that number of market makers and analysts that are covering a stock rise with syndicate size. This suggests that the quality of underwriter market making and analyst coverage are likely to be less important to larger issuers, who benefit from greater investor interest. Consistent with this, Altinkilic (2006) reports that the market making component in SEO underwriting spreads is lower for larger firms.

Ljungqvist et al. (2004) document that analysts' recommendations relative to the consensus are positively associated with investment banking relationships and brokerage pressure, but negatively associated with the presence of institutional investors in the firm being followed. The latter result is especially strong when there are more institutions holding larger blocks in the firm, and for firms whose institutional holdings are concentrated in the hands of the largest institutional investors. They conclude that presence of institutional investors (who are primary customers of the analysts' services) provides an incentive mechanism for the analysts not to succumb to pressure to provide favorable opinions on their employers' investment banking clients and to boost brokerage business. Ljungqvist, Marston, and Wilhelm (2006) find optimistic analyst reports don't help underwriters win SEO assignments. Instead, they find that analysts' reputation, lending relationships and bond underwriting increase the bank's chances of winning underwriting assignments.

Ellis, Michaely, and O'Hara (2004) report that underwriters with continuing issuer relationships tend to charge lower fees, have optimistic analyst forecasts and are active in writing analyst reports. Banks competing for new SEO assignments often take actions in advance of an underwriting assignment: add analyst coverage, make optimistic analyst forecasts, do not compete on fees and do not become more active in market making services. Banks gaining new SEO assignments move quickly to: add analyst coverage, issue optimistic forecasts and increase their market making presence. Banks facing a weakened or terminated issuer relationship tend to reduce their analyst coverage, eliminate the positive bias in analyst forecasts, but do not reduce their market making services. They conclude that investment banks compete for follow-on equity offering underwriting business along multiple-dimensions (such as fees, underpricing discount, analyst coverage, market making, debt relationship, and overall reputation), and that underwriters who deliver on all these dimensions are retained by firms, and can be viewed as providing superior overall service to the issuer.

Fernando, Gatchev, and Spindt (2005) develops and empirically tests a model of firm-underwriter selection, where high (low) quality underwriters tend to sign contracts with high (low) quality issuers. In their empirical tests, they find that issuers and underwriters will associate with different partners for subsequent offerings if changes in issuer quality and/or underwriter reputation are large enough, suggesting that the association of issuers and underwriters is transactional rather than relationship-based. However, Kim, Palia, and Saunders (2005a) report evidence that the frequency of low (high) quality issuers using high (low) quality underwriters is as frequent as high (low) quality issuers employing high (low) quality issuers, which appears to be strong evidence against the Fernando, Gatchev and Spindt model.

Gande, Puri, and Saunders (1999) was the first study to examine the competitive effects of commercial bank entry into the corporate debt underwriting market. They find that underwriter spreads and ex-ante yields have declined significantly following commercial bank entry in the market, consistent with commercial bank underwriters with prior lending relationships with issuers having an information advantage over investment banks. They show that the reduction in underwriter spreads and ex-ante yields is strongest among lower rated and smaller debt issues, where commercial banks have underwritten a relatively greater proportion of these issues (as compared to investment banks). They also show that bank entry has tended to decrease market concentration, suggesting that commercial bank entry generally has had a pro-competitive effect. However, whether this is a short-term rather than a long-term effect is yet to be determined. Narayanan, Rangan, and Rangan (2004) study commercial bank entry into the equity underwriting market and report that commercial banks are increasing their roles as lead managers in equity underwriters, though they usually participate as a co-lead manager with an experienced investment bank.

Using a sample of SEOs from 1996–2001, Drucker and Puri (1989) finds that when a financial intermediary concurrently lends to an issuer and underwrites the firm's SEO, the issuer benefits through lower financing costs, receiving lower underwriter fees and lower loan yield spreads. This is particularly true for non-investment grade issuers, for whom the informational economies of scope are likely to be large. They show that concurrent lending also helps underwriters build relationships, increasing the probability of receiving future business. Specifically, they show that issuers with prior lending relationships receive lower underwriter spreads, while an underwriter with a prior lending relationship with an issuer is more likely to receive its subsequent underwriting assignments.

Wu and Kwok (2003) study global IPOs and the effects of competition by examining the pricing of global initial public offerings made by U.S. companies as compared to purely domestic offerings. They find that global participation significantly reduces underpricing (on average by four percentage points), and that underpricing is negatively related to the proportion of shares allocated to foreign investors. They conclude that U.S. companies time their global offerings when foreign demand for U.S. shares is high. Cornett, Davidson, and Rangan (1996) investigated the effects of Rule 415 on the level of competition in the investment banking industry and find that it has weaken the competitive position of the smaller underwriters.

### 3.7. Rights and standby offerings

Since the 1950s, rights and standby offerings are used with less frequency in the U.S. However, they are still commonly employed by some regulated financial firms. Utilities, REITS, closed-end funds and conversions of mutual thrifts or insurance companies to stock charter are examples of right issuers discussed in the literature, e.g., Singh (1997), Khorana, Wahal, and Zenner (2002), Higgins, Howton, and Howton (2003), Howe and Shilling (1988), Masulis (1987). More recently, there has been a resurgence of the used of rights offers beyond utilities and financial firms by financial distressed industrial firms as reported by Heron and Lie (2004) and Ursel (2006).

### 3.8. Shelf registered offerings

In 1983 the SEC gave final approval to Rule 415, a new regulation that allowed security issuance under an expedited registration process. This option was only available to larger publicly listed firms. Bhagat, Marr, and Thompson (1985) studied direct and indirect flotation costs (underwriting fees and other expenses and underpricing) for a small sample of syndicated firm commitment and shelf issues found that shelf offerings have lower flotation costs than traditional book building method.

Sherman (1999) develops a model of underwriter certification and the effect of shelf registrations. She concludes that shelf registrations increase underwriter competition and reduce the quality of their due diligence investigations. Blackwell, Marr, and Spivey (1990) examine whether shelf issues reduce underwriters due diligence investigations and results in higher underpricing. They report that underwriter spreads vary with issuer quality and that weaker issuers have to pay a premium relative to firms using a firm commitment offering. Denis (1991) reported that most industrial security issuers used shelf offerings primarily for debt securities, which have much lower due diligence concerns. Denis (1993) finds that firms that use shelf registrations some of the time, also have lower non-shelf SEOs flotation costs. Thus, the inference about the cost saving associated with using shelf registration was thrown into question. However, Dennis also notes the low frequency of shelf registered SEOs is consistent with there not being a cost advantage.

More recently, shelf registration was expanded in 1992 to universal shelf issues, which allows the offering to be either debt or equity. This change is likely to intensify underwriter competition. Since the rule change, universal shelf registrations have dominated equity shelf registrations. Moreover, a greater portion of universal shelf issues result in equity offerings. Autore, Kumar, and Shome (2004) revisited the issue of flotation costs and the impact of shelf registration. They report that shelf issues of SEOs have overtaken non-shelf issues as the dominant flotation method beginning in 2001 for NYSE, Amex and Nasdaq listed firms. They report that 85 percent of shelf registrations result in no subsequent offer. They find that shelf issues have lower costs and greater timing flexibility. These results hold up after adjusting for the self-selection bias highlighted in the early Denis (1991) study. They also separately study universal shelf issues that result in an SEO. They note that shelf issues create valuable options that become more valuable under more volatile market conditions. Bethel and Krigman (2004) re-examine the question of reduced due diligence in shelf issues. They report that high asymmetric information issuers experience high discounts from using the shelf registration mechanism, which explains why this mechanism isn't more widely employed.

### *3.9. Over-allotment options, warrants and other direct expenses*

*Over-allotment options*. A second component of underwriter compensation is an over-allotment option, which is a warrant to buy an additional 10–15 percent of the offering at the same price as the SEO/IPO. The typical over-allotment option has a maximum life of 30 days. Underwriters can use these options to lower their risk exposure in a firm-commitment underwriting contract. This underwriter hedging activity in the IPO market is the focus of a study by Aggarwal (2000). She finds that underwriters exercise over-allotment options to cover short positions created by underwriters over-selling securities in public offerings when the after-market stock price rises relative to the offering price. She also finds that underwriters buy shares in the after-market to cover short positions when the stock price falls to the offering price or lower.

Over-allotment options can alternatively be viewed as valuable short term warrants held by underwriters that allow them to purchase up to an additional 15 percent of an undervalued offering at the underwriter's discount from the public offer price. Little research is available on the value of these options, with the exception of an early study by Hansen, Fuller and Janjigian (1987), who examine over-allotment options in SEOs of industrial firms. They estimated the value of the typical over-allotment option to be 1 percent of the offer's gross proceeds. They also report that about half their offer sample had over-allotment options. Using a logit model, they find that over-allotment options are more frequent in offers with smaller dollar size, larger relative size, greater stock and market return variances and more retail oriented (strong broker system) underwriters. In the IPO market, Lee et al. (1996) report that virtually all U.S. issues include over-allotment options and nearly all are for 15 percent of the original issue size and are issued at-the-money. Further, about 60 percent of the options are either partially or fully exercised, with the vast majority fully exercised.

*Warrants as additional underwriter compensation*. Several studies by Ng and Smith (1996) and Dunbar (1995) investigate the use and importance of warrants as an additional element of underwriter compensation in SEOs. Controlling for the selectivity imbedded in the choice of using warrants as added underwriter compensation with a logit model, they find that warrant use reduces the overall flotation costs of SEOs. Since warrants are less valuable when the underlying stock is overvalued, the credibility of smaller and less well known underwriters is increased when they accept warrants as compensation. This can reassure investors who could otherwise question the credibility of less reputable underwriters, thus lowering the average SEO underpricing necessary to sell these issues.

*Other direct flotation expenses*. The analysis of the other expenses such as registration and listing fees, legal and accounting and printing expenses is fairly limited. Smith (1977) finds for firm commitment SEOs that other direct expenses average about 1.15 percent of the offer price. He also examines the determinants of these other direct expenses and finds them to be functions of flotation method and offer size, measured by gross proceeds. His evidence documented a strong economy of scale effect in direct total flotation costs, with the smallest offerings having total direct costs ranging from 14 to 15 percent and the largest offerings having a total cost of less than 4 percent. Altinkilic and Hansen (2000) argue that a large fraction of these fees (85 percent) are a variable cost. There is little added information on other expenses.

Eckbo and Masulis (1992) estimate the determinants of direct flotation costs (sum of underwriting fees and other expenses). They find that on average direct flotation costs average over 6 percent for industrial issues and 4.25 percent for utility issues. They also report that they have a non-linear relationship to size (−), percentage change in outstanding shares (+ for industrials), log of holdings per shareholder (−), prior stock return standard deviation (+) and an indicator of underwritten firm commitments and standby offers (+).

Habib and Ljungqvist (2001) analyze the relationship of out-of-pocket expenses plus underwriting fees (which they term "promotion costs") and underpricing. They develop a model that assumes that the issuer makes decisions to minimize the wealth loss of going public, which includes the cost of underpricing and the promotion costs. They predict that promotion costs increase with the portion of the IPO that represents insider selling (size of secondary offer), the relative offer size and uncertainty. In testing their model they take account the endogeneity of underpricing, promotion costs and underwriter rank. They find that promotion costs are positively related to the estimated relative offer size, estimated proportion of insider sales and several risk proxies, namely underwriting fees and the log of sales while they are negatively related to gross proceeds and firm age. These results support the predictions of their model.

*Other flotation costs of rights*. Rights offerings are generally used only in SEOs. As noted earlier, a rights offer involves issuing short lived in-the-money warrants to existing shareholders on a pro-rata basis. This issue method differs substantially from a firm commitment method and has several potentially large indirect issue costs, which are borne by the issuer and its shareholders.[18]

(1) *Capital gains taxes*. In a rights offer, shareholders who do not wish to purchase shares of the issue must sell their rights (or subscribe and sell the shares) in order to avoid losing the value of their subscription rights or warrants. These sales are subject to capital gains taxes, which are increasing in the subscription price discount, discouraging large discounts.

[18] This discussion is partially drawn from Eckbo and Masulis (1995).

(2) *Stock liquidity and transaction costs of reselling rights*. The resale of rights by current shareholders takes place on organized exchanges, entailing dealer spreads and brokerage fees. Since shareholders avoid these costs when the firm employs an underwriter to sell its new shares, a rights offer carries an added transaction cost disadvantage for shareholders uninterested in exercising their warrants. Kothare (1997) argues that rights issuers have typically high ownership concentration, and a rights offering tends to increase concentration. The result is a higher adverse selection effect associated with buying the stock (or the rights), which Kothari finds raises the stock's bid–ask spread and this reduced liquidity is likely to lower the stock's market price.

(3) *Arbitrage activity and the risk of rights offer failure*. Investors can use rights as warrants to hedge their short sale positions in a firm's stock. This encourages increased short selling of the stock, but as additional short positions are opened, the stock price will tend to be depressed as resulting sell orders rise (at least within the bid–ask spread). Thus, between the announcement of rights offer terms and offer expiration, this short-selling activity tends to keep the stock price down, reducing the attractiveness of exercising rights for most stockholders. This creates additional uncertainty for issuers as to the ultimate rights offer subscription level.

(4) *Anti-dilution clauses and wealth transfers to convertible security holders*. If a firm has convertible securities or warrants outstanding with anti-dilution clauses in place, then issuing rights at discounts can trigger automatic reductions in conversion rates of these securities as discussed in Kaplan (1965) and Myhal (1990). These anti-dilution clauses are likely to result in improved positions for the convertible security holders, shifting wealth away from the common stock holders who are the residual claimants. As a result, there is an added incentive for firms with convertible securities outstanding to avoid issuing rights at deep discounts.

### *3.10. Market microstructure effects*

Seasoned public offers of common stock have important impacts on the secondary market in which the common stock trades.[19] The typical firm commitment offer involves a large increase in shares outstanding along with a large increase in the number of stockholders and a reduction in management and blockholder percentage ownership. As a result, one would anticipate that there would be major increases in trading volume, changes in bid–ask spread and depth, increased insider trading at the end of the lock-up period, and possibly major changes in price volatility after the public offering. One would also expect similar effects on secondary market trading of corporate bonds following subsequent bond offerings of similar seniority and duration bonds.

Theories of bid–ask spread determination are based on adverse selection and inventory cost considerations. These theories predict that if trading volume rises and price

[19] Parts of this section are drawn from Eckbo and Masulis (1995).

volatility falls, then bid–ask spreads will also fall since the expected costs of market making decline. The SEO announcement per se can also lower the asymmetric information about the firm's stock price borne by market makers, which would cause bid–ask spreads to drop further.

Amihud and Mendelson (1986) develop a valuation model of security pricing that assumes that investors have a positive preference for liquidity measured by percentage bid–ask spread. They derive a model of security pricing where the expected return is a positive and concave function of bid–ask spread. Amihud and Mendelson (1988) extend the implications of the model and present evidence that liquidity is an important determinant of security value. They argue that managers seeking to maximize current stockholder wealth should take market liquidity into account when making corporate financing decisions. Thus, in deciding whether to make an SEO and in choosing the flotation method, liquidity implications need to be taken into account. A further implication is that the negative adverse selection effect of the offer announcement can be partially offset by the positive liquidity effect.

Lease, Masulis, and Page (1991) explore the market microstructure effects of firm commitment SEOs for NYSE and AMEX listed firms. They document that share trading volume rises substantially and that price volatility falls subsequent after the SEO. Not surprisingly, both dollar bid–ask spreads and percentage spreads fall significantly after the seasoned public offering, consistent with inventory cost and adverse selection cost models of bid–ask spread determination. They also report that trading volume and price volatility fall between the announcement and the offer dates, while bid–ask spreads drop, but not to the level observed subsequent to the public offer. This is suggestive of a modest increase in liquidity following the SEO announcement and a significant improvement after the SEO. Not withstanding the improvement in stock liquidity, Altinkilic and Hansen (2006) report that on average issuer stocks experience an abnormal negative return of 2.6 percent over the week prior to the SEO. They also find that this effect cannot be explained alone by a short term price reversal effect in the immediate post-SEO period and suggest that this is due to a negative information effect related to the underwriting process.

Tripathy and Rao (1992) examine the market microstructure effects of SEOs for NASDAQ listed firms. They split their sample into large and small capitalization stocks and that larger stocks have increases in bid–ask spread over a 60 day period prior to an SEO announcement, which is followed by decreases in spread over the next 43 days. In contrast, small stocks experience increases in spread from 80 days prior to the announcement through 20 days after the announcement. Focusing on the public offering date, they find that the bid–ask spreads of large stocks decrease over the 20 days prior to the offering and decrease even more over the 20 days following the offer. Spreads of small stocks increase over the 20 days prior to the offering, but then decrease beginning just before the offering through 20 days after.

Masulis and Shivakumar (2002) separately investigates the speed of price reactions measured in 15 minute intervals to SEO announcements by NYSE/Amex and Nasdaq listed stocks. They report that Nasdaq listed stocks react more quickly to these announcements (by about an hour) and attribute it the differences in the organizational structure of the NYSE/Amex and Nasdaq market places. They find evidence consistent with NYSE/Amex limit order books and market opening mechanisms slowing price reactions to news. They also report a large number of trading halts (21%) on the NYSE around daytime SEO announcements, while there are very few on Nasdaq.

Stock offers can also cause temporary biases in daily stock returns by disrupting normal buy-sell order flow in the secondary market. Lease, Masulis, and Page (1991) document that around the public offer dates of SEOs stock returns are biased downward due to the loss of purchase orders to the temporary primary market in the stock. One result is that stock transaction prices tend to occur at the lower ask quote, rather than at the midpoint of the bid and ask, which generates an apparent fall in the stock price. There is also evidence that market makers may lower their quotes in this period due to a positive imbalance in their inventory position resulting from the predominance of sell orders at this time. Lease, Masulis and Page find that using the closing bid–ask average rather than the closing transaction prices eliminates the statistical significance of the drop and reduces by more than half the average negative offer date return.

Several more recent studies explore the impacts of market microstructure on securities issuance. Presumably, as lead underwriter they have better knowledge of potential buyers and sellers, which should give them a competitive advantage in market making immediately after the IPOs, especially for larger orders. Ellis, Michaely, and O'Hara (2000, 2004) report that the typical lead underwriter is highly active as a market maker immediately following the IPO, but that this role diminishes over the following year. Corwin, Harris, and Lipson (2004) examine IPOs listed on the NYSE and report that initial buy-side liquidity is higher for IPOs with high quality underwriters, large syndicates, low insider sales and high pre-market demand (offer is priced at or above the maximum filing range price), while sell-side liquidity is higher for IPOs that represent a large fraction of outstanding shares and have low pre-market demand (offer is priced at or below the minimum filing range price). Limit order trading is very weak on the first day of trading, though there is an unusual number of limit buy orders submitted at the offer price for cold IPOs, which are likely to be underwriter stabilization bids. They also report that pre-opening order flow is a good predictor of first day prices and are reflected in the opening price set by the specialist. Field, Cao, and Hanka (2004) study the effects of lock-up expirations on IPO stocks and find that substantial increases in insider trading by officers and directors in almost 25 percent of cases do not adversely affect stock liquidity. They find only a 3 percent increase in effective bid–ask spreads that lasts only about one week, while depth and trading activity substantially improve.

Mola and Loughran (2004) studies the effects of market microstructure factors on SEO underpricing, along with the effects of underwriter competition. They find that the offer price discount is positively related to relative offer price, a tech indicator, gross spread and a top tier analyst indicator and negatively related to a utility indicator, log of share price, a high underwriter reputation indicator and an integer offer price indicator. Mola and Loughran conclude that changing issuer composition toward smaller, riskier Nasdaq listed issuers and increasing underwriter market power measured in terms of underwriter market share, underwriter reputation and analyst quality can explain this phenomenon. As discussed earlier in the SEO underpricing section, Corwin (2003) also explores many of these issues. In addition, Butler, Grullon, and Weston (2005b) find that the underwriter spreads are negatively related to a wide range of stock liquidity measures, while Altinkilic (2006) reports that spreads are directly related to market making effort.

### *3.11. Miscellaneous offerings*

#### *3.11.1. Global offerings*

Global issues are often sold through an ADR or GDR mechanism to minimize foreign exchange issues for foreign investors. Under these mechanisms, a depository bank holds the original stock and issues new shares that are denominated in local currency and pays cash dividends in the local currency. Global offerings by U.S. firms generally use the GDR mechanism. The supply of ADRs or GDRs can be expanded or contracted by the depository bank purchasing more shares of stock or selling back some of its shareholdings with the creation or redemption of a like number of claims to these shares through the issuance or redemption of ADRs/GDRs. Foreign issuers selling shares in the U.S. must register their securities under Rule 144A as is discussed in greater detail below.

The implication of cross listing of its stock on firm value is studied by Doidge, Karolyi, and Stulz (2004). They argue that cross-listing in the United States helps controlling shareholders of foreign firms commit to limit their expropriation of minority shareholders, since U.S. security laws are stricter than most other jurisdictions. They also argue that cross-listing increases the ability of these firms to raise equity capital at more attractive terms, allowing the firms to take advantage of their growth opportunities. They show supporting evidence in that foreign companies with shares cross-listed in the U.S. had market to book ratios (at the end of 1997) that were 16.5 percent higher than that of non-cross listed firms from the same country, and that growth opportunities are more highly valued for firms that cross-list from countries with weaker investor rights (also, see LaPorta, Lopez-de-Silanes, and Shleifer, 1999).[20]

Ljungqvist, Jenkinson, and Wilhelm (2003) examine the tradeoff between investor demand estimation methods (book building versus fixed-price) and the costs associated with hiring an underwriter for initial public offerings (IPOs). Book building conditions the final issue price on market demand conditions, whereas in case of a fixed-price method, shares are priced first and then later put up for subscription. Using a dataset containing 2,143 IPOs by issuers from 65 countries outside the United States during January 1992–July 1999, they show that book building, when used in combination with U.S. banks (as underwriters) and U.S. investors, can reduce underpricing significantly relative to fixed-price offerings or book building efforts by other banks. They attribute

[20] For a more recent survey of the literature on cross-listings, see Karolyi (2006).

this result to the fact that because of their longer book building experience, U.S. banks are more likely to have access to key institutional investors and may be in a better position to reward investors dynamically for their information revelation. Interestingly, they show that for most issuers, the gains associated with lower underpricing outweighed the additional costs associated with hiring U.S. banks, such as the 7 percent gross spread that is typically paid when U.S. banks are involved (see Chen and Ritter, 2000, and Hansen, 2001).

Wu and Kwok (2003) study the underpricing, underwriting fees and direct expenses of global IPOs. They report that global offers significantly reduce underpricing by 4 percentage points relative to purely domestic IPOs. The result can not be explained by potential selection bias in the offering decision. Underpricing is found to be a decreasing function of the relative size of the global tranche. They also find that global offers are more likely as the prior performance of the U.S. stock market rises. They also find no evidence of differences in underwriting spreads or other expenses.

Bruner, Chaplinsky, and Ramchand (2004) examine the direct and indirect costs of raising equity capital for a sample of 293 first-time foreign IPOs in the United States (i.e., these companies did not have their stock traded in a domestic market or other foreign market prior to the IPO) and compare the costs to those of U.S. IPOs. They conclude that in general foreign IPOs experienced approximately the same capital raising costs as the U.S. IPOs, with the exception of foreign firms with strong investor demand and upward revisions to offer prices that incurred a smaller underpricing than that of U.S. IPOs.

Chaplinsky and Ramchand (2004) analyze the choice between issuing public and private (under Rule 144A) debt by foreign firms. They conclude that SEC Rule 144A, which permitted firms to raise capital (in terms of both debt and equity) from qualified institutional buyers without requiring registration of these securities or compliance with U.S. GAAP, has resulted in the Rule 144A debt market replacing the public debt market in terms of number and volume of foreign debt issuers, especially for high-yield and non-rated issues.

### *3.11.2. Convertible securities and warrants issuance*

Convertible debt and equity securities can be viewed as a method of issuing stock in the future, contingent on the issuer's financial conditional improving. As such, these securities are very similar to issuing warrants plus straight debt or preferred stock. These securities are often issued by privately held firms, which are raising capital from venture capitalists. These convertible securities are generally convertible preferred stock with an automatic conversion into common stock if the firm goes public. Unlike public issues of convertible securities, these privately placed equity issues generally carry powerful governance rights and may also have the feature that on conversion to common stock, the liquidation rights of the preferred issue may not have to be relinquished.

Public offerings of convertible securities are frequently convertible debt or straight debt with detachable warrants. These convertible securities are generally issued out-of-the-money with American exercise rights over most or all of the security's life. Many of these securities are also callable, which is a method that allows the issuer to force the in-the-money convertible securities and warrants to convert their securities to common stock. Also, typical convertible securities held by venture capitalists automatically convert to common stock at the time of an IPO. Lastly, many convertible securities are not protected against cash dividends, which can again create incentives on the part of the option holder to exercise their conversion rights early, so as to avoid the stock price fall associated with the ex-dividend effect. Mayers (1998) argues that firms with significant real options can benefit from issuing convertible securities that don't have to be exercised until after the real options are exercised. This is similar to staged financing in the private equity market. Mayers finds that prior to calls of convertible bonds, firms exhibit increases in capital expenditures and new long term debt financing, consistent with the exercise of important real options.

There have been a variety of studies of convertible debt, convertible preferred stock and warrant issue including: Brennan and Schwartz (1982), Stein (1992), Nyberg (1995), Kang and Lee (1996), Mayers (1998), Lewis, Rogalski, and Seward (1998), Byoun and Moore (2003), Korkeamaki and Moore (2004), and Brick, Palmon, and Patro (2004). Most of these studies have focused on offering methods, offering frequencies and announcement effects. A few of these studies have also examined components of flotation costs.

### *3.11.3. Private placements of equity and convertibles*

Wruck (1989) was first to study private placements of equity by publicly listed firms. She documented that these negotiated sales of equity by large NYSE listed firms had a positive mean announcement effect of 4.5 percent on the issuer's stock price unlike the average negative announcement effects of public offerings of stock. She analyzes the changes in shareholder ownership and concentration and documents that a private placement on average increases the voting power of the dominant blockholder and reduces the voting power of management. She finds that the change in stock value is strongly correlated with the change in ownership concentration. Sales that afterwards give the blockholder under 5 percent or more than 25 percent ownership have positive effects, while intermediate blocks result in negative effects. Moreover, sales that result in a change in control or an increase in management shareholdings have a negative effect. She argues that increasing shareholder concentration often increases shareholder wealth by improving firm efficiency and alignment of interests with outside shareholders, but at times can adversely effect outside shareholder wealth, when it is likely that substantial firm resources are diverted to private benefit.

In a follow up study, Herzel and Smith (1993) examine private placements by primarily smaller Nasdaq listed firms. They document that private placements are sold on average at substantial discount of 20 percent relative to public offerings. They argue that this underpricing is to compensate private placement investors for their investigation costs prior to investing, while the positive announcement effect reflects the positive information effect associated with a sophisticated institutional investors agreeing to purchase shares, rather than improved monitoring of management by blockholders. Hertzel and Smith also report that institutional investment declines in private placement firms.

Wu (2004) examines the identity of private placement investors. She reports that private placement firms have higher asymmetric information than firms that rely on public offerings based on issuer age, lack of venture capital backing, fewer institutional investors and wider bid–ask spreads and coverage by fewer analysts. Also, she finds that private placement investors who engage in more intensive monitoring (i.e., venture capitalists and pensions funds) are not increasing their holdings in these firms after the private placements. This result is inconsistent with increase monitoring of management after the private placement. Finally, discounts on private placements sold to managers are higher than those when managers are not involved. These discounts are also higher when managers' initial holdings are lower. These last two results are consistent with management self-dealing. Wu also reports that private placement investors are typically passive, which is consistent with the evidence of Barclay, Holderness, and Sheehan (2005).

Gomes and Phillips (2005) examine a comprehensive sample of 13,000 private and public security issues of debt, convertibles and common stock by publicly listed firms. They find that in the recent 2000–2003 period private issues exceed public issues. Gomes and Phillips report that publicly listed firms with higher levels of asymmetric information (measured by analysts' earnings forecast errors or dispersion in earnings forecasts) are more likely to issue debt in the public market, while they are more likely to issue riskier equity and convertible securities in the private capital market. They also find that smaller public firms with higher risk, lower profitability and good investment opportunities are more likely to issue equity and convertible securities privately, while public equity issues are more likely for firms experiencing a stock price rise in the prior year relative to a benchmark portfolio.

More recently, a new type of private placements of equity by public companies (PIPES) has become popular, especially with small and medium size companies. The PIPE market originated with the SEC adoption of Regulation S in 1990, which permitted U.S. issuers to sell unregistered shares to foreign investors at any price in off shore markets without first registering them with the SEC or publicly disclosing them. In 1996, the SEC modified its rules to require issuers to report the sale of Reg S shares and required investors to hold these shares for a year. To gain greater liquidity, issuers typically registered the PIPE shares with the SEC via a shelf registration within 30 days of closing of the deal. The securities typically become effective 90 days after registration.

There are two major types of PIPEs. There are traditional PIPEs that are fixed number of shares or a convertible with a fixed strike price, which can be sold at a discount through private negotiations and there is a more recent innovation called structured PIPEs. Structured PIPEs represent convertible securities having variable strike prices that decline if the underlying stock prices decline beyond a specified interval. A structured PIPE allows investors to convert into a larger number of shares if the stock price declines, thereby giving investors significant downside protection.[21] Not surprisingly, Brophy, Sialm, and Ouimet (2005) report that younger firms with weak performance in industries with high growth rates and risk levels (i.e., greater adverse selection) are the primary issuers. The typical investors in PIPEs are hedge funds.

#### *3.11.4. Unit offerings in IPOs and SEOs*

Unit offers involve the issue of a combination of common stock and warrants by an issuer. One potential advantages of selling units rather than shares is that when an issuer is very risky the market is apt to overestimate its leverage and its return volatility, which causes its warrants to be overvalued, while the stock is apt to be undervalued. The result of selling a unit is that these two effects are combined and become partially offsetting, which means firms sell the unit offers at closer to its true market value. This is similar to Brennan and Schwartz (1982) argument for why firms issue convertible securities. Warrants also give investors more time before committing to buy equity, which acts as a credible signal that the issuer holds no negative proprietary information about the firm's value. Taking into account the callability of many warrant and convertible issues and the cost of financial distress Stein (1992) argues that this can be a backdoor means of selling more equity, when the market over-estimates the adverse selection risk associated with the issuer. He finds that firms with intermediate levels of risk should issue convertibles. Unit offers of SEOs have been studied by Schultz (1993), Chemmanur and Fulghieri (1997) and Byoun (2004).

### *3.12. Conflicts of interest in the security offering process*

Recently there has been a stream of new research exploring potential conflicts of interest by decision makers in the security offering process. These conflicts are sometimes between managers and securityholders, and in other cases between underwriters and either security investors or security issuers. A key question is whether these potential conflicts are large enough to alter the security underwriting process to a measurable degree and if so, do any underwriter customers suffer any serious financial consequences. A second important question is whether there are significant economic benefits from combining underwriting and other financial services.

One major concern is that at least some managers make security issuance, pricing, and underwriting decisions to benefit themselves, rather than their shareholders. Managers can accomplish this by issuing underpriced securities to friends and family, or capturing side payments from underwriters, for instance through underwriter allocations of other firms' underpriced IPOs, often called spinning or receiving new stock options exercisable at the IPO offer price, which represent valuable in-the-money options. Studies that explore this line of research include Jung, Kim, and Stulz (1996) who

[21] A similar security is studied by Hillion and Vermalen (2004). They investigate floating rate convertible debt, which adjusts the conversion ratio for stock price drops.

tests whether firms undertaking SEOs when facing poor growth opportunities, measured by market to book ratios, are experiencing agency conflicts between managers and shareholders. Consistent with this hypothesis, they find that some firms with poor growth opportunities do undertake SEOs and that these firms have more negative announcement effects. Ljungqvist and Wilhelm (2003) finds that managers participating in friends and family programs and not making secondary offerings are more apt to have underpriced IPOs. Datta, Iskandar-Datta, and Raman (2005) presents indirect evidence that on average SEO announcement effects are positively related to managers' equity based compensation, so greater equity based compensation is associated with less negative announcement effects. Kim and Purnanandam (2006) reports a similar finding.

Turning to the management compensation effects of IPOs, Lowry and Murphy (2006) examines whether IPOs are underpriced more because managers obtain more valuable stock options with lower strike prices (set at the offer price) when new stock option plans are established at the IPO date. They find no evidence of a positive relation between underpricing and IPO option grants, which does not support a serious conflict of interest effect.

A second avenue of concern is that underwriters may have conflicts of interest with their customers due to joint production of underwriting and other financial services including brokerage, market making, security analysis, venture capitalist investing, lending and asset management, to name a few. Many researchers have investigated whether the joint production of these services creates serious conflicts of interest or whether there are significant economies of scale or scope realized from sharing financial information produced in the course of performing one or more of these services. Since financial service providers need timely information about customers' financial strength, joint production of information or sharing of this information can be particularly cost efficient.

Of all of these related services, the area that has elicited the most research interest is security analysis by underwriting firms. Underwriters seek to reduce the time and expense of selling a security offer and to lower their risk of offer failure, and the question is whether these incentives dominate the analyst's reputation concerns near security offering dates, causing sell-side analysts to hype these issues through overly optomistic earnings forecasts and investment recommendations. Michaely and Womack (1999) report evidence of such a bias. However, more recent evidence does not support this finding. Kadan et al. (2005) report that after the 2002 NYSE and NASD rules regulating sell side analyst's investment banking relationships, there is no evidence that analysts issue optimistic earnings forecasts. However, these same analysts remain reluctant to recommend selling stocks that their investment banking arms are underwriting. Other studies that find affiliated analysts do not make more optimistic earnings forecasts includes: Kolasinski and Kothari (2004), O'Brien, McNichols, and Lin (2005), Barber et al. (2005), Agrawal and Chen (2004) and Ljungqvist, Marston, and Wilhelm (2006).

There is a stream of literature including Puri (1994, 1996, 1999), Gande et al. (1997), Schenone (2004), Chaplinsky and Erwin (2005), Drucker and Puri (1989) and Li and

Masulis (2006) examining situations where lenders are also debt or equity underwriters. The basic concern is that underwriters who are also lenders have incentives to underwrite weak security issues to strengthen the financial condition of borrowers. These studies generally find no evidence supporting a significant conflict of interest effect.

Another potential underwriter conflict of interest with IPO investors occurs when IPO underwriters are also venture investors since venture investors realize substantial financial benefits when their portfolio firms complete IPOs. Several recent studies by Li and Masulis (2005, 2006) examine whether underwriters alter their underwriting and pricing decisions when they have venture investments in these issuers. However, they find no evidence to support underwriters weakening their underwriting standards to improve the returns on their venture investments.

## 4. The flotation method choice

In this section, we examine the firm's choice of issue method. We start with the so-called rights offer paradox first observed by Smith (1977). The paradox highlights the fact that a focus on *direct* issue costs alone fails to adequately explain the near disappearance of the rights offer method for large, publicly traded corporations in the U.S. We then examine how observed flotation method choices may minimize issue costs under asymmetric information and survey the empirical evidence on announcement effects of security offerings as a function of the flotation method.

### *4.1. The paradoxical decline in the use of rights*

With symmetric information between corporate insiders and outside investors, standard economic theory predicts a preference for the relatively inexpensive uninsured rights offer method for floating seasoned equity. Nevertheless, Table 10 shows that as of the mid-1970s, publicly listed companies in the U.S. have virtually abandoned the rights issue method in favor of firm commitment underwritten offerings.[22] Furthermore, this phenomenon is not restricted to U.S. offerings. Ursel and Trepanier (2001) show a strong trend towards declining use of rights and increasing use of public offerings in Canada 1970–1985. The trend away from rights is also evident in Japan: Table 11 shows a dramatic decline in rights offerings after the mid-1990s. Slovin, Sushka, and Lai (2000) report that uninsured rights represents a small fraction of total SEOs by British firms listed on the London Stock Exchange. In Hong Kong, rights are also now in a minority (Wu, Wang, and Yao, 2005). Bøhren, Eckbo, and Michalsen (1997) present evidence that issuers on the Oslo Stock Exchange have moved from uninsured rights

[22] A corporation's charter originally stipulates that shareholders have the first right of refusal (preemptive right) to purchase new equity issues. Thus, abandoning the rights method requires a shareholder vote in favor of eliminating the preemptive right. Such charter amendments became popular among U.S. publicly traded firms in the early 1970s, preceding the move towards firm commitment offerings. See also Bhagat (1983).

Table 10
SEOs by NYSE- and AMEX-listed firms, classified by flotation method (FC = firm commitments, Stand = standby rights), 1935–1955 and 1963–1981[a]

| | Total issues | | | | Industrial issues | | | | Utility issues | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total | FC | Stand | Right | Total | FC | Stand | Right | Total | FC | Stand | Right |
| 1935 | 6 | 1 | 3 | 2 | 5 | – | 3 | 2 | 1 | 1 | – | – |
| 1936 | 37 | 11 | 17 | 9 | 37 | 11 | 17 | 9 | – | – | – | – |
| 1937 | 40 | 15 | 18 | 7 | 39 | 15 | 17 | 7 | 1 | – | 1 | – |
| 1938 | 5 | 2 | – | 3 | 4 | 1 | – | 3 | 1 | 1 | – | – |
| 1939 | 13 | 6 | 3 | 4 | 8 | 5 | 3 | – | 2 | 1 | – | 1 |
| 1940 | 18 | 9 | 4 | 5 | 13 | 7 | 4 | 2 | 3 | 2 | – | 1 |
| 1941 | 9 | 1 | 3 | 5 | 9 | 5 | 3 | 1 | 6 | 2 | – | 4 |
| 1942 | 1 | 1 | – | – | 1 | 1 | – | – | – | – | – | – |
| 1943 | 14 | 8 | 5 | 1 | 13 | 7 | 5 | 1 | 1 | 1 | – | – |
| 1944 | 23 | 13 | 9 | 1 | 22 | 12 | 9 | 1 | 1 | 1 | – | – |
| 1945 | 52 | 23 | 18 | 11 | 45 | 20 | 15 | 10 | 7 | 3 | 3 | 1 |
| 1946 | 110 | 73 | 24 | 13 | 96 | 65 | 21 | 10 | 14 | 8 | 3 | 3 |
| 1947 | 53 | 27 | 12 | 14 | 29 | 19 | 5 | 5 | 24 | 8 | 7 | 9 |
| 1948 | 61 | 20 | 20 | 21 | 28 | 11 | 9 | 8 | 33 | 9 | 11 | 13 |
| 1949 | 79 | 27 | 30 | 22 | 14 | 7 | 5 | 2 | 65 | 20 | 25 | 20 |
| 1950 | 84 | 35 | 31 | 18 | 30 | 16 | 9 | 5 | 54 | 19 | 22 | 13 |
| 1951 | 131 | 61 | 49 | 21 | 63 | 40 | 16 | 7 | 68 | 21 | 33 | 14 |
| 1952 | 131 | 66 | 43 | 22 | 71 | 41 | 20 | 10 | 60 | 25 | 23 | 12 |
| 1953 | 120 | 55 | 47 | 18 | 43 | 28 | 11 | 4 | 77 | 27 | 36 | 14 |
| 1954 | 101 | 51 | 33 | 17 | 51 | 36 | 11 | 4 | 50 | 15 | 22 | 13 |
| 1955 | 113 | 44 | 56 | 13 | 56 | 29 | 25 | 2 | 57 | 15 | 31 | 11 |
| 1935–1955 | 1,201 | 549 | 425 | 227 | 677 | 376 | 208 | 93 | 525 | 179 | 217 | 129 |
| 1963 | 12 | 2 | 6 | 4 | 5 | 1 | 3 | 1 | 7 | 1 | 3 | 3 |
| 1964 | 17 | 8 | 6 | 3 | 8 | 4 | 3 | 1 | 9 | 4 | 3 | 2 |
| 1965 | 20 | 5 | 9 | 6 | 11 | 5 | 4 | 2 | 9 | 0 | 5 | 4 |
| 1966 | 27 | 12 | 12 | 3 | 17 | 7 | 8 | 2 | 10 | 5 | 4 | 1 |
| 1967 | 26 | 12 | 9 | 5 | 17 | 9 | 4 | 4 | 9 | 3 | 5 | 1 |
| 1968 | 44 | 26 | 9 | 9 | 31 | 20 | 4 | 7 | 13 | 6 | 5 | 2 |
| 1969 | 42 | 24 | 15 | 3 | 22 | 13 | 7 | 2 | 20 | 11 | 8 | 1 |
| 1970 | 49 | 36 | 10 | 3 | 22 | 18 | 2 | 2 | 27 | 18 | 8 | 1 |
| 1971 | 84 | 65 | 15 | 4 | 44 | 40 | 2 | 2 | 40 | 25 | 13 | 2 |
| 1972 | 81 | 68 | 11 | 2 | 29 | 27 | 1 | 1 | 52 | 41 | 10 | 1 |
| 1973 | 58 | 50 | 6 | 2 | 12 | 10 | 1 | 1 | 46 | 40 | 5 | 1 |
| 1974 | 53 | 47 | 4 | 3 | 6 | 5 | 0 | 1 | 47 | 42 | 3 | 2 |
| 1975 | 89 | 79 | 8 | 1 | 20 | 19 | 1 | 0 | 69 | 60 | 8 | 1 |
| 1976 | 93 | 88 | 3 | 1 | 30 | 29 | 1 | 0 | 63 | 59 | 3 | 1 |
| 1977 | 65 | 62 | 3 | 0 | 2 | 2 | 0 | 0 | 63 | 60 | 3 | 0 |
| 1978 | 90 | 86 | 3 | 1 | 25 | 23 | 2 | 0 | 65 | 63 | 1 | 1 |
| 1979 | 85 | 81 | 2 | 2 | 21 | 20 | 0 | 1 | 64 | 61 | 2 | 1 |

(*Continued on next page*)

Table 10
(*Continued*)

| | Total issues | | | | Industrial issues | | | | Utility issues | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total | FC | Stand | Right | Total | FC | Stand | Right | Total | FC | Stand | Right |
| 1980 | 162 | 157 | 2 | 3 | 87 | 86 | 0 | 1 | 75 | 71 | 2 | 2 |
| 1981 | 152 | 149 | 1 | 2 | 64 | 63 | 0 | 1 | 88 | 86 | 1 | 1 |
| 1963–1981 | 1,249 | 1,057 | 134 | 57 | 473 | 401 | 43 | 29 | 776 | 656 | 92 | 28 |
| 1982–2003 | 8,708 | 8,375 | 333[b] | – | 8,241 | 7,912 | 329[b] | – | 467 | 463 | 4[b] | |

[a]The information is from Stevenson (1957) (1935–1955), Eckbo and Masulis (1992) (1963–1981), and SDC (1982–2003). Stevenson (1957) lists common stock issues with proceeds over $1 million appearing in Sullivan and Cromwell Issuer Summaries 1933–1950 and in The Commercial and Financial Chronicle 1950–1955. Eckbo and Masulis (1992) base their sample on the Wall Street Journal Index, the Investment Dealer's Digest, and Moody's Industrials and Utilities Manuals. Their sample excludes simultaneous offers of debt/preferred stock/warrants, combination primary/secondary stock offerings, cancelled or postponed offers, and non-U.S. issues. The SDC sample shown for the period 1982–2003 includes issues on exchanges other than NYSE and AMEX.
[b]The SDC does not provide sufficient information to separate uninsured rights offerings from rights with standby underwriting. Thus, all rights are reported under the standby category in this table.

to standbys over the past two decades. A similar time trend is evident in the study of French SEOs by Gajewski and Ginglinger (2002).[23] Overall, as concluded by Eckbo and Masulis (1995) and Armitage (1998) as well, there appears to be an international trend away from rights. This trend coincides with substantial growth in listed firms' total equity size.

As discussed in Section 3 above, the uninsured rights method has by far the lowest direct costs. Thus, it appears that issuers in the U.S. and increasingly elsewhere are selecting the most expensive equity flotation method. Therein lies the rights offer paradox. Resolution of the paradox requires identifying indirect costs of rights that are of sufficient economic magnitude to make the total (direct and indirect) costs of firm commitment offerings the lowest for nearly all large, publicly traded industrial issuers in the U.S. We identified some of these indirect costs in Section 3. Eckbo and Masulis (1992) argue that a potentially large indirect cost emanates from adverse selection in the rights issue market. We discuss why information asymmetries may drive issuers away from the rights method next.

[23] Cronqvist and Nilsson (2005) report that uninsured rights are more frequent than uninsured rights over their sample period but do not show the time trend.

Table 11
Equity security issues by firms listed on the Tokyo Stock Exchange, 1956–2003[a]

| Year | Rights offerings | | Public offerings | | Preferred stocks | | Private placements | | Exercise of warrants | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) |
| 1956 | 294 | 157 | 36 | 4 | – | – | 11 | 2 | – | – | 341 | 164 |
| 1957 | 292 | 199 | 40 | 5 | – | – | 10 | 1 | – | – | 342 | 205 |
| 1958 | 147 | 160 | 30 | 5 | – | – | 3 | 0 | – | – | 180 | 165 |
| 1959 | 158 | 153 | 60 | 10 | – | – | 3 | 0 | – | – | 211 | 183 |
| 1960 | 275 | 331 | 100 | 35 | – | – | 4 | 1 | – | – | 379 | 387 |
| 1961 | 465 | 632 | 224 | 80 | 1 | 0.4 | 6 | 1 | – | – | 695 | 712 |
| 1962 | 554 | 587 | 171 | 20 | – | – | 9 | 3 | – | – | 734 | 609 |
| 1963 | 508 | 410 | 157 | 38 | – | – | 8 | 7 | – | – | 673 | 453 |
| 1964 | 434 | 623 | 85 | 4 | – | – | 14 | 3 | – | – | 533 | 631 |
| 1965 | 95 | 115 | 19 | 1 | – | – | 8 | 3 | – | – | 122 | 117 |
| 1966 | 173 | 202 | 34 | 1 | – | – | 24 | 8 | – | – | 231 | 212 |
| 1967 | 190 | 194 | 68 | 5 | – | – | 13 | 4 | – | – | 261 | 202 |
| 1968 | 201 | 303 | 80 | 10 | – | – | 12 | 2 | – | – | 293 | 315 |
| 1969 | 300 | 447 | 145 | 55 | – | – | 14 | 5 | – | – | 469 | 506 |
| 1970 | 316 | 538 | 203 | 138 | – | – | 18 | 5 | – | – | 537 | 681 |
| 1971 | 220 | 409 | 147 | 84 | – | – | 24 | 44 | – | – | 391 | 637 |
| 1972 | 180 | 284 | 275 | 665 | – | – | 43 | 92 | – | – | 498 | 1,041 |
| 1973 | 177 | 344 | 256 | 565 | – | – | 45 | 30 | – | – | 478 | 939 |
| 1974 | 214 | 244 | 193 | 277 | – | – | 31 | 23 | – | – | 438 | 544 |
| 1975 | 166 | 771 | 103 | 222 | – | – | 16 | 8 | – | – | 285 | 1,001 |

(*Continued on next page*)

Table 11
(*Continued*)

| Year | Rights offerings | | Public offerings | | Preferred stocks | | Private placements | | Exercise of warrants | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) |
| 1976 | 102 | 180 | 181 | 500 | 1 | 12 | 11 | 9 | – | – | 294 | 689 |
| 1977 | 120 | 291 | 238 | 604 | – | – | 48 | 29 | – | – | 406 | 923 |
| 1978 | 86 | 267 | 195 | 565 | – | – | 62 | 84 | – | – | 313 | 897 |
| 1979 | 64 | 262 | 229 | 629 | – | – | 42 | 63 | – | – | 325 | 953 |
| 1980 | 34 | 90 | 218 | 881 | – | – | 28 | 81 | – | – | 280 | 1,052 |
| 1981 | 67 | 494 | 249 | 1,396 | – | – | 20 | 37 | – | – | 336 | 1,928 |
| 1982 | 45 | 224 | 209 | 1,103 | 1 | 30 | 14 | 21 | 4 | 2 | 272 | 1,349 |
| 1983 | 18 | 135 | 72 | 472 | – | – | 23 | 165 | 18 | 30 | 131 | 802 |
| 1984 | 23 | 91 | 128 | 821 | 1 | 6 | 18 | 68 | 39 | 66 | 208 | 1,043 |
| 1985 | 40 | 183 | 103 | 506 | – | – | 18 | 33 | 70 | 137 | 231 | 859 |
| 1986 | 27 | 69 | 76 | 400 | – | – | 18 | 30 | 118 | 373 | 235 | 673 |
| 1987 | 26 | 436 | 99 | 1,394 | – | – | 22 | 109 | 241 | 1,074 | 388 | 3,013 |
| 1988 | 40 | 787 | 157 | 2,582 | – | – | 23 | 104 | 316 | 1,309 | 536 | 4,782 |
| 1989 | 32 | 726 | 227 | 5,830 | – | – | 22 | 102 | 436 | 2,190 | 718 | 8,849 |
| 1990 | 39 | 825 | 121 | 1,975 | – | – | 21 | 315 | 397 | 678 | 578 | 3,792 |
| 1991 | 40 | 218 | 27 | 126 | – | – | 19 | 104 | 309 | 360 | 395 | 808 |
| 1992 | 20 | 111 | 3 | 4 | – | – | 22 | 102 | 127 | 203 | 172 | 420 |
| 1993 | 9 | 48 | 4 | 7 | – | – | 14 | 150 | 184 | 617 | 211 | 823 |
| 1994 | 2 | 10 | 17 | 137 | 1 | 100 | 8 | 239 | 180 | 451 | 208 | 936 |
| 1995 | 12 | 96 | 8 | 33 | 1 | 50 | 19 | 160 | 118 | 299 | 158 | 638 |

(*Continued on next page*)

Table 11
(*Continued*)

| Year | Rights offerings | | Public offerings | | Preferred stocks | | Private placements | | Exercise of warrants | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) | No. of issues | Amount raised (¥ bils.) |
| 1996 | 9 | 337 | 36 | 305 | 5 | 539 | 20 | 219 | 187 | 674 | 257 | 2,074 |
| 1997 | 9 | 73 | 26 | 128 | 2 | 224 | 19 | 370 | 88 | 368 | 144 | 1,162 |
| 1998 | 1 | 0.3 | 12 | 284 | 5 | 471 | 35 | 696 | 35 | 88 | 88 | 1,540 |
| 1999 | – | – | 35 | 371 | 27 | 7,012 | 86 | 2,445 | 74 | 262 | 222 | 10,090 |
| 2000 | 2 | 8 | 36 | 573 | 5 | 137 | 56 | 972 | 94 | 111 | 193 | 1,798 |
| 2001 | 3 | 32 | 18 | 1,201 | 6 | 228 | 71 | 567 | 92 | 38 | 190 | 2,067 |
| 2002 | – | – | 21 | 156 | 40 | 1,029 | 79 | 502 | 82 | 276 | 222 | 1,963 |
| 2003 | 3 | 2 | 40 | 573 | 75 | 2,537 | 103 | 234 | 126 | 40 | 347 | 3,385 |
| 2004 | 4 | 4 | 80 | 754 | 55 | 1,411 | 142 | 624 | 241 | 104 | 522 | 2,900 |

Source: The Tokyo Stock Exchange Fact Book, 2005.
[a]The table includes foreign issues.

### 4.2. Adverse selection and current shareholder takeup

Myers and Majluf (1984) provide the first analytical approach to the equity issue decision under asymmetric information. It is useful to recap the setting of their primary model:

- The firm's objective is to maximize the full-information (long-run) value of current shareholders' claim on the firm.
- The firm knows the true value $a$ of its assets in place while outside investors know only the probability distribution over $a$.
- The firm needs to sell equity to raise a cash amount of $I$ dollars in order to finance a short-lived investment project with a commonly known net present value of $b$.
- The equity issue is sold using a simple flotation method: a direct offering to the public with no mechanism (such as an underwriter) for communication between the issuer and outside investors, and with no participation in the issue by current shareholders.

A key insight of Myers and Majluf (1984) is that the cost of selling undervalued stock may exceed $b$, causing the undervalued firm to forego the investment project rather than issue and invest.[24] The cost of this underinvestment drives a demand for more expensive flotation methods designed to reduce the information asymmetry between the issuer and outside investors. The cost may also induce the firm to turn to its own shareholders for additional equity capital. In sum, the Myers and Majluf (1984) setting provide a useful starting point for thinking about how undervalued firms may use alternative flotation methods to reduce costly information asymmetry.

For example, Wruck (1989) and Herzel and Smith (1993) suggest that some high-quality issuers avoid public issues in favor of private placements. In a private placement, the issuer may directly compensate the investor for costs of due diligence and quality inspection by selling the issue at a discount relative to the issue's market price. If the private placement investor holds on to the newly created block of shares, there may also be long-term benefits in terms of increased monitoring of the issuing firm's management.[25] Firms may also turn to underwriters for quality certification. Baron (1982), Booth and Smith (1986), Beatty and Ritter (1986), Titman and Trueman (1986) and Eckbo and Masulis (1992) all presume that underwriters have some ability and incentive to evaluate the extent to which the issuer's stock may be overpriced, and to avoid selling overpriced shares to the public. The incentive may emanate from an underwriter's risk of loss of reputation, or its risk of legal liability (e.g., Tinic, 1988; Blackwell, Marr, and Spivey, 1990).

[24] Dybvig and Zender (1991) argue that an appropriately structured managerial compensation contract would eliminate this underinvestment problem. Similarly, Admati and Pfleiderer (1994) point out that in a firm that has only investors who hold a fixed fraction of all its securities, management seeks to maximize shareholder wealth by always investing in positive NPV projects.

[25] It may also be the case that entrenched managers prefer a private placement. The offering price discount may be used as compensation to a friendly "white knight" investor for allowing management to maintain private benefits of control (see also Zwiebel, 1995). We return to this issue below.

Eckbo and Masulis (1992) generalize the Myers–Majluf framework by explicitly allowing current shareholder participation in the issue via a rights offer. Moreover, they introduce noisy but informative quality certification in the form of underwriting (standbys or firm commitment contracts). These refinements allow a realistic representation of the most commonly used flotation methods, and they result in a number of interesting predictions not available from Myers and Majluf (1984). In particular, as discussed in more detail in the subsequent section, the set of circumstances in which one expects a negative market reaction to equity issue announcements is considerably smaller.

To illustrate the shareholder takeup model, let $k \in [1, 0]$ denote the exogenously given and observable fraction of the issue that is taken up by current shareholders.[26] Moreover, let $C(k)$ denote total issue costs, which is the sum of direct costs $d$ and expected wealth transfer to outside investors. As in Eckbo and Norli (2004), the expected profits $\pi$ from issuing and investing can be written

$$\begin{aligned}\pi &= b - C(k)\\ &= b - d - \frac{I(1-k)\,[(a+b+I-d)-P]}{P},\end{aligned} \tag{1}$$

where $P$ is the post-issue secondary market price of the issuer. $P$ is determined by investors' equilibrium beliefs about $a$. In a separating equilibrium, $P$ equals the full-information value of the post-issue company ($P = a + b + I - d$), with issue profits of $\pi = b - d$. In a pooling equilibrium, however, undervalued firms experience a positive wealth transfer as $P < a + b + I - d$.[27]

Equation (1) shows how the magnitude of any wealth transfer cost is attenuated by shareholder takeup $k$. Essentially, shareholder takeup acts like a form of financial slack. If $k = 1$, $\pi = b - d$ and the wealth transfer cost is zero, even if the market undervalues the stock ($P < a + b + I - d$). If $k < 1$, which means that some shareholders in a rights offer will sell their rights to outside investors rather than subscribe, adverse selection costs are positive for undervalued firms *even if the rights offer is expected to be fully subscribed in the end*. If the firm uses an uninsured rights offer when $k = 0$, current shareholders sell *all* the rights, and the entire issue is sold to outside investors. This is a worst-case scenario in terms of wealth-transfer costs: since there is no quality certification, uninsured rights generate the same potential for wealth transfers associated with the direct offer mechanism in Myers and Majluf (1984).

Eckbo and Masulis (1992) argue that their shareholder takeup model resolves the rights offer paradox: high-quality issuers gravitate towards flotation methods that minimize the potential for wealth transfer costs. Their key insight is to show that the wealth

[26] Because the fraction $k$ reflects individual shareholder wealth constraints, it is in part exogenous to the firm. $k$ is observable through subscription precommitments (published in the issue prospectus), and through the rights trading activity (trades occur when current shareholder do not want to participate).

[27] As discussed in Eckbo and Norli (2004), the profit function in equation (1) presumes that the offering price $P_0$ is set consistent with market beliefs $P_0 = P$. Thus, this function ignores the possibility of using an offering price discount to convey information. We return to the offering discount below.

transfer cost associated with an uninsured rights offer increases as $k$ decreases. eventually making it optimal to add quality certification in a standby offering. As $k$ approaches zero, it is optimal to abandon rights altogether, despite the low *direct* cost of rights. In sum, the optimal flotation method choice depends on $k$. It follows that, around the world, firms gradually avoid uninsured rights in response to a gradual reduction in the willingness of wealth-constrained shareholders to keep funding corporate growth. This is consistent with the time-trend away from uninsured rights evidenced in Table 10 and Table 11, as average firm size also increases over time. It is also consistent with the fact that smaller private firms, and firms listed on smaller international stock exchanges, still use rights today.

Under the shareholder takeup model of Eckbo and Masulis (1992), the cross-sectional variation in the use of rights is driven by factors that affect individual shareholders wealth constraints and incentives. These factors include personal wealth and degree of risk aversion, the magnitude of private benefits of control, and the availability of substitute mechanisms for maintaining control benefits (e.g., restricted voting share and pyramidal ownership structures). Regulatory changes, and changes in the issue-technology also plays a role. For example, Ursel and Trepanier (2001) present some evidence that the decline in Canadian rights issues to some extent coincides with regulatory changes—such as the expanded use of short-form prospectuses and shelf registration procedures—which lead to an increase in the relative costs of rights.

Eckbo and Norli (2004) extend the analysis of Eckbo and Masulis (1992) by formalizing a sequential, multistage issue game in which issuers at each stage have access to a menu of flotation methods. At the start of the game, issuer have access to uninsured rights, rights with standby underwriting, and private placements.[28] Consistent with the evidence in Table 6, the direct issue cost $d$ is assumed to be lowest for uninsured rights. The standby underwriter and private placement investor perform noisy but informative quality certification. If, say, the private placement investor rejects purchasing the issuer based on its private evaluation, then the issuer either decides not to issue or moves on and decides between the remaining flotation methods in the next issue subgame. Thus, firms select among entire issue *strategies* and not just among individual flotation methods.

Eckbo and Norli (2004) show that there exists an equilibrium 'pecking order' of flotation methods in their issue game which depends on $k$. Figure 3 illustrates with a numerical example this pecking order.[29] The horizontal axis plots shareholder takeup $k$. The vertical axis plots total expected issue cost $C(k)$ for each of three alternative issue strategies. $C(k)$—which is linear in $k$—incorporates the issuer's participation constraint (equation (1) above), so these are equilibrium strategies. Denote a particular issue strategy as $\{x\}$. The steepest line in Figure 3 is for the "move straight to uninsured rights and

[28] One could substitute firm commitment underwriting for private placement without altering the basic model insights. Eckbo and Norli (2004) use private placements as their empirical laboratory is the Oslo Stock Exchange where uninsured rights, standby rights and private placements are the only observed flotation methods.
[29] See the Appendix of (Eckbo and Norli, 2004) for details of the parameter values.

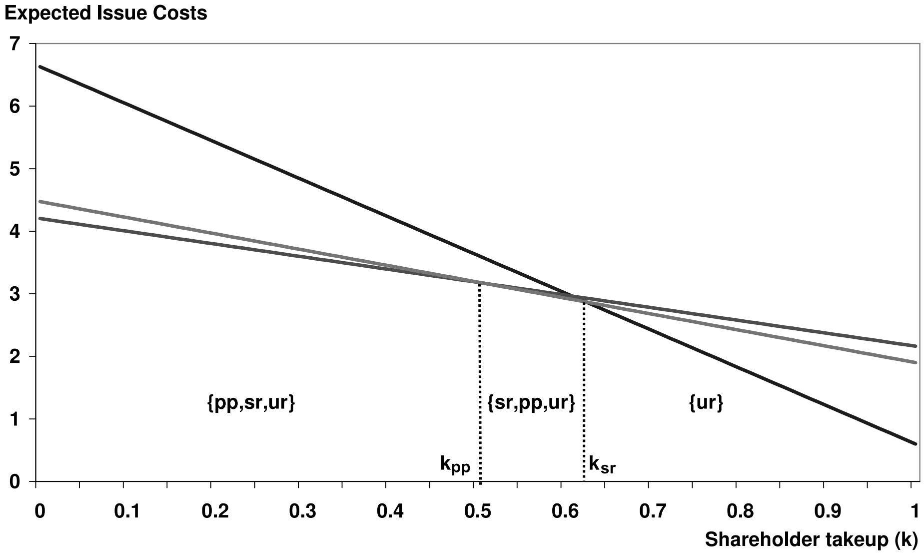



Fig. 3. Illustration of the flotation method pecking order. Source: Eckbo and Norli (2004). The horizontal axis plots shareholder takeup $k$. The vertical axis plots total expected issue cost $C(k)$ for each of three alternative issue strategies. $C(k)$ incorporates the issuer's participation constraint. The steepest line is $C(k)$ for the "move straight to uninsured rights and issue" strategy $\{ur\}$. The middle line is $C(k)$ for the strategy "start with standby rights, and if rejected try private placement, and if rejected again, sell the issue using uninsured rights" $\{sr, pp, ur\}$. The third and most horizontal line is $C(k)$ for the "start with a private placement, and if rejected try a standby rights, and if rejected again, sell the issue using uninsured rights" strategy $\{pp, sr, ur\}$. The critical values of $k$ are denoted $k_{pp}$ and $k_{sr}$. The optimal issue strategy is one that minimizes $C(k)$ conditional on $k$, i.e., the inner envelope of the three separate cost curves. Thus, it is an equilibrium for all issuers with shareholder takeup less than the critical value of $k_{pp} = 0.51$ to attempt a private placement first. When $k$ is between $k_{pp} = 0.51$ and $k_{sr} = 0.62$, the equilibrium strategy is to attempt a standby rights offering first, while all issuers with $k$ greater than $k_{sr} = 0.62$ go directly to the uninsured rights offer.

issue" strategy $\{ur\}$. The middle line is for the strategy "start with standby rights, and if rejected try private placements, and if rejected again issue using uninsured rights" $\{sr, pp, ur\}$. The third and most horizontal line is for the strategy "start with private placement, and if rejected try standby rights, and if rejected again issue using uninsured rights" $\{pp, sr, ur\}$. The critical values of $k$ that separates these strategies are denoted $k_{pp}$ and $k_{sr}$. The optimal issue strategy is one that minimizes $C(k)$, i.e., the inner envelope of the three separate cost curves. Thus, in Figure 3, it is an equilibrium for all issuers with shareholder takeup less than the critical value of $k_{pp} = 0.51$ to attempt a private placement first. When $k$ is between $k_{pp} = 0.51$ and $k_{sr} = 0.62$, the equilibrium strategy is to attempt a standby rights offering first, while all issuers with $k$ greater than $k_{sr} = 0.62$ go directly to the uninsured rights offer.

A central implication of this pecking order is that the probability of an issuer switching from rights to underwritten offer increases as $k$ decreases even if a rights offer is expected to be fully subscribed with the help of outside investors. There is growing evidence to support this prediction. Eckbo and Masulis (1992) and Singh (1997) report that the average level of shareholder takeup in U.S. rights offers is greater in uninsured rights offers than in standbys. Eckbo and Masulis (1992) also find that firms obtain substantial levels of subscription precommitments from large shareholders prior to selecting the uninsured rights method, with few such precommitments in standby rights. Information on subscription precommitments are published in the offering prospectus and are empirically useful in predicting $k$. As reviewed in Section 3 above, there is also substantial evidence more generally that flotation costs are lower for firms with greater ownership concentration, which are also the firms that tend to have greater values for $k$.

Internationally, where the rights method is much more prevalent, there is also substantial evidence consistent with a key role for shareholder takeup $k$. Bøhren, Eckbo, and Michalsen (1997) and Cronqvist and Nilsson (2005) study rights offers on the Oslo and Stockholm stock exchanges, respectively, and use the trading volume in rights to directly measure $k$.[30] They find that rights are more likely to be selected the greater the value of $k$. Moreover, Bøhren, Eckbo, and Michalsen (1997) show that the probability of switching from uninsured rights to standby rights declines with $k$, as predicted by Eckbo and Masulis (1992) and Eckbo and Norli (2004). Slovin, Sushka, and Lai (2000) find that the level of subscription levels is similar in standbys and uninsured rights in the U.K.[31] In their sample of French SEOs, Gajewski and Ginglinger (2002) report a greater ownership concentration for uninsured rights issuers than for standby rights issuers, and the lowest ownership concentration for underwritten public offerings. They also report that share allocations not taken up by the issuer's blockholders is much larger for underwritten public offerings than for uninsured rights and standbys. Using annual data on share ownership in Italy, Bigelli (1998) report that insiders' level of shareownership remains stable through the year of a rights offering, which is consistent with a high value of $k$.

### *4.3. Predicting the market reaction to issue announcements*

Table 12 summarizes the empirical predictions of the adverse selection, shareholder takeup and pecking order theories for the stock market reaction to issue announcements as a function of the flotation method. Table 12 is restricted to models in which the firm considers issuing common stock only. The choice between different types of securities—the capital structure choice—is covered in several other chapters throughout this Handbook, and has also been previously reviewed by Harris and Raviv (1991).

Let *AR* denote the announcement-induced abnormal stock return of the issuer. We first discuss predictions for *AR* of models with only a *single* flotation method, of which Myers and Majluf (1984) is the most prominent. These models provide a useful starting

[30] Rights are traded on stock exchanges. If rights trade only once (sold by a current shareholder to an outside investor), then the trading volume in rights measure $1 - k$ directly.

[31] However, it is not clear from their study whether their "takeup" variable reflects total rights-subscription levels or only subscriptions by current shareholders.

Table 12
Predicted market reaction $AR$ to SEO announcements as a function of the flotation method choice

| Study | Model specifics | Model implications for $AR$ |
|---|---|---|
| Myers and Majluf (1984) | Direct sale to public with no communication between firm and market. Current shareholders are passive bystanders to issue (they neither purchase new nor sell old shares). Managers maximize current shareholders' claim on firm, which amounts to maximizing the intrinsic (full-information) value of this claim | Separating equilibrium: $AR_{do} < 0$. Ceteris paribus, $AR_{do}$ is more negative the greater the risk that the security is overvalued by market prior to the issue announcement. A pooling equilibrium ($AR_{do} = 0$) is more likely the greater the ratio $b/E(a)$ |
| Krasker (1986) | Myers and Majluf (1984) but with varying investment size $I$ | In the separating equilibrium, $AR_{do} < 0$ and more negative the greater is $I$ |
| Heinkel and Schwartz (1986) | Issuers choose between uninsured rights, standbys and firm commitment offerings. Standbys is the most expensive flotation method and provide perfect quality *certification*. Firm commitment is simply a direct sale to market with no certification | Highest-quality issuers select standbys, intermediate-quality issuer select uninsured rights, while lowest quality issuers select firm commitments. $AR_{fc} < AR_{ur} < 0$ and $AR_{sr} > 0$ |
| Giammarino and Lewis (1988) | Myers and Majluf (1984) but with an intermediary 'financier' who may reject the issue | Semi-separating equilibrium with $AR_{fc} > 0$ |
| Eckbo and Masulis (1992) | Myers and Majluf (1984) but allowing current shareholder takeup of the (exogenous) fraction $k$ of the issue, and informative but noisy quality certification by underwriters. Single-stage flotation method game | Optimal flotation method choice depends on $k$: Separating equilibrium where no low-$k$ firms select uninsured rights. Adverse selection greatest for firm commitments, lowest for uninsured rights, with standbys in between: $AR_{fc} < AR_{sr} < AR_{ur} \leqslant 0$ |
| Cooney and Kalay (1993); Wu and Wang (2005, 2006a) | Myers and Majluf (1984) but with possible managerial overinvestment ($b < 0$) | Separating equilibrium with $AR_{do} > 0$ due to prior market uncertainty about $b < 0$ |
| Bøhren, Eckbo, and Michalsen (1997) | Eckbo and Masulis (1992) with uninsured rights and standbys only, but with varying underwriter quality certification ("effectiveness") | High-$k$ issuers select uninsured rights regardless of firm quality, so no adverse selection ($AR_{sr} = 0$). Adverse selection in standbys if underwriter "ineffective" ($AR_{sr} < 0$), but positive selection if underwriter "effective" ($AR_{sr} > 0$) |

(*Continued on next page*)

Table 12
(*Continued*)

| Study | Model specifics | Model implications for $AR$ |
|---|---|---|
| Eckbo and Norli (2004) | Eckbo and Masulis (1992) but with a multistage issue game. Private placement replaces firm commitments. If issuer is rejected by the private placement investor or the standby underwriter, it moves to the next subgame consisting of the remaining flotation method choices | Equilibrium where issuers pool over entire issue *strategies*, but where some issuers are rejected by the noisy quality inspection. High-$k$ firms prefer the issue strategy $\{ur\}$ which implies $\{AR_{ur} = 0\}$. Intermediate-$k$ firms prefer the strategy $\{sr, pp, ur\}$ which implies $\{AR_{sr} > 0, AR_{pp} = 0, AR_{ur} < 0\}$. Low-$k$ issuers prefer $\{pp, sr, ur\}$ implying $\{AR_{pp} > 0, AR_{sr} = 0, AR_{ur} < 0\}$ |

In all the models below, the firm knows the true value of its assets in place $a$ while shareholders and outside investors only know the probability distribution over $a$. The firm needs to sell equity (no debt allowed) to raise the amount $I$ required to invest in a short-lived project with net present value $b$. The models differ in their assumptions about managerial objectives and availability of flotation methods. $AR_{do}$, $AR_{ur}$, $AR_{sr}$, $AR_{fc}$, $AR_{pp}$ denote the market reactions to "direct offering", "uninsured rights", "standby rights", "firm commitment", and "private placement", respectively. In Eckbo and Norli (2004), an issue strategy such as the one denoted $\{pp, sr, ur\}$ means "try private placement first, if rejected, try standby rights, if rejected again, do uninsured rights".

point for understanding the effects of adverse selection per se. We then turn to models where the firm is allowed to select from a menu of commonly used flotation methods, either in single-stage or in multi-stage (sequential) games.

### 4.3.1. Models with a single flotation method

Recall that the setting of Myers and Majluf (1984) is a direct equity sale to the public. Current shareholders are assumed to be passive and there are no mechanism for quality certification. In their separating equilibrium, some undervalued firms prefer not to sell shares, which implies that the pool of issuing firms is overpriced ex ante. The market therefore discounts issuers' stock price in response to news of the offer ($AR < 0$). Alternatively, in their pooling equilibrium, the value of $b$ is sufficiently large for all firms to issue, which implies that the issue announcement conveys no new information to the market ($AR = 0$). Ceteris paribus, in their separating equilibrium, $AR$ is more negative the greater the ex ante risk that the security is overvalued by the market. The latter implication helps distinguish the Myers and Majluf (1984) adverse selection model from a signaling model such as Miller and Rock (1985), in which external financing conveys negative information per se, regardless of the potential for security mispricing.

Strictly speaking, tests of the Myers and Majluf (1984) prediction $AR \leqslant 0$ requires a sample of direct equity sales to the public. As direct sales are rare events, no such experiment has been reported to date. Existing studies draw from the set of available flotation methods, which in U.S. studies is predominantly firm commitment offerings, while rights offerings dominate throughout the rest of the world. The subsequent theoretical work represents attempts to refine the single-flotation-method environment of Myers and Majluf (1984) in various ways, adding predictive power in samples dominated by more complex flotation methods.

Krasker (1986) allows the size of the investment project—and therefore the required financing amount $I$—to vary across firms. He derives a separating equilibrium in which greater amounts $I$ implies greater adverse selection, so $AR$ is more negative the greater the amount raised in the offering.

Giammarino and Lewis (1988) introduces a simple bargaining game between the issuer and an uninformed financial intermediary. The purpose is to examine the implications of allowing the purchaser of the issue to reject the offering (which never happens in Myers and Majluf (1984)). The issuer suggests an offer price that is either "high" or "low", and the financier accepts or rejects the offer. In their semi-pooling equilibrium, the high-value type always suggests a high offer price, while the financier randomizes between accepting and rejecting the high offer price, but always accepts a low offer price. The information content of the issue announcement depends on which issuer type is most eager to finance the project, measured by the ratio of assets in place to post-issue value. If the low-value type is more eager, it will find a way to avoid being rejected too often by the financier. This is accomplished by randomizing between the low price (which is always accepted by the financier) and the high price. In this equilibrium, the low-value type ends up being revealed in the separating part of the equilibrium, so $AR < 0$. Conversely, when the high-value type is relatively more eager to obtain financing, the equilibrium implies $AR > 0$. This latter equilibrium does not exist in a setting such as Myers and Majluf (1984) where $b$ is constant across issue types, since then the low-value type will always be the most eager to obtain financing.

Cooney and Kalay (1993) and Wu and Wang (2005, 2006a) allow managers to overinvest ($b < 0$). In Cooney and Kalay (1993), it is possible for a firm with overvalued stock to issue stock to invest in negative NPV projects, while a firm with undervalued stock may still issue stock to avoid loosing very profitable NPV investment opportunities. Thus, in their model equity issuance has two effects, a negative signal about current assets in place and a positive signal about new investment opportunities, where either effect can dominate. In Wu and Wang (2005, 2006a) the overinvestment is introduced by explicitly assuming that managers enjoy a certain level of private benefits of control. In both papers, there is ex ante uncertainty about whether or not an issuer will try to fund a negative NPV project. They show that this type of uncertainty may produce a positive equilibrium market reaction to some equity issues. The positive reaction reflects the surprise when firms issue to fund projects with a greater value of $b$ than expected.

The following numerical example illustrates a positive issue surprise effect by simply adding the shareholder takeup parameter $k$ to the original Myers and Majluf (1984) model. Suppose the market does not know $k$ ex ante, but believes that $k = 0$. Moreover, it is common knowledge that the firm's assets in place $a$ may be in one of two equally likely states: "high" with $a = \$150$ or "low" with $a = 50$. In both states, the project NPV is $b = 20$. With $k = 0$ (which means we are back in the Myers–Majluf model), it follows that the firm in this example will only issue if it is in the low state.[32] This implies a pre-issue stock price $p^-$ which reflects an underinvestment discount (capitalizing the value of the project only in the low state): $p^- = (150 \times 0.5 + (50 + 20) \times 0.5) = \$110$. If the firm announces a stock issue and reveals $k = 0$, the post-issue price will be $p^+ = (50 + 20) = \$70$. In this case, the firm sells the fraction $(100/210) \times 100 = 48\%$ of the firm in order to raise \$100, generating a market reaction of $AR = 100 \times (70 - 110)/110 = -36\%$.

However, suppose the issuer surprises the market by revealing $k = 1$ through the offering process. Since $k = 1$ implies that the firm prefers to issue in both states (there is no wealth transfer to outside investors), there is pooling and the issue announcement carries no information about the true state. Still, the announcement causes the market to eliminate the underinvestment discount, now capitalizing the value of the project in *both* states: $p^+ = \$120$ and $AR = 9\%$. In this example, new information revealing a high value of $k$ reverses market expectations from a separating equilibrium to a pooling equilibrium, resulting in $AR > 0$.

It is clear from the above that the implied market reaction to issue announcements may be negative, zero, or positive in information settings that represent simple refinements of the original Myers and Majluf (1984) setup—even when preserving their single flotation method environment. We next describe predictions emanating from models allowing for a choice between several flotation methods.

### *4.3.2. Modelling the flotation method choice*

In the first model of the flotation method choice, Heinkel and Schwartz (1986) allow issuers to choose between uninsured rights, standby rights and 'firm commitment' offerings. In their model, uninsured rights carry a risk of offering failure, while standby rights and firm commitment offers fully guarantee the offering proceeds. The standby underwriter fully reveals the issuer type while the firm commitment underwriter is uninformed. In equilibrium, the highest-valued issuers select standbys, intermediate-value issuer select uninsured rights, while the lowest-valued issuers select firm commitment offers. Thus, this model predicts $AR_{fc} < AR_{ur} < 0$ and $AR_{sr} > 0$.

In the Heinkel and Schwartz (1986) model, the quality certification in a standby rights offer makes this a more expensive flotation method than firm commitment offerings,

[32] Note that $k = 0$ still means that the firm could put on a fully subscribed rights offer. However, in such a rights offer, every subscriber would be a new shareholder.

which is counterfactual (Table 6). Moreover, there is no explicit role for current shareholder takeup. Eckbo and Masulis (1992) offers a menu of flotation methods which allows shareholder takeup ($k$) and informative but noisy quality certification by underwriters in both standbys and firm commitment offerings. As discussed above (equation (1)), shareholder takeup reduces the size of the offering to outside investors, acting like financial slack in Myers and Majluf (1984). In equilibrium, high-$k$ firms select uninsured rights with little or no adverse selection, intermediate-$k$ firms select standby rights, while low-$k$ firms select firm commitments. They predict that $AR_{fc} < AR_{sr} < 0$ and that $AR_{ur} \approx 0$.

Building on Eckbo and Masulis (1992), Bøhren, Eckbo, and Michalsen (1997) model two flotation methods: uninsured rights and standbys. They refine the empirical prediction on announcement returns by varying the effectiveness of the underwriter in detecting overpriced issues. As in Eckbo and Masulis (1992) all high-$k$ issuers select uninsured rights which results in $AR \approx 0$. Moreover, in an equilibrium with "ineffective" underwriter certification, some overvalued issuers decide to risk the certification process, leading to adverse selection in the pool of low-$k$ firms selecting standby rights offerings, so $AR_{sr} < 0$. However, in an equilibrium with "effective" underwriters some low-$k$ firms prefer not to issue rather than risk being detected by the quality certification process, so the standby pool exhibits positive selection and $AR > 0$.

Eckbo and Norli (2004) is the first model to allow a sequential flotation method choice. As discussed above, they prove the existence of a sequential pooling equilibrium in which issuers pool over entire issue strategies. Pooling results when the issue profits $\pi$ in equation (1) is non-negative for both high-value and low-value firms. The issue methods are private placement, standby rights and uninsured rights. Both the private placement investor and the standby underwriter perform an informative but noisy quality inspection and may reject the issue. Recall the definition of an issue strategy, e.g., $\{pp, sr, ur\}$ which means "try a private placement first, if rejected try standby rights, and if rejected again do an uninsured rights offer". Although issuers pool over issue *strategies*, they may eventually end up using different flotation methods due to randomness in the quality inspection process. The predictions for the market reaction are as follows:

**Eckbo and Norli (2004)—Pecking order.** *Suppose $k$ is known ex ante and that issuers follow the pecking order illustrated in Figure 3. Let "high $k$" mean $k \in [k_{sr}, 1]$, "medium $k$" mean $k \in [k_{pp}, k_{sr}]$ and "low $k$" mean $k \in [0, k_{pp}]$. It is part of a sequential pooling equilibrium for high-$k$ issuers to select the strategy $\{ur\}$, for medium-$k$ issuers to select the strategy $\{sr, pp, ur\}$ and for low-$k$ issuers to choose $\{pp, sr, ur\}$. The associated market reaction AR to the issue announcement is as follows*:

| | $k$ high | $k$ medium | $k$ low |
|---|---|---|---|
| Uninsured rights: | $AR_{ur} = 0$ | $AR_{ur} < 0$ | $AR_{ur} < 0$ |
| Standby rights: | off-equilibrium | $AR_{sr} > 0$ | $AR_{sr} = 0$ |
| Private placement: | off-equilibrium | $AR_{pp} = 0$ | $AR_{pp} > 0$ |

The intuition behind these predictions for $AR$ is as follows. Starting with the first line (uninsured rights), firms with high $k$ prefer to issue using the relatively low-cost uninsured rights method. Since there is no inspection, there is also no information conveyed by the issue decision, thus $AR_{ur} = 0$. Firms with medium and low $k$ values prefer quality inspection (Figure 3). Thus, issuers of uninsured rights with medium or low $k$ have necessarily been rejected twice by the inspection, so $AR_{ur} < 0$.

Second, in the line for the standby rights method, medium-$k$ issuers prefer standbys, creating a positive market reaction ($AR_{sr} > 0$) due to the positive inspection result. Low-$k$ issuers prefer private placement (Figure 3). Thus, low-$k$ issuers that issue using standbys have been rejected by the private placement inspection before accepted by the standby underwriter inspection. From the market's point of view, these two inspection results cancel out, so there is no new information and $AR_{sr} = 0$. Similarly, in the line for the private placement method, medium-$k$ issuers that use private placement have first been rejected by the standby underwriter, thus $AR_{pp} = 0$. Low-$k$ issuers prefer private placement (Figure 3), so the successful inspection result implies $AR_{pp} > 0$.

We now turn to a summary of the international evidence on SEO announcement returns, and then draw inferences about the theoretical predictions above.

### 4.4. *Evidence on issue announcement returns*

Abnormal returns are typically measured over the two-day window $[-1, 0]$ ending with the public announcement date (day 0), or over the three-day window $[-1, +1]$. Abnormal return to issuer $i$ on day $t$ is typically defined using a simple market model:

$$\gamma_{it} \equiv r_{it} - E(r_{it}) = r_{it} - (\alpha_i + \beta_i r_{mt}), \tag{2}$$

where $r_{it}$ is the daily stock return in excess of the risk-free rate, $r_{mt}$ is the daily excess return on the value-weighted CRSP market return, and $\alpha$ and $\beta$ are estimated during some pre-event period. For event windows containing multiple periods, the cumulative abnormal return is found by adding daily abnormal returns. With the market model estimation, it is important not to "contaminate" the estimate of $\alpha$ with the well-known average stock price runup over the year prior to the typical U.S. stock issue. If this runup is treated as "normal" then the estimate of $\alpha$ will be overstated, resulting in a downward bias in the estimated abnormal return $\gamma$. One solution to this problem is to estimate the market model parameters using post-issue stock returns.

Some studies estimate $\gamma$ directly by means of a conditional market model,

$$r_{it} = \alpha_i + \beta_i r_{mt} + \gamma_i d_t + \epsilon_{it}, \tag{3}$$

where $d_{it}$ is a dummy variable that takes on a value of 1 during the event window and zero otherwise, and $\epsilon_{it}$ is the regression error term. If the event dummy $d_t$ takes on a value of one over $\omega$ days in the event window, then the cumulative abnormal return over the event window is $\omega\gamma_i$.[33]

[33] See Thompson (1985, 1995) for details of this event-study approach.

The studies form average abnormal returns across a sample of $N$ issues as $AR_t \equiv (1/N)\sum_i^N \gamma_{it}$ and report tests of the hypothesis that $AR_t = 0$. Statistical significance is inferred using either a $t$-statistic for the average, or a $z$-statistic

$$z_t = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \frac{\gamma_{it}}{\sigma_i}, \tag{4}$$

where $\sigma_i$ is the time series estimate of the standard error of $\gamma_{it}$.[34] For large sample size $N$, this $z$-statistic has a standard normal distribution under the null hypothesis of a zero average abnormal return.

We have organized the evidence on average announcement effects to security offerings in three tables. Table 13 covers studies of SEOs by U.S. firms, classified by the flotation method. Table 14 show international evidence on SEOs, again by flotation methods. We separate U.S. from international studies as the international evidence show very different results than that of U.S. studies. Third, Table 15 show the announcement effect of straight and convertible debt offerings by U.S. firms.

### 4.4.1. Market reaction to SEOs in the U.S.

In this section, we highlight four main conclusions from the U.S. evidence. As surveyed by Eckbo and Masulis (1995), the perhaps most striking finding of papers published in the 1980s is the significantly negative market reaction to firm commitment offerings by U.S. firms. These papers are shown in Panel (a) of Table 13. For brevity, the table pools results for industrial and utility issuers—although it is well known that the market reaction to industrial issuers is more negative than for utility offerings. For example, while the two-day average abnormal return averages about −2% across the two issuer types (using sample-size weights), it averages about −3% for industrials and −1% for utilities (Asquith and Mullins, 1986; Masulis and Korwar, 1986; Mikkelson and Partch, 1986; Korajczyk, Lucas, and McDonald, 1990; Hansen and Crutchley, 1990; Eckbo and Masulis, 1992). The lower market reaction to utilities is consistent with adverse selection arguments as utilities generally have less discretion than industrial companies in timing the issue to short-term overvaluation. The regulatory process reduces discretion to time the market, either by slowing the issue approval process or by forcing the firm to issue at times determined in part by the incentives of the regulator.

In 1985, the Wall Street Journal changed its reporting system for SEO announcements with the effect of making it more costly to collect accurate issue announcement dates for broad, representative samples.[35] This, combined with the very strong inferences made from the earlier studies, probably explains why there is a drop in the

[34] Some studies report $t$-statistics using a cross-sectional estimate of the standard error. See Kothari and Warner (2007), Chapter 1 of this volume, for a discussion of various event-study procedures.

[35] Jung, Kim, and Stulz (1996): "Before 1985, the WSJ reports on equity issues as a regular news item. From 1985, most of the information on new issues is reported in the 'new securities issues column' which

Table 13
Average market reaction ($AR$, %) to announcements of seasoned equity offerings (SEOs) by U.S. firms, classified by flotation method

| Study | Sample size | Sample period | $AR$ (%) |
|---|---|---|---|
| (a) Firm commitments: $N = 15{,}017$; $AR_{fc} = -2.22$* | | | |
| Asquith and Mullins (1986) | 392 | 1963–1981 | −1.6* |
| Masulis and Korwar (1986) | 972 | 1963–1980 | −1.85* |
| Mikkelson and Partch (1986) | 80 | 1972–1982 | −3.56* |
| Kalay and Shimrat (1987) | 455 | 1970–1982 | −3.36* |
| Korajczyk, Lucas, and McDonald (1990) | 1,285 | 1974–1983 | −2.94* |
| Hansen and Crutchley (1990) | 109 | 1975–1982 | −3.65* |
| Eckbo and Masulis (1992) | 1,057 | 1963–1981 | −2.0* |
| Jegadeesh, Weinstein, and Welch (1993)[a] | 411 | 1980–1989 | −1.16* |
| Slovin, Sushka, and Bendeck (1994)[a] | 175 | 1973–1988 | −2.87* |
| Denis (1994) | 435 | 1977–1990 | −2.49* |
| Bayless and Chaplinsky (1996) | 1,884 | 1968–1990 | −2.3* |
| Altinkilic and Hansen (2003) | 1,703 | 1990–1997 | −2.23* |
| Bethel and Krigman (2004) | 2,592 | 1992–2001 | −2.01* |
| Heron and Lie (2004) | 3,658 | 1980–1998 | −2.5* |
| D'Mello, Schlingemann, and Subramaniam (2005) | 1,621 | 1982–1995 | −1.87* |
| (b) Private placements: $N = 2{,}830$; $AR_{pp} = 2.45$* | | | |
| Wruck (1989) | 99 | 1979–1985 | 1.89* |
| Herzel and Smith (1993) | 106 | 1980–1987 | 1.72* |
| Hertzel et al. (2002) | 619 | 1980–1996 | 2.4* |
| Chaplinsky and Haushalter (2003) | 1,050 | 1995–2000 | 3.49* |
| Krishnamurthy et al. (2005) | 397 | 1983–1992 | 1.43* |
| Barclay, Holderness, and Sheehan (2005) | 559 | 1979–1997 | 1.7* |
| (c) Uninsured rights: | | | |
| Eckbo and Masulis (1992) | 53 | 1963–1981 | −0.59 |
| (d) Standby rights: $N = 349$; $AR_{sr} = -1.33$* | | | |
| Hansen (1988) | 102 | 1964–1986 | −2.4* |
| Eckbo and Masulis (1992) | 128 | 1963–1981 | −0.70* |
| Singh (1997) | 63 | 1963–1985 | −1.07* |
| Heron and Lie (2004) | 56 | 1980–1998 | −1.10 |
| (e) Shelf offerings: $N = 1{,}851$; $AR_{sh} = -0.66$* | | | |
| Bhagat, Marr, and Thompson (1985) | 93 | 1982–1983 | −0.81* |
| Moore, Peterson, and Peterson (1986)[b] | 84 | 1982–1983 | −1.10* |
| Denis (1991)[c] | 40 | 1982–1986 | −1.00* |
| Heron and Lie (2004) | 256 | 1980–1998 | −1.30* |

(*Continued on next page*)

Table 13
(*Continued*)

| Study | Sample size | Sample period | *AR* (%) |
|---|---|---|---|
| Bethel and Krigman (2004)[b] | 747 | 1992–2001 | −0.24 |
| Bethel and Krigman (2004) | 391 | 1992–2001 | −1.27* |
| Autore, Kumar, and Shome (2004)[c,d] | 156 | 1990–2003 | −1.16 |

The table focuses on studies that use daily stock return to measure the SEO announcement effect *AR*, and where the flotation method may be reasonably deduced from the sample selection criteria. The sample must include primary offerings, possibly in combination with secondary equity offerings. Some studies measure *AR* over the two-day window $[-1, 0]$ while others use a three-day window $[-1, +1]$, and the table does not make a distinction between these. Some studies also separate out industrials from utilities, and when they do, we report results averaged across both issuer types. The *AR* in the panel heading is the average across the studies in the panel, weighted by the respective sample sizes. The superscript * indicates that the *AR* is significantly different from zero at the 1% level.

[a]Sample is restricted to the first SEO following the IPO.

[b]The event day is the shelf registration day (not the offering announcement).

[c]Sample is restricted to firms that issue *both* shelf and nonshelf registered shares.

[d]This abnormal return is the sum of the abnormal returns around the registration and offering dates.

number of studies of SEO announcement effects after 1986. However, more recently, Jegadeesh, Weinstein, and Welch (1993), Slovin, Sushka, and Bendeck (1994), Denis (1994), Bayless and Chaplinsky (1996), Bethel and Krigman (2004), Heron and Lie (2004), and D'Mello, Schlingemann, and Subramaniam (2005) all confirm that the market reaction to firm commitment offerings in the U.S. is on average negative and about −2%. Overall, over the period 1963–1995 and using a sample-weighted average, the market reacted to a firm commitment equity offering announcement by discounting the second-hand market price of the issuer's shares, resulting in a statistically significant $AR_{fc} = -2.22\%$.

A second striking result from the 1980s is the finding of Wruck (1989) of a significantly positive two-day market reaction of 1.9% to 128 announcements of equity private placements. The type of security sold in her private placements includes primarily common stock (101 cases) but also preferred stock, convertible preferred stock, and warrants. Thus, Wruck's sample has a different equity security composition than the studies of firm commitment SEOs. As shown in Panel (b) of Table 13, several recent studies using substantially expanded samples confirm her finding of a significantly positive announcement effect. These include Herzel and Smith (1993), Hertzel et al. (2002),

contains mostly offering information. Hence, the event dates since 1985 reflect issues that are more likely to be anticipated because the announcement of an equity issue is typically made earlier (by days or weeks) via news-wire services than the WSJ listing. This biases the abnormal return estimate".

Chaplinsky and Haushalter (2003), Krishnamurthy et al. (2005), and Barclay, Holderness, and Sheehan (2005).[36] Over the period 1979–2000, the sample-weighted average market reaction to private placements is a significantly positive $AR_{pp} = 2.45\%$.

A third important finding is that selling SEOs via the rights method appears to affect the market reaction to the issue announcement, relative to that of both firm commitments and private placements. This impact was first demonstrated by Eckbo and Masulis (1992) who examine both uninsured rights and standbys (in addition to firm commitments), and is evident also in studies examining standbys only, such as Hansen (1988), Singh (1997) and Heron and Lie (2004). As shown in panels (c) and (d) of Table 13, uninsured rights are met with a neutral market reaction—$AR_{ur} = 0.59\%$—whereas standbys elicit a significantly negative market reaction on average. The market reaction to standbys is smaller than the size of the negative market reaction to firm commitment SEOs. Over the period 1963–1998, the sample-weighted average abnormal return to standby announcements is a statistically significant $AR_{sr} \approx -1.33\%$.

Fourth, Bhagat, Marr, and Thompson (1985), Moore, Peterson, and Peterson (1986), and Denis (1991) report that early users of the shelf-registration method for offering shares (under SEC Rule 415) experienced a significantly negative market reaction of about $-1\%$. As discussed above (Table 3), the number of shelf-registered SEOs peaked in 1982 and 1983, almost disappeared in the period 1984–1991, and then picked up again, with a relatively large number occurring over the period 1997–2003. The announcement effect of this later period is reflected in the results reported by Heron and Lie (2004), Bethel and Krigman (2004), and Autore, Kumar, and Shome (2004). These more recent studies confirm the basic conclusion from the early sample period: despite the greater timing discretion afforded shelf-registered issuers, the average market reaction is no more negative for shelf issues than for non-shelf firm commitment offerings. Over the period 1982–2003, the sample-weighted average market reaction to the announcements of shelf-registered SEOs is small, but statistically significant: $AR_{sh} = -0.66\%$.

### *4.4.2. Market reaction to SEOs internationally*

At the time of the survey of Eckbo and Masulis (1995) there were relatively few studies reporting the market reaction to security offerings internationally. With the exception of Japan (Table 11), rights issues (uninsured or standbys) are still the norm in smaller equity markets. Table 14 summarize the findings of international studies of SEOs where the flotation methods is reported to be either uninsured rights, standbys, private placements, firm commitments or a foreign offering using either American (ADR) or global (GDR) drawing rights. Note that Table 14 is restricted to studies that show results for each flotation method separately, eliminating, e.g., studies that pool uninsured and

[36] Chaplinsky and Haushalter (2003) report that on average announcement returns are positive for traditional PIPE issuers and negative for structured PIPE issuers. Brophy, Sialm, and Ouimet (2005) also report positive announcement effect for PIPE issuers.

Table 14
Average market reaction ($AR$, %) to announcements of seasoned equity offerings (SEOs) internationally, classified by flotation method

| Country | Study | Sample size | Sample period | $AR$ (%) |
|---|---|---|---|---|
| (a) Uninsured rights: $N = 484$; $AR_{ur} = 0.70$ | | | | |
| Korea | Kang (1990) | 89 | 1984–1988 | 0.95% |
| Greece | Tsangarakis (1996) | 55 | 1981–1990 | 3.97* |
| Norway | Bøhren, Eckbo, and Michalsen (1997) | 74 | 1980–1993 | 1.55* |
| Italy | Bigelli (1998) | 82 | 1980–1994 | 0.79 |
| U.K. | Slovin, Sushka, and Lai (2000) | 20 | 1986–1994 | −4.96* |
| France | Gajewski and Ginglinger (2002) | 57 | 1986–1996 | −1.11* |
| Sweden | Cronqvist and Nilsson (2005) | 107 | 1986–1999 | 0.19 |
| (b) Standby rights: $N = 1{,}201$; $AR_{sr} = -1.32^*$ | | | | |
| Japan | Kang and Stulz (1996) | 28 | 1985–1991 | 2.21* |
| Norway | Bøhren, Eckbo, and Michalsen (1997) | 114 | 1980–1993 | −0.23 |
| France | Gajewski and Ginglinger (2002) | 140 | 1986–1996 | −0.74* |
| Norway | Eckbo and Norli (2004) | 143 | 1980–1996 | −0.58 |
| Sweden | Cronqvist and Nilsson (2005) | 53 | 1986–1999 | 0.72 |
| U.K. | Burton, Lonie, and Power (1999)[a] | 37 | 1989–1991 | −7.76* |
| U.K. | Slovin, Sushka, and Lai (2000)[b] | 200 | 1986–1994 | −2.90* |
| Hong Kong | Wu and Wang (2006b) | 180 | 1989–1997 | −3.37* |
| (c) Private placement: $N = 691$; $AR_{pp} = 3.12^*$ | | | | |
| U.K. | Slovin, Sushka, and Lai (2000)[c] | 76 | 1986–1994 | 3.31* |
| Norway | Eckbo and Norli (2004) | 136 | 1980–1996 | 1.39 |
| Sweden | Cronqvist and Nilsson (2005) | 136 | 1986–1999 | 7.27* |
| Japan | Kato and Schallheim (1993) | 76 | 1974–1988 | 4.28* |
| Japan | Kang and Stulz (1996) | 69 | 1985–1991 | 3.88* |
| Hong Kong | Wu, Wang, and Yao (2005) | 99 | 1989–1997 | 1.97* |

standby rights in a single sample. Eckbo and Masulis (1995) survey several of these pooled rights and standby samples, including Marsh (1979) (U.K.), Loderer and Zimmermann (1987) (Switzerland), Hietala and Loyttyniemi (1991) (Finland), and Dehnert (1991) (Australia). The main conclusion in the earlier survey was that "the average market reaction is typically positive for uninsured rights and small, but negative for standbys" (p. 1046). They do not report studies on firm commitment offerings internationally.

The evidence summarized in Table 14 goes further. Starting with uninsured rights offerings in Panel (a), uninsured rights offers are associated with a neutral or positive market reaction in smaller markets such as Greece, Norway and Sweden, but a negative market reaction in larger markets such as France and the U.K. Tsangarakis (1996) and Bøhren, Eckbo, and Michalsen (1997) report a significantly positive market reaction

Table 14
(*Continued*)

| Country | Study | Sample size | Sample period | *AR* (%) |
|---|---|---|---|---|
| (d) Firm commitments: $N = 1{,}064$; $AR_{fc} = 1.10^{*}$ | | | | |
| Japan | Kang and Stulz (1996) | 185 | 1985–1991 | 0.51* |
| Japan | Cooney, Kato, and Schallheim (2003) | 555 | 1974–1991 | 0.72*[d] |
| France | Gajewski and Ginglinger (2002) | 18 | 1986–1996 | −0.33 |
| Hong Kong | Wu, Wang, and Yao (2005) | 306 | 1989–1997 | 1.93* |

In the panel headings, *N* is the aggregate sample size across all studies in the panel, and *AR* is sample-weighted average market reaction. The superscript * indicates that the *AR* is significantly different from zero at the 1% level. The table is restricted to studies that (1) use daily stock return to measure the SEO announcement effect *AR* and (2) report the announcement effect by individual f lotation method. For example, studies that pool uninsured and standby rights in one sample are excluded. Some studies measure *AR* over the two-day window $[-1, 0]$ while others use a three-day window $[-1, +1]$, and the table does not make a distinction between these. Some studies also separate out industrials from utilities, and when they do, we report result averaged across both issuer types.

[a]The authors do not indicate whether their rights sample is standbys or uninsured rights. However, judging from the sample frequency, and the information in Slovin, Sushka, and Lai (2000), we placed this study in the standby category.

[b]In 111 of the 200 cases, shareholder takeup is greater than 90%. For these cases, the announcement period return is reported to be −0.33% and statistically insignificant.

[c]The authors refer to these as "placings" or "bought deals" that increases shareholder dispersion.

[d]The event day is the board meeting date.

for Greece and Norway, respectively, while Cronqvist and Nilsson (2005) report an insignificant market reaction to uninsured rights offers in Sweden. Slovin, Sushka, and Lai (2000) reports a significantly negative market reaction to U.K. uninsured rights offers, while Gajewski and Ginglinger (2002) report a significantly negative market reaction in France also. The sample-weighted cross-country average is however a non-negative and statistically insignificant $AR_{ur} = 0.70\%$.

It should be noted that the cross-country average may hide important country-specific institutional effects, which often motivates a study of foreign issues. Thus, although Slovin, Sushka, and Lai (2000) report results for a relatively small sample (20) of uninsured rights, the significantly negative market reaction may emanate from economically important unique institutional characteristics of the London Stock Exchange. A similar argument goes for the negative effect for the 57 uninsured rights offers in France studied by Gajewski and Ginglinger (2002). We return to this issue below.

Second, Panel (b) of Table 14 shows that standby offering are met with a positive market reaction in Japan (Kang and Stulz, 1996), a neutral market reaction in Norway and Sweden (Bøhren, Eckbo, and Michalsen, 1997; Cronqvist and Nilsson, 2005), and a negative market reaction in the U.K. (Burton, Lonie, and Power, 1999; Slovin, Sushka, and Lai, 2000) and a negative announcement effect in Hong Kong (Wu and Wang, 2006b). The sample-weighted average market reaction to standbys is significantly negative: $AR_{sr} = -1.32\%$.

Third, the market reaction to private placements is consistently positive and large across countries. The largest reported impact is in Sweden, where Cronqvist and Nilsson (2005) report a market reaction of 7.2% across 136 placements, followed by Japan with approximately 4% (Kato and Schallheim, 1993; Kang and Stulz, 1996), and the U.K. with 3.3% (Slovin, Sushka, and Lai, 2000). Significantly positive effects are also reported for private placements in Hong Kong (Wu, Wang, and Yao, 2005) and Norway (Eckbo and Norli, 2004). The sample-weighted average market reaction across these private placement studies is a significant $AR_{pp} = 3.12\%$, which is close in magnitude to the average market reaction to private placements in the U.S.

Fourth, with the exception of Japan and France, the relatively expensive firm commitment underwriting method has not yet spread internationally. Both Kang and Stulz (1996) and Cooney, Kato, and Schallheim (2003) report a small but statistically significant, positive average market reaction for Japan, while Gajewski and Ginglinger (2002) reports a statistically insignificant market reaction to firm commitment offerings in France. The sample-weighted average is an insignificant $AR_{fc} = 1.10\%$. Whether this surprising result holds up in samples of Japanese SEOs after 1992, as well as internationally as other countries start to adopt the firm commitment method, remains an interesting issue for future research.

Finally, while not shown in Table 14, recent papers have studied the average market reaction when firms announce foreign exchange listings—either foreign firms in the U.S. via American Depository Rights (ADRs) and U.S. firm globally via Global Depository rights (GDRs). Chaplinsky and Ramchand (2000) compare the stock price reactions of 349 global equity issues (involving a simultaneous sale of common equity at the same offer price in the U.S. market and one or more international markets) with 459 domestic equity issues that are sold exclusively in the U.S. market during 1986–1995. They find that all else equal, the negative stock price reaction that accompanies equity issues is reduced by 0.8 percent on average for global offers compared to domestic offers of similar size, issued during the same time period.[37]

Subsequent papers have confirmed the finding of Chaplinsky and Ramchand (2000) that firms announcing global issues have a lower stock price reaction as compared to announcements of domestic (U.S.) equity issues. For example, Wu and Kwok (2002) find that announcements of global equity issues result in a percentage point lower stock price reaction relative to comparable domestic issues. Errunza and Miller (2003) document that global equity offerings of foreign firms after their initial cross listing in the United States have a reduced stock price reaction (less by 1.5 percent) as compared to stock price reaction to SEOs of similar firms on the local exchanges.

[37] This result is based on a Heckman two-step procedure to adjust for selection bias. They also find that the announcement effect is more favorable as the number of new foreign investors rises. Also see Foerster and Karolyi (2000) for information on ADR SEOs.

Table 15
Average market reaction ($AR$, %) to announcements of debt offerings by U.S. firms

| Study | Sample size | Sample period | $AR$ (%) |
|---|---|---|---|
| (a) Stock price reaction to straight debt offerings: $N = 3{,}041$; $AR_{sd} = -0.22$ | | | |
| Dann and Mikkelson (1984) | 150 | 1969–1979 | −0.37* |
| Mikkelson and Partch (1986) | 171 | 1972–1982 | −0.23 |
| Eckbo (1986) | 648 | 1964–1981 | −0.10 |
| Hansen and Crutchley (1990) | 188 | 1975–1982 | 0.11 |
| Shyam-Sunder (1991) | 297 | 1980–1984 | −0.11 |
| Chaplinsky and Hansen (1993) | 245 | 1974–1984 | 0.05 |
| Johnson (1995) | 129 | 1977–1983 | 0.32 |
| Jung, Kim, and Stulz (1996) | 276 | 1977–1984 | −0.09 |
| Howton, Howton, and Perfect (1998) | 937 | 1983–1993 | −0.50* |
| (b) Stock price reaction to convertible debt offerings: $N = 307$; $AR_{cd} = -1.8^*$ | | | |
| Dann and Mikkelson (1984) | 132 | 1969–1979 | −2.30* |
| Mikkelson and Partch (1986) | 33 | 1972–1982 | −1.97* |
| Eckbo (1986) | 75 | 1964–1981 | −1.25* |
| Hansen and Crutchley (1990) | 67 | 1975–1982 | −1.45* |

In the panel headings, $N$ is the aggregate sample size across all studies in the panel, and $AR$ is sample-weighted average market reaction. The superscript * indicates that the $AR$ is significantly different from zero at the 1% level. The table focuses on studies that use daily stock return to measure the SEO announcement effect $AR$, and where the flotation method may be reasonably deduced from the sample selection criteria. Some studies measure $AR$ over the two-day window $[-1, 0]$, while others use a three-day window $[-1, +1]$, and the table does not make a distinction between these. Some studies also separate out industrials from utilities, and when they do, we report results averaged across both issuer types.

### *4.4.3. Market reaction to corporate debt offerings*

The basic adverse selection argument of Myers and Majluf (1984) strongly suggests that the market reaction to security offerings should be smaller the lower the risk that the security is overpriced. This implication is also a basic motivation for the financing pecking order of Myers (1984). Given the predictable contractual payment stream embedded in a debt contract—protected by bankruptcy law—the risk of market mispricing is almost certainly lower for a corporate debt instrument than for common stock. Thus, the market reaction to debt issues should therefore be smaller than for equity.

Table 15 lists studies reporting the stock-price announcement effect of straight and convertible debt offerings by U.S. firms. In Panel (a), the overall evidence is of a statistically insignificant market reaction to straight debt issuances. Dann and Mikkelson (1984) report a significantly negative average abnormal stock return of −0.37%, while Howton, Howton, and Perfect (1998) also report significantly negative market reaction of −0.50% over the two-day announcement period. However, the average market reaction is small and insignificant in all of the subsequent studies by Mikkelson and Partch (1986), Eckbo (1986), Hansen and Crutchley (1990), Shyam-Sunder (1991), Chaplinsky and Hansen (1993), Johnson (1995), Jung, Kim, and Stulz (1996) and Howton, Howton, and Perfect (1998). The sample-weighted average across all of the studies is a statistically insignificant $AR_{sd} = -0.22\%$.

Straight debt issues are to some extent predictable as the maturity date approaches and the firm needs to refinance. Bayless and Chaplinsky (1991), Chaplinsky and Hansen (1993) and Jung, Kim, and Stulz (1996) develop models to predict whether an issuer will choose to sell a public issue debt or equity. Chaplinsky and Hansen (1993) examine issuers of public debt and find that issues have substantial predictability and that issuers have significantly lower earnings, significantly higher investment growth and debt refinancing needs in the years immediately preceding and following the offering. Gomes and Phillips (2005) examine private and public security issuance activity by publicly listed firms. They find that firms with higher levels of asymmetric information measured by analysts' earnings forecast errors or dispersion in earnings forecasts are less likely to issue common stock or convertibles relative to debt in the public capital markets, but these firms are more likely to issue equity and convertibles over debt in the private capital market. They also find that smaller public firms with higher risk, lower profitability and good investment opportunities are more likely to issue equity and convertibles privately, while firms experiencing stock price rise in the prior year relative to a benchmark portfolio are more likely to issue equity in the public market.

Since announcement returns represents only the unanticipated portion of the total price effect, this raises the question of whether partial anticipation explains the largely insignificant market reaction to straight debt issues in Panel (a) of Table 15. Eckbo (1986) addresses this issue by partitioning his sample according to the stated purpose of the issue (refunding versus finding of investment program), and according to risk (bond ratings). Presumably, the degree of market anticipation is lower the riskier the debt issue, and if the purpose is to fund new investment opportunities. However, Eckbo (1986) reports that none of the subsample results sorted in this fashion indicate a significant market reaction. Shyam-Sunder (1991) also find no effect of bond risk on announcement returns, as measured by bond ratings. Bayless and Chaplinsky (1991) develop a forecasting model for a firm's debt versus equity issuance choice and find larger announcement effects when a security that is not expected is issued. For example, debt issue announcements when an equity issue was expected have a positive 1% average abnormal stock return (1 day). Chaplinsky and Hansen (1993) partition the debt sample according to stated purpose of the issue, and find that the market reaction is insignificant except in the sample of 68 issues with "no purpose specified" where it is a significantly negative $-0.63\%$. Overall, there are few indications that the evidence in Panel (a) is significantly affected by partial anticipation. However, this remains a topic for future research.

Finally, Panel (b) of Table 15 lists studies that report the stock market reaction to convertible debt offerings. Since convertibles are a hybrid of straight bonds and warrants, the risk of overpricing (of the warrant) is greater than for straight bonds. They are

Table 16
Summary of sample-weighted average market reaction (*AR*, %) to security offerings (aggregate sample size and sample period in parentheses)

| Type of offering | U.S. | Foreign |
|---|---|---|
| A. SEOs | | |
| Uninsured rights | $AR_{ur} = -0.59$ (53; 1963–1981) | $AR_{ur} = 0.70$ (484; 1980–1999) |
| Standby rights | $AR_{sr} = -1.33^*$ (349; 1963–1998) | $AR_{sr} = -1.32^*$ (1,201; 1980–1999) |
| Private placements | $AR_{pp} = 2.45^*$ (2,830; 1979–2000) | $AR_{pp} = 3.12^*$ (691; 1974–1999) |
| Firm commitments | $AR_{fc} = -2.22^*$ (15,017; 1963–2001) | $AR_{fc} = 1.10^*$ (1,064; 1974–1997) |
| Shelf offerings | $AR_{sh} = -0.66^*$ (1,851; 1980–2003) | n.a. |
| B. Debt offerings | | |
| Straight debt | $AR_d = -0.24$ (2,615; 1964–1993) | n.a. |
| Convertible debt | $AR_{cd} = -1.82^*$ (307; 1964–1982) | n.a. |

The *AR* reported in this table also appear in the panels headings in Tables 13, 14 and 15. The reported *AR* weighs each individual study in the panel with its sample size. Superscript * indicates statistical significance at the 1% level.

also less predictable than straight debt offerings. So, the expectation is that convertibles will be met with a stronger market reaction than straight debt issues. Dann and Mikkelson (1984), Mikkelson and Partch (1986), Eckbo (1986), and Hansen and Crutchley (1990) all report negative and statistically significant market reactions to convertible debt offerings. The sample-weighted average abnormal return is a statistically significant $AR_{cd} = -1.82\%$.

### 4.5. Implications of the announcement-return evidence

For convenience, the sample-weighted averages reported in these tables are summarized in Table 16. The significant price reaction to security offerings leaves little doubt that these corporate events typically convey significant new information to the market. As such, the evidence provides generic support for models of the issue decision that presume some form of asymmetric information between the issuer and the market.

What is more difficult to determine, of course, is the precise content of the new information that the market is reacting to. We discuss some possible inferences below. These are the result of cross-sectional analysis of the announcement effect, often performed using multivariate regressions with the announcement effect *AR* as dependent variable. The expected profits from issuing and investing shown in equation (1), and the various theoretical models listed in Table 12 suggest a link between *AR* and a set of characteristics:

$$AR = f(m, C, k, q, \beta, \sigma, I, b/a, P), \quad m = ur, sr, fc, pp, \tag{5}$$

where the parameters are the flotation method choice ($m \in [ur, sr, fc, pp]$), direct and indirect issue costs ($C$), expected shareholder takeup of the issue ($k$), signal quality or the informativeness of the available issue-quality certification technology ($q$), private benefits of control ($\beta$), the ex ante risk that the security is overpriced ($\sigma$), growth as given by the size of the project's investment amount ($I$) and the size of the project's NPV relative to the value of assets in place ($b/a$), and market beliefs about the nature of firms' equilibrium flotation strategies. These beliefs imply an issue market price of $P$, which in some equilibria are lower than the true, intrinsic value, resulting in an undervaluation cost-component in $C$.

A caveat before proceeding with the results: it should be noted that the explanatory power of regressions of the type in equation (5) as reported in the literature is uniformly low, almost always less than 10%. More seriously, these cross-sectional regressions are typically estimated using linear estimators (such as OLS). Eckbo, Maksimovic, and Williams (1990) show that linear estimators (such as OLS and GLS) are biased and inconsistent when the issuer self-selects the timing of the event (in this case security issue) and derive a consistent, non-linear estimator.[38] Some studies (e.g., Bøhren, Eckbo, and Michalsen, 1997) report results with the nonlinear estimator, while others (e.g., Eckbo and Masulis, 1992) report that key inferences are unchanged when using OLS. Moreover, the potential for bias is smaller for utilities that are constrained by the regulatory process. However, for the vast majority of studies reporting cross-sectional regressions, the magnitude of the bias introduced by self-selection is largely unknown.

*Adverse Selection and growth opportunities*. In Myers and Majluf (1984), the market prices firms correctly only on average, causing some highly undervalued firms to avoid dilutive equity issues. Here, the information content is simply the adverse selection revealed by the firm's willingness to issue (separating equilibrium). The negative average market reactions to SEOs sold *to the market* in the U.S., such as in standby rights and firm commitment offerings, is consistent with this generic framework. Moreover, as pointed out by Eckbo and Masulis (1992), equity issues that are purchased by current shareholders (i.e., *not* sold to the market) results in pooling and therefore do not convey information. This prediction is also supported by the evidence on uninsured rights in Table 16, both in the U.S. and internationally.

The adverse selection model also implies that the market reaction to equity offerings should be more negative the greater the issue size (Krasker, 1986) and the greater the ex ante uncertainty that the issue is overpriced. The uncertainty hypothesis is supported by the evidence that debt offerings are met with little or no market reaction, while convertible debt offerings produce a negative effect that is only about half the size of the

[38] This issue is surveyed extensively in Li and Prabhala (2007) (Chapter 2 of this volume).

average market reaction to SEOs. Convertibles are a hybrid between debt and equity, and a convertible debt offering may be viewed by the market as a delayed equity issue (Stein, 1992). The uncertainty hypothesis is also supported by the finding that the market reaction to equity issues by regulated utilities is much smaller (though still significant) than the average market reaction to industrial issuers. The regulatory process required for a utility to issue equity reduces the issuer's discretion to time the issue to periods where the market is overvaluing the stock.

The evidence on the effect of issue size on the market reaction is mixed. While Jung, Kim, and Stulz (1996) find no relationship to issue size, Masulis and Korwar (1986), Korajczyk, Lucas, and McDonald (1990), and Bayless and Chaplinsky (1996) find a significantly negative relation between the announcement-induced abnormal return and the size of the offer.

As is evident from equation (1) in Section 4.2, the firm's incentive to issue is greater the greater the investment project's NPV ($b$). If $b$ is sufficiently large relative to the value of assets in place $a$, then the firm will issue even it the shares are undervalued by the market. If $b$ is sufficiently large relative to $a$ for all firms, there is no adverse selection (pooling equilibrium) and no adverse market reaction to the issue announcement. However, in a separating equilibrium (with adverse selection), the market reaction will be more favorable the greater the ratio $b/a$. Since the value of $b$ is unobservable to the econometrician, studies have used the issuer's B/M ratio or Tobin's $Q$ as a proxy for "growth". The evidence is mixed: while Jung, Kim, and Stulz (1996) report a significantly positive relation between the market reaction to equity announcements and B/M ratios, several studies fail to find a significant relation (Barclay and Litzenberger, 1988; Dierkens, 1991; Pilotte, 1992; Denis, 1994).

*Shareholder takeup*. In Eckbo and Masulis (1992), shareholder takeup $k$ simply acts like financial slack. The greater $k$, the smaller the issue sold to the market, and the lower the scope for wealth transfer from outside investors. Thus, the greater $k$, the smaller the market reaction to the issue announcement. In the notation of Table 16, the prediction is $AR_{fc} < AR_{sr} < 0$ and $AR_{ur} \approx 0$. This prediction is supported by the evidence on U.S. offerings: $AR_{fc} = -2.2\%$, $AR_{sr} = -1.3\%$ (both significantly different from zero and significantly different from each other), and $AR_{ur} = -0.6\%$ (not significant). There is also direct evidence that the takeup parameter $k$ is highest in uninsured rights offerings, lowest in firm commitments, with standbys in between. Thus, the evidence supports the hypothesis that expected shareholder takeup affects the flotation method choice under adverse selection.[39]

*Quality certification*. In the vernacular of Eckbo and Masulis (1992), Bøhren, Eckbo, and Michalsen (1997), and Eckbo and Norli (2004), the significantly negative market

[39] Bøhren, Eckbo, and Michalsen (1997) and Cronqvist and Nilsson (2005) provide direct evidence on $k$. Generally speaking, the value of $k$ depends on shareholder (personal) wealth constraints and demand for diversification by risk-averse investors. Moreover, $k$ is likely to reflect the presence (if any) of individual shareholders' private benefits of control.

reaction to standbys and firm commitment offerings indicate that the signal quality of the underwriter certification technology only partially reveals the issuer's true quality. With perfect revelation and firm-value-maximization on the part of the issuing firms, the market reaction would be non-negative.[40] Thus, the evidence favors models that presume some form of imperfection in the underwriter's quality certification.

*Shareholder monitoring*. A private placement offers opportunities and incentives for communication between the issuer and the private placement investor which may alleviate ex ante investor nervousness with the possibility that the offer is overpriced. This may induce *positive* selection in the pool of private placement issuers. This is consistent with the evidence. As summarized in Table 16, the typical private placement offering of equity generates a significantly positive market reaction, with $AR_{pp} = 2.5\%$ in the U.S. and $AR_{pp} = 3.1\%$ internationally.

What is the nature of the positive information? Wruck (1989) and Herzel and Smith (1993) suggest that the positive announcement effect reflects the fact that the firm is willing to subject itself to increased monitoring and certification by a large, private placement investor. A positive announcement effect if also predicted by the variant of the Myers and Majluf (1984) model developed by Cooney and Kalay (1993) and Wu and Wang (2005), where managers are allowed to select value-decreasing investment projects. Cronqvist and Nilsson (2005) and Wu and Wang (2005) argue that large shareholders prefer a rights issue over a private placement in order to protect private benefits of control. Cronqvist and Nilsson (2005) conclude that family-controlled firms in Sweden avoid issue methods that dilute control benefits. Wu, Wang, and Yao (2005) and Wu and Wang (2006b) reach a similar conclusion after studying control-diluting placements and rights issues in Hong Kong. Thus, the selection of private placement carries a positive signal relative to a rights offer, which is also consistent with the evidence.

Do private placements in fact lead to increased monitoring? Empirically, Barclay, Holderness, and Sheehan (2005) conclude that there is little direct evidence of monitoring activities by private placement investors in the U.S. If this is in fact true, then the positive announcement effect of private placements represents positive information about the issuer per se, perhaps due to the certification role played by the private placement investor (Eckbo and Norli, 2004).

*Managerial earnings expectations*. Ross (1977) develops a model in which the firm's issue decision reflects private managerial information about the firm's future earnings prospects. Managers face personal bankruptcy costs and prefer to issue equity over debt when they have private information indicating a future decline in earnings, and vice versa for debt issues. This model implies a negative market reaction to an equity issue and a positive market reaction to a debt issue. While the empirical evidence is consistent with the first part of this prediction, the evidence contradicts the second part. The

[40] As discussed in Section 3 above, the focus of the underwriter is typically on certifying the existence and value ($b$) of the investment project, the validity of the firm's accounting statements, the firm's strategic plans, etc.

market reaction to straight debt offerings summarized in Table 16 is not statistically significantly different from zero. As shown by Eckbo (1986), even large debt issues—where the stated use of the proceeds is to fund the firm's investment program—do not elicit a positive market response.

*Wealth transfer to bondholders*. Holding the firm's investment policy constant, an equity issue reduces the risk of the firm's outstanding debt. However, it is unlikely that this effect explains much of the empirical evidence. While studies of bond returns in response to equity issues are difficult due to data constraints, Kalay and Shimrat (1987) find that equity issues on average cause bond prices to *fall* rather than increase. Moreover, as indicated above, there is little if any evidence that large debt issues cause equity prices to rise. In sum, the wealth transfer hypothesis is inconsistent with the evidence.

### *4.6. Signaling and the rights offer discount*

Heinkel and Schwartz (1986) presents a model in which relatively high-quality uninsured rights issuers signal their quality to the market by *lowering* the rights offer discounts. They assume that a failed rights offer is costly for all issuers. Suppose there are two issuer types, "high" and "low", and let the two firms have the same ex ante market price $P$ (before the rights offer announcement). The low type has a greater probability than the high type of experiencing a stock price reduction over the fixed rights offer period (say, four weeks) before the rights expire. If the rights subscription price $P_0$ is set close to $P$, the rights are expected to trade close to zero, and the probability that the offer will fail (because the stock price drops) is greatest for the low-value type. In the separating equilibrium considered by Heinkel and Schwartz (1986), the high-value firm signals its type by reducing the rights offer discount.

Alternatively, one may use a signaling models such as that of John and Williams (1985) to generate a *positive* impact of a rights offer discount, opposite to Heinkel and Schwartz (1986). As discussed by Hietala and Loyttyniemi (1991) and Bigelli (1998), in some European countries, a rights offer sometimes produces an increase in dividend yield. For example, if the rights offer does not affect the firm's dollar dividend per share, and the rights offer subscription price is set at a discount from the pre-offer stock price, then the dividend as a percent of the post-offer share price increases as the share price falls due to the discounted sale of shares. For a given dollar dividend, the increase in dividend yield is proportional to the discount in the rights offer price. The dividend yield will increase as long as the dividend per share is reduced by less than the share-split effect of the rights offer discount. A positive signaling effect of the dividend implication of a rights offer discount also reduces the expected cost of offering failure, as it increases the probability that the rights will be in the money at the expiration date.

We are aware of four studies that report evidence on the information content of rights offer discounts. First, with their sample of U.S. rights offers, Eckbo and Masulis (1992) regress the offering-day abnormal stock return (which in the U.S. contains the market reaction to the news of the offering price) on offer-specific characteristics, including the discount and the flotation method. The estimated coefficient on the discount is insignificantly different from zero whether the issuer is an industrial firm or a public utility. The lack of a significant impact of the discount holds whether or not they account for subscription precommitments in uninsured rights. This is important because greater levels of subscription precommitments lower the risk of rights offer failure, thus reducing the signaling effect of the discount itself. Overall, they find no support for the proposition that the rights offer discount signals information (positive or negative) to the market about the true value of the issuer.

Second, using Norwegian standbys and uninsured rights offerings, Bøhren, Eckbo, and Michalsen (1997) also examine the information content of the rights offer discounts. In contrast to rights offerings in the U.S., Norwegian issuers are required to set the rights offer price a minimum of three weeks prior to the beginning of the rights offer period. With a minimum rights offer period in Norway of two weeks, this means that the issuer (and standby underwriter) must forecast the issuer's secondary market price at least five weeks ahead when determining the optimal offer price. The longer prediction period probably increases the risk of offering failure relative to the U.S., making the Norwegian rights offers a relatively powerful laboratory for examining signaling effects. They fail to find a statistically significant effect of the offer price discount on the market reaction to rights offer announcements.

Third, with a sample of U.S. utility standby rights offerings, Singh (1997) report that abnormal stock returns over the "rights settlement period" (i.e., the period from the day before the offer price release day and the following six days) are positively correlated with the offering price discount.[41] Since his sample includes fully guaranteed rights offerings only, there are no failure costs, so the signaling argument of Heinkel–Schwartz does not apply.

Finally, focusing specifically on dividend implications of rights issues, Bigelli (1998) reports a dividend-yield increase in more than 80% of his sample of Italian rights offers. He finds that the average market reaction to rights offer announcements is positive, and positively related to the subscription price discount. This is inconsistent with Heinkel and Schwartz (1986) but consistent with separating equilibria in which unanticipated dividend increases have information content. Further research is needed to establish whether dividend increases associated with rights issues have information content also in other issue markets.

[41] Singh (1997) reports that there are on average 30 trading days between the first public announcement of the standby and the "price release date" (the date on which the market first learns of the actual subscription price). The price release date typically coincides with the date of the price amendment of the offering prospectus, which is also typically the start of the offering.

## 5. Security offerings and market timing

Consider a company that faces a steady stream of new projects. In the standard corporate finance textbook, projects are executed if they have a positive net present value. If the owner of the project needs external financing, capital markets will provide the needed funds and the type of security has no effect on the project's value. In this setting, there is no room for timing a security offering. However, Graham and Harvey (2001) present survey evidence that suggests that managers are concerned about the appropriate timing of equity issues. Moreover, the stylized facts concerning the stock price dynamics around SEOs (a stock price runup prior to the issue, a negative market reaction to the announcement of the issue, and long-run returns that appear low compared to similar firms) seems to indicate that managers are timing these issues around periods of temporary overvaluation.

This section reviews various models that focus on explaining the timing of seasoned equity offerings. Prior to the mid 1990s, the low long-run stock returns were not commonly known. Thus, papers written prior to this period focused on explaining the stock price runup and the negative average announcement effect. Later models also had to explain post-issue stock price performance patterns. We discuss three classes of models: one based on rational market pricing, another with some non-rational agents, and finally a statistical model of "pseudo-timing".

### *5.1. Timing theories with rational market pricing*

As discussed in Section 4, information asymmetry between managers and investors may create an incentive for managers to time an equity issue. Some undervalued firms will forgo profitable projects because the dilution costs of issuing undervalued equity borne by existing shareholders are too high relative to the project's profitability. Other undervalued firms will only issue if the project can be financed with debt. Myers (1984) builds on this insight and suggests that there is a financing choice pecking order in which firms only use equity as a last resort.

Korajczyk, Lucas, and McDonald (1992) and Choe, Masulis, and Nanda (1993) develop models of dynamic adverse selection that imply a relationship between equity issue activity and, respectively, firm specific information releases and the business cycle. The model of Korajczyk, Lucas, and McDonald (1992) predicts clustering of equity issues after information releases (especially quarterly and annual financial reports). Choe, Masulis, and Nanda (1993) observe that during periods of economic expansions, corporate investment opportunities are more profitable, and thus, adverse selection costs are lower. In these models, managers time the sale of equity offers to periods when information asymmetries are less severe. Bayless and Chaplinsky (1996) report that equity issues tend to cluster in periods with smaller average announcement effects. They interpret this pattern as evidence that issuers timing equity offerings to periods with lower levels of asymmetric information.

The model of Lucas and McDonald (1990) departs from other models of adverse selection in that they allow the firm's investment opportunity to be postponed. This gives undervalued firms an incentive to postpone an issue until the stock price is higher relative to the manager's valuation based on proprietary information. This implies that empirically we should obverse more equity issues following bull markets.

Projects that can be postponed as the firm waits for more favorable market conditions to issue equity can be viewed as real options. Carlson, Fisher, and Giammarino (2005, 2006) present a real option model with rational agents that can explain the stock price dynamics around seasoned equity offerings. We discuss these models in more detail below.

### *5.1.1. Adverse selection and the business cycle*

In Choe, Masulis, and Nanda (1993), an adverse selection argument similar to Myers and Majluf (1984) is developed where firms choose between issuing debt and equity across business cycle expansions and contractions, where firms receive non-deferrable profitable investment opportunities, and they must issue debt or equity securities to pursue them.[42] If a firm issues debt, investors will demand either protective covenants or a price discount for anticipated asset substitution risk once the debt is issued. This imposes a debt issuance cost on all issuers. On the other hand, firms with undervalued equity will only issue equity when the dilution cost from selling undervalued stock is less than or equal to the debt issuance cost.[43] In the aggregate, the marginal equity issuer will find the dilution cost of issuing undervalued equity is just equal to the cost of debt issuance and will be indifferent to issuing debt or equity. All other firms will find that one of the two securities will dominate due to their lower issuance costs. Also, if a firm issues equity, then the market knows that the equity was not substantially underpriced, because if it was the firm would have issued debt. Thus, an equity announcement should be greeted with a negative price reaction because investors now know that the firms issuing equity are drawn from a less desirable distribution that is truncated from above and the opposite is true for firms issuing debt.

Choe, Masulis, and Nanda observe that corporate investment opportunities are typically more profitable in periods of economic expansions than during contractions. This can reduce the dilution effect of equity issuance, though the cost of debt issuance is relatively insensitive to the point in the business cycle when an offer occurs. In economic expansions it is common knowledge that the average firm issuing equity will be more profitable and the marginal equity issuer will need to be more underpriced ex ante, if its equity dilution effect is to equate to the debt issuance cost. In addition, all less underpriced firms will prefer to issue equity. Thus, fewer firms will choose to issue debt over

[42] Parts of this section are drawn from Eckbo and Masulis (1995).

[43] The dilution cost of issuing equity is assumed to be more than offset by the profits of the investment opportunity or else no investment would take place.

equity. As more profitable and more underpriced firms find it optimal to equity finance, the equity offer announcement effect (the adverse selection effect for the average equity issuer) is reduced, lowering the issuance cost of equity. Thus in economic expansions, the model predicts a smaller equity offer announcement effect and an rise in the relative frequency of equity offers.[44]

Consistent with the prior prediction, both Moore (1980) and Choe, Masulis, and Nanda (1993) find empirical evidence that the frequency of equity offers relative to debt offers rises in expansions, while at the same time the magnitude of the negative stock price reaction to firm commitment equity offer announcements decreases. In contrast, debt issues are insensitive to this equity issue mispricing effect. The evidence in Choe, Masulis, and Nanda (1993), Marsh (1982) and Taggart (1977) indicates that the number of straight debt offers does not fall in economic contractions and may in fact rise if interest rates also fall with the contraction. This latter effect may in part reflect debt refinancing activities in these periods.

The model of Choe, Masulis, and Nanda (1993) also predicts that the adverse selection effect increases as investor uncertainty concerning the value of assets in place rises. Schwert (1989) documents that stock price volatility varies over the business cycle, increasing during recessions.[45] Controlling for the effect of the business cycle, Choe, Masulis, and Nanda (1993) find that the relative frequency of equity issues is significantly negatively related to the issuer's daily stock return variance, which gives further empirical support to their adverse selection framework.

Several other hypotheses concerning the timing of equity offers can be extended to a business cycle environment. For example, under Myers (1984)'s pecking order hypothesis, firms are viewed as preferring to finance projects internally if possible, otherwise to issue low risk debt and to issue equity only as a last resort. Imposing an arbitrary limit on firm leverage, the timing of equity issues is affected by business cycle downturns that reduce internal sources of funds and raise leverage by lowering asset values, thereby making equity offers more attractive. However, this equity issuance scenario is inconsistent with the evidence found in Choe, Masulis, and Nanda (1993).

Another hypothesis is based on debt-equity wealth transfers predicted by Galai and Masulis (1976) and Jensen and Meckling (1976) to occur when leverage is unexpectedly revised. If a firm issues equity, thus lowering its leverage, debtholders gain since their risk premium continues to be paid in full, while their risk bearing falls. This tends to discourage management seeking to maximize shareholder wealth from undertaking equity offers, except when leverage has become unacceptably high. In economic contractions, debtholders bear greater risk and expect greater risk premiums. So in downturns, equity offers cause leverage to fall more, resulting in larger reductions in debt risk-bearing and

[44] If less profitable investment projects or projects with varying profitability are assumed, then the model predicts in economic expansions that fewer undervalued firms will forego equity financing because of their project's greater profitability.

[45] Schwert links this volatility increase to increases in operating leverage, which is likely to be positively related to investor uncertainty concerning the value of assets in place.

greater debtholder wealth gains. Thus, there are greater costs to equity issues in economic downturns, leading to a lower predicted frequency of equity offers and a more negative stock price reaction. However, the predicted positive price reaction of outstanding debt to equity offers under the wealth transfer hypothesis is not observed by Kalay and Shimrat (1987).

In the Stulz (1990) model of free cash flow, debt issuance becomes more attractive when a firm's free cash flow increases. In economic contractions, if earnings decline less sharply than capital spending, which is typically the case, then free cash flow can increase, which increases the attractiveness of debt offerings. The cost of debt issuance in the Stulz model is underinvestment in profitable projects, but this would tend to be less of a problem in economic downturns. Thus, debt issuance would appear to be predicted to rise in contractions under the Stulz model, which is contrary to the evidence in Marsh (1982) and Taggart (1977), but somewhat supported by the evidence reported by Choe, Masulis, and Nanda (1993). This prediction is also supported by the evidence found in Jung, Kim, and Stulz (1996), who observe that firms with relatively good investment opportunities measured by the market to book ratio, are significantly more likely to issue equity over straight debt.

Lucas and McDonald (1990) develop a dynamic model of the equity issuance process that predicts a greater frequency of equity issuance following a general stock market increase. They show that since firm's with temporarily underpriced stock have an incentive to postpone an offering until the stock price is higher, the resulting average pre-announcement price path of these issuing firms will be upward sloping. On the other hand, firms with temporarily overpriced stock will issue equity immediately as new investment opportunities arise. If the arrival of investment projects is uncorrelated with a firm's price history, then the average pre-equity offering announcement price path of temporarily overvalued stocks will be flat. As a result, the average preannouncement price path of all issuing firms will be upward sloping, as is typically observed in samples of firm commitment equity offers. Lucas and McDonald also argue that the market reaction to an equity issue announcement will be more negative for firms with higher pre-announcement period stock price gains, which is supported by the regression results of Masulis and Korwar (1986), Korajczyk, Lucas, and McDonald (1990), Eckbo and Masulis (1992), and Jung, Kim, and Stulz (1996).

As discussed in Section 4, Eckbo and Masulis (1992) point out that increased shareholder participation in equity issues reduces the incentives of firms with undervalued equity to postpone their offers since current shareholders capture part of any underpricing. At one extreme, when current shareholders purchase the entire issue (shareholder takeup $k = 1$), the firm issues immediately regardless of its current degree of underpricing. Thus, in a sample of issuers where the average level of shareholder participation is known to be large, the Eckbo and Masulis (1992) model predicts that there should be little or no stock price runup prior to the issue announcement. This prediction is supported by their evidence of little or no runup prior to an uninsured rights offer announcement, a modest positive runup prior to standby offer announcement and a larger positive runup effect prior to a firm commitment underwritten offer announcement.

Another hypothesis that predicts variation in the relative frequency of equity and debt offers over the business cycle is the belief of many practitioners that management prefers debt issuance when interest rates are historically low and prefers to issue stock when its price is historically high, regardless of whether this is caused by relatively low equity risk premiums or relatively high expected cash flows.[46] Since stock market prices tend to reflect future economic prospects, this hypothesis tends to predict increases in equity offers in economic expansions, when equity prices are relatively high and debt issues in economic contractions, when interest rates are also low.[47] These predictions are consistent with the evidence in Marsh (1982) and Taggart (1977), but only partially consistent with the evidence in Choe, Masulis, and Nanda (1993).

Bayless and Chaplinsky (1991) explore the effects of both firm-specific and macroeconomic variables on the security issue choice. The macroeconomic variables include the prior 3 month performance of the stock market (S&P 500), 3 month change in the Treasury bill interest rate and a corporate default premium. They find larger announcement effects when a security that is not expected is issued. Korajczyk and Levy (2003) also explore the effects of macroeconomic conditions and financial constraints on the security issue choice. They report that financially unconstrained firms act in a significantly different manner from financially constrained firms, which are defined as firms not paying cash dividends, not making net equity or debt repurchases and having a market to book ratio of greater than one. The lagged macroeconomic variables that they examine are: the term spread, the default spread and a three month equity market return. They find that unconstrained firms issue activity is significantly affected by macroeconomic variables, while for constrained firms, this is not the case, except for the lagged stock market return. They also find that equity issuance is more likely when the lagged three month average of two-day SEO announcement returns is less negative and when the issuer's prior one year abnormal stock returns is higher. Korajczyk and Levy also estimate firm target leverage and then use deviations from it as another explanatory variable for the security issue choice decision and find that a leverage deficit leads to a significant increase in debt issuance. Lastly, they report that target leverage is counter-cyclical for the unconstrained firms, while it is pro-cyclical for the constrained firms. Their results suggest that researchers should be concerned with whether an issuing firm is financially constrained or not and they should also consider including macroeconomic variables as controls in their analysis of offering announcement effects.

#### *5.1.2. Optimal investments and equity offerings*

As pointed out by Carlson, Fisher, and Giammarino (2005, 2006), it is commonly assumed that investments in risky projects will increase asset risk. Moreover, this assumption is difficult to square with the observation that post SEO long-run stock returns

[46] See, for example, the survey of CFOs by Graham and Harvey (2001).

[47] This equity issuance effect can also be reinforced when warrants and convertible securities are outstanding, since a rise in the stock price can push these options into-the-money and also make conversion forcing calls attractive for many firms.

are low compared to the stock returns of similar non-issuing firms (also shown in Section 5.3 below). However, they argue that this observation follows naturally when projects are viewed as options on the cash flow potentially generated by the project.

When project execution is flexible in time, a project becomes a real option. Managers can time the starting time of the project to maximize the value of the firm. An option to grow the company through execution of the project is a levered claim. The required return on a levered claim is higher than the required return on an unlevered claim on the same assets. Exercising the real option, i.e., making the investment necessary to start the projects, unlevers the claim. Thus, when firms grow they convert real options into assets in place. The assets may be risky, but an option on these assets is even riskier. Thus, when projects are financed using seasoned equity, the model predicts that realized returns on average should be lower after a SEO. This does not happen because the SEO is timed, but rather because there has been a fundamental shift in the riskiness of the firm's assets. Since growth options only are exercised when they move sufficiently in-the-money, the model also explains the pre-issue stock price runup.

In the model of Carlson, Fisher, and Giammarino (2005, 2006) the required return is endogenous and depends (among other things) on the optimally timed investment decisions made by the firm. If the expected return is assumed to be time varying but exogenous, more projects will become profitable as the discount rate drops. This will increase investments and lead some firms to raise capital. Thus, time varying expected returns predict that stock prices will rise prior to equity issues and that returns will be lower after the issue. Pastor and Veronesi (2005) develop a model of IPO waves along these lines. Their model predicts that IPOs should cluster and that such IPO waves should be preceded by high market return and followed by low market return.

The relationship between investments and stock return was first formalized by Cochrane (1991). In a production based asset pricing model, Cochrane shows that a firm's investment return (the rate of return obtained on the marginal real investment) should be equal to the stock return. Thus, when the real investment level is high, the marginal return on invested capital is low, and stock returns should be correspondingly low. Cochrane (2005) interprets this argument as a first-differenced version of $Q$-theory of investment. Zhang (2005) develops the $Q$-theoretical argument further. Zhang focuses on time varying expected return and shows how $Q$-theory, among other things, implies that firms conducting a SEO should have lower post-issue returns than otherwise similar firms. Lyandres, Sun, and Zhang (2005) explore the investment based explanation for the low long-run stock returns of SEO firms. They find the investment to asset ratios of SEO firms are about twice as large as the investment to asset ratios of non-issuing firms. Thus, under the $Q$-theory of investment, the expected return of SEO firms should be lower than the expected return for non-issuing firms.

In sum, the investment based theories predict that subsequent to an SEO, a firm will have lower market risk and thus, lower expected rates of return. This offers a potential explanation for the finding, discussed in detail in Section 5.3 below, that stock returns are relatively low—but not necessarily *abnormally* low—following SEOs or IPOs. It also suggests that matching an equity-issuing firm with a non-issuing firm based on size and book-to-market ratio alone may be insufficient as a control for systematic risk. Such a match ignores the lower risk caused by the issuer's investment activity, and may lead to spurious evidence of "abnormal" post-issue returns.

#### *5.1.3. Pseudo market timing*

Schultz (2003) proposed pseudo market timing as another rational market explanation for the weak long-run stock returns observed after equity issues. The premise for the pseudo market timing hypothesis is that more firms issue equity as stock prices increase. It is irrelevant for the hypothesis why this happens, but, any of the rational theories discussed above could be the reason for increased issue activity as stock prices increases. Regardless of why the number of issues increases, the long-run performance has nothing to do with manager's predicting future returns. Schultz (2003) shows that if firms tend to issue stock after stock price increases (for whatever reason), on average issues will be followed ex post by underperformance. The reason is simple. Consider IPOs and suppose expected one-period returns are zero for all periods and all IPOs. Moreover, the return distribution is a bimodal $+10\%$ and $-10\%$ in each period. Let there be a single IPO at time zero. If the return in period one is $-10\%$, there will be no new IPOs at time one. Alternatively, suppose the return in period one is $+10\%$ and that there are four IPOs in this period. Now, compute the one-period abnormal buy-and-hold return for these two equally likely sample paths. It is $2\%$ for the "up" sample and $-10\%$ for the "down" sample, with an equally weighted average of $-4\%$. Schultz (2003) refers to this result as "pseudo market timing" because it may easily be confused by the researcher with real forecasting ability on the part of issuing firms' managers.

Several authors have explored to what extent pseudo market timing can explain the low return observed after IPOs. Dahlquist and de Jong (2004), Viswanathan and Wei (2004), and Ang, Gu, and Hochberg (2005) argue that pseudo market timing only is a potential explanation for the low post issue return when samples are small. Based on simulation experiments, all papers conclude that pseudo market timing is highly unlikely to be the main explanation for the low post issue stock market returns. The simulation experiments assume a stationary event generating process. Schultz (2004) show that one cannot reject a null that IPOs follow a nonstationary process and goes on to argue that, although pseudo market timing is a small sample problem, it is likely to be important in practice. Note that Schultz (2003)'s pseudo-timing argument also holds in principle for other security issuances, and in particular for SEOs where the matched-firm technique also have produced evidence of long-run underpricing by issuing firms (discussed below).

### *5.2. Timing theories with non-rational market pricing*

#### *5.2.1. Timing of firm-specific returns*

The timing hypothesis ("windows-of-opportunity") builds on the notion that investors are overly optimistic about the prospects of issuing firms, and as a consequence prices do not fully incorporate managerial incentives to time equity issues. This results in initial overpricing of issuing firms and a subsequent long-run underperformance when investors correct this initial mispricing over time.

The overconfidence hypothesis of Daniel, Hirshleifer, and Subrahmanyam (1998) is closely related, but is derived in a formal model and carries some explicit empirical predictions. The overconfidence hypothesis is based on the assumption that investors are overconfident about the precision of their private information, but not about the precision of public information. Overweighting private information relative to public information causes underreaction to new *public* information. Thus, the theory predicts that discretionary corporate events (such as equity issues) associated with abnormal announcement period returns, on average should be followed by long-run abnormal performance of the same sign as the average announcement period abnormal return, and there should be a positive correlation between announcement period abnormal returns and post-offer long-run abnormal returns.

Several empirical papers have explored different aspects of the timing and overconfidence hypotheses. Teoh, Welch, and Wong (1998) look at discretionary accruals in the years around an equity offering. The idea is that if investors are overly optimistic about the prospect of firms issuing equity, they would be willing to buy more shares and pay higher prices for them. As a result, issuing firms have incentives to cultivate this optimism by reporting inflated earnings before an equity offer. Both papers find evidence of earnings management prior to SEOs. For example, Teoh, Welch, and Wong (1998) find that although cash flows from operations on average decline prior to the SEOs, the reported discretionary accruals cause earnings to peak around the offer dates. Moreover, the amount of discretionary accruals prior to the seasoned equity offering is negatively related to the post-issue long-run stock return performance. The authors view this as evidence in favor of timing and overly optimistic investors. However, this issue is not settled as Shivakumar (2000) produces contradictory evidence using the specification of Teoh, Welch, and Wong (1998).

Cornett, Mehran, and Tehranian (1998) employ a direct test of the relationship between the incentive to time an issue and the subsequent stock return performance. They study voluntary and involuntary SEOs by commercial banks. Capital regulations in the banking industry state that banks are not allowed to have total capital ratios below a certain level. If the total capital ratio falls below the regulated lower bound, a bank may need to issue new equity to raise their capital ratio. Cornett, Mehran, and Tehranian (1998) define an involuntary SEO as an issue by a bank with capital ratio close to or below the required minimum ratio. If timing is driving the long-run underperformance of SEOs, we should expect to see less or no underperformance for involuntary issues. The results support the timing hypothesis, showing no abnormal three-year post issue stock return performance for the involuntary issues, while the voluntary issues show significant underperformance.

Brous, Datar, and Kini (2001) perform another test of the timing and overconfidence hypotheses. They argue that if managers are timing equity issues and investors systematically underreact to the issue announcements, we should expect to see that investors are disappointed when firms convey their post-issue earnings. That is to say, post-issue earnings announcement on average should be associated with negative stock price reactions. However, their results show no evidence of abnormal stock price reactions to the earnings announcements.

Kang, Kim, and Stulz (1999) tests the overconfidence hypothesis using data on Japanese public and private equity offerings. The non-negative announcement period abnormal return to Japanese equity offerings supports the view that equity offerings are regarded as good news in Japan. Nonetheless, they document post-issue negative long run abnormal performance. Taken at face value, this is evidence goes against the overconfidence hypothesis, but is consistent with investment based theories of equity issuance.

### *5.2.2. Timing the market*

Baker and Wurgler (2000) document that the proportion of equity in total new issues, termed "the equity share", is negatively correlated with future aggregate equity market returns. For example, when the equity share was in its top historical quartile, the average market return in the following year was $-6\%$. This could suggest that managers are able to time the market component of their company's returns. However, Baker, Ruback, and Wurgler (2007) is cautious about this interpretation. They suggest that: "A more plausible explanation is that broad waves of investor sentiment lead many firms to be mispriced in the same direction at the same time. Then, the *average* financing decision will contain information about the *average* (i.e., market level) mispricing, even though individual managers are perceiving and responding only to their *own* firm's mispricing".

Butler, Grullon, and Weston (2005a) question that timing ability or investor sentiment explain the predictive power of the equity share. They suggest that the apparent ability to time the market can be understood as a form of aggregate "pseudo market timing". They point out that on an ex-post basis equity share value tends to be high around market peaks and low around market troughs. Thus, it is the tendency to issue equity when prices are high that leads to a spurious relationship between equity share and future stock returns when measured ex post. They go on to argue that if equity tends to be issued when current prices are high, then equity issuance activity should go down during unexpected market declines—making pre-shock equity issuance look relatively high and post-shock equity issuance look relatively low. Thus, aggregate pseudo market timing should be most pronounced around market shocks. This prediction is supported by evidence that the predictive ability of the equity share is driven by the Great Depression (1920–1931) and the 1973–1974 Oil Crisis.

The main point in Butler, Grullon, and Weston (2005a) is that pseudo market timing can appear as real timing ability in small samples. Baker, Taliaferro, and Wurgler (2004) show that this problem extends to all time-series predictive regressions based on managerial decision variables. Moreover, it is a special case of the small sample bias studied by, among others, Stambaugh (1986, 1999).[48] For example, when a financial ratio such as book-to-market is used as a predictive variable, it will “pseudo-time” the market since the book-to-market ratio is hard-wired to rise as the market falls. There is an extensive literature on how to estimate the bias that this causes in predictive regressions. Using simulations, Baker, Taliaferro, and Wurgler (2004) report that pseudo-timing accounts for less than two percent of the predictive power of the equity share. However, the role of the pseudo-timing when the econometrician also allows for a non-stationary economic environment remains to be determined.

The debate about what causes the apparent ability of firms to time their equity issues to periods that are followed by low market returns is still inconclusive. Rational explanations along the lines of Carlson, Fisher, and Giammarino (2005, 2006) and Pastor and Veronesi (2005) are interesting and consistent with the arguments and results of several papers that empirically investigate long-run performance following security offerings. Next we turn to an in depth review of this long-run stock return literature.

### *5.3. Evidence on long-run post-issue stock returns*

Stocks generate surprisingly low returns over holding periods of 2–5 years following an equity issue date, as first shown for SEOs by Stigler (1964) and later reconfirmed and extended to IPOs by Ritter (1991) and more recent SEOs by Loughran and Ritter (1995). As discussed above, to some researchers, this long-run return evidence challenges the efficient markets hypotheses and motivates the development of behavioral asset pricing models. Responding to this challenge, Brav and Gompers (1997), Brav, Geczy, and Gompers (2000), Eckbo, Masulis, and Norli (2000), Eckbo and Norli (2005), and Lyandres, Sun, and Zhang (2005) present large-sample evidence that the low post-issue return pattern is consistent with standard multi-factor pricing models, and tend to be concentrated in small growth stocks with active investment programs. Thus, the low post-issue returns may be a manifestation of the more general finding in Fama and French (1992) that small growth stocks tend to exhibit low returns during the post-1963 period, or simply reflect the fact that asset pricing models have especially poor explanatory power for small growth stocks.

However, the proper interpretation of the low long-run returns following security issuances remains an unsettled issue. Ritter (2003) states that “the long-run performance evidence shows that in general the market underreacts to the [equity issue] announcements” (p. 262). Given the importance of the long-run performance evidence for the overall question of corporate timing and market efficiency, we provide a detailed review of the long-run performance evidence following IPOs, SEOs as well as corporate debt issues. We also report new updated abnormal return estimates of issuer abnormal returns

[48] See Baker, Taliaferro, and Wurgler (2004) for a more extensive list of papers that have studied this small sample bias.

based on security offerings made over the 1980–2001 sample period, and compare these to the extant literature.

*5.3.1. Sample selection*

The choice of sample period generally affects the statistical significance of reported abnormal return estimates.[49] Shorter sample periods reduce statistical power, while different sample periods have varying exposure to the problem of cross-correlation of overlapping holding-period returns (discussed extensively by Kothari and Warner (2007) in Chapter 1 of this volume). The literature uses security offer samples from as early as 1961 (Mitchell and Stafford, 2000) and as late as 2003 (Lyandres, Sun, and Zhang, 2005), with the bulk of the existing studies sampling from the 1980s and the early 1990s. The primary data source after 1980 is SDC, while earlier samples typically are found by searching the Wall Street Journal for issue announcements or relying on the SECs now defunct Registered Offerings of Securities database. Stock returns are almost always drawn from CRSP Daily Stock Price and Returns database.

Some authors exclude issues by public utilities on the grounds that the regulatory agencies make utility issues relatively predictable. Utility issues occurred on relatively frequent basis in the 1970s, and again as a result of deregulations in the late 1990s (Eckbo, Masulis, and Norli, 2000). As discussed above, the market reaction to SEOs is significantly smaller for utility issuers than for industrial issuers. Thus, it matters whether the utility issues are pooled in the long-run performance analysis. It is also customary to exclude issuers with stock price less than $5, as well as unit offerings and simultaneous offerings of other securities. Issues by foreign corporations, closed-end funds, unit investment trusts, and real estate investment trusts are also customarily excluded. Moreover, most studies require data on book value of equity, taken from Compustat, which further reduces sample size.

Our sample selection for the long-run analysis below is as follows. We start with the overall sample of 80,627 security issues from Section 2.3 above. Recall that this sample already ensures that the issuing firm is found on the CRSP tape for the relevant period. We then exclude the following issues using information from SDC: (1) ADRs and GDRs, (2) simultaneous offerings of debt and equity, (3) simultaneous offerings of international issues, (4) unit offerings, (5) offers with missing SDC information on offering proceeds, and (6) offerings after year 2000. The last restriction ensures five years of post-issue stock return data. These six criteria reduces the total sample to 54,283. We then apply restrictions specific to CRSP: (7) CRSP share code must be either 10 or 11 (ordinary common shares), (8) the issuer must be listed on NYSE/AMEX/Nasdaq,

[49] Figure 3 in Eckbo and Norli (2005) presents a striking illustration of the impact of sample period on the average holding period return. Due to the slump in the stock market in the mid-1970s, a study of long-run returns following IPOs (which starts with the first Nasdaq IPOs in 1973), will easily conclude that the IPO portfolio underperform the *risk-free* rate if the sample period ends prior to the mid-1980s.

and (9) information on market value of equity must be available. This results in a total sample of 44,986.

The breakdown of the total sample of 44,986 offerings across different types of security offerings is shown below. The second column of numbers indicates the sample size when we also require the issuer to have Compustat information on equity book-to-market ratio (B/M). The latter constraint is imposed when we identify non-issuing firms matched on B/M.[50]

Sample for the survey's long-run analysis ($N = 44{,}986$)

| Security type | Total | B/M available |
|---|---|---|
| IPO | 5,907 | 5,403 |
| SEO | 6,698 | 6,285 |
| Private placement of equity | 506 | 506 |
| Preferred equity | 1,530 | 1,412 |
| Convertible debt | 1,157 | 897 |
| Private placement of debt | 9,584 | 8,584 |
| Public straight debt issue | 18,447 | 17,360 |

We start the abnormal return analysis using the matched firm technique which requires B/M information. We then report the results of risk adjustments using factor regressions of portfolios of issuing firms.

### *5.3.2. Cumulative buy-and-hold returns for issuers versus matched firms*

The typical buy-and-hold experiment involves buying the issuing firm's stock in the month following the issue month, and holding the stock for a period of three to five years or until delisting, whichever comes first. In a sample of $N$ issues, the average return over a holding period of $T$ months is computed as the average cumulative ($T$-period) return, also referred to as $\overline{\text{BHR}}$ (for "buy-and-hold return"):

$$\overline{\text{BHR}} \equiv \frac{1}{\omega_i} \sum_{i=1}^{N} \left[ \prod_{t=\tau_i}^{T_i} (1 + R_{it}) - 1 \right], \tag{6}$$

where $R_{it}$ denotes the return to stock $i$ over month $t$, and $\omega_i$ is stock $i$'s weight in forming the average holding-period return ($\omega_i = 1/N$ when equal-weighting). The

[50] Book value is defined as "the Compustat book value of stockholders equity, plus balance sheet deferred taxes and investment tax credits (if available), minus the book value of preferred stock. Depending on availability, we use the redemption, liquidation, or par value (in that order) to estimate the value of preferred stock" (Fama and French, 1993, p. 8). If available on Compustat, the issuer book value of equity is also measured at the end of the year prior to the issue year. If this book value is not available, we use the first available book value on Compustat starting with the issue year and ending with the year following the issue year. On average, the first available book value is found 6.1 months after the offer date. Brav and Gompers (1997) look a maximum of 12 months ahead for book values while Brav, Geczy, and Gompers (2000) look a maximum of 18 months ahead.

effective holding period for stock $i$ is $T_i$, where $T_i$ in the analysis below is either five years or the time until delisting or the occurrence of a new SEO, whichever comes first. Kothari and Warner (1997), Barber and Lyon (1997) and Lyon, Barber, and Tsai (1999) provide simulation-based analyses of the statistical properties of test statistics based on long-run return metrics such as BHR. In Chapter 1 of this volume, Kothari and Warner (2007) survey the main statistical conclusions from this analysis.[51]

The matched-firm technique equates the *expected* return to issuing firms with the *realized* return to a non-issuing firm, usually matched on firm characteristics such as industry, size and book-to-market ratio. The abnormal or unexpected return BHAR is then

$$\mathrm{BHAR}_{\mathrm{Issuer}} \equiv \mathrm{BHR}_{\mathrm{Issuer}} - \mathrm{BHR}_{\mathrm{Matched\,firm}}. \tag{7}$$

Table 17 shows average five-year buy-and-hold returns following security offerings by U.S. firms that took place over the period 1980 through 2000, classified by the type of issuer.[52] As in Eckbo, Masulis, and Norli (2000) and Eckbo and Norli (2005), the matched firms are selected from all CRSP-listed companies at the end of the year prior to the issue-year and that are not in our sample of issuers for a period of five years prior to the offer date. We first select the subset of firms that have equity market values within 30% of the equity market value of the issuer. This subset is then ranked according to book-to-market ratios. The size and book-to-market matched firm is the firm with the book-to-market ratio, measured at the end of the year prior to the issue year, that is closest to the issuer's ratio. Matched firms are included for the full five-year holding period or until they are delisted, whichever occurs sooner. If a match delists, a new match is drawn from the *original* list of candidates described above.

[51] An alternative to $\overline{\mathrm{BHR}}$ is to estimate the average monthly return to a strategy of investing in the stocks of issuers and hold these for up to $T$ periods. The $T$-period return would then be formed as the *cumulative average* (portfolio) return, or

$$\overline{\mathrm{CMR}} \equiv \prod_{t=\tau}^{T}\left[1 + \frac{1}{\omega_t}\sum_{i=1}^{N_t} R_{it}\right] - 1.$$

As noted by Kothari and Warner (2007), depending on the return generating process, the statistical properties of $\overline{\mathrm{BHR}}$ and $\overline{\mathrm{CMR}}$ can be very different. Notice also that while $\overline{\mathrm{CMR}}$ represents the return on a feasible investment strategy, $\overline{\mathrm{BHR}}$ does not. You obtain $\overline{\mathrm{CMR}}$ by investing one dollar in the first security issue at the beginning of the sample period, and then successively rebalancing this initial investment to include subsequent issues as they appear (and $N$ increases), all with a $T$-period holding period. In contrast, $\overline{\mathrm{BHR}}$ is formed in event time—and thus presumes prior knowledge of the magnitude of $N$. Thus, estimates of $\overline{\mathrm{CMR}}$ are better suited than estimates of $\overline{\mathrm{BHR}}$ to address the question of whether investors have an incentive to take advantage of a potential market mispricing of security issues. Most of the empirical studies using the matched firm technique report results based on $\overline{\mathrm{BHR}}$, which we follow here. In the subsequent section, however, we discuss portfolio benchmark returns based on asset pricing models, which uses the return concept $\overline{\mathrm{CMR}}$ on a monthly basis, i.e., without the $T$-period cumulation.

[52] Utilities are firms with CRSP SIC codes in the interval [4910, 4939].

Table 17
Five-year buy-and-hold stock percent returns ($\overline{\text{BHR}}$) for U.S. issuers and size- and book-to-market matched control firms, 1980–2000

| | | Equally-weighted $\overline{\text{BHR}}$ | | | | Value-weighted $\overline{\text{BHR}}$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| Type of security issued | $N$ | Issuer | Match | Diff | $p(t)$ | Issuer | Match | Diff | $p(t)$ |
| A. Issues by industrial firms ($N = 20{,}262$) | | | | | | | | | |
| Initial public offerings | 5,018 | 35.7 | 53.8 | −18.0 | 0.010 | 52.8 | 67.6 | −14.8 | 0.208 |
| Seasoned equity offerings | 4,971 | 49.9 | 79.5 | −29.7 | 0.000 | 79.8 | 105.7 | −26.0 | 0.026 |
| Private placement of equity | 506 | 13.0 | 57.1 | −44.1 | 0.000 | 31.1 | 54.1 | −23.0 | 0.223 |
| Preferred equity | 379 | 43.8 | 96.7 | −52.9 | 0.000 | 79.1 | 113.6 | −34.5 | 0.238 |
| Convertible debt | 897 | 46.5 | 86.9 | −40.4 | 0.006 | 46.5 | 83.6 | −37.1 | 0.068 |
| Private placement of debt | 4,228 | 76.0 | 89.2 | −13.2 | 0.002 | 87.0 | 97.0 | −10.0 | 0.282 |
| Straight debt | 4,263 | 77.6 | 94.6 | −17.0 | 0.000 | 71.2 | 88.0 | −16.8 | 0.000 |
| B. Issues by banks and financial institutions ($N = 16{,}521$) | | | | | | | | | |
| Initial public offerings | 385 | 71.7 | 51.1 | 20.6 | 0.154 | 112.1 | 50.7 | 61.5 | 0.233 |
| Seasoned equity offerings | 655 | 98.3 | 98.3 | 0.0 | 0.999 | 75.6 | 73.4 | 2.3 | 0.870 |
| Preferred equity | 573 | 104.6 | 72.1 | 32.5 | 0.000 | 59.5 | 48.3 | 11.1 | 0.310 |
| Private placement of debt | 3,478 | 138.0 | 86.5 | 51.5 | 0.000 | 102.4 | 50.4 | 52.0 | 0.000 |
| Straight debt | 11,430 | 116.0 | 76.9 | 39.2 | 0.000 | 88.2 | 40.5 | 47.8 | 0.000 |
| C. Issues by public utilities ($N = 3{,}664$) | | | | | | | | | |
| Seasoned equity offerings | 659 | 116.3 | 135.4 | −19.1 | 0.012 | 100.6 | 132.9 | −32.3 | 0.010 |
| Preferred equity | 460 | 79.4 | 103.0 | −23.5 | 0.000 | 70.4 | 85.1 | −14.7 | 0.104 |
| Private placement of debt | 878 | 87.2 | 95.2 | −8.0 | 0.270 | 44.0 | 70.6 | −26.6 | 0.002 |
| Straight debt | 1,667 | 75.0 | 92.9 | −17.9 | 0.000 | 63.7 | 80.7 | −17.0 | 0.001 |

Buy-and-hold percent returns are defined as:

$$\overline{\text{BHR}} \equiv \omega_i \sum_{i=1}^{N} \left[ \prod_{t=\tau_i}^{T_i} (1 + R_{it}) - 1 \right] \times 100.$$

When equal-weighting, $\omega_i \equiv 1/N$, and when value-weighting, $\omega_i = MV_i/MV$, where $MV_i$ is the issuer's common stock market value (in 1999 dollars) at the start of the holding period and $MV = \sum_i MV_i$. The abnormal buy-and-hold returns shown in the column marked "Diff" represent the difference between the $\overline{\text{BHR}}$ in the "Issuer" and "Match" columns. The rows marked "$N$" contain number of issues. The $p$-values for equal-weighted abnormal returns are $p$-values of the $t$-statistic using a two-sided test of no difference in average five-year buy-and-hold returns for issuer and matching firms. The $p$-values for the value-weighted abnormal returns are computed using $U \equiv \omega' x/(\sigma\sqrt{\omega'\omega})$, where $\omega$ is a vector of value weights and $x$ is the corresponding vector of differences in buy-and-hold returns for issuer and match. Assuming that $x$ is distributed normal $N(\mu, \sigma^2)$ and that $\sigma^2$ can be consistently estimated using $\sum_i \omega_i (x_i - \bar{x})^2$, where $\bar{x} = \sum_i \omega_i x_i$, $U$ is distributed $N(0, 1)$.

Table 17 shows issuers on average underperform their matched firms when $\overline{\text{BHR}}$ is formed using equal-weights. For industrial issuers (Panel A), the five-year difference in the buy-and-hold returns of issuers and matched firms ranges from −52.0% for preferred equity placements ($N = 379$) to −13.2% for private placements of debt ($N = 4{,}228$). For IPOs ($N = 5{,}018$), the difference in buy-and-hold returns is −18.0% and −29.7% for SEOs ($N = 4{,}971$). Straight debt issues ($N = 4{,}263$) are associated with a difference in $\overline{\text{BHR}}$ of −17.0% while the return difference is 40.4% for convertible debt issues ($N = 897$). All return differences are statistically difference from zero at the one percent level.

Going from equal-weighting to value-weighting the returns alters the results dramatically. With value-weights, none of the differences are statistically different from zero at the one percent level, with the exception of *straight debt* issues ($p$-value of 0.000). Moreover, SEOs underperform their matched firms with a $p$-value of 0.026. Since value-weighting gives additional weight to above-average successful firms (relative to equal-weighting), the reduction in underperformance is expected. However, the fact that straight debt issuers in the value-weighted category reliably underperform matched firms while most other equity-type of issues do not is surprising.

Turning to security issuers by banks and financial institutions, there is no evidence of underperformance and some evidence of significant overperformance relative to the matched firms. With equal-weighting, financial issuers outperform matched firms when issuing preferred equity ($N = 573$) and straight debt placed either publicly ($N = 11{,}430$) or privately ($N = 3{,}478$). Value-weighting has almost no impact on the performance measure, except that preferred equity is no longer associated with abnormal performance relative to the matched firms.

As shown in the third panel of Table 17, issues by public utility companies produce underperformance similar to that of industrial issuers. The exception is private placements of equity ($N = 878$) which produces statistically insignificant underperformance for the equal-weighted buy-and-hold measure. Private placements do, however, significantly underperform using the value-weighted measure, as do SEOs ($N = 659$) and issuers of straight debt ($N = 1{,}667$).

Table 18 lists published studies that present evidence on buy-and-hold returns for several of the security sales in Table 17. For IPOs, and consistent with the results in Table 17, the studies of Brav, Geczy, and Gompers (2000), Ritter and Welch (2002) and Eckbo and Norli (2005) show insignificant abnormal returns over both three-year and five-year time horizons. For SEOs, the studies with the largest samples are Jegadeesh (2000), Brav, Geczy, and Gompers (2000), Eckbo, Masulis, and Norli (2000) and Clarke, Dunbar, and Kahle (2001). These show evidence of significant negative performance (3-year or 5-year), ranging from −4% to −34%. This is consistent with the −30% abnormal buy-and-hold return for the SEOs in Table 17. There is also negative, relative performance following private placements of equity (Hertzel et al., 2002; Krishnamurthy et al., 2005). Interestingly, Krishnamurthy et al. (2005) show that investors who participate in the private placement discount realize a normal post-issue, long-run performance.

Turning to debt offerings, with the exception of Eckbo, Masulis, and Norli (2000), there is consistent evidence of negative performance following convertible debt issues

Table 18
Average difference in equal-weighted buy-and-hold returns for U.S. issuers ($\overline{\mathrm{BHR}}_i$) and size- and book-to-market matched control firms ($\overline{\mathrm{BHR}}_m$)

| Study | Issuer type | Sample size | Sample period | Holding period | $\overline{\mathrm{BHR}}_i - \overline{\mathrm{BHR}}_m$ |
|---|---|---|---|---|---|
| A. IPOs | | | | | |
| Brav and Gompers (1997) | All | 3,407 | 1972–1992 | 5 yrs | 1.9%[a] |
| Brav and Gompers (1997) | All | 934 | 1972–1992 | 5 yrs | 16.5%*[b] |
| Brav, Geczy, and Gompers (2000) | All | 3,501 | 1975–1992 | 5 yrs | 6.6% |
| Ritter and Welch (2002) | All | 6,249 | 1980–2001 | 3 yrs | −5.1% |
| Eckbo and Norli (2005) | All | 5,365 | 1972–1998 | 5 yrs | −2.4% |
| B. SEOs | | | | | |
| Spiess and Affleck-Graves (1995) | All | 1,247 | 1975–1989 | 3 yrs | −22.8%* |
| Lee (1997) | All | 1,513 | 1976–1990 | 3 yrs | −20.3%*[c] |
| Jegadeesh (2000) | All | 2,992 | 1970–1993 | 5 yrs | −34.3%* |
| Brav, Geczy, and Gompers (2000) | All | 3,775 | 1975–1992 | 5 yrs | −26.3%* |
| Eckbo, Masulis, and Norli (2000) | Ind | 3,851 | 1964–1995 | 5 yrs | −23.2%* |
| Kahle (2000) | Ind | 1,739 | 1981–1992 | 3 yrs | −14.7%* |
| Clarke, Dunbar, and Kahle (2001) | All | 3,092 | 1984–1996 | 3 yrs | −14.3%*[d] |
| Clarke, Dunbar, and Kahle (2001) | All | 174 | 1984–1996 | 3 yrs | −3.3%*[e] |
| C. Private placements of equity | | | | | |
| Hertzel et al. (2002) | All | 591 | 1980–1996 | 3 yrs | −23.8* |
| Krishnamurthy et al. (2005) | All | 275 | 1983–1992 | 3 yrs | −38.4%*[f] |
| Krishnamurthy et al. (2005) | All | 273 | 1983–1992 | 3 yrs | −1.24%[g] |
| D. Straight debt offerings | | | | | |
| Spiess and Affleck-Graves (1999) | All | 392 | 1975–1989 | 5 yrs | −14.3% |
| Kahle (2000) | Ind | 523 | 1981–1992 | 3 yrs | −9.5% |
| Eckbo, Masulis, and Norli (2000) | Ind | 981 | 1964–1995 | 5 yrs | −11.2% |
| Eckbo, Masulis, and Norli (2000) | Util | 348 | 1964–1995 | 5 yrs | −10.4%* |
| Butler and Wan (2005) | Ind | 799 | 1975–1999 | 5 yrs | −24.0%*[h] |

(*Continued on next page*)

(Lee and Loughran, 1998; Spiess and Affleck-Graves, 1999; Kahle, 2000; Lewis, Rogalski, and Seward, 2001). For straight debt offerings, however, the literature shows insignificant long-run performance (Spiess and Affleck-Graves, 1999; Kahle, 2000, and industrial issuers in Eckbo, Masulis, and Norli, 2000). This contrasts with the results in Table 17 where debt issuers significantly underperform non-issuing matched firms. While the magnitudes of the abnormal returns are similar for straight debt issues in Table 17 and Table 18, the much larger sample size in Table 17 appears to provide greater precision, causing the null of zero abnormal performance to be rejected at the 0.1% level or better.

Measurement problems aside, underperformance following straight debt issues represents an enigma: there is little adverse selection as the choice of debt over equity is

Table 18
(*Continued*)

| Study | Issuer type | Sample size | Sample period | Holding period | $\overline{\text{BHR}}_i - \overline{\text{BHR}}_m$ |
|---|---|---|---|---|---|
| E. Convertible debt offerings | | | | | |
| Spiess and Affleck-Graves (1999) | All | 400 | 1975–1989 | 5 yrs | −37.0%* |
| Lee and Loughran (1998) | All | 986 | 1975–1990 | 5 yrs | −30.4%* |
| Eckbo, Masulis, and Norli (2000) | Ind | 459 | 1964–1995 | 5 yrs | −16.1% |
| Kahle (2000) | Ind | 527 | 1981–1992 | 3 yrs | −18.1%* |
| Lewis, Rogalski, and Seward (2001) | All | 566 | 1979–1990 | 5 yrs | −26.5%* |
| Butler and Wan (2005) | Ind | 303 | 1975–1999 | 5 yrs | −24.0%*[h] |

Buy-and-hold percent returns are defined as:

$$\overline{\text{BHR}} \equiv \frac{1}{N}\sum_{i=1}^{N}\left[\prod_{t=\tau_i}^{T_i}(1+R_{it})-1\right]\times 100.$$

Superscript * indicates significantly different from zero at the 1% level.
[a]Sample of non-venture-backed IPOs.
[b]Sample of venture-backed IPOs.
[c]Sample of primary issues. Matching firms are matched on size, book-to-market and prior annual return.
[d]Sample of completed SEOs.
[e]Sample of cancelled SEOs.
[f]Returns to non-participating investors (who do not buy shares in the private placement).
[g]Returns to participating investors (those who also capture the discount in the offering).
[h]Abnormal returns are insignificant when also matching on liquidity.

often thought to be associated with managerial beliefs that the firm's future earnings prospects are good (e.g., Ross, 1977). So why would debt issuers underperform non-issuing firms matched on size and B/M? Moreover, why would this underperformance be close to the magnitude for SEOs? The answer may reflect a combination of statistical problems with buy-and-hold return $\overline{\text{BHR}}$, as well as the matched firm technique producing the wrong benchmark for measuring the true systematic risk of issuing firms. In the subsequent section, we address this issue by measuring abnormal performance to issuing firms using both a monthly return horizon and a risk adjustment emanating from factor regressions.

Eckbo and Norli (2005) also examine the frequency of company delistings from the stock exchange due to bankruptcy/liquidation over the five-year period following IPOs. The idea is that low post-issue returns may be driven by a greater exit due to bankruptcy/liquidation compared to the rate for the matched firms. However, they find no evidence that the rate of bankruptcy/liquidations (or delisting due to takeover) differs across issuer and their matches.

### *5.3.3. Average monthly abnormal returns using factor pricing regressions*

In this section, we use empirical asset pricing models to generate portfolio expected returns. An asset pricing model is estimated using monthly returns, with the intercept term in the pricing model (also referred to as "Jensen's alpha" from Jensen (1968), or simply $\alpha$) as the measure of the average monthly abnormal return. The most commonly used empirical asset pricing models in this literature are of the multi-factor (APT) type in general, and the three-factor model of Fama and French (1993) in particular.[53]

The factor pricing analysis proceeds as follows. Let $r_{pt}$ denote the return on issuer–portfolio $p$ in excess of the risk-free rate, and assume that expected excess returns are generated by a $K$-factor model,

$$E(r_{pt}) = \beta_p' \lambda, \tag{8}$$

where $\beta_p$ is a $K$-vector of risk factor sensitivities (systematic risks) and $\lambda$ is a $K$-vector of expected risk premiums. The return generating process can be written as

$$r_{pt} = E(r_{pt}) + \beta_p' f_t + e_{pt}, \tag{9}$$

where $f_t$ is a $K$-vector of risk factor shocks and $e_{pt}$ is the portfolio's idiosyncratic risk with expectation zero. The factor shocks are deviations of the factor realizations from their expected values, i.e., $f_t \equiv F_t - E(F_t)$, where $F_t$ is a $K$-vector of factor realizations and $E(F_t)$ is a $K$-vector of factor expected returns.

Regression equation (9) requires specification of $E(F_t)$, which is generally unobservable. To get around this issue, it is common to replace the raw factors $F$ with factor mimicking portfolios. Specifically, consider the excess return $r_{kt}$ on a portfolio that has unit factor sensitivity to the $k$th factor and zero sensitivity to the remaining $K - 1$ factors. Since this portfolio must also satisfy equation (8), it follows that $E(r_{kt}) = \lambda_k$. Thus, when substituting a $K$-vector $r_{Ft}$ of the returns on factor-mimicking portfolios for the raw factors $F$, equations (8) and (9) imply the following regression equation in terms of observables:

$$r_{pt} = \beta_p' r_{Ft} + e_{pt}. \tag{10}$$

Equation (10) generates portfolio $p$'s returns, and inserting a constant term $\alpha_p$ yields the alpha measure of abnormal return.

We estimate alphas using two models which include the Fama and French (1993) factors as well as two additional characteristics-based risk factors:

$$r_{pt} = \begin{cases} \alpha_p + \beta_1 \,\mathrm{RM} + \beta_2 \,\mathrm{SMB}_t + \beta_3 \,\mathrm{HML}_t + e_t, \\ \alpha_p' + \beta_1 \,\mathrm{RM} + \beta_2 \,\mathrm{SMB}_t + \beta_3 \,\mathrm{HML}_t + \beta_4 \,\mathrm{UMD} + \beta_5 \,\mathrm{LMH} + e_t, \end{cases} \tag{11}$$

where $r_{pt}$ is the excess return to an equal-weighted portfolio of issuers, RM is the excess return on the CRSP value weighted market index. SMB and HML are the Fama and

[53] See, e.g., Connor and Korajczyk (1995) and Ferson (2003) for extensive surveys of multifactor models.

French (1993) size and book-to-market factors. UMD is a momentum factor inspired by Carhart (1997) and constructed as the returns difference between the one-third highest and the one-third lowest CRSP performers over the past 12 months. LMH is the Eckbo and Norli (2005) turnover factor, defined as a portfolio long in low-turnover stocks and short in high-turnover stocks.

The alpha estimates are reported in Table 19 for equity issuers, and Table 20 for debt issuers. As first reported by Eckbo and Norli (2005), the estimated coefficients on the turnover factor LMH tend to be both a greater and more significant than the coefficients on the momentum factor UMD. When the coefficient on LMH is significant, the extended model increases the regression $R^2$ marginally above the Fama–French model. Moreover, when significant, the estimated coefficients on both UMD and LMH are typically negative, indicating that issuers tend to be relatively liquid, growth stocks.

When using the Fama–French model, the alphas are significant and negative for private placements of equity (panel F of Table 19) and for private placements of straight debt (panels D and F in Table 20). However, the alpha estimates are insignificant in all samples when using the extended model. There is ample evidence that the momentum factor UMD helps explain the cross-section of expected stock returns. Evidence that the turnover factor LMH is also priced is found in Eckbo and Norli (2002 and 2005). Assuming UMD and LMH are indeed priced risk factors, then the results in Table 19 and Table 20 fail to reject the hypothesis of zero post-issue abnormal performance.

Table 21 shows the alpha estimates reported in much of the literature that uses factor regressions to estimate post-issue abnormal performance. For IPOs, and with the exception of non-venture-backed IPOs studied by Brav and Gompers (1997), the alphas are statistically insignificantly different from zero (Brav, Geczy, and Gompers, 2000; Ritter and Welch, 2002; Eckbo and Norli, 2005). For SEOs, and with the exception of Jegadeesh (2000), all large-sample studies (3,000+ SEOs) also report insignificant alphas. These include Brav, Geczy, and Gompers (2000), Eckbo, Masulis, and Norli (2000), and Lyandres, Sun, and Zhang (2005). For portfolios of SEOs, the Fama–French model tend to produce larger (and sometimes significant) alphas than extended models adding UMD, LMH and, most recently, the investment factor of Lyandres, Sun, and Zhang (2005). Overall, assuming these factors are priced, the null of zero abnormal post-SEO performance is not rejected.

Finally, studies of debt issues also find alphas that are indistinguishable from zero. The largest sample is found in Eckbo, Masulis, and Norli (2000), who study a total of 1,329 straight debt issues and 459 convertible debt offerings, report insignificant alpha estimates for both types of debt issues. Spiess and Affleck-Graves (1999) report significantly negative alphas for a constrained sample of debt issuers, where issues by a given company that take place within five years of each other are excluded. However, Butler and Wan (2005) show that adding a liquidity factor (much like the turnover factor of Eckbo and Norli (2005) produces insignificant alpha estimates also for the type of restricted sample used by Spiess and Affleck-Graves (1999). Thus, again assuming

Table 19
Monthly abnormal equal-weighted portfolio return ($\alpha_p$) following IPOs, SEOs, and equity private placements (PPEs), 1980–2000

| $\alpha_p$ | RM | SMB | HML | UMD | LMH | $R^2$ |
|---|---|---|---|---|---|---|
| A. Sample of 5,128 IPOs by industrial firms | | | | | | |
| −0.16 (0.492) | 1.14 (0.000) | 1.17 (0.000) | −0.29 (0.006) | | | 0.838 |
| 0.25 (0.416) | 0.95 (0.000) | 1.03 (0.000) | −0.27 (0.020) | −0.19 (0.061) | −0.53 (0.006) | 0.863 |
| B. Sample of 779 IPOs by banks and financial firms | | | | | | |
| −0.10 (0.695) | 1.09 (0.000) | 0.75 (0.000) | 0.61 (0.000) | | | 0.616 |
| 0.03 (0.922) | 1.06 (0.000) | 0.74 (0.000) | 0.60 (0.000) | −0.10 (0.278) | −0.06 (0.726) | 0.618 |
| C. Sample of 5,127 SEOs by industrial issuers | | | | | | |
| −0.18 (0.167) | 1.20 (0.000) | 0.92 (0.000) | −0.11 (0.057) | | | 0.923 |
| 0.18 (0.125) | 1.04 (0.000) | 0.80 (0.000) | −0.09 (0.073) | −0.17 (0.000) | −0.45 (0.000) | 0.949 |
| D. Sample of 878 SEOs by banks and financial firms | | | | | | |
| −0.16 (0.378) | 1.12 (0.000) | 0.52 (0.000) | 0.77 (0.000) | | | 0.720 |
| −0.09 (0.650) | 1.10 (0.000) | 0.51 (0.000) | 0.77 (0.000) | −0.05 (0.421) | −0.05 (0.650) | 0.720 |
| E. Sample of 693 SEOs by public utilities | | | | | | |
| 0.06 (0.744) | 0.62 (0.000) | 0.05 (0.374) | 0.65 (0.000) | | | 0.458 |
| −0.08 (0.644) | 0.74 (0.000) | 0.15 (0.008) | 0.61 (0.000) | 0.01 (0.829) | 0.34 (0.002) | 0.481 |
| F. Sample of 506 PPEs by industrial issuers | | | | | | |
| −0.48 (0.066) | 1.15 (0.000) | 1.14 (0.000) | −0.37 (0.001) | | | 0.783 |
| −0.04 (0.884) | 1.03 (0.000) | 1.11 (0.000) | −0.40 (0.000) | −0.32 (0.000) | −0.21 (0.178) | 0.811 |

Starting in February 1980, a firm is added to the portfolio in the month following the month of the IPO and held for five years or until delisting (if sooner). The IPO sampling stops in 12/2000 while the abnormal return estimation ends in December 2002. Abnormal returns are estimated using the following asset pricing model:

$$r_{pt} = \alpha_p + \beta_1 \,\mathrm{RM} + \beta_2 \,\mathrm{SMB}_t + \beta_3 \,\mathrm{HML}_t + \beta_4 \,\mathrm{UMD} + \beta_5 \,\mathrm{LMH} + e_t,$$

where $r_{pt}$ is the portfolio excess return, RM is the excess return on the CRSP value weighted market index, SMB and HML are the Fama and French (1993) size and book-to-market factors, UMD is a momentum factor constructed as the returns difference between the one-third highest and the one-third lowest CRSP performers over the past 12 months, and LMH is the Eckbo and Norli (2005) turnover factor (a portfolio long in low-turnover stocks and short in high-turnover stocks). The coefficients are estimated using OLS. Standard errors are computed using the heteroskedasticity consistent estimator of White (1980). The numbers in parentheses are $p$-values. $R^2$ is the adjusted $R$-squared.

liquidity is a priced risk factor, one cannot reject the null of zero abnormal performance following debt offerings by U.S. firms.[54]

[54] Brav et al. (2005) examine institutional lender pricing of (private) loans to equity-issuing firms. They report lower loan yields for equity-issuers relative to non-issuing firms. This is further evidence consistent with the proposition that the relatively low post-issue equity returns reflect lower risk.

Table 20
Monthly abnormal equal-weighted portfolio return ($\alpha_p$) following public (SDOs) and private (PPDs) offerings of straight debt, 1980–2000

| $\alpha_p$ | RM | SMB | HML | UMD | LMH | $R^2$ |
|---|---|---|---|---|---|---|
| A. Sample of 4,546 SDOs by industrial issuers | | | | | | |
| −0.16 (0.116) | 1.12 (0.000) | 0.10 (0.100) | 0.43 (0.000) | | | 0.887 |
| 0.04 (0.674) | 1.05 (0.000) | 0.06 (0.217) | 0.42 (0.000) | −0.13 (0.000) | −0.16 (0.018) | 0.906 |
| B. Sample of 12,191 SDOs by banks and financial firms | | | | | | |
| 0.04 (0.820) | 1.32 (0.000) | −0.05 (0.469) | 0.68 (0.000) | | | 0.798 |
| 0.21 (0.233) | 1.28 (0.000) | −0.07 (0.354) | 0.66 (0.000) | −0.13 (0.009) | −0.08 (0.486) | 0.807 |
| C. Sample of 1,710 SDOs by public utilities | | | | | | |
| −0.03 (0.865) | 0.65 (0.000) | −0.11 (0.093) | 0.70 (0.000) | | | 0.444 |
| −0.18 (0.387) | 0.79 (0.000) | 0.02 (0.796) | 0.64 (0.000) | −0.01 (0.774) | 0.43 (0.001) | 0.472 |
| D. Sample of 4,730 PPDs by industrial issuers | | | | | | |
| −0.29 (0.021) | 1.18 (0.000) | 0.48 (0.000) | 0.43 (0.000) | | | 0.887 |
| 0.06 (0.654) | 1.04 (0.000) | 0.40 (0.000) | 0.42 (0.000) | −0.21 (0.000) | −0.31 (0.000) | 0.931 |
| E. Sample of 3,931 PPDs by banks and financial firms | | | | | | |
| −0.08 (0.691) | 1.44 (0.000) | 0.28 (0.004) | 0.65 (0.000) | | | 0.770 |
| 0.13 (0.543) | 1.32 (0.000) | 0.19 (0.057) | 0.67 (0.000) | −0.07 (0.165) | −0.33 (0.030) | 0.780 |
| F. Sample of 923 PPDs by public utilities | | | | | | |
| −0.29 (0.021) | 1.18 (0.000) | 0.48 (0.000) | 0.43 (0.000) | | | 0.887 |
| −0.24 (0.319) | 0.80 (0.000) | 0.05 (0.529) | 0.66 (0.000) | −0.03 (0.708) | 0.28 (0.052) | 0.444 |

Starting in February 1980, a firm is added to the portfolio in the month following the month of the SDO and held for the minimum of five years and its delisting date. The SDO sampling stops in 12/2000 while the abnormal return estimation ends in December 2002. Abnormal returns are estimated using the following asset pricing model:

$$r_{pt} = \alpha_p + \beta_1 \,\mathrm{RM} + \beta_2 \,\mathrm{SMB}_t + \beta_3 \,\mathrm{HML}_t + \beta_4 \,\mathrm{UMD} + \beta_5 \,\mathrm{LMH} + e_t$$

where $r_{pt}$ is the portfolio excess return, RM is the excess return on the CRSP value weighted market index, SMB and HML are the Fama and French (1993) size and book-to-market factors, UMD is a momentum factor constructed as the returns difference between the one-third highest and the one-third lowest CRSP performers over the past 12 months, and LMH is the Eckbo and Norli (2005) turnover factor (a portfolio long in low-turnover stocks and short in high-turnover stocks). The coefficients are estimated using OLS. Standard errors are computed using the heteroskedasticity consistent estimator of White (1980). The numbers in parentheses are $p$-values. $R^2$ is the adjusted $R$-squared.

## 5.4. Robustness issues

The matched-firm technique discussed above uses firm characteristics (size and B/M) to adjust for priced risks, while the factor regression approach uses a set of prespeci-

Table 21
Average monthly abnormal equal-weighted portfolio return ($\alpha$) for three-to-five year holding periods following securities offerings by U.S. firms

| Study | Issuer type | Sample size | Sample period | Holding period | $\alpha$ |
|---|---|---|---|---|---|
| A. IPOs | | | | | |
| Brav and Gompers (1997) | All | 3, 407 | 1972–1992 | 5 yrs | −0.49%*[a] |
| Brav and Gompers (1997) | All | 934 | 1972–1992 | 5 yrs | 0.09%[b] |
| Brav, Geczy, and Gompers (2000) | All | 3, 501 | 1975–1992 | 5 yrs | −0.19% |
| Ritter and Welch (2002) | All | 6, 249 | 1973–2001 | 3 yrs | −0.21% |
| Eckbo and Norli (2005) | All | 5, 365 | 1972–1998 | 5 yrs | 0.40%[c] |
| Eckbo and Norli (2005) | All | 5, 365 | 1972–1998 | 5 yrs | 0.18%[d] |
| B. SEOs | | | | | |
| Jegadeesh (2000) | All | 2, 992 | 1970–1993 | 5 yrs | −0.31%* |
| Brav, Geczy, and Gompers (2000) | All | 3, 775 | 1975–1992 | 5 yrs | −0.19% |
| Eckbo, Masulis, and Norli (2000) | Ind | 3, 315 | 1964–1995 | 5 yrs | −0.05%[d] |
| Eckbo, Masulis, and Norli (2000) | Ind | 3, 315 | 1964–1995 | 5 yrs | −0.14%[e] |
| Eckbo, Masulis, and Norli (2000) | Util | 880 | 1964–1995 | 5 yrs | −0.13%[d] |
| Bayless and Jay (2003) | Ind | 1, 239 | 1971–1995 | 5 yrs | −0.54%* |
| Krishnamurthy et al. (2005) | All | 1, 477 | 1983–1992 | 3 yrs | −0.36%* |
| Eckbo and Norli (2005) | Ind | 1, 704 | 1964–1995 | 5 yrs | −0.03%[c] |
| Lyandres, Sun, and Zhang (2005) | All | 6, 122 | 1970–2003 | 3 yrs | 0.02%[f] |
| D'Mello, Schlingemann, and Subramaniam (2005) | All | 1, 621 | 1982–1995 | 3 yrs | −0.31%* |
| C. Private placements of equity | | | | | |
| Hertzel et al. (2002) | All | 619 | 1980–1996 | 3 yrs | −1.18%* |
| Krishnamurthy et al. (2005) | All | 276 | 1983–1992 | 3 yrs | −0.77%* |
| D. Straight debt offerings | | | | | |
| Spiess and Affleck-Graves (1999) | All | 392 | 1975–1989 | 5 yrs | −0.29%* |
| Eckbo, Masulis, and Norli (2000) | Ind | 981 | 1964–1995 | 5 yrs | −0.10% |
| Eckbo, Masulis, and Norli (2000) | Util | 348 | 1964–1995 | 5 yrs | −0.22% |
| Butler and Wan (2005) | Ind | 799 | 1975–1999 | 5 yrs | −0.18%[g] |

(*Continued on next page*)

fiedportfolios as proxies for pervasive risks. Either approach suffers from potential "bad model" problems in terms of representing the true asset pricing model. Since tests for abnormal returns are always a joint test of the risk factors assumed to generate expected return, it is therefore useful to provide information on the sensitivity of abnormal return estimates to alternative model specifications. Moreover, factor regressions may suffer from non-stationarity in the estimated parameters that may be predictable using publicly available information. Also, Loughran and Ritter (2000) point out that the factor mimicking portfolios used in the regressions for estimating alphas contain issuing firms, and they argue that this "contamination" may reduce the power of the tests.

Table 21
(*Continued*)

| Study | Issuer type | Sample size | Sample period | Holding period | $\alpha$ |
|---|---|---|---|---|---|
| E. Convertible debt offerings | | | | | |
| Spiess and Affleck-Graves (1999) | All | 400 | 1975–1989 | 5 yrs | −0.31* |
| Eckbo, Masulis, and Norli (2000) | Ind | 459 | 1964–1995 | 5 yrs | −0.31% |

The table reports the time-series estimate of the constant term $\alpha$ resulting from regressing the excess return on a portfolio of issuing firms on a set of pricing factors in an empirical asset pricing model. The issuer portfolio is formed using equal-weights. The issuer's stock typically enters the portfolio in the month following the issue month, and is held from three to five years. Superscript * indicates that the $\alpha$ is statistically significantly different from zero at the 1% level.
[a]Sample of non-venture-backed IPOs.
[b]Sample of venture-backed IPOs.
[c]Pricing model with Fama–French, momentum and liquidity factors.
[d]Pricing model with macroeconomic risk factors.
[e]Pricing model with Fama–French factors.
[f]Pricing model with Fama–French, momentum and investment factors.
[g]Pricing model with Fama–French and liquidity factors.

Eckbo, Masulis, and Norli (2000) examine these robustness issues for their sample of SEOs and debt issues. Below, we discuss their approach, repeat their analysis using our data, and draw qualitative inferences.[55] Overall, this discussion serves to illustrate that the main conclusion of zero long-run abnormal performance following issue-activity is robust.

### *5.4.1. Alternative and omitted risk factors*

*The matched-firm technique*. The matched-firm technique produces evidence of abnormal post-issue stock returns while the factor regression approach does not. This raises the question of whether the characteristics-based matched-firm technique omits priced risk factors. To check this, Eckbo, Masulis, and Norli (2000) and Eckbo and Norli (2005) estimate the abnormal return (alpha) to a zero-investment portfolio that is long in issuer stocks and short in matched firms. This portfolio controls for any omitted risk factor with identical factor betas across issuer and matched firm, effectively combining the two standard matched-firm and asset pricing techniques.

To illustrate, suppose the true set of risk factors is given by the vector $F$, and that only a subset $F_1$ of this vector is included in the regression model, with the complement vector $F_2$ omitted. Let $I$ denote issuer and $M$ matched firm. The "issuer–match"

[55] Detailed results are available upon request.

regression is then

$$r_I - r_M = (\alpha_I - \alpha_M) + (\beta_{1I} - \beta_{1M})F_1 + \epsilon, \tag{12}$$

where $\epsilon = (\beta_{2I} - \beta_{2M})F_2 + u$, where $u$ is a white noise error term. The definition of a "good match" is that $\beta_I$ is close to $\beta_M$. For example, if the size and B/M matching often used in the literature in fact produces a good match, then you expect the "issuer–match" regression to have both a small alpha and values of beta close to zero. Alternatively, if the matching technique fails to control for important risk factors, then the zero-investment "issuer–match" portfolio will contain significant factor loadings.

Eckbo, Masulis, and Norli (2000) (SEOs and debt offerings), Eckbo and Norli (2005) (IPOs), and this survey (all issue categories) all lead to the conclusion that the zero-investment portfolio exhibit significant factor loadings in the extended Fama–French model, but that the alpha of this portfolio is not significantly different from zero. This is consistent with the proposition that the technique of matching on size and B/M is insufficient to control for important risk exposures of the issuing firms. Lyandres, Sun, and Zhang (2005) reach a similar conclusion for their sample of SEOs after performing a three-way sort of size, B/M and investment intensity.

*Alternative factor structures*. Eckbo, Masulis, and Norli (2000) use a model with six prespecified macro factors: the value-weighted CRSP market index, and factor mimicking portfolios for the return spread between Treasury bonds with 20-year and one-year maturity, the return spread between 90-day and 30-day Treasury bills, the seasonally adjusted percent change in real per capita consumption of nondurable goods, the difference in the monthly yield change on BAA-rated and AAA-rated corporate bonds, and unexpected inflation.[56] This six-factor model produce regression $R^2$ similar to the Fama–French model, and the alphas are uniformly indistinguishable from zero.

Eckbo, Masulis, and Norli (2000) also report alpha estimates when the time series of the demeaned, raw macroeconomic factors is used rather than factor-mimicking portfolios. Raw macro factor shocks are interesting in part because they are not affected by stock market mispricing (if any). Also, factor-mimicking portfolios contain measurement error vis-à-vis the true risk factors, which raw factors avoid. On the other hand, there is measurement error induced by the demeaned raw macroeconomic factors themselves. It is difficult to determine a priori which of the two sources of measurement error is most severe (and thus whether factor mimicking is superior).[57] In any event, the alpha estimates remain insignificantly different from zero, though somewhat larger in absolute value than those for regressions based on factor-mimicking portfolios.[58]

[56] These factors also appear in, Ferson and Harvey (1991), Evans (1994), Ferson and Korajczyk (1995), Ferson and Schadt (1996), and Eckbo and Smith (1998) among others.

[57] Factor mimicking portfolios are required when estimating risk premiums (denominated in returns).

[58] Eckbo, Masulis, and Norli (2000) report that a similar conclusion emerges when alpha is estimated using factors extracted from the covariance matrix of returns using the principal components approach of Connor and Korajczyk (1988). Although principal component factors do not have intuitive economic interpretations, they provide yet another factor structure useful for sensitivity analysis.

### 5.4.2. Time-varying factor loadings

Nonstationary factor loadings may produce (i) significant performance in subperiods, (ii) predictable changes in factor loadings which affect the alpha estimates, and (iii) significant effect of using value-weighted instead of equal-weighted issuer portfolios.

*Nonstationarities*. Eckbo, Masulis, and Norli (2000) examine holding periods of between one and five years. For example, with a two-year holding period, firms enter the SEO issuer portfolio as before, but exit after only two years (or at a subsequent security offer or delisting, whichever occurs earlier). This serves to check whether any subperiod abnormal performance is washed out in the averaging of returns over the five-year holding period. The conclusion emerging from the analysis of one-to-five-year holding periods remain the same: none of the alphas are significantly different from zero.

Eckbo, Masulis, and Norli (2000) also reestimate alphas using factor-mimicking portfolios that are continuously updated. That is, the portfolio weights are constructed using a rolling estimation period where the factor loadings are reestimated every month. This rolling estimation procedure relaxes the stationarity assumption on the factor-mimicking weights. The alphas are again all insignificant.

*Predictable changes in factor loadings*. Eckbo, Masulis, and Norli (2000) and Eckbo and Norli (2005) reexamine the null hypothesis of zero abnormal performance using a conditional factor model framework.[59] They follow Ferson and Schadt (1996) and assume that factor loadings are linearly related to a set of $L$ known information variables $Z_{t-1}$:

$$\beta_{1pt-1} = b_{p0} + B_{p1} Z_{t-1}. \tag{13}$$

Here, $b_{p0}$ is a $K$-vector of "average" factor loadings that are time-invariant, $B_{p1}$ is a $(K \times L)$ coefficient matrix, and $Z_{t-1}$ is an $L$-vector of information variables (observables) at time $t-1$. The product $B_{p1} Z_{t-1}$ captures the predictable time variation in the factor loadings. After substituting equation (13) into equation (10), the return-generating process becomes

$$r_{pt} = b'_{p0} r_{Ft} + b'_{p1}(Z_{t-1} \otimes r_{Ft}) + e_{pt}, \tag{14}$$

where the $KL$-vector $b_{p1}$ is $\mathrm{vec}(B_{p1})$ and the symbol $\otimes$ denotes the Kronecker product.[60] This factor model is estimated after adding a constant term $\alpha_p$, which equals zero under the null hypothesis of zero abnormal returns. The information variables in $Z_{t-1}$ include the lagged dividend yield on the CRSP value-weighted market index, the lagged 30-day Treasury bill rate, and the lagged values of the credit and yield curve spreads, BAA–AAA and TBILLspr, respectively. The alpha estimates all remain insignificantly different from zero.

[59] A survey of conditional factor model econometrics is found in Ferson (1995).

[60] The operator $\mathrm{vec}(\cdot)$ vectorizes the matrix argument by stacking each column starting with the first column of the matrix.

*Value-weighted issuer portfolios*. The results reported above are based on equal-weighted issuer portfolios. With value-weights, relatively successful firms gradually increase their portfolio weights. If the relatively low return of issuers is driven by "losers", then value-weighting increases the average portfolio return and possibly the abnormal performance parameter alpha. The literature is fairly unanimous on this issue: alphas with value-weighted issuer portfolios appears less negative than for equal-weighted portfolios, and they sometimes provide evidence of issuer *over*performance relative to matched firms.

### 5.4.3. Issue-purged factors

Loughran and Ritter (2000) argue that it is counterproductive to generate factor-mimicking portfolios without excluding security issuers from the stock universe. Inclusion of security issuers in the factor portfolios results in the factor regressions having the same firm on both sides of the regression (albeit with a small weight in the factor portfolio). They argue that this substantially reduces power to detect abnormal return via the estimated alpha.

Note that, under the null hypothesis of zero abnormal performance, purging the factor-mimicking portfolios for ex post issuing firms biases the tests in favor of finding a significant alpha. This, of course, means that failing to reject the null hypothesis even with purged factor portfolios *a fortiori* supports the market efficiency hypothesis over the market over/underreaction proposition.

Eckbo, Masulis, and Norli (2000) report that, on average, 11.1% of the firms in the factor-mimicking portfolios also make SEOs during the subsequent five years. They purge their factors by eliminating a firm from the factor-mimicking portfolios if the firm issued equity (primary offerings) over the previous five years. Lyandres, Sun, and Zhang (2005) also report results based on purged factors. The main conclusion of both studies is zero abnormal returns when using issuer-purged factor regressions.

## 6. Conclusions and issues for future research

The economics of security offerings has generated considerable empirical research interest over the past two decades. This survey alone identifies more than 280 studies largely restricted to public seasoned security offerings for cash—and we have surely missed some. In addition, there are a large number of related studies discussed in other surveys in this Handbook, including those on IPO underpricing (Ljungqvist, 2007), security swaps associated with corporate takeovers and restructurings (Betton, Eckbo, and Thorburn, 2007; Eckbo and Thorburn, 2007; Hotchkiss et al., 2007), stock compensation to employees (Aggarwal, 2007), private equity (Gompers, 2007), and credit markets (Drucker and Puri, 1989). In all of these settings, the issuer faces both direct and indirect flotation costs that depend on (1) constraints imposed by security regulations, (2) the range of available flotation method choices, (3) underwriter competition, (4) information asymmetries between issuer and outside investors, and (5) the efficiency of market pricing. This survey discusses each of these five determinants of flotation costs. Several findings emerge, as well as new questions for future research, some of which are discussed below.

Public security offerings for cash are vulnerable to conflicts of interests. These conflicts have created rationales for substantial regulatory protections of investors and requirements on issuers. The legal requirements are designed to ensure that investors receive adequate information disclosure and they limit the "aggressive" marketing by the issuer. In general, legal systems, tax codes and securities regulations and the treatment of investors of a country are likely to have a significant bearing on the level of security offering activity. In the U.S., major regulatory milestones include the Securities Act of 1933 (establishing issue registration and disclosure rules), the Securities Exchange Act of 1934 (requiring periodic public disclosures via annual 10-K, quarterly 10-Q and occasional 8-K statements), the move to adopt generally accepted accounting standards (GAAP), the introduction in 1982 of "shelf registration" rules for relatively low-risk issuers (SEC Rule 415), registration exemptions aimed at reducing regulatory costs and improving the liquidity of privately placed securities by privately held companies and foreign issuers in 1990 (SEC Rule 144A), the establishment of Self Regulatory Authorities (NYSE, NASD) who impose various listing requirements and regulate many activities of broker-dealers and underwriters, and most recently, the creation as of December 2005 of a new category of issuers called "well-known seasoned issuers". These issuers are given automatic shelf registration status and may have oral or written communications with investors before during and after the offering process.

Looking internationally, there has been an increase in disclosure regulation and increased regulation and enforcement of insider trading activity. Moreover, parallel to U.S. securities regulation developments, similar national regulatory authorities are developing around the globe. In 1998, the International Organization of Securities Commissions (IOSC)—a global organization of national security regulators-adopted a comprehensive set of objectives and principles of securities regulation, which today are recognized by the world financial community as international benchmarks for all markets.

Additional research is needed to increase our understanding of the impacts of national securities laws, corporation laws and bankruptcy laws for firm issuance decisions. More cross country analyses could help in this regard. Moreover, we need a better understanding of the effects of political processes on these critical legal statutes. How does political corruption influence issuance costs and security issuance choice? How strong are the financial incentives of the dominant economic powers in a nation to limit potential competition through restrictions on capital market development and what are the most effective mechanisms for overcoming these effects? How important are particular reforms that reduce the barriers to global capital market activity in promoting national financial and economic development?

Regulatory changes provide interesting laboratories for examining empirically the exogenous determinants of issue costs and issuers' choice of security and flotation methods. On the one hand, the large increase in the aggregate amount of securities offerings over the past 25 years suggest that the stricter disclosure requirements has had a positive effect on firms' incentives to issue securities. However, additional analysis is needed of the specific effects of the new SEC securities regulations on disclosure requirements, shelf registrations and the creation of "well-known seasoned issuers". Do these regulatory changes have a significant effect on flotation costs, the choice of offering methods, the types of securities issued and timing of offerings? Is there evidence that these new rules lower asymmetric information between issuers and investors? How do the new regulations affect the frequency of foreign security issues in the U.S.?

One of the early regulatory experiments that financial economists studied was SEC Rule 415, known as shelf registration. This regulatory change was designed to lower issue costs. As we show in Table 3, only fifteen percent of the SEOs by U.S. firms employ the shelf registration procedure (half of the debt issuers use shelf registration). SEO shelf offerings tend to be relatively large-but infrequent. The apparent reluctance to take advantage of the relatively low-cost shelf registration procedure is puzzling. It is possible that shelf registration exacerbates adverse selection in issue markets, and is therefore selected only by relatively transparent firms (where the information asymmetry is relatively low). Such self-selection of the issue method suggests that the market reaction to shelf issues should be no lower than the market reaction to traditional non-shelf (underwritten) issues, which is broadly consistent with the reported empirical evidence.

As a general matter, the field would benefit from further analysis of the endogeneity of the choice of security offered and flotation method. The existing literature generally adjusts for endogeneity using predictive models of the issuer's choice of securities and issue method with very modest explanatory power. In estimating such a model, we need to know to what extent are the types of securities issued, their flotation costs and issuance method affected by issuer investment and financing characteristics, asset structure, capital structure, industry identity and the issuer's corporate governance? We also need better predictive models of an issuer's choice of security to sell. Hence, there is a need for further theoretical and empirical research to improve the explanatory power of these predictive models. After which, we need to re-evaluate the robustness of the major results in the prior literature.

Another important regulatory experiment is the 2002 enactment of Sarbanes–Oxley. This landmark legislation has imposed substantial corporate governance constraints and obligations on publicly held companies, preempting state corporation law in a number of areas. A number of the interesting questions are raised by the law. What are the effects of Sarbanes–Oxley on domestic and foreign issuers of securities in the U.S.? How does this law affect auditor independence and the reliability of auditor certification of the financial statements or the market reaction to news of issuer–auditor disagreements? How does this law change the likelihood of earnings restatements and shareholder reactions to new financial statements? What is the importance of board of directors' powers relative to shareholder powers and the potential benefits of giving shareholders stronger voting rights and control rights in determining the security issuance decision and the costs of security issuance? We need a clearer understanding of how security contract characteristics can be altered to better align the interests of different classes of securities and to protect against the extraction of private benefits of controls by managers.

Turning to specific determinants of issue costs, we survey a large body of empirical research on the underwriting function in general and on the determinants of underwriter compensation more specifically. The field continues to only partially understand the effects of asymmetric information between issuers/underwriters and outside investors and use of various institutional mechanisms to limit this effect such as the right to renege on primary offering buy orders, restrictions on short selling by underwriters, restrictions on short selling by investors and lock-up provisions on insiders, the use of overallotment options, the choice of auditor, price stabilization, shareholder suits against issuers and underwriters, the effects of new SEC disclosure regulations and how important are certain accounting rules.

How important is security liquidity to flotation costs and how can this liquidity be improved cheaply? How important is it to have short selling opportunities or an active option market for the stock? Do these opportunities increase security price volatility and does this increase the costs of liquidity? To what extent do various information producers such as financial analysts, bond rating agencies, auditors, market makers/exchanges that report bids, asks and transaction prices, and investment bank fairness opinions reduce heterogeneous expectations among investors and increase securities trading and their liquidity? There is also a need to further investigate of the degree of interdependence of underpricing, underwriting spreads, out of pocket expenses and the probability of offer withdrawal and why these relationships appear to vary qualitatively by type of security, which is somewhat puzzling.

Another important question is how underwriter competition is impacted by the entry of commercial banks and foreign financial institutions. What are the fundamental services offered by underwriters and how do these services enhance share liquidity in the primary and secondary markets and what are the impacts on security prices? Further analysis is needed on the impacts of investment banking competition, and the interrelationship of underwriting services for debt and equity offerings with M&A advisory services.

How does learning take place in security contract innovation in the private equity market (venture capital term sheets), private placement market and public security markets. For example, how have bond covenants, and microfinancing mechanisms evolved? To what extent are innovations triggered by widely covered scandals, which broadcast problems in existing contracting technology? Are there spillover effects in contracting technology across security markets and across countries? What determines the speed of technology transfers?

Empirical research in this area is constrained by the availability and reliability of databases within the reach of university budgets. One important area that is under-studied because of a lack of data is corporate bond issue activity. We know very little about the flotation process for corporate bonds. What are its unique institutional features of the corporate bond offering process? However, new databases will soon be available in this area allowing researchers to investigate many interesting questions. What are the determinants of flotation costs and how is it impacted by bond seniority, collateral, affiliated company guarantees, maturity, sinking funds, call protection, and the instrument's liquidity and interest rate volatility and changes in the issuer's capital structure and financial condition?

We reconfirm the empirical fact—first established by Mikkelson and Partch (1986)—that public seasoned equity issues for cash (that is, SEOs) are rare corporate financing events. Eckbo and Norli (2006) report that for a sample of 6,000+ IPOs from the period 1980–2005, about half of the IPO firms undertake *no* public follow-on offering over the remainder of the sample period (regardless of the security type), and only one-quarter follow on with a SEO. The low issuance activity is relevant for the more general question of firms' capital structure choice, and for a pecking order theory in particular.

Fama and French (2005) show that including employee compensation and equity swaps in mergers and acquisitions in a broader definition of seasoned equity issues leads to the conclusion that the typical firm issues equity every year. They view this high frequency of equity issues as evidence against the Myers (1984) pecking order. However, it is questionable whether the type of information asymmetry assumed in Myers and Majluf (1984)—which motivates Myers (1984)'s pecking order—is relevant for employee stock repurchases and option holdings. Also, equity swaps to finance mergers and acquisitions introduce two-sided information asymmetry, which can under some reasonable conditions place equity at the *top* of a (modified) pecking order. Clearly, additional research on the theoretical placement of equity swaps in a pecking order, as well as on the trade-off between debt and equity issues is required, before we can have confidence in the ability (or lack thereof) of the pecking order to explain the nature of and motivations for firms' issuing behavior.

There is a large empirical literature providing estimates of the market reaction to security issue announcements, both in the U.S. and internationally. This market reaction is interesting in part because it shows a significant equity price dilution affect, even for issuers who hire reputable underwriters to market their shares. This evidence is broadly consistent with primary issue markets being characterized by adverse selection. Research extending the basic intuition provided by Myers and Majluf (1984) adverse selection model has shown that the amount of price dilution also depends on the degree to which the issuer's own shareholders participate in the issue (in a rights offer), the existence of strong investment opportunities as well as on the sequential nature of the issuer's flotation method choice. It is also important to recognize that the Myers and Majluf model assumes strong management alignment of interest with old shareholders, which may or may not be the case. The various equilibria from these adverse selection models predict a negative, zero or positive market reaction to SEOs, which points to the importance of using carefully "controlled experiments" when testing more generalized theories of issuing behavior, e.g., such as the pecking order.

The literature on announcement effects represents such "controlled experiments" and has produced several interesting findings. The typical firm commitment underwritten offering in the U.S. is met with a statistically significant negative market reaction of close to $-2\%$, which represents a dilution in dollar terms equal to approximately 15%

of the proceeds of the typical SEO. If one views this dilution effect as an issue cost (which is arguably the case), then it swamps even relatively high firm commitment underwriting costs. Differential average market reactions across issues and issuer types are also important. The market reaction is less negative for regulated utilities, for smaller issues, for less risky securities such as debt, for issue methods that involve preemptive rights, for shelf offerings, and for private placements (which tend to elicit a positive market reaction).

These empirical regularities are broadly consistent with the predictions of separating equilibria reflecting adverse selection in issue markets. For SEOs internationally, where the equity flotation method typically involves preemptive rights, the empirical evidence is also largely consistent with theories of adverse selection. Samples of foreign issues are interesting both because they allow a study of rights (which have largely disappeared in the U.S.), and because they provide greater variation in institutional and ownership characteristics of issuing firms. We expect future studies of foreign security issues to contribute substantially towards our understanding of the economics of the issuance process.

The survey ends with a review of the empirical literature on post-issue stock returns—so-called "long-run" performance studies—and we complement this literature with our own performance estimates. The key theoretical question in this literature is whether firms are able to exploit their private information at the expense of outside investors. In the vernacular of Loughran and Ritter (1995), are firms able to time their equity issues to temporary "windows of opportunity", when it is possible to sell overpriced equity to new investors? Do investors who purchase and hold the new shares through the subsequent price correction period realize a negative risk-adjusted (abnormal) holding-period stock return?

The literature is in substantial agreement that the average realized two-to-five-year holding period (raw) returns following equity issues is significantly lower than the average return realized by non-issuing firms matched to have similar size and book-to-market value. We show that this result also holds for security issues beyond SEOs and IPOs, such as private equity issues, and issues of straight and convertible debt. The extant evidence that issuers underperform non-issuing matched firms appears convincing. The controversy starts when one interprets this underperformance as a measure of abnormal returns to issuers. In the jargon of asset pricing theory, the difference between the return to issuers and non-issuing matched firms is a measure of abnormal (or unexpected) returns, only if the two types of firms have identical exposures to priced risk factors. A number of studies have shown that the assumption of equal risk exposures is unlikely to hold.

Recent research also indicates that security issuers often exercise large real investment options around the same time. Theory predicts that converting investment options to assets in place should cause risk profiles—and therefore issuers' expected returns—to fall. This has the effect of making their initial "matching firms" too risky in the port-issue period. This mismatch causes the benchmark expected returns of the "matching" firms to be too high and thus, the long term performance of issuers is biased downward.

This discussion points to the futility of using non-issuing firms matched on size and book-to-market ratio to benchmark risk. This may not be surprising when one considers that issuers self-select both the timing and type of security to issue. The similarity in firm size and book-to-market ratio notwithstanding, firms that decide to issue and invest are likely to be in a different economic state and at different points in their life cycle than firms that either do not invest or use internal equity to finance investment.

The empirical asset pricing approach allows a more consistent and plausible way of identifying and correcting for the true risk exposures of issuers. While we lack a unified asset pricing theory with a priori identifiable factors, there is ample evidence that large portfolios that in addition to market risk captures firm characteristics such as equity size, book-to-market ratio, return momentum and (perhaps) liquidity, explain a significant portion of the cross-section of expected stock return. Using these portfolios as risk factors, the difference between the average returns to issuers and non-issuing matched firms become negligible. Thus, the joint hypothesis of the risk model and market efficiency in pricing new securities issues cannot be rejected at conventional levels of confidence. We provide a broad update of this result across several types of new issues, such as public and private placements of equity and different types of debt issues. Overall, this part of the survey leads us to conclude that the long-run performance literature to date fails to provide systematic evidence in favor of behavioral models of either issuer or market behavior.

Research on security offerings continues to advance rapidly. It is currently being strongly influenced by advances in asset pricing theories, market microstructure, optimal capital structure and financing theories, theories of corporate governance, agency and optimal contracting. The development of new databases on security offerings outside the U.S. and of various fixed income and hybrid securities in the U.S. and elsewhere is also stimulating new empirical research on security offerings. At the same time, researchers are incorporating more institutional features regarding laws, regulations, taxes and political considerations into their analyses of the security offering process. The end result is a much richer understanding of the complexities of the security offering process and how much we still need to learn.

## References


Admati, A.R., Pfleiderer, P., 1994. Robust financial contracting and the role of venture capitalists. Journal of Finance 49, 371–402.

Aggarwal, R., 2000. Stabilization activities by underwriters after initial public offerings. Journal of Finance 55, 1075–1103.

Aggarwal, R., 2007. Executive compensation. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 2. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 17, forthcoming.

Agrawal, A., Chen, M., 2004. Analyst conflicts of interest and research quality. Working Paper, University of Alabama.

Altinkilic, O. 2006. Underwriter compensation and market making in SEOs. Working Paper, University of Pittsburg.

Altinkilic, O., Hansen, R.S., 2000. Are there economies of scale in underwriting fees? Evidence of rising external financing costs. Review of Financial Studies 13, 191–218.

Altinkilic, O., Hansen, R.S., 2003. Discounting and underpricing in seasoned equity offers. Journal of Financial Economics 69, 285–324.

Altinkilic, O., Hansen, R.S. 2006. The puzzling wealth loss around equity offerings: Is it merely price pressure or is there more information? Working Paper, University of Pittsburgh.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid–ask spread. Journal of Financial Economics 17, 223–249.

Amihud, Y., Mendelson, H., 1988. Liquidity and asset prices: Financial management implications. Financial Management 17, 5–15.

Ang, A., Gu, L., Hochberg, Y.V., 2005. Is IP underperformance a Peso problem? Working Paper, Columbia University and Cornell University.

Armitage, S., 1998. Seasoned equity offers and rights issues: A review of the evidence. European Journal of Finance 4, 29–59.

Asquith, P., Mullins Jr., D.W., 1986. Equity issues and offering dilution. Journal of Financial Economics 15, 61–89.

Autore, D., Kumar, R., Shome, D., 2004. The revival of shelf-registered corporate equity offerings. Working Paper, Virginia Tech.

Baker, M., Ruback, R.S., Wurgler, J., 2007. Behavioral corporate finance. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 4 (this volume).

Baker, M., Taliaferro, R., Wurgler, J., 2004. Predictive regressions based on managerial decision variables: Is there a small sample bias? Journal of Finance, in press.

Baker, M., Wurgler, J., 2000. The equity share in new issues and aggregate stock returns. Journal of Finance 55, 2219–2257.

Barber, B.M., Lehavy, R., McNichols, M., Trueman, B., 2005. Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations. Journal of Accounting Research, in press.

Barber, B.M., Lyon, J.D., 1997. Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. Journal of Financial Economics 43, 341–372.

Barclay, M.J., Holderness, C.G., Sheehan, D.P., 2005. Private placements and managerial entrenchment. Working Paper, University of Rochester, Boston College, and Penn State University.

Barclay, M.J., Litzenberger, R.H., 1988. Announcement effects of new equity issues and the use of intraday price data. Journal of Financial Economics 21, 71–99.

Baron, D.P., 1982. A model of the demand of investment banking advising and distribution services for new issues. Journal of Finance 37, 955–976.

Bayless, M., Chaplinsky, S., 1991. Expectations of security type and the information content of debt and equity offers. Journal of Financial Intermediation 1, 195–214.

Bayless, M., Chaplinsky, S., 1996. Is there a window of opportunity for seasoned equity issuance? Journal of Finance 51, 253–278.

Bayless, M., Jay, N.R., 2003. Is the performance of firms following seasoned equity issues anomalous? Journal of Banking and Finance 27, 1273–1296.

Beatty, R.P., Ritter, J.R., 1986. Investment banking, reputation, and the underpricing of initial public offerings. Journal of Financial Economics 15, 213–232.

Beatty, R.P., Welch, I., 1996. Issuer expenses and legal liability in initial public offerings. Journal of Law and Economics 39, 545–602.

Benveniste, L.M., Spindt, P.A., 1989. How investment bankers determine the offer price and allocation of new issues. Journal of Financial Economics 24, 343–362.

Benveniste, L.M., Wilhelm, W., 1990. A comparative analysis of IPO proceeds under alternative regulatory environments. Journal of Financial Economics 28, 173–207.

Bethel, J.E., Krigman, L., 2004. Unallocated shelf registration: Why doesn't everybody use it? Equity issuance, managerial choice, and market reaction. Working Paper, Babson College.

Betton, S., Eckbo, B.E., Thorburn, K.S., 2007. Corporate takeovers. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 2. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 15, forthcoming.

Bhagat, S., 1983. The effect of pre-emptive right amendments on shareholder wealth. Journal of Financial Economics 12, 287–310.

Bhagat, S., Marr, M.W., Thompson, G.R., 1985. The rule 415 experiment: Equity markets. Journal of Finance 40, 1385–1401.

Bigelli, M., 1998. The quasi-split effect, active insiders and the Italian market reaction to equity rights issues. European Financial Management 32, 185–206.

Blackwell, D.W., Marr, M.W., Spivey, M.F., 1990. Shelf registration and the reduce due diligence argument: Implications of the underwriter certification and the implicit insurance hypothesis. Journal of Financial and Quantitative Analysis 25, 245–259.

Boehmer, E., Fishe, R.P., 2003. Who receives IPO allocations? An analysis of 'regular investors'. Working Paper, University of Miami.

Bøhren, Ø., Eckbo, B.E., Michalsen, D., 1997. Why underwrite rights offerings? Some new evidence. Journal of Financial Economics 46, 223–261.

Booth, J.R., Chua, L., 1996. Ownership dispersion, costly information, and IPO underpricing. Journal of Financial and Economics 41, 291–310.

Booth, J.R., Smith II, R.L., 1986. Capital raising, underwriting, and the certification hypothesis. Journal of Financial Economics 15, 261–281.

Brau, J.C., Fawcett, S.E., 2006. Initial public offerings: An analysis of theory and practice. Journal of Finance 61, 399–436.

Brau, J.C., Lambson, V.E., McQueen, G., 2005. Lockups revisited. Journal of Financial and Quantitative Analysis 40, 519–530.

Brav, A., Geczy, C., Gompers, P.A., 2000. Is the abnormal return following equity issuances anomalous? Journal of Financial Economics 56, 209–249.

Brav, A., Gompers, P.A., 1997. Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and non-venture capital-backed companies. Journal of Finance 52, 1791–1821.

Brav, A., Gompers, P.A., 2003. The role of lock-ups in initial public offerings. Review of Financial Studies 16, 1–29.

Brav, A., Michaely, R., Roberts, M., Zarutskie, R., 2005. Evidence on the tradeoff between risk and return for IPO and SEO firms. Working Paper, Duke University, Cornell University, and University of Pennsylvania.

Brennan, M.J., Schwartz, E.S., 1982. The case for convertibles. Chase Financial Quarterly 1, 27–46.

Brick, I.E., Palmon, O., Patro, D.K., 2004. Stock price response to calls of convertible bonds: Still a puzzle. Working Paper, University of Newark.

Brophy, D.J., Sialm, C., Ouimet, P.P., 2005. Pipe dreams? The performance of companies issuing equity privately. Working Paper, University of Michigan.

Brous, P.A., Datar, V., Kini, O., 2001. Is the market optimistic about the future earnings of seasoned equity offering firms? Journal of Financial and Quantitative Analysis 36, 141–168.

Brown, S.J., Warner, J.B., 1985. Use of daily stock returns: The case of event studies. Journal of Financial Economics 14, 3–31.

Bruner, R., Chaplinsky, S.J., Ramchand, L., 2004. US-Bound IPOs: Issue costs and selective entry. Financial Management 33, 39–60.

Buhner, T., Kaserer, C., 2002. The structure of external financing costs and the economies of scale view: New evidence from seasoned equity offerings in Germany. European Financial Management 8, 315–338.

Burch, T., Nanda, V., Warther, V., 2005. Does it pay to be loyal? An empirical analysis of underwriting relationships and fees. Journal of Financial Economics 77, 673–699.

Burton, B.M., Lonie, A.A., Power, D.M., 1999. Does the issue method influence the market reaction to seasoned equity offer announcements? Applied Economic Letters 6, 459–462.

Butler, A.W., Grullon, G., Weston, J.P., 2005a. Can managers forecast aggregate market returns? Journal of Finance 60, 965–988.

Butler, A.W., Grullon, G., Weston, J.P., 2005b. Stock market liquidity and the cost of issuing equity. Journal of Financial and Quantitative Analysis 40, 331–348.

Butler, A.W., Wan, H., 2005. Stock market liquidity and the long-run stock performance of debt issuers. Working Paper, University of Texas at Dallas and University of South Florida.

Byoun, S., 2004. Stock performance following seasoned stock-warrant unit offerings. Journal of Business 77, 75–100.

Byoun, S., Moore, W.T., 2003. Shares vs warrant-share units: Evidence from seasoned equity offerings. Journal of Corporate Finance 9, 575–590.

Cai, N., Ramchand, L., Warga, A., 2004. The pricing of equity IPOs that follow public debt offerings. Financial Management 33, 5–26.

Carhart, M.M., 1997. On the persistence in mutual fund performance. Journal of Finance 52, 57–82.

Carlson, M., Fisher, A., Giammarino, R., 2005. Corporate investment and asset price dynamics: Implications for the cross-section of returns. Journal of Finance 59, 2577–2603.

Carlson, M., Fisher, A., Giammarino, R., 2006. Corporate investment and asset price dynamics: Implications for SEO event studies and long-run performance. Journal of Finance 61, 1009–1034.

Carter, R.B., Dark, F.H., Singh, A.K., 1998. Underwriter reputation, initial returns, and the long-run performance of IPO stocks. Journal of Finance 53, 253–311.

Chaplinsky, S.J., Erwin, G.R., 2005. Great expectations: Banks as equity underwriters. Working Paper, University of Virginia.

Chaplinsky, S.J., Hansen, R.S., 1993. Partial anticipation, the flow of information and the economic impact of corporate debt sales. Review of Financial Studies 6, 709–732.

Chaplinsky, S.J., Haushalter, D., 2003. Financing under extreme uncertainty: Evidence from pipes. Working Paper, University of Virginia.

Chaplinsky, S.J., Ramchand, L., 2000. The impact of global equity offerings. Journal of Finance 55, 2767–2790.

Chaplinsky, S.J., Ramchand, L., 2004. The impact of SEC rule 144a on corporate debt issuance. Journal of Business 77, 1073–1097.

Chemmanur, T.J., Fulghieri, P., 1997. Why include warrants in new equity issues? A theory of unit IPOs. Journal of Financial and Quantitative Analysis 32, 1–24.

Chen, H.-C., Ritter, J.R., 2000. The seven percent solution. Journal of Finance 55, 1105–1131.

Choe, H., Masulis, R.W., Nanda, V., 1993. Common stock offerings across the business cycle: Theory and evidence. Journal of Empirical Finance 1, 3–31.

Clarke, J., Dunbar, C., Kahle, K.M., 2001. Long-run performance and insider trading in completed and cancelled seasoned equity offerings. Journal of Financial and Quantitative Analysis 36, 415–430.

Cochrane, J.H., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. Journal of Finance 46, 209–237.

Cochrane, J.H., 2005. Financial markets and the real economy. Working Paper, University of Chicago.

Connor, G., Korajczyk, R.A., 1988. Risk and return in an equilibrium apt: Application of a new methodology. Journal of Financial Economics 21, 255–289.

Connor, G., Korajczyk, R.A., 1995. The arbitrage pricing theory and multifactor models of asset returns. In: Jarrow, R.A., Maksimovic, V., Ziemba, W.T. (Eds.), Handbooks in Operation Research and Management Science Series. North-Holland, Amsterdam, Chapter 4, pp. 87–144.

Cooney, J.W., Kalay, A., 1993. Positive information from equity issue announcement. Journal of Financial Economics 33, 149–172.

Cooney, J.W., Kato, K., Schallheim, J.S., 2003. Underwriter certification and Japanese seasoned equity issues. Review of Financial Studies 16, 949–982.

Cornelli, F., Goldreich, D., 2001. Bookbuilding and strategic allocation. Journal of Finance 56, 2337–2369.

Cornelli, F., Goldreich, D., 2003. Bookbuilding: How informative is the order book? Journal of Finance 58, 1415–1443.

Cornett, M.M., Davidson III, W.N., Rangan, N., 1996. Deregulation in investment banking: Industry concentration following rule 415. Journal of Banking and Finance 20, 85–113.

Cornett, M.M., Mehran, H., Tehranian, H., 1998. Are financial markets overly optimistic about the prospects of firms that issues equity? Evidence from voluntary versus involuntary equity issuance by banks. Journal of Finance 53, 2139–2159.
Corwin, S., 2003. The determinants of underpricing for seasoned equity offers. Journal of Finance 58, 2249–2279.
Corwin, S., Harris, J.H., Lipson, M.L., 2004. The development of secondary market liquidity for NYSE-listed IPOs. Journal of Finance 59, 2339–2374.
Corwin, S., Schultz, P., 2005. The role of IPO syndicates: Pricing, information production and underwriter competition. Journal of Finance 60, 443–486.
Cotter, J.F., Chen, A., Kao, L., 2004. Underwriter price stabilization of seasoned equity offerings: The evidence from transactions data. Academia Economic Papers 32, 53–81.
Cronqvist, H., Nilsson, M., 2005. The choice between rights offerings and private equity placements. Journal of Financial Economics 78, 375–407.
Dahlquist, M., de Jong, F., 2004. Pseudo market timing: Fact or fiction? Working Paper, Swedish Institute for Financial Research.
Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53, 1839–1885.
Dann, L.Y., Mikkelson, W.H., 1984. Convertible debt issuance, capital structure change and financing-related information: Some new evidence. Journal of Financial Economics 13, 157–186.
Dasgupta, S., Hansen, R.G., 2007. Auctions in corporate finance. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 3 (this volume).
Datta, S., Iskandar-Datta, M., Patel, A., 1997. The pricing of initial public offers of corporate straight debt. Journal of Finance 52, 379–396.
Datta, S., Iskandar-Datta, M., Raman, K., 2005. Executive compensation structure and corporate equity financing decisions. Journal of Business 78, 1859–1889.
Dechow, P.M., Dichev, I.D., 2002. The quality of accruals and earnings: The role of accrual estimation errors. Accounting Review 77, 35–59.
Dehnert, J., 1991. The determinants of the size of equity issue announcement effects. Unpublished Dissertation Proposal, Australian School of Management, University of New South Wales.
Denis, D.J., 1991. Shelf registration and the market for seasoned equity offerings. Journal of Business 64, 189–212.
Denis, D.J., 1993. The cost of equity issues since rule 415: A closer look. Journal of Financial Research 16, 77–78.
Denis, D.J., 1994. Investment opportunities and the market reaction to equity offerings. Journal of Financial and Quantitative Analysis 29, 159–177.
Dierkens, N., 1991. Information asymmetry and equity issues. Journal of Financial and Quantitative Analysis 26, 181–198.
D'Mello, R., Schlingemann, F., Subramaniam, V., 2005. Are there monitoring benefits to institutional ownership? Evidence from seasoned equity offerings. Working Paper, Wayne State University, University of Pittsburgh, and Tulane University.
Doidge, C., Karolyi, G.A., Stulz, R., 2004. Why are foreign firms listed in the U.S. worth more? Journal of Financial Economics 71, 205–238.
Drucker, S., Puri, M., 1989. On the benefits of concurrent lending and underwriting. Journal of Finance 60, 2762–2799.
Drucker, S., Puri, M., 2007. Banks in capital markets. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 5 (this volume).
Dunbar, C.G., 1995. The use of warrants as underwriter compensation in initial public offerings. Journal of Financial Economics 38, 59–78.
Dunbar, C.G., 2000. Factors affecting investment bank initial public offering market share. Journal of Financial Economics 55 (1), 3–41.

Dybvig, P.H., Zender, J.F., 1991. Capital structure and dividend irrelevance with asymmetric information. Review of Financial Studies 4, 201–219.

Eckbo, B.E., 1986. Valuation effects of corporate debt offerings. Journal of Financial Economics 15, 119–151.

Eckbo, B.E., Giammarino, R.M., Heinkel, R.L., 1990. Asymmetric information and the medium of exchange in takeovers: Theory and tests. Review of Financial Studies 3, 651–675.

Eckbo, B.E., Maksimovic, V., Williams, J., 1990. Consistent estimation of cross-sectional models in event studies. Review of Financial Studies 3, 343–365.

Eckbo, B.E., Masulis, R.W., 1992. Adverse selection and the rights offer paradox. Journal of Financial Economics 32, 293–322.

Eckbo, B.E., Masulis, R.W., 1995. Seasoned equity offerings: A survey. In: Jarrow, R.A., Maksimovic, V., Ziemba, W.T. (Eds.), Finance. In: Handbooks in Operations Research and Management Science Series. Elsevier/North-Holland, Amsterdam, Chapter 31, pp. 1017–1072.

Eckbo, B.E., Masulis, R.W., Norli, Ø., 2000. Seasoned public offerings: Resolution of the 'new issues puzzle'. Journal of Financial Economics 56, 251–291.

Eckbo, B.E., Norli, Ø., 2002. Pervasive liquidity risk. Working Paper, Tuck School of Business at Dartmouth.

Eckbo, B.E., Norli, Ø., 2004. The choice of SEO selling mechanism: Theory and new evidence. Working Paper, Tuck School of Business at Dartmouth.

Eckbo, B.E., Norli, Ø., 2005. Liquidity risk, leverage and long-run IPO returns. Journal of Corporate Finance 11, 1–35.

Eckbo, B.E., Norli, Ø., 2006. Corporate investment, risk dynamics, and the timing of post-IPO security offerings. Working Paper, Tuck School at Dartmouth and the Norwegian School of Management.

Eckbo, B.E., Smith, D.C., 1998. The conditional performance of insider trades. Journal of Finance 53, 467–498.

Eckbo, B.E., Thorburn, K.S., 2007. Financial restructurings. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 2. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 16, forthcoming.

Edelen, R.M., Kadlec, G.B., 2005. Issuer surplus and the partial adjustment of IPO prices to public information. Journal of Financial Economics 77, 347–373.

Ellis, K., Michaely, R., O'Hara, M., 2000. When the underwriter is the market maker: An examination of trading in the IPO after-market. Journal of Finance 55, 1039–1074.

Ellis, K., Michaely, R., O'Hara, M., 2004. Competition in investment banking: Proactive, reactive, or retaliatory? Working Paper, University of California.

Errunza, V.R., Miller, D.P., 2003. Valuation effects of seasoned global equity offerings. Journal of Banking and Finance 27, 1611–1623.

Evans, M.D., 1994. Expected returns, time-varying risk, and risk premia. Journal of Finance 49, 655–679.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 43, 3–56.

Fama, E.F., French, K.R., 2004. New lists: Fundamentals and survival rates. Journal of Financial Economics 73, 229–269.

Fama, E.F., French, K.R., 2005. Financing decisions: Who issues stock? Journal of Financial Economics 76, 549–582.

Fernando, C.S., Gatchev, V.A., Spindt, P.A., 2005. Wanna dance? How firms and underwriters choose each other. Journal of Finance 60, 1713–1757.

Ferson, W.E., 1995. Theory and empirical testing of asset pricing models. In: Jarrow, R.A., Maksimovic, V., Ziemba, W.T. (Eds.), Finance, Handbooks in Operation Research and Management Science. Elsevier/North-Holland, Amsterdam, Chapter 31, pp. 145–200.

Ferson, W.E., 2003. Tests of multifactor pricing models, volatility bounds and portfolio performance. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance: Financial Markets and Asset Pricing. In: Handbooks in Economics Series. Elsevier/North-Holland, Amsterdam, Chapter 12, pp. 743–802.

Ferson, W.E., Harvey, C.R., 1991. The variation of economic risk premiums. Journal of Political Economy 99, 385–415.
Ferson, W.E., Korajczyk, R.A., 1995. Do arbitrage pricing models explain the predictability of stock returns? Journal of Business 68, 309–349.
Ferson, W.E., Schadt, R.W., 1996. Measuring fund strategy and performance in changing economic conditions. Journal of Finance 51, 425–461.
Field, L.C., Cao, C., Hanka, G., 2004. Liquidity consequences of lockup expirations. Journal of Financial and Quantitative Analysis 39, 25–46.
Field, L.C., Hanka, G., 2001. The expiration of IPO share lockups. Journal of Finance 56, 471–500.
Foerster, S.R., Karolyi, G.A., 2000. The long-run performance of global equity offerings. Journal of Financial and Quantitative Analysis 35 4, 499–523.
Frank, M.Z., Goyal, V.K., 2003. Testing the pecking order theory of capital structure. Journal of Financial Economics 67, 217–248.
Frank, M.Z., Goyal, V.K., 2007. Trade-off and pecking order theories of debt. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 2. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 12, forthcoming.
Gajewski, J., Ginglinger, E., 2002. Seasoned equity issues in a closely held market: Evidence from France. European Finance Review 6, 291–319.
Galai, D., Masulis, R.W., 1976. The option pricing model and the risk factor of stock. Journal of Financial Economics 3, 53–81.
Gande, A., Puri, M., Saunders, A., 1999. Bank entry, competition, and the market for corporate securities underwriting. Journal of Financial Economics 54, 165–185.
Gande, A., Puri, M., Saunders, A., Walter, I., 1997. Bank underwriting of debt securities: Modern evidence. Review of Financial Studies 10, 1175–1202.
Giammarino, R.M., Lewis, T., 1988. A theory of negotiated equity financing. Review of Financial Studies 1, 265–288.
Gomes, A., Phillips, G., 2005. Why do public firms issue private and public securities? Working Paper, University of Pennsylvania and University of Maryland.
Gompers, P., 2007. Venture capital. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance. In: Handbooks in Finance Series, Vol. 1. Elsevier/North-Holland, Amsterdam, Chapter 9 (this volume).
Gompers, P., Lerner, J., 1999. Conflict of interest in the issuance of public securities: Evidence from venture capital. Journal of Law and Economics 42, 1–28.
Graham, J.R., Harvey, C.R., 2001. The theory and practice of corporate finance: Evidence from the field. Journal of Financial Economics 60, 187–243.
Habib, M., Ljungqvist, A., 2001. Underpricing and entrepreneurial wealth losses in IPOs: Theory and evidence. Review of Financial Studies 14, 433–458.
Hanley, K.W., 1993. The underpricing of initial public offerings and the partial adjustment phenomenon. Journal of Financial Economics 34, 231–250.
Hansen, R.S., 1986. Evaluating the costs of a new equity issue. Midland Corporate Finance Journal, 42–52. Spring.
Hansen, R.S., 1988. The demise of the rights issue. Review of Financial Studies 1, 289–309.
Hansen, R.S., 2001. Do investment banks compete in IPOs? The advent for the 7% plus contract. Journal of Financial Economics 59, 313–346.
Hansen, R.S., Crutchley, C., 1990. Corporate earnings and financings: An empirical analysis. Journal of Business 63, 347–371.
Hansen, R.S., Fuller, B.R., Janjigian, V., 1987. The over allotment option and equity financing flotation costs: An empirical investigation. Financial Management 16, 24–32.
Hansen, R.S., Pinkerton, J.M., 1982. Direct equity financing: A resolution of a paradox. Journal of Finance 37, 651–665.
Hansen, R.S., Torregrosa, P., 1992. Underwriter compensation and corporate monitoring. Journal of Finance 47, 1537–1555.

Harris, M., Raviv, A., 1991. The theory of capital structure. Journal of Finance 46, 297–355.
Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.
Heinkel, R., Schwartz, E.S., 1986. Rights versus underwritten offerings: An asymmetric information approach. Journal of Finance 41, 1–18.
Heron, R.A., Lie, E., 2004. A comparison of the motivations for and the information content of different types of equity offerings. Journal of Business 77, 605–632.
Hertzel, M., Lemmon, M., Linck, J., Rees, L., 2002. Long-run performance following private placements of equity. Journal of Finance 57, 2595–2617.
Herzel, M., Smith, R.L., 1993. Market discounts and shareholder gains for placing equity privately. Journal of Finance 48, 459–485.
Hietala, P., Loyttyniemi, T., 1991. An implicit dividend increase in rights issues: Theory and evidence. Working Paper, INSEAD.
Higgins, E.J., Howton, S.D., Howton, S.W., 2003. An analysis of the closed–end fund seasoned equity offerings. Journal of Financial Research 26, 243–257.
Hillion, P., Vermalen, T., 2004. Death spiral convertibles. Journal of Financial Economics 71 2, 381–415.
Hotchkiss, E., John, K., Mooradian, R., Thorburn, K.S., 2007. Bankruptcy and financial distress. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 2. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 14, forthcoming.
Hovakimian, A., Hovakimian, G., Tehranian, H., 2004. Determinants of target capital structure: The case of dual debt and equity issues. Journal of Financial Economics 71, 517–540.
Howe, J., Shilling, J., 1988. Capital structure theory and REIT security offerings. Journal of Finance 43, 983–993.
Howton, S.D., Howton, S.W., Perfect, S.B., 1998. The market reaction to straight debt issues: The effects of free cash flow. Journal of Financial Research 21, 219–228.
Jagannathan, R., Sherman, A., 2006. Why do IPO auctions fail? Working Paper, Northwestern University.
Jegadeesh, N., 2000. Long-run performance of seasoned equity offerings: Benchmark errors and biases in expectations. Financial Management 29, 5–30.
Jegadeesh, N., Weinstein, M., Welch, I., 1993. An empirical investigation of IPO returns and subsequent equity offerings. Journal of Financial Economics 34, 153–175.
Jenkinson, T., Jones, H., 2004. Bids and allocations in European IPO bookbuilding. Journal of Finance 59, 2309–2338.
Jensen, M.C., 1968. The performance of mutual funds in the period 1945–1964. Journal of Finance 23, 389–416.
Jensen, M.C., Meckling, W., 1976. Theory of the firm: Managerial behavior, agency costs, and capital structure. Journal of Financial Economics 3, 305–360.
Jensen, M.R., Pugh, W.N., 1995. The market reaction to cancelled equity offerings. Financial Review 30, 83–113.
John, K., Williams, J., 1985. Dividends, dilution and taxes: A signaling equilibrium. Journal of Finance 40, 1053–1070.
Johnson, S.A., 1995. Dividend payout and the valuation effects of bond announcements. Journal of Financial and Quantitative Analysis 30, 407–423.
Jung, K., Kim, Y.-C., Stulz, R., 1996. Timing, investment opportunities, managerial discretion, and the security issue decision. Journal of Financial Economics 42, 159–185.
Kadan, O., Madureira, L., Wang, R., Zach, T., 2005. Conflicts of interest and stock recommendations—the effects of the global settlement and related regulations. Working Paper, Washington University.
Kadlec, G.B., Loderer, C.F., Sheehan, D.P., 1994. Issue day effects for common stock offerings: Causes and consequences. Working Paper, Penn State University.
Kahle, K.M., 2000. Insider trading and the long-run performance of new security issues. Journal of Corporate Finance 6, 25–53.
Kalay, A., Shimrat, A., 1987. Firm value and seasoned equity issues: Price pressure, wealth redistribution, or negative information. Journal of Financial Economics 19, 109–126.

Kang, H., 1990. Effects of seasoned equity offerings in Korea on shareholder's wealth. In: Rhee, S., Chang, R. (Eds.), Pacific-Basin Capital Markets Research. North-Holland, Amsterdam.

Kang, J.-K., Kim, Y.-C., Stulz, R., 1999. The underreaction hypothesis and the new issues puzzle: Evidence from Japan. Review of Financial Studies 12, 519–534.

Kang, J.-K., Lee, Y.W., 1996. The pricing of convertible debt offerings. Journal of Financial Economics 41 1, 231–248.

Kang, J.-K., Stulz, R., 1996. How different is Japanese corporate finance? An investigation of the information content of new security issues. Review of Financial Studies 9, 109–139.

Kaplan, S.A., 1965. Piercing the corporate boilerplate: Anti-dilution clauses in convertible securities. University of Chicago Law Review 33, 1–31.

Karolyi, G.A., 2006. The world of cross-listings and cross-listings of the world: Challenging conventional wisdom. Review of Finance 10, 99–152.

Kaserer, C., Kraft, M., 2003. How issue size, risk, and complexity are influencing external financing costs: German IPOs analyzed from an economies of scale perspective. Journal of Business Finance & Accounting 30, 479–512.

Kato, K., Schallheim, J.S., 1993. Private equity financing in Japan and corporate grouping (keiretsu). Pacific-Basin Finance Journal 1, 287–307.

Khorana, A., Wahal, S., Zenner, M., 2002. Agency conflicts in closed-end funds: The case of rights offerings. Journal of Financial and Quantitative Analysis 37, 177–200.

Kim, D., Palia, D., Saunders, A., 2005a. Are initial returns and underwriting spreads in equity issues complements or substitutes? Working Paper, Rutgers University.

Kim, D., Palia, D., Saunders, A., 2005b. The long run behavior of equity underwriting spreads. Working Paper, Rutgers University.

Kim, E.H., Purnanandam, A., 2006. Why does the stock market react negatively to SEO announcements? Working Paper, University of Michigan.

Kim, K.A., Shin, H.-H., 2004. The puzzling increase in the underpricing of seasoned equity offerings. Financial Review 39, 343–365.

Kolasinski, A., Kothari, S.P., 2004. Investment banking and analyst objectivity: Evidence from forecasts and recommendations of analysts affiliated with M&A advisors. Working Paper, MIT.

Korajczyk, R.A., Levy, A., 2003. Capital structure choice, macroeconomic conditions and financial constraints. Journal of Financial Economics 68, 75–109.

Korajczyk, R.A., Lucas, D.J., McDonald, R.L., 1990. Understanding stock price behavior around the time of equity issues. In: Hubbard, R. (Ed.), Asymmetric Information, Corporate Finance and Investment. University of Chicago Press, Chicago.

Korajczyk, R.A., Lucas, D.J., McDonald, R.L., 1992. Equity issues with time-varying asymmetric information. Journal of Financial and Quantitative Analysis 27, 397–417.

Korkeamaki, T.P., Moore, W.T., 2004. Convertible bond design and capital investment: The role of call provisions. Journal of Finance 59, 391–405.

Kothare, M., 1997. The effect of equity issues on ownership structure and stock liquidity: A comparison of rights and public offers. Journal of Financial Economics 43, 131–148.

Kothari, S.P., Warner, J.B., 1997. Measuring long-horizon security price performance. Journal of Financial Economics, 301–339.

Kothari, S.P., Warner, J.B., 2007. Econometrics of event studies. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. Elsevier/North-Holland, Amsterdam, Chapter 1 (this volume).

Krasker, W.S., 1986. Stock price movements in response to stock issues under asymmetric information. Journal of Finance 41, 93–106.

Krigman, L., Shaw, W.H., Womack, K.L., 2001. Why do firms switch underwriters? Journal of Financial Economics 60, 245–284.

Krishnamurthy, S., Spindt, P., Subramaniam, V., Woidtke, T., 2005. Does investor identity matter in equity issues? Evidence from private placements. Journal of Financial Intermediation 14, 210–238.

LaPorta, R., Lopez-de-Silanes, F., Shleifer, A., 1999. Corporate ownership around the world. Journal of Finance 54, 471–518.
Lease, R., Masulis, R.W., Page, J., 1991. An investigation of market microstructure impacts on event study returns. Journal of Finance 46, 1523–1536.
Lee, G., Masulis, R.W., 2006. Seasoned equity offerings: Quality of accounting information and expected flotation costs. Working Paper, Vanderbilt University.
Lee, I., 1997. Do firms knowingly sell overvalued equity? Journal of Finance 52, 1439–1466.
Lee, I., Lochhead, S., Ritter, J.R., Zhao, Q., 1996. The costs of raising capital. Journal of Financial Research 19, 59–74.
Lee, I., Loughran, T., 1998. Performance following convertible bond issuance. Journal of Corporate Finance 4, 185–207.
Lewellen, K., 2006. Risk, reputation and IPO support. Journal of Finance 61, 613–663.
Lewis, C., Rogalski, R., Seward, J.K., 1998. Agency problems, information asymmetrics and convertible debt security design. Journal of Financial Intermediation 7, 32–59.
Lewis, C.M., Rogalski, R.J., Seward, J.K., 2001. The long-run performance of firms that issue convertible debt: An empirical analysis of operating characteristics and analyst forecast. Journal of Corporate Finance 7, 447–474.
Li, K., Prabhala, N.R., 2007. Self-selection models in corporate finance. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 2 (this volume).
Li, X., Masulis, R.W., 2005. Pre-IPO investments by financial intermediaries: Certification or moral hazard? Working Paper, University of Miami and Vanderbilt University.
Li, X., Masulis, R.W., 2006. Do venture investments by financial institutions affect the IPO underwriting process? Working Paper, Vanderbilt University.
Ljungqvist, A., 2007. IPO underpricing. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. In: Handbooks in Finance Series. Elsevier/North-Holland, Amsterdam, Chapter 7 (this volume).
Ljungqvist, A., Jenkinson, T., Wilhelm, W.J., 2003. Global integration in primary equity markets: The role of U.S. banks and U.S. investors. Review of Financial Studies 16, 63–99.
Ljungqvist, A., Marston, F., Starks, L.T., Wei, K.D., Yan, H., 2004. Conflicts of interest in sell–side research and the moderating role of institutional investors. Working Paper, New York University and CEPR.
Ljungqvist, A., Marston, F., Wilhelm, W., 2006. Competing for security underwriting mandates: Banking relationships and analyst recommendations. Journal of Finance 61, 301–340.
Ljungqvist, A., Wilhelm, W.J., 2003. IPO pricing in the dot-com bubble. Journal of Finance 58, 723–752.
Loderer, C.F., Zimmermann, H., 1987. Stock offerings in a different institutional setting: The swiss case. Journal of Banking and Finance 12, 353–377.
Loughran, T., Ritter, J.R., 1995. The new issues puzzle. Journal of Finance 50, 23–51.
Loughran, T., Ritter, J.R., 2000. Uniformly least powerful tests of market efficiency. Journal of Financial Economics 55, 361–389.
Loughran, T., Ritter, J.R., 2002. Why don't issuers get upset about leaving money on the table in IPOs? Review of Financial Studies 15, 413–443.
Loughran, T., Ritter, J.R., Rydqvist, K., 1994. Initial public offerings: International insights. Pacific-Basin Finance Journal 2, 165–199.
Lowry, M., Murphy, K., 2006. Executive stock options and IPO underpricing. Working Paper, Pennsylvania State University.
Lowry, M., Schwert, G.W., 2002. IPO market cycles: Bubbles or sequential learning? Journal of Finance 57, 1171–1198.
Lowry, M., Shu, S., 2002. Litigation risk and IPO underpricing. Journal of Financial Economics 65, 309–335.
Lucas, D.J., McDonald, R.L., 1990. Equity issues and stock price dynamics. Journal of Finance 45, 1019–1043.
Lyandres, E., Sun, L., Zhang, L., 2005. Investment-based underperformance following seasoned equity offerings. Working Paper, Rice University and University of Rochester.

Lyon, J.D., Barber, B.M., Tsai, C.-L., 1999. Improved methods for tests of long-run abnormal stock returns. Journal of Finance 54, 165–201.

Marsh, P., 1979. Equity rights issues and the efficiency of the U.K. stock market. Journal of Finance 34, 839–862.

Marsh, P., 1982. The choice between equity and debt: An empirical study. Journal of Finance 37, 121–144.

Masulis, R., 1987. Changes in ownership structure: Conversions of mutual savings and loans to stock charter. Journal of Financial Economics 18, 29–60.

Masulis, R., Shivakumar, L., 2002. Does market structure affect the immediacy of stock price responses to news? Journal of Financial and Quantitative Analysis 37, 617–648.

Masulis, R.W., Korwar, A.N., 1986. Seasoned equity offerings: An empirical investigation. Journal of Financial Economics 15, 91–118.

Mayers, D., 1998. Why firms issue convertible bonds: The matching of financial and real investment options. Journal of Financial Economics 47, 83–102.

Megginson, W., Netter, J., 2001. From state to market: A survey of empirical studies on privatization. Journal of Economic Literature 39, 321–389.

Megginson, W., Weiss, K., 1991. Venture capital certification in initial public offerings. Journal of Finance 46, 879–903.

Meidan, D., 2004. A re-examination of price pressure around seasoned equity offerings. Working Paper, Northwestern University.

Michaely, R., Womack, K.L., 1999. Conflict of interest and the credibility of underwriter analyst recommendations. Review of Financial Studies 12, 653–686.

Mikkelson, W.H., Partch, M.M., 1986. Valuation effects of security offerings and the issuance process. Journal of Financial Economics 15, 30–60.

Miller, M.H., Rock, K., 1985. Dividend policy under asymmetric information. Journal of Finance 40, 1031–1051.

Mitchell, M.L., Stafford, E., 2000. Managerial decisions and long-term stock price performance. Journal of Business, 287–329.

Mola, S., Loughran, T., 2004. Discounting and clustering in seasoned equity offering prices. Journal of Financial and Quantitative Analysis 39, 1–3.

Moore, G., 1980. Business Cycle, Inflation and Forecasting. NBER Studies in Business Cycles 24. Ballinger Publishing Co., Cambridge, MA.

Moore, N.H., Peterson, D.R., Peterson, P.P., 1986. Shelf registrations and shareholder wealth: A comparison of shelf and traditional equity offerings. Journal of Finance 41, 451–463.

Mullineaux, D., Roten, I., 2005. Equity underwriting spreads at commercial bank holding companies and investment banks. Journal of Financial Services Research 27, 243–258.

Myers, S.C., 1984. The capital structure puzzle. Journal of Finance 39, 575–592.

Myers, S.C., Majluf, N.S., 1984. Corporate financing and investment decisions when firms have information that investors do not have. Journal of Financial Economics 13, 187–221.

Myhal, P., 1990. Some observations on the usual anti-dilution provisions. Canadian Business Law Journal 17, 283–302.

Narayanan, R.P., Rangan, K.P., Rangan, N.K., 2004. The role of syndicate structure in bank underwriting. Journal of Financial Economics 72, 555–580.

Ng, C.K., Smith, R.L., 1996. Determinants of contract choice: The use of warrants to compensate underwriters of seasoned equity issues. Journal of Finance 51, 363–380.

Nyberg, S., 1995. Reciprocal shareholding and takeover deterrence. International Journal of Industrial Organization 13, 355–372.

O'Brien, P., McNichols, M., Lin, H.-W., 2005. Analyst impartiality and investment banking relationships. Journal of Accounting Research 43, 623–650.

Officer, D., Smith, R., 1986. Announcement effects of withdrawn security offerings: Evidence on the wealth redistribution hypothesis. Journal of Financial Research 9, 229–238.

Pastor, L., Veronesi, P., 2005. Rational IPO waves. Journal of Finance 60, 1713–1757.

Pilotte, E., 1992. Growth opportunities and the stock price response to new financing. Journal of Business 65, 371–394.

Puri, M., 1994. The long term default performance of bank underwritten security issues. Journal of Banking and Finance 18, 397–418.

Puri, M., 1996. Commercial banks in investment banking: Conflict of interest or certification role? Journal of Financial Economics 40, 373–401.

Puri, M., 1999. Commercial banks as underwriters: Implications for the going public process. Journal of Financial Economics 54, 133–163.

Ritter, J.R., 1991. The long-run performance of initial public offerings. Journal of Finance 47, 3–27.

Ritter, J.R., 2003. Investment banking and security issuance. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance: Corporate Finance. In: Handbooks in Economics Series. Elsevier/North-Holland, Amsterdam, Chapter 5, pp. 254–304.

Ritter, J.R., Welch, I., 2002. A review of IPO activity, pricing, and allocations. Journal of Finance 57, 1795–1828.

Ross, S., 1977. The determination of financial structure: The incentive signalling approach. Bell Journal of Economics 8, 23–40.

Safieddine, A., Wilhelm, W., 1996. An empirical investigation of short-selling activity prior to seasoned equity offerings. Journal of Finance 51, 729–749.

Schenone, C., 2004. The effect of banking relations on the firm's IPO underpricing. Journal of Finance 59, 2903–2958.

Schultz, P., 1993. Unit initial public offerings: A form of staged financing. Journal of Financial Economics 34, 199–229.

Schultz, P., 2003. Pseudo market timing and the long-run underperformance of IPOs. Journal of Finance 58, 483–517.

Schultz, P., 2004. Pseudo market timing and the stationarity of the event-generating process. Working Paper, University of Notre Dame.

Schwert, G.W., 1989. Why does stock market volatility change over time. Journal of Finance 44, 1115–1153.

Sherman, A.E., 1999. Underwriter certification and the effect of shelf registration on due diligence. Financial Management 28, 5–19.

Sherman, A.E., Titman, S., 2002. Building the IPO order book: Underpricing and participation limits with costly information. Journal of Financial Economics 65, 3–29.

Shivakumar, L., 2000. Do firms mislead investors by overstating earnings before seasoned equity offerings? Journal of Accounting and Economics 29, 339–371.

Shultz, P., 2003. Pseudo market timing and the long-run underperformance of IPOs. Journal of Finance 58, 483–507.

Shyam-Sunder, L., 1991. The stock price effect of risky versus safe debt. Journal of Financial and Quantitative Analysis 26, 549–558.

Shyam-Sunder, L., Myers, S.C., 1999. Testing static trade-off against pecking order models of capital structure. Journal of Financial Economics 51, 219–244.

Singh, A.K., 1997. Layoffs and underwritten rights offers. Journal of Financial Economics 43, 105–130.

Slovin, M.B., Sushka, M.E., Bendeck, Y.M., 1994. Seasoned common stock issuance following an IPO. Journal of Banking and Finance 18, 207–226.

Slovin, M.B., Sushka, M.E., Lai, W.L., 2000. Alternative flotation methods, adverse selection, and ownership structure: Evidence from seasoned equity issuance in the U.K. Journal of Financial Economics 57, 157–190.

Smith Jr., C.W., 1977. Alternative methods for raising capital: Rights versus underwritten offerings. Journal of Financial Economics 5, 273–307.

Smith Jr., C.W., 1986. Investment banking and the capital acquisition process. Journal of Financial Economics 15, 3–29.

Spiess, D.K., Affleck-Graves, J., 1995. Underperformance in long-run stock returns following seasoned equity offerings. Journal of Financial Economics 38, 243–267.

Spiess, D.K., Affleck-Graves, J., 1999. The long-run performance of stock returns following debt offerings. Journal of Financial Economics 54, 45–73.
Stambaugh, R.F., 1986. Bias in regressions with lagged stochastic regressors. Working Paper, University of Chicago.
Stambaugh, R.F., 1999. Predictive regressions. Journal of Financial Economics 54, 375–421.
Stein, J.C., 1992. Convertible bonds as backdoor equity financing. Journal of Financial Economics 32, 3–21.
Stevenson, H.W., 1957. Common stock financing. Michigan Business Report No. 29, Bureau of Business Research, School of Business Administration, University of Michigan.
Stigler, G., 1964. Public regulation of the securities markets. Journal of Business 37, 117–142.
Stulz, R.M., 1990. Managerial discretion and optimal financing policies. Journal of Financial Economics 26, 3–27.
Taggart, R.A., 1977. A model of corporate financing decisions. Journal of Finance 32, 1467–1485.
Teoh, S.H., Welch, I., Wong, T.J., 1998. Earnings management and underperformance of seasoned equity offerings. Journal of Financial Economics 49, 63–99.
Thompson, R., 1985. Conditioning the return-generating process on firm-specific events: A discussion of event study methods. Journal of Financial and Quantitative Analysis 20, 151–168.
Thompson, R., 1995. Empirical methods in event studies in corporate finance. In: Jarrow, R.A., Maksimovic, V., Ziemba, W.T. (Eds.), Finance, Handbooks in Operation Research and Management Science, vol. 9. North-Holland, Amsterdam, Chapter 29, pp. 963–992.
Tinic, S.M., 1988. Anatomy of initial public offerings of common stock. Journal of Finance 43, 789–822.
Titman, S., Trueman, B., 1986. Information quality and the valuation of new issues. Journal of Accounting and Economics 8, 159–172.
Tripathy, N., Rao, R.P., 1992. Adverse selection, spread behavior, and over-the-counter seasoned equity offerings. Journal of Financial Research 15, 39–56.
Tsangarakis, N.V., 1996. Shareholder wealth effects of equity issues in emerging markets: Evidence from rights offerings in Greece. Financial Management 25, 21–32.
Ursel, N., 2006. Rights offerings and corporate financial condition. Financial Management, 31–52.
Ursel, N., Trepanier, D.J., 2001. Securities regulation reform and the decline of rights offerings. Canadian Journal of Administrative Sciences 18, 77–86.
Viswanathan, S., Wei, B., 2004. Endogenous events and long run returns. Working Paper, Duke University.
White, H., 1980. A heteroskedasticity-consistent covariance matrix and estimator and a direct test for heteroscedasticity. Econometrica 48, 817–838.
Wruck, K.H., 1989. Equity ownership concentration and firm value: Evidence from private equity financings. Journal of Financial Economics 23, 3–28.
Wu, C., Kwok, C.C.Y., 2002. Why do U.S. firms choose global equity offerings? Financial Management 31, 47–65.
Wu, C., Kwok, C.C.Y., 2003. The pricing of global and domestic initial public offerings by U.S. companies. Journal of Banking and Finance 27, 1167–1184.
Wu, X., Wang, Z., 2005. Equity financing in a Myers–Majluf framework with private benefits of control. Journal of Corporate Finance 11, 915–945.
Wu, X., Wang, Z., 2006a. The choice of seasoned equity flotation method under asymmetric information about private benefits of control. Working Paper, City University of Hong Kong.
Wu, X., Wang, Z., 2006b. Seasoned equity issues and ownership concentration in a closely held market. Working Paper, City University of Hong Kong.
Wu, X., Wang, Z., Yao, J., 2005. Understanding the positive announcement effects of private equity placements: New insights from Hong Kong data. Review of Finance 9, 385–414.
Wu, Y.L., 2004. The choice of equity-selling mechanisms. Journal of Financial Economics 74, 93–119.
Yeoman, J., 2001. The optimal spread and offering price for underwritten securities. Journal of Financial Economics 62, 169–198.
Zhang, L., 2005. Anomalies. Working Paper, University of Rochester.
Zwiebel, J., 1995. Corporate conservatism and relative compensation. Journal of Political Economy 103, 1–25.

*Chapter 7*

# IPO UNDERPRICING*

ALEXANDER LJUNGQVIST[1]

*Leonard N. Stern School of Business, New York University, 44 West 4th Street, Suite 9-160, New York, NY 10012, USA*
*e-mail: aljungqv@stern.nyu.edu*

## Contents



* Thanks for helpful comments go to Martijn Cremers, Espen Eckbo, Roger Edelen, David Goldreich, Tim Jenkinson, Ron Masulis, Jay Ritter, Ann Sherman, Seha Tinic, and William J. Wilhelm.

[1] Address for correspondence: Stern School of Business, New York University, Suite 9-160, 44 West Fourth Street, New York, NY 10012-1126. Phone 212-998-0304. Fax 212-995-4220.

*Handbook of Corporate Finance, Volume 1*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01007-5*

## Abstract

When companies go public, the equity they sell in an initial public offering tends to be underpriced, resulting in a substantial price jump on the first day of trading. The underpricing discount in the United States averaged more than 20% during the 1990s, implying that firms left considerable amounts of money on the table. What explains this phenomenon?

This chapter reviews the principal theories that have been proposed to explain IPO underpricing and discusses the empirical evidence. Theories of underpricing can be grouped under four broad headings: asymmetric information, institutional, control, and behavioral. The key parties to an IPO transaction are the issuing firm, the bank underwriting and marketing the deal, and the new investors. Asymmetric information models assume that one of these parties knows more than the others, and that the resulting information frictions give rise to underpricing in equilibrium. Institutional theories focus on three features of the marketplace: litigation, banks' price stabilizing activities once trading starts, and taxes. Control theories argue that underpricing helps shape the shareholder base so as to reduce intervention by outside shareholders once the company is public. Finally, behavioral theories assume the presence of 'irrational' investors who bid up the price of IPO shares beyond true value.

Broadly speaking, the empirical evidence supports the view that information frictions have a first-order effect on underpricing. At the same time, the enormous variation in the extent of underpricing over time raises doubt in some people's mind whether information-based explanations on their own can account for the huge amounts of money left on the table in hot markets, such as the internet bubble of 1998–2000. Arising from this debate, there is continued interest in behavioral explanations, cross-country tests that exploit interesting institutional differences, conflicts of interest within investment banks, and the use of auctions to market and price IPOs.

**Keywords**

initial public offerings, underpricing, investment banks, asymmetric information, behavioral finance

## 1. Introduction

Going public marks an important watershed in the life of a young company. It provides access to public equity capital and so may lower the cost of funding the company's operations and investments. It also provides a venue for trading the company's shares, enabling its existing shareholders to diversify their investments and to crystallize their capital gains from backing the company—an important consideration for venture capitalists. The act of going public itself shines a spotlight on the company, and the attendant publicity may bring indirect benefits, such as attracting a different caliber of manager. At the same time, the company acquires new obligations in the form of transparency and disclosure requirements, and becomes accountable to a larger group of relatively anonymous shareholders who will tend to vote with their feet (by selling the shares) rather than assist the company's decision-makers in the way a venture capitalist might.

Most companies that go public do so via an initial public offering of shares to investors. IPOs have interested financial economists for many decades. Early writers, notably Logue (1973) and Ibbotson (1975), documented that when companies go public, the shares they sell tend to be underpriced, in that the share price jumps substantially on the first day of trading. Since the 1960s, this 'underpricing discount' has averaged around 19% in the United States, suggesting that firms leave considerable amounts of money on the table. Underpricing has tended to fluctuate a great deal, averaging 21% in the 1960s, 12% in the 1970s, 16% in the 1980s, 21% in the 1990s, and 40% in the four years since 2000 (reflecting mostly the tail-end of the late 1990s internet boom).[2] Clearly, underpricing is costly to a firm's owners: shares sold for personal account are sold at too low a price, while the value of shares retained after the IPO is diluted. In dollar terms, IPO firms appear to leave many billions 'on the table' every year in the U.S. IPO market alone.

This remarkable empirical regularity inspired a large theoretical literature in the 1980s and 1990s trying to rationalize why IPOs are underpriced. The resulting theoretical models in turn have been confronted with the data over the past fifteen years or so. This chapter will outline the main theories of IPO underpricing and discuss the empirical evidence.

Theories of underpricing can be grouped under four broad headings: asymmetric information, institutional reasons, control considerations, and behavioral approaches. The best established of these are the asymmetric information based models. The key parties to an IPO transaction are the issuing firm, the bank underwriting and marketing the deal, and investors. Asymmetric information models assume that one of these parties knows more than the others. Baron (1982) assumes that the bank is better informed about demand conditions than the issuer, leading to a principal-agent problem

[2] Underpricing averages are based on data available on Jay Ritter's website (http://bear.cba.ufl.edu/ritter/ipodata.htm).

in which underpricing is used to induce optimal selling effort. Welch (1989) and others assume that the issuer is better informed about its true value, leading to an equilibrium in which higher-valued firms use underpricing as a signal. Rock (1986) assumes that some investors are better informed than others and so can avoid participating in overvalued IPOs. The resulting winner's curse experienced by uninformed investors has to be countered by deliberate underpricing. Finally, Benveniste and Spindt (1989) assume that underpricing compensates better-informed investors for truthfully revealing their information before the issue price is finalized, thus reducing the expected amount of money left on the table.

Institutional theories focus on three features of the marketplace: litigation, banks' price stabilizing activities once trading starts, and taxes. Control theories argue that underpricing helps shape the shareholder base so as to reduce intervention by outside investors once the company is public. Behavioral theories assume either the presence of 'irrational' investors who bid up the price of IPO shares beyond true value, or that issuers suffer from behavioral biases causing them to put insufficient pressure on the underwriting banks to have underpricing reduced.

Broadly speaking, the empirical evidence supports the view that information frictions (including agency conflicts between the issuing company and its investment bank) contribute to IPO underpricing. The evidence regarding institutional theories is more mixed, not least because we still observe underpricing in countries where litigation, price stabilization, and taxes play no role in the IPO market. Control theories are relatively new and the final word is still out on their plausibility. Behavioral approaches, finally, are at present still in their infancy, though what evidence is available is generally consistent both with the presence of overoptimistic investors and with behavioral biases among the decision-makers at IPO firms.

The empirical IPO literature has become increasingly sophisticated, focusing on testing specific hypotheses or entire models, sometimes in a structural econometric fashion, rather than simply describing the phenomenon of underpricing or correlating it with more or less ad hoc variables. The move towards more sophisticated, theory-led tests is a very positive development. As we will see, it has on more than one occasion led to received wisdom being overturned.

In addition to becoming more sophisticated econometrically, the empirical IPO literature has also increasingly recognized the importance and power of the institutional framework within which IPOs are conducted. To provide a benchmark, consider the way the typical IPO is conducted in the U.S. Having chosen an investment bank to lead-manage its IPO, the company first files a registration (or S-1) statement with the Securities and Exchange Commission, containing descriptive and accounting information about the company's history, business model, performance, and so on. The S.E.C. vets the information for misstatements and omissions, a process which takes several weeks. Once the S.E.C. declares the offer 'effective', the investment bank introduces the company to institutional investors on a so called 'road show'. The managers pitch the company's investment case, and the investors provide feedback in the form of more or less explicit, but always non-binding, indications of interest. On the basis of these indications of interest, which are recorded in a 'book', and the state of the market, the investment bank proposes an offer price to the company. Once priced, investors are asked to confirm their indications of interest, shares are allocated, and a few hours later, trading begins. This process is known as bookbuilding.

The precise details of the institutional framework potentially have a bearing on the efficiency of the capital-raising process. For instance, regulatory constraints imposed on the bank conducting the deal concerning the pricing or allocation of IPO shares can influence the extent of underpricing, as can the way pricing-relevant information is gathered, aggregated, and paid for. This recognition has recently sparked another trend: interest in the IPO experience of countries other than the U.S. Despite the fact that IPO practices appear to become more homogeneous around the world (see Ljungqvist, Jenkinson, and Wilhelm, 2003), institutional frameworks differ in ways that allow sharper tests of theoretical predictions. The United Kingdom, for example, is interesting for the fact that integrated (one-stop-shop) securities houses familiar from Wall Street compete with financial intermediaries that specialize in either corporate finance advice or stockbroking, but do not perform both functions. What services the intermediary offers very likely affects the internal conflicts of interest it is subject to. Or take Taiwan. The Taiwan Stock Exchange does not permit bookbuilding and instead operates a discriminatory-price auction system that prices IPOs based on investors' bids, and investors pay what they bid. This would seem a suitable way to price IPOs from a revenue-maximization point of view, except that the market regulator in Taiwan also imposes various constraints on the auction process which typically lead to widespread underpricing.

The empirical IPO literature is now fairly mature—the main stylized facts have been established, and most theories have been subjected to rigorous empirical testing. We know that IPOs are underpriced and that the extent of underpricing, and the number of companies going public, fluctuates over time. Broadly speaking, there is a large body of evidence supporting the view that information frictions (including agency conflicts between the issuing company and its investment bank) have a first-order effect on underpricing. Still, there is continued interest in at least four areas: behavioral approaches to explain why the extent of underpricing varies over time, peaking during the recent 'dot-com bubble'; tests exploiting cross-country differences in institutional frameworks; work shedding light on the allegedly conflicted behavior of investment banks during the stock market boom of the late 1990s; and the potential for using auction mechanisms to price and allocate IPOs.[3]

[3] There is surprisingly little literature on IPO auctions, especially regarding the potential costs and benefits of moving from bookbuilding to auctions for pricing IPOs. Jagannathan and Sherman (2006) surveys the international experience of using IPO auctions in a large number of countries, concluding that auctions have fallen out of favor in the last ten or 15 years. Derrien and Womack (2002) show that in France, where issuers can choose between bookbuilding and auctions, auctions are associated with lower and less variable underpricing than are bookbuilding IPOs.

Within the available space, it is impossible to do justice to all theoretical and empirical contributions. Therefore, I have focused my discussion on the main "milestone" papers that have shaped the way I think about this literature. Inevitably, this reflects my tastes. Notable surveys embodying somewhat different tastes include Ritter and Welch (2002) and Ritter (2003).

## 2. Evidence of underpricing

Underpricing is estimated as the percentage difference between the price at which the IPO shares were sold to investors (the offer price) and the price at which the shares subsequently trade in the market. In well-developed capital markets and in the absence of restrictions on how much prices are allowed to fluctuated by from day to day, the full extent of underpricing is evident fairly quickly, certainly by the end of the first day of trading, and so most studies use the first-day closing price when computing initial underpricing returns. Using later prices, say at the end of the first week of trading, typically makes little difference.

In less developed capital markets, or in the presence of 'daily volatility limits' restricting price fluctuations, aftermarket prices may take some time before they equilibrate supply and demand. The Athens Stock Exchange, for instance, specified daily volatility limits of plus or minus eight percent during the 1990s. Thus for many underpriced IPOs, the first-day return would equal 8% by force of regulation. In such cases, it makes more sense to measure underpricing over a longer window.

In the U.S. and increasingly in Europe, the offer price is set just days (or even more typically, hours) before trading on the stock market begins. This means that market movements between pricing and trading are negligible and so usually ignored. But in some countries (for instance, Taiwan and Finland), there are substantial delays between pricing and trading, and so it makes sense to adjust the estimate of underpricing for interim market movements.

As an alternative to computing percentage initial returns, underpricing can also be measured as the (dollar) amount of 'money left on the table'. This is defined as the difference between the aftermarket trading price and the offer price, multiplied by the number of shares sold at the IPO. The implicit assumption in this calculation is that shares sold at the offer price could have been sold at the aftermarket trading price instead—that is, that aftermarket demand is price-inelastic.

Figures 1–3 provide evidence of underpricing in a range of countries. The U.S. probably has the most active IPO market in the world, by number of companies going public and by the aggregate amount of capital raised. Over long periods of time, underpricing in the U.S. averages between 10 and 20 percent, but as Figure 1 shows, there is a substantial degree of variation over time. There are occasional periods when the average IPO is *over*priced, and there are (more frequent) periods when waves of companies go public at quite substantial discounts to their aftermarket trading value. In

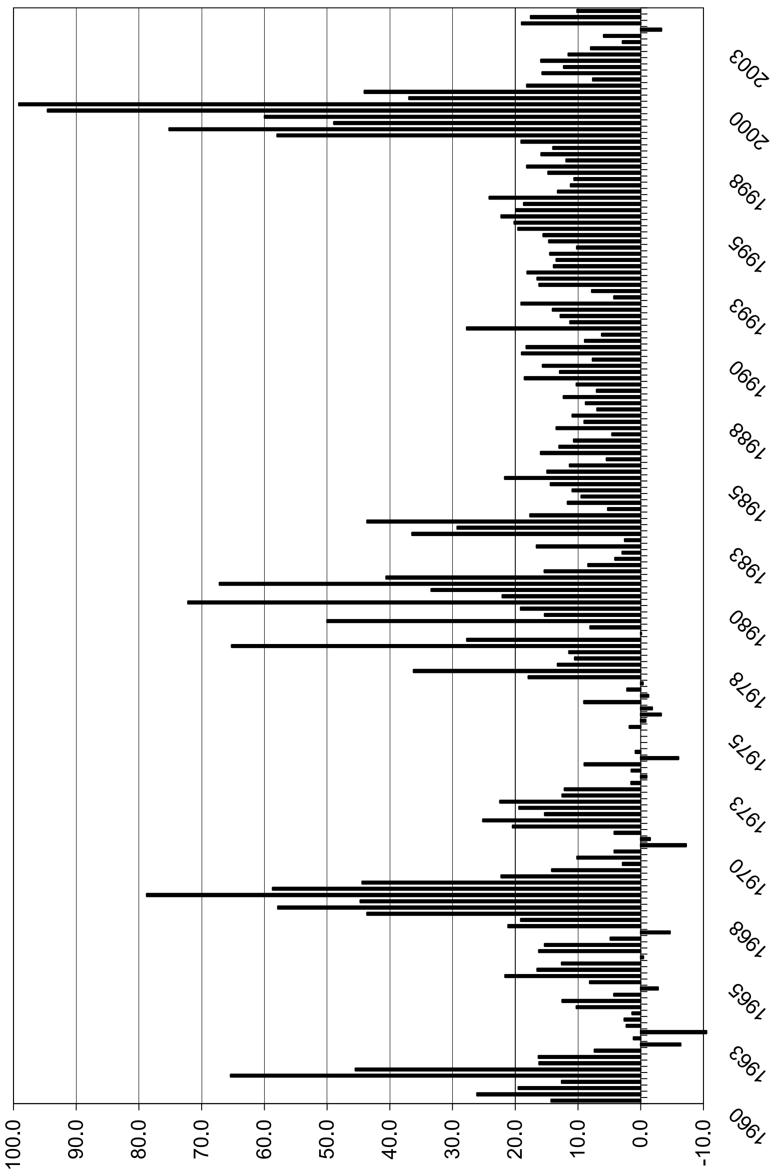



Fig. 1. Initial IPO returns in the United States, 1960 to 2003. The figure reports quarterly equal-weighted average initial IPO returns in % for 14,906 IPOs completed in the United States between 1960 and 2003, calculated as the first-day closing price over the IPO offer price less one. Source: Jay Ritter. Data used by permission.

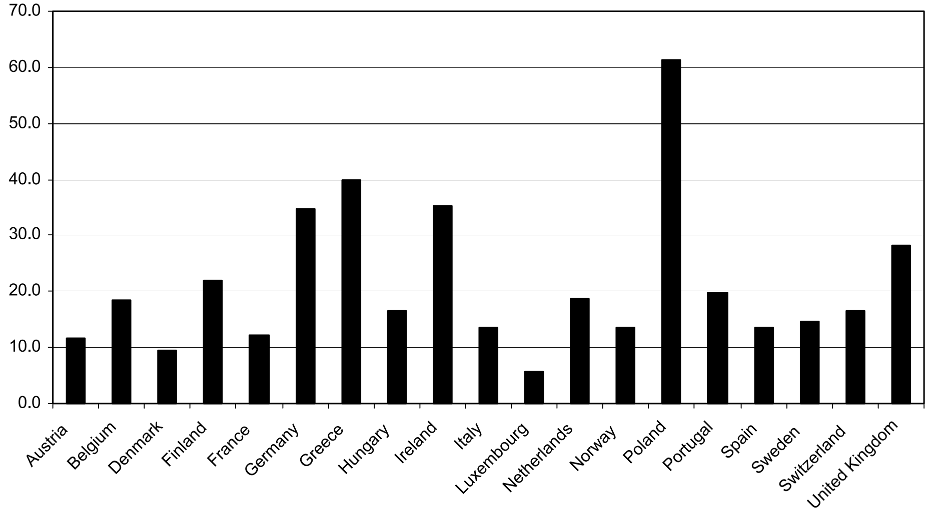



Fig. 2. Initial IPO returns in Europe, 1990 to 2003. The figure reports equal-weighted average initial IPO returns in % for 19 European countries, calculated as the aftermarket trading price over the IPO offer price less one. Aftermarket trading prices are measured on the first day of trading in all countries except France and Greece, where they are measured on the fifth day of trading due to daily volatility limits. IPOs are identified by the author using a range of sources including national stock exchanges, Thomson Financial's SDC global new issue database, Dealogic's Equityware, and news searches. Due to cross-listings, some companies go public outside their home country. The figure shows initial IPO returns by country of *listing*. Aftermarket trading prices are mostly from Datastream, with missing data hand filled from news searches. Between 1990 and 2003, 4,079 IPOs were completed in the 19 countries shown in the figure. This breaks down as follows: Austria (83), Belgium (102), Denmark (69), Finland (70), France (679), Germany (583), Greece (301), Hungary (54), Ireland (22), Italy (158), Luxembourg (5), Netherlands (77), Norway (167), Poland (214), Portugal (33), Spain (47), Sweden (180), Switzerland (68), and the United Kingdom (1,167). Source: author's calculations.

1999 and 2000, for instance, the average IPO was underpriced by 71% and 57%, respectively. In dollar terms, U.S. issuers left an aggregate of $62 billion on the table in those two years alone. Such periods are often called 'hot issue markets'. Given these vast amounts of money left on the table, it is surprising that issuers appear to put so little pressure on underwriters to change the way IPOs are priced. A recent counterexample, however, is Google's IPO which unusually for a U.S. IPO, was priced using an auction.

Figures 2 and 3 report average initial IPO returns for 19 European countries over the period 1990–2003, and for eight Asia-Pacific and eight Latin American countries over the period 1990–2001. Clearly, the extent of underpricing varies from country to country. For instance, it is markedly lower in France than in Germany, and higher in Asia than in Latin America. It is likely that these cross-country differences are at least in part related to differences in the institutional framework within which IPOs are priced and allocated.

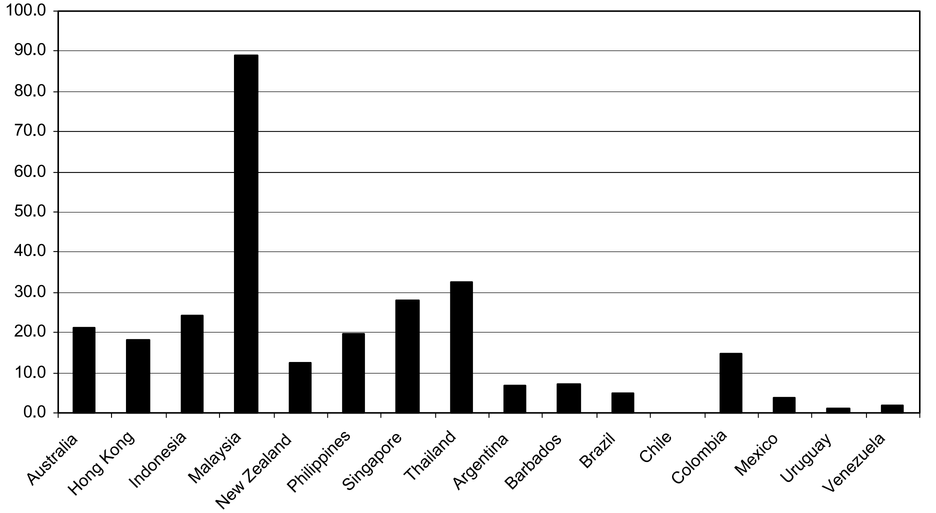



Fig. 3. Initial IPO returns in Asia-Pacific and Latin America, 1990 to 2001. The figure reports equal-weighted average initial IPO returns in % for eight Asian-Pacific and eight Latin American countries, calculated as the aftermarket trading price over the IPO offer price less one. Aftermarket trading prices are measured on the first day of trading. IPOs are identified by the author using a range of sources including national stock exchanges, Thomson Financial's SDC global new issue database, Dealogic's Equityware, and news searches. Due to cross-listings, some companies go public outside their home country. The figure shows initial IPO returns by country of *listing*. Aftermarket trading prices are mostly from Datastream, with missing data hand filled from news searches. Between 1990 and 2001, 2,716 IPOs were completed in the 16 countries shown in the figure. This breaks down as follows: Australia (633), Hong Kong (523), Indonesia (213), Malaysia (506), New Zealand (51), Philippines (91), Singapore (313), Thailand (251), Argentina (25), Barbados (1), Brazil (13), Chile (7), Colombia (3), Mexico (79), Uruguay (1), and Venezuela (6). Source: author's calculations.

## 3. Asymmetric information models

### *3.1. The winner's curse*

The key parties to an IPO transaction are the issuing firm, the bank underwriting and marketing the deal, and the investors buying the stock. Asymmetric information models of underpricing assume that one of these parties knows more than the others. Perhaps the best-known asymmetric information model is Rock's (1986) winner's curse, which is an application of Akerlof's (1970) lemons problem. Rock assumes that some investors are better informed about the true value of the shares on offer than are investors in general, the issuing firm, or its underwriting bank. Informed investors bid only for attractively priced IPOs, whereas the uninformed bid indiscriminately. This imposes a 'winner's curse' on uninformed investors: in unattractive offerings, they receive all the shares they have bid for, while in attractive offerings, their demand is partly crowded out by the informed. Thus, the return uninformed investors earn conditional on receiving an allocation is below the simple average underpricing return shown in Section 2. In the extreme case, the uninformed are rationed completely in underpriced IPOs and receive 100 percent allocations in overpriced IPOs, resulting in average returns that are negative.

When conditional expected returns are negative, uninformed investors will be unwilling to bid for IPO allocations, so the IPO market will be populated only with (equally) informed investors. Rock assumes that the primary market is dependent on the continued participation of uninformed investors, in the sense that informed demand is insufficient to take up all shares on offer even in attractive offerings.[4] This requires that conditional expected returns are non-negative so that the uninformed at least break even.[5] In other words, all IPOs must be underpriced in expectation. This does not remove the allocation bias against the uninformed—they will still be crowded out by informed investors in the most underpriced offerings—but they will no longer (expect to) make losses on average, even adjusted for rationing. Note that it is not rationing *per se* that necessitates underpricing; it is instead the bias in rationing, with uninformed investors expecting more rationing in good than in bad offerings.

Rock's model requires one more assumption. Collectively, firms seeking to go public benefit from underpricing, because it is the key to ensuring the continued participation in the IPO market of the uninformed, whose capital is needed by assumption. Individually, on the other hand, underpricing is clearly costly to a firm going public. This creates an incentive for an individual firm to free-ride by underpricing too little. Beatty and Ritter (1986) argue that as repeat players, investment banks have an incentive to ensure that new issues are underpriced by enough lest they lose underwriting commissions in the future. Investment banks thus coerce issuers into underpricing. Of course, they cannot underprice too much for fear of losing underwriting market share.

### *3.1.1. Testable implications and evidence*

*Adjusted for rationing, uninformed investors earn zero initial returns. Informed investors' conditional returns just cover their costs of becoming informed.*

[4] This ad hoc assumption is actually unnecessary, because a situation where everyone is informed is not in fact an equilibrium. Imagine that all remaining investors are informed. Only attractively priced IPOs will succeed and all others will fail for lack of buyers. But then, assuming that becoming informed is costly, this creates an incentive to stay uninformed and to free-ride on the information of the other investors instead. The investor would simply bid for IPO shares indiscriminately, receiving shares in the attractive IPOs but not in the unattractive ones (which will still fail)—clearly a profitable strategy. Since every investor faces the same incentive, no one would choose to become informed, so unattractive offerings would no longer fail. But if no one is informed, there is an incentive to become informed, in order to avoid the unattractive IPOs. So a situation in which no one is informed is not an equilibrium either, unless becoming informed is prohibitively expensive.

[5] How realistic is the assumption that issuers must pay for the uninformed investors' participation in an offering? If, as Rock asserts, the resources of the informed are limited, the uninformed could simply invest through the informed investors, in exchange for a fee, to avoid the mistake of buying into overpriced issues. (Renaissance Capital Corporation, for instance, manages a mutual fund called 'IPO Plus Aftermarket Fund'.) This is one of the reasons why investment funds exist in the first place: there are economies of scale in becoming informed.

At the heart of the winner's curse model is the idea that, if properly adjusted for rationing, uninformed investors' abnormal returns are zero, on average—that is, just enough to ensure their continued participation in the market. This implication has been tested extensively in the context of countries that impose strict allocation rules. The earliest study is Koh and Walter's (1989) analysis of Singapore, where during the 1970s and 1980s oversubscribed IPOs were allocated by random ballot. Thus two investors bidding for the same number of shares had an equal chance of receiving an allocation. Using data on 66 IPOs, Koh and Walter show that the likelihood of receiving an allocation was negatively related to the degree of underpricing, and that average initial returns fall substantially, from 27% to 1%, when adjusted for rationing.

Levis (1990) conducts a similar analysis for the U.K. Though now no longer in regular use, the preferred IPO method in the U.K. until the early 1990s was the 'offer for sale', which required that allocations be pro-rated in the event of over-subscription. The unconditional average degree of underpricing for the 123 IPOs in Levis' sample is 8.6%, but this declines to 5.14% or less for medium-sized and small applications conditional on being allocated stock. Thus while rationing reduces the initial returns among small investors, it does not drive them down to zero. Keloharju (1993) provides similar evidence for Finland, though he also shows that investors placing *large* orders lose money on an allocation-weighted basis. In Israel, this latter finding seems to hold true more generally: uninformed IPO investors do not appear to break even at all. Amihud, Hauser, and Kirsh (2003) find that uninformed investors earned a *negative* allocation-weighted initial return in Israel in the early 1990s, of $-1.2\%$ on average.

Whether the informed investors' conditional underpricing return just covers the cost of their information production is harder to test in the absence of data on the cost of becoming informed. Of course, the sheer magnitude of money left on the table in certain periods and certain countries documented in Section 2 strongly suggests it is unlikely that underpricing solely compensates investors for becoming informed.

How severe is the allocation bias in practice? The answer depends on who is informed and who is not, a distinction that mostly defies precise empirical testing. Several studies have looked at institutional versus retail investors. Needless to say, it cannot be ruled out that the information asymmetry is most severe *within* groups, rather than between institutional and retail investors. Nevertheless, this approach has yielded some interesting insights. Hanley and Wilhelm (1995), for example, show that there is little difference in the size of allocations institutions receive in underpriced and overpriced issues. Thus institutions do not appear to cherry-pick the best offerings. Aggarwal, Prabhala, and Puri (2002), on the other hand, find that institutional investors earn greater returns on their IPO allocations than do retail investors, largely because they are allocated more stock in those IPOs that are most likely to appreciate in price.

*Underpricing is lower if information is distributed more homogeneously across investor groups.*

Rock's (1986) winner's curse model turns on information heterogeneity among investors. Michaely and Shaw (1994) argue that as this heterogeneity goes to zero, the winner's curse disappears and with it the reason to underprice. By focusing on a segment of the IPO market in which heterogeneity is likely to be low, this prediction can be tested. According to Michaely and Shaw, institutional investors largely avoid IPOs of master limited partnership (MLPs), for a variety of tax reasons. If the informed investors are mainly institutions, and retail investors are mainly uninformed, information heterogeneity among investors in MLPs should be low. Consistent with this prediction, Michaely and Shaw show that average underpricing among 39 MLP IPOs completed between 1984 and 1988 is −0.04%. For comparison, underpricing among non-MLP IPOs over the same time period averaged 8.5%.

*The greater is ex ante uncertainty, the higher is expected underpricing*.

A key empirical implication, due to Ritter (1984) and formalized in Beatty and Ritter (1986), is that underpricing should increase in the ex ante uncertainty about the value of the IPO firm. Beatty and Ritter provide the following intuition. An investor who decides to engage in information production implicitly invests in a call option on the IPO, which will be exercised if the 'true' price exceeds the strike price, that is, the price at which the shares are offered. The value of this option increases in the extent of valuation uncertainty. Thus, more investors will become informed the greater the valuation uncertainty. This raises the required underpricing, since an increase in the number of informed investors aggravates the winner's curse problem.

This hypothesis has received overwhelming empirical support, though it is worth noting that all other asymmetric-information models of IPO underpricing reviewed later in this chapter also predict a positive relation between initial returns and ex ante uncertainty. Thus, most empirical studies of IPO underpricing face the challenge of controlling for ex ante uncertainty, whatever theory they are trying to test. The various proxies that have been used in the literature loosely fall into four groups: company characteristics, offering characteristics, prospectus disclosure, and aftermarket variables.

Popular proxies based on company characteristics include age (Ritter, 1984; Megginson and Weiss, 1991; Ljungqvist and Wilhelm, 2003, and others), measures of size such as log sales (Ritter, 1984), or the industry the company is from (Benveniste et al., 2003). Among offering characteristics, a popular proxy for valuation uncertainty is gross proceeds. However, Habib and Ljungqvist (1998) show that, as a matter of identities, underpricing is strictly decreasing in gross proceeds even when holding uncertainty constant.[6] This clearly makes it unsuitable as a proxy for valuation uncertainty. Other proxies include the number of uses of IPO proceeds as disclosed in the prospectus (Beatty and Ritter, 1986) and the number of risk factors listed in the prospectus (Beatty and Welch, 1996). However, in the absence of rules standardizing what uses and risks must be disclosed, it is unclear whether variation in these measures reflects underlying differences in uncertainty or merely in drafting. A potentially more promising approach

[6] Essentially, this follows because IPO proceeds are positively correlated with the number of newly issued shares, whereas the post-IPO share price is negatively correlated with that number because of dilution.

might be to identify specific uses or risk factors that, if present, indicate higher uncertainty. Ljungqvist and Wilhelm (2003), for instance, argue that firms intending to use their IPO proceeds mainly to fund "operating expenses" rather than investment or debt repayment are potentially more risky. Finally, aftermarket variables such as trading volume (Miller and Reilly, 1987) or volatility (Ritter, 1984, 1987) rely on information which was not in fact available at the time of the IPO. Indeed, it is even possible that such variables are endogenous to the outcome of the IPO. For instance, heavily underpriced IPOs tend to generate more investor interest and so more after-market trading, with the causation running from underpricing to after-market trading behavior rather than the other way around.

*Underwriters that underprice too much (too little) lose business from issuers (investors).*

Consistent with Beatty and Ritter's (1986) claim that underwriters coerce issuers into underpricing to prevent uninformed investors leaving the IPO market, Nanda and Yun (1997) find that overpricing (but not high levels of underpricing) lead to a decrease in the lead underwriter's own stock market value, whereas moderate levels of underpricing are associated with an increase in stock market value, perhaps indicating that underwriters can extract quid pro quo benefits from investors to whom they allocate moderately underpriced shares. In a similar vein, Dunbar (2000) finds that banks subsequently lose IPO market share if they either underprice or overprice too much, squarely supporting Beatty and Ritter's claim.

*Underpricing can be reduced by reducing the information asymmetry between informed and uninformed investors.*

As underpricing represents an involuntary cost to the issuer, there are clear incentives to reduce the information asymmetry and the resulting adverse selection problem between informed and uninformed investors. Habib and Ljungqvist (2001) generalize the notion that issuers have an incentive to reduce underpricing, and model their optimal behavior. They argue that if issuers can take costly actions that reduce underpricing, they will do so up to the point where the marginal cost of reducing underpricing further just equals the marginal benefit. This marginal benefit is not measured by underpricing itself, but by the reduction in the issuer's wealth loss that underpricing implies. Wealth losses and underpricing are not the same: compare an issuer who floats a single share with one who floats the entire company. Clearly the latter's wealth would suffer much more from underpricing, giving him a stronger incentive to take costly actions to reduce underpricing. Using data for a large sample of IPOs completed on Nasdaq in the early 1990s, Habib and Ljungqvist find that issuers optimize, in the sense that spending an additional dollar on reducing underpricing would reduce wealth losses by 98 cents at the margin—resulting in a net benefit that is statistically zero.

A specific way to reduce the informational asymmetry is to hire a prestigious underwriter (Booth and Smith, 1986; Carter and Manaster, 1990; Michaely and Shaw, 1994) or a reputable auditor (Titman and Trueman, 1986). By agreeing to be associated with an offering, prestigious intermediaries "certify" the quality of the issue. For instance, if reputation capital is valuable, prestigious banks will refrain from underwriting low-quality issuers. The information content of the firm's choice of intermediaries may therefore reduce investors' incentives to produce their own information, which in turn will mitigate the winner's curse.

The empirical evidence on this point is mixed. Early studies, focusing on data from the 1970s and 1980s, have tended to find a negative relation between various measures of underwriter reputation and initial returns. Carter and Manaster (1990) provide a ranking of underwriters based on their position in the 'tombstone' advertisements in the financial press that follow the completion of an IPO. This ranking, since updated by Jay Ritter, is much used in the empirical IPO literature. Megginson and Weiss (1991) measure underwriters' reputation instead by their market share, and this approach too is widely used. In practice, results are typically not very sensitive to the choice of underwriter reputation measure.

Results are, however, highly sensitive to the period studied. Beatty and Welch (1996), who use data from the early 1990s, show that the sign of the relation has flipped since the 1970s and 1980s, such that more prestigious underwriters are now associated with *higher* underpricing. This has sparked a debate, still ongoing, about the causes of this shift. One hypothesis, favored by Loughran and Ritter (2004), is that banks have begun to underprice IPOs strategically, in an effort to enrich themselves or their investment clients. Another is that top banks have lowered their criteria for selecting IPOs to underwrite, resulting in a higher average risk profile (and so higher underpricing) for their IPOs.

Habib and Ljungqvist (2001) argue that part of the shift may be due to endogeneity biases. Issuers don't choose underwriters randomly, nor do banks randomly agree which companies to take public (see Fernando, Gatchev, and Spindt, 2005, for further analysis of the latter point). Thus the choices we actually observe are presumably made by optimizing agents. Moreover, issuers likely base their choices, at least in part, on the underpricing they expect to suffer. This leads to endogeneity bias when regressing initial returns on underwriter choice. For instance, a company that is straightforward to value will expect low underpricing, and so has little to gain from the greater certification ability of a top bank. A high-risk issuer, on the other hand, will expect substantial underpricing in the absence of a prestigious underwriter. Taking this into account, Habib and Ljungqvist show that the sign flips back to being negative even in the 1990s.

### *3.2. Information revelation theories*

Over the past decade, the strict pro-rata allocation rules that give rise to Rock's (1986) winner's curse have given way in many countries to bookbuilding methods which give underwriters wide discretion over allocations. Bookbuilding involves underwriters eliciting indications of interest from investors which are then used in setting the price. If—as Rock assumes—some investors are better informed than either the company or other investors, eliciting their information before setting the price becomes one of the key tasks for the investment bank taking a company public.

However, in the absence of inducements, revealing positive information to the underwriter is not incentive-compatible. Doing so would, presumably, result in a higher offer price and so a lower profit to the informed investor. Worse still, there is a strong incentive to actively misrepresent positive information—that is, to claim that the issuer's future looks bleak when it doesn't—to induce the underwriter to set a lower offer price. The challenge for the underwriter is therefore to design a mechanism that induces investors to reveal their information truthfully, by making it in their best interest to do so.

Benveniste and Spindt (1989), Benveniste and Wilhelm (1990), and Spatt and Srivastava (1991) show that bookbuilding can, under certain conditions, be such a mechanism. After collecting investors' indications of interest, the bank allocates no (or only a few) shares to any investor who bid conservatively. This mitigates the incentive to misrepresent positive information: doing so results in exclusion from the IPO. Investors who bid aggressively and so reveal favorable information, on the other hand, are rewarded with disproportionately large allocations of shares. The more aggressive are investors' bids, the more the offer price is raised. However, to ensure truth-telling the allocations have to involve underpriced stock. If the underwriter left no money on the table, truthful reporting would again not be incentive-compatible.

It follows that imposing constraints on the underwriter's allocation discretion can interfere with the efficiency of this mechanism. For instance, requiring that a certain fraction of the shares be allocated to retail investors, as is common in parts of Europe and Asia, reduces underwriters' ability to target allocations at the most aggressive (institutional) bidders and so may force them to rely more on price than on allocations to reward truth-telling. This hurts the issuing firm: underpricing all shares by \$1 but skewing allocations so that co-operative investors reap most of the underpricing profits is preferable to having to underprice all shares by \$2 to generate the same dollar reward for co-operative investors on smaller allocations.

Even though their IPOs are underpriced, issuers benefit from these arrangements. Bookbuilding allows them to extract positive information and raise the offer price in response—even though the price will rise further in the after-market because some money has to be left on the table. Thus the price revision over the course of bookbuilding and the first-day underpricing return are positively correlated. This is often referred to as the 'partial adjustment' phenomenon (Hanley, 1993). Cross-sectionally, the more positive the information (and so the greater the incentive to withhold it), the more money has to be left on the table.

If underwriters and institutional investors deal with each other repeatedly in the IPO market, the cost of information acquisition can be reduced. In a repeated game, investors must weigh the one-off gain from lying against the possibility of being excluded from not only the current but all future IPOs managed by this underwriter. This change to the incentive compatibility constraint implies that banks that are more active in the IPO market have a natural advantage in pricing IPOs: their larger IPO deal flow allows them to obtain investors' cooperation more cheaply than less active underwriters could.

A second advantage of repeated interaction is that is allows underwriters to ‘bundle’ offerings across time. To ensure continued access to lucrative IPOs in the future, investors will from time to time buy poorly received IPOs, as long as the loss they suffer in any given IPO does not exceed the present value of future rents they expect to derive from doing business with the underwriter. This leads to an important implication for the allocation patterns we expect to see. Underwriters should treat regular investors more favorably than occasional investors even when the latter bid more aggressively into the book than the former. This follows because the value of the bank’s underwriting activities depends more on the future cooperation of regular investors than on being able to price any given IPO more fully.

### *3.2.1. Extensions*

The Benveniste and Spindt (1989) paradigm has been extended in numerous ways. Benveniste and Wilhelm (1990) investigate its interaction with Rock’s (1986) winner’s curse. If bookbuilding succeeds in extracting the informed investors’ private information, the informational asymmetry among investors will be reduced. This, in turn, reduces the winner’s curse and thus the level of underpricing required to ensure uninformed investors break even. As argued earlier, regulatory constraints on allocation decisions, common outside the U.S., reduce the effectiveness of bookbuilding, because they undermine underwriters’ ability to reward informed investors for truth-telling. Such constraints can therefore weaken underwriters’ ability to reduce the winner’s curse, again resulting in higher underpricing.[7]

Giving underwriters discretion over allocation decisions is not the only way to lower information acquisition costs. Generally, any tool that allows the underwriter to more directly and exclusively target the reward at those investors who reveal their private information can reduce the overall cost of information acquisition, to the benefit of issuers. One such tool, proposed by Benveniste, Busaba, and Wilhelm (1996), is the promise of selective price support—effectively, a put option offered selectively to co-operative investors. In many countries underwriters intervene in the after-market to prevent prices from falling below the offer price. Empirical evidence suggests this ‘money-back guarantee’ benefits large investors especially, who are likely to be the type of investors underwriters seek to involve in the bookbuilding process.[8]

Busaba, Benveniste, and Guo (2001) show that underwriters can reduce the required extent of underpricing if the issuer has a credible option to withdraw the offering. Downplaying positive information increases the likelihood that the issuer will withdraw, which reduces an investor’s gain from misrepresenting positive information. This in turn reduces the reward required to induce truthful revelation. Consistent with this

[7] Note that here the *existence* of underpricing is due to asymmetric information and a winner’s curse, while institutional factors affect the level/extent of underpricing.

[8] I will discuss price support more fully in Section 4.2.

prediction, James and Wier (1990) find that companies that have secured lines of credit before their IPOs (and thus have a more credible threat to withdraw) experience lower underpricing.

In the Benveniste and Spindt framework, investors incur no cost in becoming informed. If information production is costly, underwriters need to decide how much information production to induce. Sherman and Titman (2002) explore this question in a setting where more information increases the *accuracy* of price discovery, resulting in a trade-off between the (issuer-specific) benefit of greater pricing accuracy and the cost of more information production.

The idea of costly information production is further investigated by Benveniste, Busaba, and Wilhelm (2002) and Benveniste et al. (2003) who link the underwriter's capacity to 'bundle' IPOs over time to the empirical observation that IPOs tend to occur in waves. The central idea is that valuation uncertainty is composed of a firm-specific and an industry component. Obtaining information about the industry component allows investors to evaluate other offerings in that industry more cheaply. Such economies of scale could result in too few firms going public, because the first firm to do so must compensate investors for their whole valuation effort, while later firms can 'free-ride' on the information production.[9] By establishing networks of regular investors, underwriters may be able to reduce this negative externality. To do so, they compensate investors for their information costs across a *sequence* of offerings. This is consistent with the observation that investment banks tend to specialize in particular industries, and that companies tend to go public in industry-specific 'waves'.

### *3.2.2. Testable implications and evidence*

The most direct tests of bookbuilding theories of IPO underpricing are Cornelli and Goldreich (2001, 2003) and Jenkinson and Jones (2004). These studies exploit proprietary datasets from two different European investment banks. The datasets contain information on the bids institutional investors submitted into the book, as well as the allocations they received. Such data are usually kept confidential, so their availability provides a rare opportunity to test information revelation theories of underpricing. Two potential drawbacks are that the sample sizes are relatively small, and that the results are bank-specific and so may not generalize to other banks. Indeed, the fact that Jenkinson and Jones' results are at odds with those of Cornelli and Goldreich, as we will see, may in large part be due to differences in the sophistication with which the two banks carry out bookbuilding.

[9] The idea that information spillovers can cause IPO clustering is explored in three papers that are not based on the Benveniste–Spindt information-acquisition framework. Booth and Chua (1996) point out that when many companies come to market, the marginal cost of information production is lower, so average underpricing falls. Mauer and Senbet (1992) argue that IPO companies that start trading in the secondary market may reduce the valuation uncertainty surrounding companies with similar technologies which are in the process of going public. Stoughton, Wong, and Zechner (2001) develop a model in which one firm's IPO provides information about industry prospects, thus causing many similar companies to go public soon after.

Cornelli and Goldreich (2001, 2003) have access to the IPO books of a leading European investment bank active in up to 37 cross-border IPOs outside the U.S., including a number of privatizations. They observe essentially two different types of bids: strike (or market) orders and price-limited bids. Unlike strike orders, price-limited bids specify a maximum price an investor is willing to pay for a given number of shares. Thus such bids arguably convey more information to the underwriter than strike orders. In the Benveniste–Spindt framework, investors submitting price-limited bids should therefore receive disproportionately larger allocations than investors submitting strike orders, and this allocation bias should become more pronounced, the more aggressive the price limit.

The results generally support the Benveniste–Spindt model. Cornelli and Goldreich (2001) find that price-limited bids receive 19 percent greater allocations than strike orders. The value of an additional price-limited bid to the underwriter should depend on how much information it has already gathered from other investors. Consistent with this conjecture, Cornelli and Goldreich show that investors submitting price-limited bids receive larger allocations when the book contains fewer limit bids. Finally, more aggressive limit bids yield larger allocations than less aggressive ones, as predicted.

Allocations are not only related to the characteristics of the bid, they are also driven by the characteristics of the bidder. Frequent bidders receive larger allocations (relative to their bid size) than infrequent bidders, consistent with the prediction that regular investors should be favored over occasional ones even when the latter bid more aggressively.

In their 2003 follow-on article, Cornelli and Goldreich ask whether limit orders do reveal pricing-relevant information. On average, final offer prices are closely related to the limit orders in the book, in particular those submitted by large and by frequent bidders. The underwriter sets the offer price close to the quantity-weighted average of the limit prices in the book. Limit bids are especially influential when they indicate a consensus among bidders. Taken together, these findings provide strong support for Benveniste and Spindt's (1989) view that bookbuilding serves to extract information from investors.

Jenkinson and Jones (2004) have data for 27 IPOs managed by a different European investment bank. The allocation and pricing decisions of this bank differ markedly from Cornelli and Goldreich's, and provide less support for bookbuilding theories of IPO underpricing. Price-limited bids are much rarer at this bank, and they are not associated with favorable allocations. The main allocation pattern this bank has in common with Cornelli and Goldreich's is that more frequent bidders are treated preferentially. Jenkinson and Jones interpret their findings as "cast[ing] doubt upon the extent of information production during the bookbuilding period".

There are many possible reasons why Jenkinson and Jones' findings look so different from Cornelli and Goldreich's, beyond uncontrollable differences in the types of deals examined. The most obvious are based on differences in the sophistication with which these two European investment banks carry out bookbuilding. First, a bank's ability to extract information is larger the more active it is in the IPO market, since a higher rate of future deal flow increases the investor's incentive to co-operate with the bank today. Since the authors have revealed the identity of their respective bank to me, I am able to confirm that Cornelli and Goldreich's bank is associated with substantially larger deal flow.

Second, Benveniste and Spindt's (1989) argument assumes that the bank has access to a set of *informed* investors whose information it seeks to elicit with the help of favorable allocations of underpriced stocks. The quality of the information it acquires is clearly related to the quality of the investors it has access to. And it is not unreasonable to assume that banks differ in the quality of their investor networks. Indeed, bids by U.S. investors comprise only 1% of the sample in Jenkinson and Jones versus 13% in Cornelli and Goldreich. In sum, it appears likely that Cornelli and Goldreich's bank is both more active and better connected and thus in a better position to extract pricing-relevant information from investors.

No corresponding bookbuilding data are available for U.S. banks. Thus, whether these European results can be generalized to the U.S. depends on how similar bookbuilding techniques are in Europe and the U.S. Ljungqvist, Jenkinson, and Wilhelm (2003) provide evidence from 65 countries showing that the quality of bookbuilding—as measured by the underpricing cost of inducing truthful information reporting—heavily depends on whether a U.S. bank lead-manages the issue and on whether U.S.-based investors are targeted. Indeed, bookbuilding by non-U.S. banks targeted at their domestic (non-U.S.) clients appears to provide no pricing advantage over fixed-price offerings completed without bookbuilding.

Controlling for the fact that issuers *choose* whether to hire U.S. banks and have their IPOs marketed to U.S. investors, Ljungqvist, Jenkinson, and Wilhelm (2003) show that underpricing is reduced by 41.6% on average when U.S. banks and U.S. investors are involved. This benefit doesn't come free: U.S. banks charge higher underwriting fees than do domestic banks. But on net, 73% of issuers would have been worse off had they chosen local banks and local investors instead, in the sense that the resulting increase in underpricing cost would have exceeded the savings on the underwriting fees. The median firm switching to the 'cheaper' strategy would have suffered a reduction in net proceeds of US$11.7 million. These findings are consistent with the prediction that access to informed (U.S.) investors favors certain U.S. investment banks.

While no other datasets have yet matched the level of detail of Cornelli and Goldreich's (2001, 2003) and Jenkinson and Jones' (2004), several studies have used *aggregate* allocation data on the fractions of an IPO allocated to institutional and retail investors, respectively. If institutions are more likely to be informed than retail investors, this allocation split can be thought of as a crude approximation of the extent to which underwriters favor informed investors in their allocation decisions.

Hanley and Wilhelm (1995), for instance, use a sample of 38 U.S. IPOs conducted by a leading (unnamed) investment bank over the period 1983–1988. IPO allocations clearly favor institutions over retail investors: institutions are allocated 66.8% of the average IPO. Cross-sectionally, institutional allocations are larger the more the offer price exceeds the midpoint of the indicative filing range established at the beginning of bookbuilding. Positive price revisions presumably follow when informed investors reveal positive information, and this is precisely when underwriters need to reward co-operative investors with favorable allocations. At the same time, however, institutions are given similar allocations in overpriced as in underpriced deals, which is consistent with the prediction that underwriters 'bundle' IPOs over time and regular investors sometimes are expected to buy 'cold' IPOs.

Aggarwal, Prabhala, and Puri (2002) analyze a more recent dataset covering 164 IPOs managed by nine different banks in 1997 and 1998. As in Hanley and Wilhelm (1995), institutional investors are allocated the lion's share of IPO stock and institutional allocations increase in the price revision relative to the filing range. Underpricing, in turn, is larger the more stock institutions were allocated. This makes sense within the Benveniste–Spindt framework, since underwriters likely use both price (i.e., underpricing) and quantity (i.e., allocation size) to ensure truthful revelation of particularly positive information.

Ljungqvist and Wilhelm (2002) depart from the previous two studies by estimating the *structural* links between IPO allocations, price revisions, and initial returns. They argue that these three variables are jointly determined, in the sense that the degree of price revision depends on how much (positive) information investors reveal, which in turn depends on their expected economic reward in the form of allocations of underpriced stock. Using aggregate allocation data from France, Germany, the U.K., and the U.S., they find that price revisions increase in institutional allocations and vice versa, and initial returns increase in price revisions but decrease in institutional allocations. The latter result suggests that constraints on the size of institutional allocations—which are widespread in France and (during the early 1990s) in the U.K.—result in underwriters relying more on price than on quantity to reward truthful revelation. This is costly to issuers, since blanket underpricing rewards both informed and uninformed bidders.

There is one key prediction of the Benveniste and Spindt (1989) framework that can be tested without proprietary bid or allocation data. Revisions in the offer price and the number of shares offered during bookbuilding likely reflect investors' level of interest and the aggregate nature of their information. An IPO for which positive information is revealed should be priced towards the upper end of the indicative price range (or if the information is particularly positive, above the range) whereas a less well received offering should be priced towards the lower end. Benveniste and Spindt's model suggests that underpricing should be concentrated among the IPOs drawing the highest level of pre-market interest. In other words, even though the underwriter adjusts the price upwards, he does so only partially, in order to leave enough money on the table to compensate informed investors for their truthful revelation. Hanley (1993) was the first to provide empirical evidence of this 'partial adjustment' phenomenon. Numerous subsequent studies have corroborated this finding, both in the U.S. and internationally.

Loughran and Ritter (2002) criticize Hanley's (1993) interpretation of the partial adjustment phenomenon, by showing that underwriters, when setting the offer price, do not fully incorporate *public* information in the form of pre-pricing returns on the market index. (See also Bradley and Jordan, 2002.) This appears to contradict the Benveniste–

Spindt (1989) framework, since public information is freely available and so there is no need to compensate investors for it by leaving money on the table. Loughran and Ritter prefer a behavioral explanation, which will be discussed more fully in Section 6.3. In short, when the IPO is doing poorly (and so the price is likely to be revised downwards), issuers bargain hard with the underwriter over the issue price. When the IPO is doing well (and so the price is likely to be revised upwards), issuers are complacent. This leads to an asymmetric relation between prior market returns and offer price revisions, at least to the extent that the state of the market correlates with how the IPO is doing.

Lowry and Schwert (2004) reexamine this question. While their findings confirm the existence of a positive and statistically significant relation between offer price revisions and pre-pricing market returns, they argue that this effect is negligible economically. Edelen and Kadlec (2005), too, reexamine Loughran and Ritter's (2002) critique, and show that the apparent asymmetry may be driven by sample selection bias. In a sample of *completed* IPOs, negative market returns have indeed no effect on offer price revisions. But negative market returns have a significant impact on the decision to withdraw the IPO. When this is taken into account using the Heckman (1979) approach, the asymmetry disappears.

Whether symmetric or asymmetric, public information appears not to be fully priced. Why not? In contrast to Loughran and Ritter (2002), Edelen and Kadlec (2005) propose a *rational* explanation, noting that issuers must trade off the proceeds from the IPO against the probability of the IPO succeeding. In the context of a search model, aggressive pricing increases the probability of failure. When comparable firms' valuations are low, the IPO is likely to generate relatively little 'surplus' for the issuer. Therefore, the issuer has little to lose if the deal fails, and pushes the underwriter to extract as high proceeds as possible, even though this implies a greater risk of the deal failing. When comparable firm valuations are high, the issuer is unwilling to risk failure because there is much to be gained from going public. In this situation, the issuer does not insist on aggressive pricing. Thus as comparable firms' valuations increase, so too does the degree of underpricing.

### 3.3. Principal-agent models

Theories of bookbuilding stress the important role of investment banks in eliciting information that is valuable in price-setting, and the benefit of giving them discretion over allocation decisions. Some authors—most prominently perhaps Loughran and Ritter (2004)—stress the 'dark side' of these institutional arrangements, by highlighting the potential for agency problems between the investment bank and the issuing firm.

A multitude of regulatory investigations following the bursting of the late 1990s 'dot-com bubble' has recently revived academic interest in agency models of IPO underpricing. For instance, the fact that underpricing represents a wealth transfer from the IPO company to investors can give rise to rent-seeking behavior, whereby investors compete for allocations of underpriced stock by offering the underwriter side-payments. Such side-payments could take the form of excessive trading commissions paid on unrelated transactions (Loughran and Ritter, 2002), an activity that Credit Suisse First Boston was fined $100 million for in 2002.[10] Or investment bankers might allocate underpriced stock to executives at companies in the hope of winning their future investment banking business, a practice known as 'spinning'. In either case, the underwriter stands to gain from deliberately underpricing the issuer's stock.

Underwriting fees are typically proportional to IPO proceeds, and thus inversely related to underpricing. This provides a countervailing incentive to keep underpricing low. But at times, it is conceivable that the bank's private benefits of underpricing greatly exceed this implied loss of underwriting fees.

The theoretical literature linking agency conflicts and IPO underpricing goes back more than 20 years. Early models focused on how a bank's informational advantage over issuing companies might allow the bank to exert sub-optimal effort in marketing and distributing the stock. If effort is not perfectly observable and verifiable, banks find themselves in a moral hazard situation when acting as the issuers' agents in selling an IPO. Baron and Holmström (1980) and Baron (1982) construct screening models which focus on the underwriter's benefit from underpricing. In a screening model, the uninformed party offers a menu or schedule of contracts, from which the informed party selects the one that is optimal given her unobserved type and/or hidden action. The contract schedule is designed to optimize the uninformed party's objective, which, given its informational disadvantage, will not be first-best optimal. An example is the various combinations of premium and deductible that a car insurer may offer in order to price-discriminate between different risks (unobservable type) or to induce safe driving (hidden action).

To induce optimal use of the underwriter's superior information about investor demand, the issuer in Baron's model delegates the pricing decision to the bank. Given its information, the underwriter self-selects a contract from a menu of combinations of IPO prices and underwriting spreads. If likely demand is low, it selects a high spread and a low price, and *vice versa* if demand is high.[11] This optimizes the underwriter's unobservable selling effort by making it dependent on market demand. Compared with the first-best solution under symmetric information, the second-best incentive-compatible contract involves underpricing in equilibrium, essentially since its informational advantage allows the underwriter to capture positive rents in the form of below-first-best effort costs.

The more uncertain the value of the firm, the greater the asymmetry of information between issuer and underwriter, and thus the more valuable the latter's services become, resulting in greater underpricing. This is a further rationalization for the empirical observation that underpricing and proxies for *ex ante* uncertainty are positively related.

[10] Source: NASD Regulation, Inc., news release dated January 22, 2002.

[11] There is empirical support for the notion of a menu of compensation contracts. Dunbar (1995) shows that issuers successfully offer underwriters a menu that minimizes offering costs by inducing self-selection.

Biais, Bossaerts, and Rochet (2002) combine the agency cost setting of Baron (1982) with Benveniste and Spindt's (1989) assumption that some investors hold pricing-relevant information worth extracting before the offer price is set. In such a setting, the investment banker could collude with the informed investors, to the potential detriment of the issuing company. Biais, Bossaerts, and Rochet derive an optimal IPO mechanism that maximizes the issuer's proceeds. In this mechanism, the IPO price is set higher the fewer shares are allocated to (uninformed) retail investors. Allocating more to institutional investors when their private signals are positive (i.e., when the IPO price should be set higher) is consistent with Benveniste and Spindt's information acquisition argument. Conversely, allocating more to retail investors when institutional investors' signals are less positive while at the same time lowering the IPO price lessens the winner's curse.

### *3.3.1. Testable implications and evidence*

In principle, issuers can mitigate agency conflicts in two ways: they can monitor the investment bank's selling effort and bargain hard over the price, or they can use contract design to realign the bank's incentives by making its compensation an increasing function of the offer price. Ljungqvist and Wilhelm (2003) provide evidence consistent with monitoring and bargaining in the U.S. in the second half of the 1990s. They show that first-day returns are lower, the greater are the monitoring incentives of the issuing firms' decision-makers (say the CEO). Monitoring incentives are taken to increase in the relevant decision-maker's equity ownership level and the number of personal shares he sells at the time of the IPO. Higher equity ownership gives the decision-maker a greater stake in the outcome of the pricing negotiations, while underpricing stock sold for personal account represents a direct wealth transfer from the decision-maker to IPO investors.

Ljungqvist (2003) studies the role of underwriter compensation in mitigating conflicts of interest between companies going public and their investment bankers. Making the bank's compensation more sensitive to the issuer's valuation should reduce agency conflicts and thus underpricing. Consistent with this prediction, Ljungqvist shows that contracting on higher commissions in a large sample of U.K. IPOs completed between 1991 and 2002 leads to significantly lower initial returns, after controlling for other influences on underpricing and a variety of endogeneity concerns. These results indicate that issuing firms' contractual choices affect the pricing behavior of their IPO underwriters. Moreover, the empirical results cannot reliably reject the hypothesis that the intensity of incentives is optimal, and so that contracts are efficient.

A potentially powerful way to test the agency models is to investigate the underpricing experience of IPOs that have little or no informational asymmetry between issuer and bank. The two most prominent cases in point involve underwriters that own equity stakes in the IPO company and situations where a company underwrites its IPO itself. Some interesting evidence along these lines is available for the U.S. Muscarella and Vetsuypens (1989) study a set of 38 self-underwritten investment bank IPOs in the 1970s and 1980s. Since issuer and underwriter are identical, there can be no agency problem.

However, these 38 investment bank IPOs appear to have been underpriced by roughly as much as other IPOs, which Muscarella and Vetsuypens interpret as contradicting the agency models.

There are only so many investment banks taking themselves public, so Muscarella and Vetsuypens' (1989) approach does not lend itself straightforwardly to large-sample testing. But over the course of the 1990s, investment banks emerged as an important pre-IPO shareholder group in many IPO companies (Ljungqvist and Wilhelm, 2003). Often, they acquired stakes in these companies indirectly, via their venture capital operations. By the year 2000, investment banks were pre-IPO shareholders in 44% of companies going public. These equity stakes should reduce their incentives to underprice the stock to the issuer's detriment, and the size of this effect should be proportional to the size of their equity stake.

The evidence reported in Ljungqvist and Wilhelm (2003) supports both these predictions. The greater the investment bank's equity holding, the lower are first-day underpricing returns. This finding contrasts with the earlier result of Muscarella and Vetsuypens (1989) that investment banks underwriting their own IPOs suffered as much underpricing as other issuers. However, the negative relation between investment bank equity holdings and underpricing does not appear to depend on whether the investment bank acted as lead underwriter. Focusing on venture-backed IPOs only, Li and Masulis (2003) also find that initial returns decrease in the size of investment banks' pre-IPO equity holdings, though in their case, the effect is more pronounced for lead underwriters than for other syndicate members.

How widespread is the self-dealing behavior alleged in recent regulatory investigations into IPO practices? In general, this is hard to address empirically. For instance, banks do not typically publish the kind of allocation data necessary to examine 'spinning'. Notwithstanding Congressional disclosure of IPO allocations to executives at WorldCom and the class action suit over spinning against eBay, Inc., the relevant data are unlikely to become available in a systematic fashion.

The link between allocations and trading commissions is potentially more readily observable. In an innovative paper, Reuter (2004) combines data on the recipients of the brokerage commissions paid by U.S. mutual funds with data on the mutual funds' equity holdings. The fund holdings data are used to approximate IPO allocations, on the assumption that funds do not trade their IPO allocations in any systematic way (that is, in a way that is correlated with the variables of interest). Reuter finds a positive relation between the commissions mutual funds paid to lead managers and the size of reported holdings in the managers' IPOs. One interpretation is that fund managers 'buy' underpriced IPO allocations with their trading commissions. Another is that underwriters allocate IPOs to clients they have strong relationships with, which includes executing much of the clients' trades.

Reuter's (2004) point estimates suggest that investment banks received 85 cents in trading commissions per dollar of underpricing gain allocated to mutual funds in 1996–1998. Assuming trading commissions were used to 'buy' underpriced IPO allocations, banks appear to have been very good at capturing the lion's share of the rent over that time period. Interestingly, however, in 1999 the point estimate falls to only 19 cents in trading commissions per dollar of underpricing gain. Thus at the height of the IPO bubble, the 'price' of underpriced IPO allocations seems to have dropped substantially. In fact, in aggregate dollar terms, almost the entire increase in money left on the table in 1999 appears to have accrued to mutual funds, with banks' revenue from trading commissions largely unchanged in 1999 compared to earlier years. This is hard to reconcile with the view that banks deliberately increased underpricing during the IPO bubble: if they did, they were curiously inept at profiting from it.

### *3.4. Underpricing as a signal of firm quality*

The final group of asymmetric information models reverses Rock's assumption regarding the informational asymmetry between issuing firms and investors. If companies have better information about the present value or risk of their future cash flows than do investors, underpricing may be used to signal the company's 'true' high value. This is clearly costly, but if successful, signaling may allow the issuer to return to the market to sell equity on better terms at a later date. In the words of Ibbotson (1975), who is credited with the original intuition for the IPO signaling literature, issuers underprice in order to 'leave a good taste in investors' mouths'. Allen and Faulhaber (1989), Grinblatt and Hwang (1989), and Welch (1989) have contributed theories with this feature.

Suppose there are two types of firms, denoted high-quality and low-quality, which look indistinguishable to investors. Firms raise equity in two stages, via an IPO and at a later date. High-quality firms have incentive to credibly signal their higher quality, in order to raise capital on more advantageous terms. Low-quality firms have incentive to mimic whatever high-quality firms do. The proposed signal in the IPO signaling models is the issue price.

With some positive probability, a firm's true type is revealed to investors before the post-IPO financing stage. This exposes low-quality issuers to the risk that any cheating on their part will be detected before they can reap the benefit from imitating the high-quality issuers' signal. This makes separation between the two types possible. Provided the risk of detection and the implied reduction in IPO proceeds are sufficiently great to deter the low-quality firms from imitating the high-quality ones, a high-quality firm can influence investors' after-market beliefs about its value by deliberately leaving money on the table at the IPO. This money is 'recouped' when the firm returns to the market at a later date. Low-quality firms refrain from mimicking the signal (i.e., from underpricing) because the risk of detection means they may not be able to recoup the cost of the signal later.

Signaling models are open to the challenge that the proposed signaling device may be dominated by other signals. Would firms really choose the underpricing signal if they had a wider range of signals to choose from? Such a range could include the choice of particularly reputable underwriters (Booth and Smith, 1986), auditors (Titman and Trueman, 1986), or venture capitalists (Megginson and Weiss, 1991; Lee and Wahal, 2004), each of whom could perform a certification-of-quality role; the quality of the board of directors, and in particular the choice of non-executive directors, who similarly would put their reputation on the line; and direct disclosure of information to IPO investors, backed by a mechanism designed to deter fraudulent disclosure (Hughes, 1986).

### *3.4.1. Testable implications and evidence*

The signaling models generate a rich set of empirical implications predicting that underpricing is positively related to the probability, size, speed, and announcement effect of subsequent equity sales.[12] In common with the other asymmetric information theories of underpricing, the signaling models also predict a positive relation between underpricing and the ex ante uncertainty about firm value. This follows because a noisier environment increases the extent of underpricing that is necessary to achieve separation.

One of the most notable empirical tests of the signaling models is due to Jegadeesh, Weinstein, and Welch (1993). Using data on IPOs completed between 1980 and 1986, Jegadeesh, Weinstein, and Welch find that the likelihood of issuing seasoned equity and the size of seasoned equity issues increase in IPO underpricing, as expected. However, they note that these statistically significant relations are relatively weak economically. For instance, the least underpriced quintile of IPOs face a 15.6% likelihood of issuing seasoned equity, compared to 23.9% in the most underpriced quintile. The results are equally consistent with a pooling equilibrium: firms pool at the IPO and reissue equity only once the market learns their true quality. Consistent with the possibility of pooling, Jegadeesh, Weinstein, and Welch find that post-IPO share price returns better explain whether a company subsequently raises equity than the degree of IPO underpricing.

As Michaely and Shaw (1994) note, the decision how much money to leave on the table and whether to reissue equity later on are not independent of each other in the signaling framework. The same logic applies to the size of any seasoned equity offering. Thus, these decisions should be modeled simultaneously. Michaely and Shaw estimate a simultaneous system using underwriter reputation to identify the underpricing equation and post-IPO performance to identify the equation modeling the size of the seasoned equity offering. The results do not support the signaling models: the decision how much to underprice is not significantly related to the reissue decision and vice versa, consistent with Jegadeesh, Weinstein, and Welch (1993).

Welch (1996) endogenizes the decision how long to wait before returning to the equity market. The longer a firm waits, the greater is the probability that nature will reveal its true value. Thus a high-quality firm can afford to wait longer, but the cost of this strategy is that it may not receive funds when it most needs them. Empirically, Welch finds that the time to SEO increases in IPO underpricing while firms that return to the market earlier do so after experiencing high post-IPO stock market returns.

[12] For a survey of seasoned equity offers more generally, see Chapter 6 by Eckbo, Masulis, and Nørli in this volume.

Usually, companies announcing seasoned equity offerings experience negative announcement-date returns. In the signaling framework, we would expect a less negative stock price reaction in response to SEO announcements by 'high-quality' companies, which under separation means companies that underpriced their IPOs by more. Both Jegadeesh, Weinstein, and Welch (1993) and Slovin, Sushka, and Bendeck (1994) find evidence consistent with this prediction.

Spiess and Pettway (1997) add an interesting observation to the empirical literature on IPO signaling models. In their data, pre-IPO shareholders sell personal shares at the IPO in half of all IPOs, and such insider selling is no less common among the more underpriced firms. This suggests that insiders at high-quality firms do not wait to realize the benefit of their underpricing signal by delaying their sales of personally held shares. Such behavior seems inconsistent with the logic of the signaling models.

## 4. Institutional explanations

We now turn to three 'institutional' explanations for IPO underpricing. First, the litigiousness of American investors has inspired a *legal insurance* or *lawsuit avoidance* hypothesis. The basic idea, which goes back at least to Logue (1973) and Ibbotson (1975), is that companies deliberately sell their stock at a discount to reduce the likelihood of future lawsuits from shareholders disappointed with the post-IPO performance of their shares. This explanation is somewhat U.S.-centric, in that underpricing is a global phenomenon, while strict liability laws are not. The risk of being sued is not economically significant in Australia (Lee, Taylor, and Walter, 1996), Finland (Keloharju, 1993), Germany (Ljungqvist, 1997), Japan (Beller, Terai, and Levine, 1992), Sweden (Rydqvist, 1994), Switzerland (Kunz and Aggarwal, 1994), or the U.K. (Jenkinson, 1990), all of which experience underpricing. Still, it is possible that lawsuit avoidance is a second-order driver of IPO underpricing.

The second institutional approach is based on the practice of price support. One of the services that underwriters provide in connection with an IPO is *price stabilization*, intended to reduce price drops in the after-market for a few days or weeks. Perhaps surprisingly, such 'price manipulation' is legal in many countries, including the U.S. (1934 Securities Exchange Act, Rule 10b-7, since replaced by Regulation M). Statistically, price stabilization results in fewer observations of overpricing, and so shifts up the observed mean initial return.

Third, there may be tax advantages to IPO underpricing. This results in a trade-off between the tax benefit and the dilution cost of underpricing. Depending on their tax situation, managers may prefer more or less underpricing.

### *4.1. Legal liability*

Stringent disclosure rules in the U.S. expose underwriters and issuers to considerable risk of litigation by investors on the grounds that material facts were mis-stated or omitted from the IPO prospectus. Lowry and Shu (2002) estimate that nearly 6 percent of companies floated in the U.S. between 1988 and 1995 subsequently were sued for violations relating to the IPO, with damages awarded to plaintiffs averaging 13.3% of IPO proceeds.

Tinic (1988), Hughes and Thakor (1992), and Hensler (1995) argue that intentional underpricing may act like insurance against such securities litigation. Lawsuits are obviously costly to the defendants, not only directly—damages, legal fees, diversion of management time, etc.—but also in terms of the potential damage to their reputation capital: litigation-prone investment banks may lose the confidence of their regular investors, while issuers may face a higher cost of capital in future capital issues. Hughes and Thakor propose a trade-off between on the one hand minimizing the probability of litigation, and hence minimizing these costs, and on the other maximizing the gross proceeds from the IPO (and thus the underwriter's commission thereon). Crucially, they assume that the probability of litigation increases in the offer price: the more overpriced an issue, the more likely is a future lawsuit. In addition, they predict that underpricing reduces not only (i) the probability of a lawsuit, but also (ii) the probability of an adverse ruling conditional on a lawsuit being filed, and (iii) the amount of damages awarded in the event of an adverse ruling (since actual damages in the U.S. are limited by the offer price).

As a point of legal fact, the amount of damages that can be awarded in lawsuits filed under Section 11 of the 1933 Securities Act increases in the difference between the offer price and the subsequent (lower) trading price. Thus, underpricing reduces the likely damages. This in turn reduces the probability of litigation assuming the size of expected damages affect class-action lawyers' incentives to file a suit.

#### *4.1.1. Testable implications and evidence*

Tinic (1988) proposes that the enactment of the 1933 Securities Act represents a regime shift that potentially allows us to test the legal liability hypothesis. Prior to the 1933 Act, the principle of *caveat emptor* largely protected issuers and investment banks against litigation risk, and so underpricing should have been low. After 1933, litigation risk should have featured more prominently when investment banks priced deals, and so underpricing should have increased. Moreover, banks with a comparative advantage at due diligence might, post-1933, feel less need to insure against lawsuits by means of underpricing, leading to a negative relation between a bank's experience and initial returns.

Tinic identifies a sample of 70 IPOs completed between 1923 and 1930 and compares their average underpricing to that of a sample of 134 IPOs completed between 1966 and 1971. As predicted, average underpricing was lower before 1933, but the difference is not particularly large: 5.2% in 1923–1930 versus 11.1% in 1966–1971. Moreover, it is well-documented that underpricing varies immensely over time (see Ibbotson and Jaffe, 1975 and Figure 1 in Section 2 of this chapter), so we cannot rule out that Tinic's results are driven by factors other than increased litigation risk. Drake and Vetsuypens (1993), for instance, show that average initial returns in the six years *after* Tinic's sample period (1972–1977) were actually lower than between 1923 and 1930. Evidence based on the enactment of the 1933 Securities Act is thus inconclusive.

Tinic also finds that more experienced underwriters were associated with lower underpricing in the post-1933 sample but not before. This is consistent with his prediction that greater due diligence skills reduce the need for underpricing as a form of protection against lawsuits. On the other hand, simple certification arguments yield the same prediction, so as a test of the legal insurance hypothesis, the relation between underwriter experience and underpricing has little power. Moreover, as discussed in Section 3.1, this relation appears to have changed sign in the 1990s (Beatty and Welch, 1996). However, it is not impossible to rationalize a positive relation within the legal insurance hypothesis: more prestigious underwriters may have deeper pockets and so are more worth suing, leading them to rely more heavily on underpricing. Evidence based on the relation between underpricing and underwriter experience thus also appears inconclusive.

A potentially more promising research avenue is to investigate the predicted negative link between underpricing and the probability of litigation, and to do so cross-sectionally. Drake and Vetsuypens (1993) study a sample of 93 IPO firms that were sued and compare them to a sample of 93 IPOs that were not sued, matched on IPO year, offer size, and underwriter prestige. Sued firms are just as underpriced as the control sample, and underpriced firms are sued more often than overpriced firms. Drake and Vetsuypens interpret these findings as inconsistent with the legal insurance hypothesis.

Lowry and Shu (2002) argue that such an *ex post* comparison misses the point because it does not truly consider the *probability* of being sued. Empirical analysis of the link between underpricing and the probability of litigation needs to be careful about the following simultaneity problem: firms choose a certain level of underpricing to reduce the probability of litigation, but the level of underpricing they choose depends on the probability of being sued. Put differently, greater underpricing reduces litigation risk, but greater litigation risk requires more underpricing.

Due to this simultaneity problem, ordinary least squares estimates are likely biased. Lowry and Shu propose a two-stage least squares approach. As identifying variables, they use prior market-index returns in the underpricing equation and the IPO firm's expected stock turnover in the litigation equation. The authors motivate these choices on the basis of prior work and economic common sense, but do not test whether they are valid[13] or strong[14] identifying variables statistically. Loughran and Ritter (2002) found a positive relation between lagged index returns and underpricing, but there is no reason

[13] A necessary and sufficient condition for instrument validity is that the system satisfy the order and rank conditions. The order condition is easy to check. It requires that the variable be correlated with the endogenous variable of the first-stage regression, but not with the endogenous variable of the second-stage regression. A variety of formal tests are available. Stock turnover appears to fail the order condition (see Lowry and Shu, 2002, Table 5, p. 329).

[14] Weak instruments may aggravate the effect of simultaneity bias, rather than solving it. To be considered strong, an instrument needs to be highly correlated with the first-stage endogenous variable. Staiger and Stock (1997) recommend a cut-off of $F = 10$. On this basis, Lowry and Shu's instruments would appear to be weak.

to expect lagged index returns to affect lawsuits many years later. This makes lagged index returns a plausible instrument for underpricing. Damages generally increase in the number of shares traded at the allegedly misleading prices, so stock turnover may be a plausible instrument for litigation risk a priori.[15]

The OLS and 2SLS estimates give rise to radically different conclusions. The OLS results suggest that underpricing decreases in the incidence of (actual) lawsuits, suggesting that firms underprice less the more often they are sued. The sign of this relation flips in the 2SLS model. Here, underpricing increases in the predicted probability of lawsuits, consistent with the lawsuit avoidance hypothesis. Interestingly, greater underpricing does not appear to have much deterrence effect: the probability of being sued does not decrease in the instrumented underpricing return, at least not at conventional significance levels.

Lowry and Shu's study is sensitive to econometric concerns, and using more careful tools than prior work it finds evidence consistent with the proposition that firms use underpricing as a form of insurance against future litigation. Unfortunately, their empirical model is not able to gauge the *economic* magnitude of this effect (because their system cannot identify all relevant parameters). They are thus unable to say if litigation risk has a first-order effect on underpricing.

### *4.2. Price stabilization*

Rather than forming a symmetric distribution around some positive mean, underpricing returns typically peak sharply at zero and rarely fall below zero. In a controversial paper, Ruud (1993) takes these statistical regularities as her starting point to argue that IPOs are *not* deliberately underpriced. Rather, IPOs are priced at expected market value but offerings whose prices threaten to fall below the offer price are stabilized in after-market trading. Such price stabilization would tend to eliminate the left tail of the distribution of initial returns, and thus lead to the appearance of a positive average price jump. Thus what we observe in the data may not be the unconditional expectation of true initial returns but the mean conditional upon underwriter intervention in the aftermarket. Estimating the unobserved unconditional mean of the return distribution in a Tobit model, Ruud finds that average (logged) first-day returns are indeed close to zero.

This largely statistical view of the origins of IPO underpricing leaves little room for economics. Why would underwriters stabilize prices in the first place? Subsequent theoretical work on price stabilization has stressed its role in reducing underpricing. Benveniste, Busaba, and Wilhelm (1996) formalize Smith's (1986) notion of price stabilization as a mechanism that 'bonds' underwriters and investors. Because their dollar fees increase in gross proceeds, underwriters have a natural incentive to raise the offer

[15] Though note that empirically, stock turnover does correlate with underpricing, violating the order condition. Strictly speaking, the system estimated in Lowry and Shu relies for identification on the functional form of the probit equation modeling litigation risk, not on the use of instrumental variables.

price. Following a bookbuilding exercise, they could, for instance, overstate investor interest and price the IPO aggressively. Clever IPO investors will recognize this adverse incentive and, in the absence of any counteracting force, may not cooperate in the bookbuilding exercise in the first place. By implicitly committing themselves to price support—which is costlier, the more the offer price exceeds 'true' share value—underwriters may convince investors that the issue will not be intentionally overpriced.

According to Benveniste, Busaba, and Wilhelm (1996), the main beneficiaries of price support should be the institutional investors who participate in bookbuilding.[16] Using the Rock (1986) framework discussed in Section 3.1, Chowdhry and Nanda (1996) instead view retail investors as the main beneficiaries of price support. Analytically, we can think of price support as a put option written by the underwriter and held by the IPO investors, in the sense that stabilizing activities put a floor under early after-market prices and thus act as insurance against price falls. This may reduce the uninformed investors' winner's curse. Indeed, price support may be a more efficient way of counteracting the winner's curse than Rock's solution that all IPOs be underpriced on average, because price support is extended in the states of the world when uninformed investors suffer the most: overpriced offerings. Underpricing, on the other hand, is a blunter instrument because (absent price discrimination) it is offered to both uninformed and informed investors.

#### *4.2.1. How widespread is price support?*

Direct evidence of price support is limited because stabilizing activities are generally notifiable, if at all, only to market regulators, and not to investors at large. Thus it is hard to identify which IPOs were initially supported, how the intensity of intervention varied over time, and at what time support was withdrawn. Most work therefore relies on indirect evidence. For instance, one might investigate after-market microstructure data for behavior indicative of price support, and relate it to the underwriter's pre-market activities such as bookbuilding. This is particularly promising on NASDAQ, where underwriters can, and usually do, become market-makers for the companies they take public.

The microstructure variables of interest are the bid–ask spreads that underwriters charge (especially compared to competing market-makers who are not part of the original IPO syndicate); who provides 'price leadership' (by offering the best bid and ask prices); who trades with whom and in what trade sizes; what risks underwriters take in the after-market; and how much inventory dealers accumulate (indicating that they are net buyers). Schultz and Zaman (1994) and Hanley, Kumar, and Seguin (1993) find microstructure evidence consistent with widespread price support, especially among weak IPOs. Using proprietary Nasdaq data that identifies the transacting parties, Ellis,

[16] After all, if retail investors provide no pricing-relevant information in the pre-market, there is no reason to reward them by offering them price support.

Michaely, and O'Hara (2000) show that the lead IPO underwriter always becomes the dominant market-maker and accumulates sizeable inventories over the first 20 trading days. Underwriters buy back substantially more stock in 'cold' offerings (those that opened below their offer prices and never recovered in the first 20 days) than in 'hot' offerings (those that never fell below their offer prices in the first 20 days). These inventory accumulation patterns are strong evidence of price support activities, and indicate that such activities persist for a perhaps surprising length of time.

Asquith, Jones, and Kieschnick (1998) use a mixture-of-distributions approach to gauge how widespread price support is. Mixture-of-distributions models assume that the observed distribution is a mixture of two (or more) normal distributions with different means and standard deviations. They tend to be useful when modeling heavily skewed empirical distributions (such as underpricing returns). The technique estimates the fraction of the observations coming from each underlying distribution along with their means and standard deviations. Imposing the assumption that the data are generated by two (and no more) underlying distributions, one for supported offerings and one for unsupported ones, they argue that about half of all U.S. IPOs appear to have been supported in 1982–1983.

#### *4.2.2. Testable implications and evidence*

From the perspective of understanding why IPOs are (or appear to be) underpriced, the main empirical questions are (1) whether price support alone can account for positive underpricing returns and, assuming it cannot, (2) what effect the presence of price support has on the level of underpricing that results.

Asquith, Jones, and Kieschnick (1998) investigate whether observed underpricing is the byproduct of price support, as Ruud proposes, or whether it may have independent causes. Using the aforementioned mixture-of-distributions approach, they estimate the average underpricing returns for the two hypothesized distributions of supported and unsupported IPOs. If Ruud is correct in saying that there is no deliberate underpricing, then the initial return distribution of unsupported offerings should have a mean of zero. This, however, is not what Asquith, Jones, and Kieschnick find. Instead, the distribution interpreted as reflecting unsupported firms has mean underpricing of about 18 percent, while the distribution interpreted as reflecting supported IPOs has zero mean underpricing.

This suggests that underpricing is caused by factors other than price support. But the apparently widespread practice of price support may still affect how underpriced an IPO ends up being. We saw earlier that both Benveniste, Busaba, and Wilhelm (1996) and Chowdhry and Nanda (1996) predict that price support reduces the need to underprice, albeit for different reasons. Benveniste, Erdal, and Wilhelm (1998) try to distinguish between the two theories' contrasting predictions regarding who benefits from price support using detailed transactions data for 504 U.S. firms floated in 1993 and 1994. They find that it is overwhelmingly large (presumably institutional) traders who execute sell orders in stabilized offerings, rather than small (presumably retail) traders.

This lends support to the view that price support is offered mainly for the benefit of institutional investors, as modeled by Benveniste, Busaba, and Wilhelm (1996).

However, what remains unclear is whether, and by how much, the provision of price support reduces the required degree of underpricing.

### *4.3. Tax arguments*

Perhaps surprisingly, underpricing may be advantageous from a tax point of view. Rydqvist (1997) explores this possibility in the context of Swedish IPOs. The argument is simple. Before 1990, Sweden taxed employment income much more heavily than capital gains. This created an incentive to pay employees by allocating appreciating assets in lieu of salaries. One such appreciating asset is underpriced stock, allocated preferentially to the firm's own employees at the IPO. In 1990, the Swedish tax authorities made underpricing-related gains subject to income tax, removing the incentive to allocate underpriced stock to employees. Underpricing then fell from an average of 41% in 1980–1989 to 8% in 1990–1994.

A similar argument is put forward by Taranto (2003). A quirk of U.S. tax laws may increase senior managers' incentive to underprice their company's IPO. Holders of managerial or employee stock options pay tax in two steps. First, when they exercise the option, they pay income tax on the difference between the strike price and 'fair market value'. Second, when they eventually sell the underlying stock they acquired at exercise, they pay capital gains tax on the difference between 'fair market value' and the sale price. Since the capital gains tax liability is deferred, and since capital gains tax rates are typically lower than income tax rates, managers prefer 'fair market value' to be as low as possible. U.S. tax law considers 'fair market value' for options exercised in conjunction with an IPO to be the offer price, rather than the price that will prevail in the market once trading begins. This then generates an incentive to underprice.[17]

While it is unlikely that tax alone can explain why IPOs are underpriced, the tax benefit from underpricing may help explain the cross-section of underpricing returns. Taranto's (2003) empirical results are generally consistent with this argument, in that they show companies to be more underpriced the more they rely on managerial and employee stock options. However, it is possible that boards award stock options to protect managers from dilution in anticipation of the underwriter underpricing the stock. Thus the direction of causation is unclear.

## 5. Ownership and control

Going public is, in many cases, a step towards the eventual separation of ownership and control. Ownership matters for the effects it can have on management's incentives to

[17] A similar argument applies to restricted stock grants. Holders of unvested restricted stock can elect to pay income tax before vesting, based on 'fair market value'. Once the stock vests and is sold, capital gains tax becomes due on the difference between 'fair market value' and the sale price.

make optimal operating and investment decisions. In particular, where the separation of ownership and control is incomplete, an agency problem between non-managing and managing shareholders can arise (Jensen and Meckling, 1976): rather than maximizing expected shareholder value, managers may maximize the expected private utility of their control benefits (say, perquisite consumption) at the expense of outside shareholders.

Two principal models have sought to rationalize the underpricing phenomenon within the context of an agency cost approach. Their predictions are diametrically opposed: while Brennan and Franks (1997) view underpricing as a means to entrench managerial control and the attendant agency costs by *avoiding* monitoring by a large outside shareholder, Stoughton and Zechner's (1998) analysis instead suggests that underpricing may be used to minimize agency costs by *encouraging* monitoring.

### *5.1. Underpricing as a means to retain control*

Brennan and Franks (1997) argue underpricing gives managers the opportunity to protect their private benefits by allocating shares strategically when taking their company public. Managers seek to avoid allocating large stakes to investors for fear that their non-value-maximizing behavior would receive unwelcome scrutiny. Small outside stakes reduce external monitoring, owing to two free-rider problems. First, because it is a public good, shareholders will invest in a sub-optimally low level of monitoring (Shleifer and Vishny, 1986). Second, greater ownership dispersion implies that the incumbent managers benefit from a reduced threat of being ousted in a hostile takeover (Grossman and Hart, 1980). The role of underpricing in this view is to generate excess demand. Excess demand enables managers to ration investors so that they end up holding smaller stakes in the business.

#### *5.1.1. Testable implications and evidence*

The principal testable implication of the Brennan–Franks model is that underpricing results in excess demand and thus greater ownership dispersion. Using detailed data on individual bids and allocations in 69 U.K. IPOs completed between 1986 and 1989, Brennan and Franks confirm that large bids are discriminated against in favor of small ones, an effect that is stronger the more underpriced and oversubscribed the IPO. However, the protection of private benefits of control may not be the only reason why managers favor greater dispersion. Booth and Chua (1996) argue that owners value a more dispersed ownership structure because it likely results in a more liquid secondary market for their shares. In Zingales (1995), a more diffuse ownership structure helps managers negotiate a higher price when selling their controlling shareholding some time after the IPO. Thus, a link between underpricing and ownership dispersion is not sufficient evidence in favor of Brennan and Franks' model.

Zingales (1995) assumes that an IPO is frequently only the first stage in a multi-period sell-out strategy which will culminate in the complete transfer of ownership and control from the original founders to new owners. Brennan and Franks, on the other hand, assume that the IPO is designed to *prevent* a transfer of control in spite of the partial transfer of ownership. Who is right? The empirical evidence is more nearly consistent with the staged-sale notion. Pagano, Panetta, and Zingales (1998) document that most Italian IPOs are followed by private sales of controlling blocks to large outside investors. Indeed, control turnover is twice as common in newly listed firms as in the universe of unlisted companies. In the U.S., control turnover in the first five years is 29 percent in IPO firms with at least five years of trading history prior to flotation and 13 percent for younger companies (Mikkelson, Partch, and Shah, 1997). Similarly, officers and directors in U.S. IPOs on average own 66 percent of equity before the IPO and 44 percent immediately afterwards, which is reduced to 29 percent over the subsequent five years, and to 18 percent ten years later (Mikkelson, Partch, and Shah, 1997).

Underpricing-induced ownership dispersion is not the only way to protect private benefits of control. An obvious alternative is to put in place takeover defenses or simply to issue non-voting stock. Field and Karpoff (2002) show that a majority of U.S. firms deploy at least one takeover defense just before going public, especially when private benefits of control appear large and internal monitoring mechanisms look weak—that is, when managers' compensation packages are unusually generous, their own equity stakes are small, and non-directors play a smaller role in corporate governance. Interestingly, however, these firms are still underpriced—though we do not know whether they are *less* underpriced than firms that choose to entrench their managers via the Brennan–Franks mechanism—so the protection of private benefits is unlikely to be the only explanation of underpricing, at least in the U.S.

Issuing non-voting shares would guarantee that managers could retain control of the company and all attendant private benefits. Whether it dominates the Brennan–Franks underpricing mechanism is an empirical matter. Non-voting shares tend to trade at lower multiples than voting shares. This voting discount could be smaller or larger than the money left on the table via underpricing. Smart and Zutter (2003) find that U.S. companies that issue non-voting stock in their IPOs are less underpriced and have higher institutional ownership after the IPO. This is consistent with the notion that non-voting stock can substitute for the Brennan–Franks mechanism. At the same time, Smart and Zutter find that non-voting IPO shares trade at lower multiples, though they do not investigate how these compare with the monetary benefit of reduced underpricing.

Arugaslan, Cook, and Kieschnick (2004) take issue with Smart and Zutter's (2003) study on econometric grounds, pointing out that the main reason why IPOs involving non-voting stock are less underpriced than voting-stock IPOs is that they are larger. Size in turn is an important determinant of institutional investors' stock selection, and may thus be driving the higher post-IPO institutional ownership Smart and Zutter observe among non-voting-stock IPOs.

Underpricing and the resulting excess demand will shield managers from outside monitoring only to the extent that outside investors do not assemble large blocks once trading has begun. Brennan and Franks (1997) suggest that such open-market purchases may not be profitable. If the market anticipates the gains that would accrue if management were monitored by a sufficiently large outside shareholder, prices will rise in response to large-scale buying. This will tend to make it unprofitable to assemble a large block of shares in the aftermarket, the more so the more diffuse the ownership structure is to start with. Empirically, however, this argument meets with little success. Field and Sheehan (2004) find next to no relation between the creation of new blocks after the IPO and the level of underpricing at the IPO.

### *5.2. Underpricing as a means to reduce agency costs*

Brennan and Franks (1997) implicitly assume that, in the wake of the separation of ownership and control, managers try to maximize their expected private utility by entrenching their control benefits. However, it could be argued that managers should actually seek to minimize, rather than maximize, their scope for extracting private benefits of control. Why? Agency costs are ultimately borne by the owners of a company, in the form of lower IPO proceeds and a lower subsequent market value for their shares. To the extent that managers are part-owners, they bear at least some of the costs of their own non-profit-maximizing behavior. If their stakes are large enough so that the agency costs they bear outweigh the private benefits they enjoy, it will be in their interest to reduce, not entrench, their discretion.

Based on this intuition, Stoughton and Zechner (1998) observe that, in contrast to Brennan and Franks, it may be value-enhancing to allocate shares to a large outside investor who is able to monitor managerial actions. Monitoring is a public good as all shareholders benefit, whether or not they contribute to its provision. Since a large shareholder will monitor only in so far as this is privately optimal (which is a function of the size of her stake), there will be too little monitoring from the point of view of both shareholders and incumbent managers. To encourage better monitoring, managers may try to allocate a particularly large stake to an investor. However, if the allocation is sub-optimally large from the investor's point of view (say, because it is not easily diversified), an added incentive may be offered in the form of underpricing. Such underpricing may not even represent an opportunity cost: in the absence of monitoring, the firm would have had to be floated at a lower price anyway, owing to outside shareholders anticipating higher agency costs.

A closer look at Stoughton and Zechner's model is constructive. The selling mechanism is modeled as a two-stage process akin to bookbuilding. In the first stage, issuers extract the demand schedule from a likely monitor and set the offer price such that this investor optimally demands a large enough number of shares to subsequently engage in effective monitoring. In the second stage, small investors are allocated shares at the same price (unless price discrimination is possible, which in practice it rarely is). Rationing is observed at this stage, as small investors would like to buy further shares at the low offer price.

Why are the predictions of Brennan and Franks and Stoughton and Zechner so different? There are at least two reasons. The first is the different institutional environments in which the models are placed. Brennan and Franks effectively model an IPO mechanism involving prices that are fixed rather than responsive to demand and shares that are allocated pro rata. Stoughton and Zechner, on the other hand, model a bookbuilding regime with discretionary allocations. In a pro-rata regime Stoughton and Zechner would have difficulty allocating enough stock to the large shareholder to ensure effective monitoring. In a bookbuilding regime, Brennan and Franks would not need to underprice as much to discriminate against large investors: absent pro rata allocation rules, the issuer (and underwriter) could simply select which investors to exclude from allocations. This illustrates the importance of the institutional assumptions made in IPO modeling.

Second, Stoughton and Zechner assume that managers internalize the agency costs they impose on outside investors, via the lower price that investors are willing to pay for the stock. This internalization is absent from the Brennan–Franks model.

The ownership and control dimension is a promising, albeit nascent, field in the study of IPO underpricing. Much more empirical evidence is needed before we can assess the validity of the theoretical contributions and before we can say whether control considerations are of first or second-order importance when offer prices are set.

## 6. Behavioral explanations

In the late 1990s initial returns increased substantially. As pointed out in Section 2, U.S. issuers left an aggregate of $62 billion on the table in 1999 and 2000 alone. Many researchers are doubtful whether informational frictions, the risk of lawsuits, or control considerations could possibly be severe enough to warrant underpricing on this scale. As a consequence, some argue we should turn to behavioral explanations for IPO underpricing. Behavioral theories assume either the presence of 'irrational' investors who bid up the price of IPO shares beyond true value, or that issuers are subject to behavioral biases and therefore fail to put pressure on the underwriting banks to have underpricing reduced. This literature is still in its infancy.[18]

The IPO market is a good setting in which to study the effect of 'irrational' investors on stock prices. IPO firms by definition have no prior share price history and tend to be young, immature, and relatively informationally opaque. Not surprisingly, therefore, they are hard to value, and it seems reasonable to assume that investors will have a wide range of priors about their market values. In Section 6.2, we will review one recent theory of IPO underpricing that builds on this assumption. In Section 6.3, we will turn to a model of behaviorally challenged managers. We begin, however, with a discussion of a model of rational 'informational cascades'.

### *6.1. Cascades*

Welch (1992) shows that 'informational cascades' can develop in some forms of IPOs if investors make their investment decisions sequentially: later investors can condition

[18] For a survey of behavioral corporate finance more generally, see Chapter 4 by Baker, Ruback, and Wurgler in this volume.

their bids on the bids of earlier investors, rationally disregarding their own information. Successful initial sales are interpreted by subsequent investors as evidence that earlier investors held favorable information, encouraging later investors to invest whatever their own information. Conversely, disappointing initial sales can dissuade later investors from investing irrespective of their private signals. As a consequence, demand either snowballs or remains low over time.

The possibility of cascades gives market power to early investors who can 'demand' more underpricing in return for committing to the IPO and thus starting a positive cascade. It is in this sense that cascades may play a role in explaining IPO underpricing. But cascades are not inevitable. In bookbuilding cascades do not develop because the underwriter can maintain secrecy over the development of demand in the book. Less underpricing is therefore required. Bookbuilding also offers the issuer the valuable option to increase the offer size if demand turns out to be high (either unconditionally, by issuing more shares, or conditionally, by giving the underwriter a so called overallotment option).[19]

If investors can communicate freely, cascades also do not form, for then investors can learn the entire distribution of signals. Yet Welch (1992) shows that issuers are better off with cascades than with free communication, because free communication aggregates all available information which maximizes the issuing company's informational disadvantage compared to investors. Moreover, preventing free communication reduces the chance that one investor's negative information becomes widely known, and so reduces the likelihood that the IPO will fail.

### *6.1.1. Testable implications and evidence*

Arguing that underwriters with national reach can more easily segment the market and so prevent communication among investors than can local or regional underwriters, Welch (1992) derives several testable implications. Most importantly, compared to locally or regionally distributed IPOs, IPOs managed by national underwriters are predicted to be less underpriced. While this implication has not been tested explicitly, it relates to the literature on the relation between underpricing and underwriter reputation discussed earlier, at least to the extent that market-share or tombstone-ranking measures of reputation correlate with the bank's geographic reach. Recall that the sign on the relation between underpricing and underwriter reputation has flipped since the 1970s and 1980s, which implies mixed support for the cascades model.

On the other hand, Welch (1992) also stresses the factors determining which issuer chooses which type of underwriter. Specifically, in the presence of fixed costs, the more risk averse and capital-constrained the issuer, the greater the benefits of national distribution. Thus the choice of underwriter is not random, implying that simple OLS estimates

[19] Overallotment options entitle the underwriter to purchase additional shares (usually 15% of the offer size) from the issuer at the IPO price. Such options are sometimes called 'green shoes'.

of the relation between underpricing and the bank's geographic reach (or underwriter reputation) must be interpreted with caution. This reinforces Habib and Ljungqvist's (2001) argument discussed in Section 3.1, albeit on the basis of a different model of IPO underpricing.

At a more basic level, Amihud, Hauser, and Kirsh's (2003) analysis of demand and allocations in Israeli IPOs supports Welch's (1992) prediction that demand is either extremely low or there is oversubscription, with few cases in between.

In conclusion, Welch's cascades model remains one of the least explored explanations of IPO underpricing.

### *6.2. Investor sentiment*

Behavioral finance is interested in the effect on stock prices of 'irrational' or 'sentiment' investors. The potential for such an effect would seem particularly large in the case of IPOs, since IPO firms are young, immature, and relatively informationally opaque and hence hard to value. The first paper to model an IPO company's optimal response to the presence of sentiment investors is Ljungqvist, Nanda, and Singh (2004). They assume some sentiment investors hold optimistic beliefs about the future prospects for the IPO company. The issuer's objective is to capture as much of the 'surplus' under the sentiment investors' downward-sloping demand curve as possible, that is, to maximize the excess valuation over the fundamental value of the stock. Flooding the market with stock will depress the price, so the optimal strategy involves holding back stock in inventory to keep the price from falling. Eventually, nature reveals the true value of the stock and the price reverts to fundamental value. That is, in the long-run IPO returns are negative, consistent with the empirical evidence in Ritter (1991) and others. This assumes the existence of short sale constraints, or else arbitrageurs would trade in such a way that prices reflected fundamental value even in the short term.

Regulatory constraints on price discrimination and inventory holding prevent the issuer from implementing such a strategy directly. Instead, the optimal mechanism involves the issuer allocating stock to 'regular' institutional investors for subsequent resale to sentiment investors, at prices the regulars maintain by restricting supply. Because the hot market can end prematurely, carrying IPO stock in inventory is risky, so to break even in expectation regulars require the stock to be underpriced—even in the absence of asymmetric information. However, the offer price still exceeds fundamental value, as it capitalizes the regulars' expected gain from trading with the sentiment investors, and so the issuer benefits from this mechanism.

#### *6.2.1. Testable implications and evidence*

The model generates a number of new and refutable empirical predictions. Most obviously, the model predicts that companies going public in a hot market subsequently underperform, both relative to the first-day price and to the offer price. Underperformance relative to the first-day price is not surprising; it follows from the twin assumptions of sentiment investors and short-sale constraints (see Miller, 1977). Underperformance relative to the offer price is a stronger prediction. It follows because the offer price exceeds fundamental value by an amount equal to the issuer's share in the surplus extracted from the sentiment investors. Purnanandam and Swaminathan (2004) lend support to the prediction that the offer price can exceed fundamental value. They show that compared to its industry peers' multiples, the median IPO firm in 1980–1997 was overpriced at the offer by 50%. Interestingly, it is the firms that are most overpriced in this sense which subsequently underperform. Cook, Jarrell, and Kieschnick (2003) refine this analysis by conditioning on hot and cold markets. They find that IPO firms trade at higher valuations only in hot markets, consistent with the spirit of the Ljungqvist, Nanda, and Singh (2004) model. Cornelli, Goldreich, and Ljungqvist (2006) use data from the grey market (the when-issued market that precedes European IPOs and that involves mostly retail traders) to show that long-run underperformance is concentrated among those IPOs whose grey market prices were particularly high. They also report evidence suggesting that grey market investors do not update their prior beliefs about the value of an IPO in an unbiased fashion.

Ofek and Richardson (2003) show that high initial returns occur when institutions sell IPO shares to retail investors on the first day, and that such high initial returns are followed by sizeable reversals to the end of 2000, when the 'dot-com bubble' eventually burst. This is precisely the pattern Ljungqvist, Nanda, and Singh (2004) predict.

At the heart of Ljungqvist, Nanda, and Singh's (2004) story is the idea that banks market IPOs and that it matters whom they target in their marketing. Cook, Kieschnick, and Van Ness (2006) find a significant positive relation between promotional activities (proxied by the number of newspaper articles mentioning the IPO firm in the prior six months) and the valuations at which IPOs are sold, which they interpret as evidence that investment bankers manage to sell overvalued IPO stock to retail investors to the benefit of the issuer and the investment bank's regular clients.

Using German data on IPO trading by 5,000 retail customers of an online broker, Dorn (2002) documents that retail investors overpay for IPOs following periods of high underpricing in recent IPOs, and for IPOs that are in the news. Consistent with the Ljungqvist, Nanda, and Singh (2004) model, he also shows that 'hot' IPOs pass from institutional into retail hands. Over time, high initial returns are reversed as net purchases by retail investors subside, eventually resulting in underperformance over the first six to 12 months after the IPO.

The model may also be able to reconcile the conflicting empirical evidence regarding the relation between underpricing and long-run performance. Ritter (1991) documents that underpricing and long-run performance are negatively related, while Krigman, Shaw, and Womack (1999) find a positive relation. In the Ljungqvist, Nanda, and Singh (2004) model, the relation is not necessarily monotonic. In particular, the relation is negative only if the probability of the hot market ending is small. If the hot market is highly likely to end, the issuer optimally reduces the offer size, implying regular investors hold smaller inventories and so require less underpricing to break even. At the same time, the reduction in offer size aggravates long-run underperformance, given the negative slope of the sentiment demand curve.

Recall from Section 3.1 that the empirical evidence on the relation between underwriter reputation and underpricing is mixed. Consistent with evidence from the 1990s (Beatty and Welch, 1996), Ljungqvist, Nanda, and Singh (2004) predict that underpricing *in*creases in underwriter reputation. Underwriters enjoying a large IPO deal flow can more easily punish regular investors who attempt to free-ride on the inventory-holding strategy by dumping their shares prematurely, before the price falls. This in turn implies that the more active banks can underwrite larger IPOs, as more inventory can be held over time. Since underpricing is compensation for the expected inventory losses in the face of a non-zero probability that the hot market will end before all inventory has been unloaded, the more active underwriters will be associated with greater underpricing.

### *6.3. Prospect theory and mental accounting*

Loughran and Ritter (2002) propose an explanation for IPO underpricing that stresses behavioral biases among the decision-makers of the IPO firm, rather than among investors. Combining prospect theory-style reference-point preferences with Thaler's (1980, 1985) notion of mental accounting, Loughran and Ritter argue that issuers fail to 'get upset' about leaving millions of dollars 'on the table' in the form of large first-day returns because they tend to sum the wealth loss due to underpricing with the (often larger) wealth gain on retained shares as prices jump in the after-market. Such 'complacent' behavior benefits the investment bank if investors engage in rent-seeking to increase their chances of being allocated underpriced stock.

Loughran and Ritter (2002) assume that the decision-maker's initial valuation beliefs are reflected in the mean of the indicative price range reported in the issuing firm's IPO registration statement. This belief serves as a reference point against which the gain or loss from (as opposed to the expected utility of) the outcome of the IPO can be assessed. The offer price for an IPO routinely differs from this reference point, either because the bank 'manipulated' the decision-maker's expectations by low-balling the price range, or in reflection of information revealed during marketing efforts directed at institutional investors. As argued earlier, offer prices appear only to 'partially adjust' (Hanley, 1993) in the sense that large positive revisions from the reference point are associated with large initial price increases from the offer price during the first day of trading. Such partial adjustment is consistent with both the Benveniste and Spindt (1989) information-acquisition model of IPO underpricing and Loughran and Ritter's complacency argument.

The decision-maker perceives a positive revision from the reference point as a wealth gain (assuming he retains shares after the IPO). At the same time, a positive initial return is perceived as a wealth loss under the assumption that shares could have been sold at the higher first-day trading price. If the perceived gain exceeds the underpricing loss, the decision-marker is satisfied with the IPO underwriter's performance at the IPO.

### 6.3.1. Testable implications and evidence

Ljungqvist and Wilhelm (2005) use the structure suggested by Loughran and Ritter's (2002) behavioral perspective to test whether the CEOs of recent IPO firms make *subsequent* decisions consistent with a behavioral measure of their perception of the IPO's outcome. Specifically, they investigate whether CEOs deemed 'satisfied' with the underwriter's performance according to Loughran and Ritter's story are more likely to hire their IPO underwriters to lead-manage later seasoned equity offerings. Controlling for other known factors, IPO firms are less likely to switch underwriters for their SEO when they were deemed 'satisfied' with the IPO underwriter's performance. Underwriters also appear to benefit from behavioral biases in the sense that they extract higher fees for subsequent transactions involving 'satisfied' decision-makers.

While these tests suggest there is explanatory power in the behavioral model, they do not speak directly to whether deviations from expected utility maximization determine patterns in IPO initial returns. More work is needed.

## 7. Concluding remarks

The empirical IPO literature is now fairly mature. We know that IPOs are underpriced in virtually all countries and that the number of companies going public and the extent of underpricing fluctuate over time. There is a large body of theoretical work explaining IPO underpricing, and most theories have been subjected to rigorous empirical testing. Broadly speaking, the empirical evidence supports the view that information frictions (including agency conflicts between the issuing company and its investment bank) have a first-order effect on underpricing. Specifically,

- The bulk of underpricing-related gains accrue to informed (or at least institutional) investors; uninformed (or at least retail) investors earn little or no excess returns from investing in IPOs.
- In the cross-section, underpricing increases in the ex ante uncertainty surrounding a firm's valuation.
- There is ample evidence suggesting that some investors are informed and that their information influences the investment bank's choice of offer price.

At the same time, the enormous variation in the extent of underpricing over time raises doubt in some people's mind whether information-based explanations on their own can account for the huge amounts of money left on the table in hot markets, such as the internet bubble of 1998–2000.

Against this background, vigorous debate continues between two broad views of what causes underpricing: the Benveniste and Spindt (1989) perspective which emphasizes the necessity of underpricing if the underwriter is to efficiently extract pricing-relevant information from better informed investors and thereby maximize the issuer's expected proceeds, and the agency view commonly associated with Jay Ritter's work which stresses the self-interested nature of investment banks. The sometimes strident tone of this debate on both sides belies the fact that the truth is probably somewhere in between. For the information-acquisition mechanism of Benveniste and Spindt to work, underwriters need to be given discretion over the way they price and allocate IPO shares. Allocation discretion, in turn, may well aggravate an agency problem between the issuer and its banker arising from the fact that bankers deal repeatedly with institutional investors but infrequently with issuers.

Arising from this debate, there is continued interest in at least four areas:
(a) behavioral approaches to explain why the extent of underpricing varies so much over time;
(b) tests exploiting cross-country differences in institutional frameworks;
(c) work shedding light on the allegedly conflicted behavior of investment banks during the market boom of the late 1990s; and
(d) the potential for using auction mechanisms to price and allocate IPOs.

## References

Aggarwal, R., Prabhala, N.R., Puri, M., 2002. Institutional allocation in initial public offerings: Empirical evidence. Journal of Finance 57, 1421–1442.
Akerlof, G., 1970. The market for "Lemons": Quality uncertainty and the market mechanism. Quarterly Journal of Economics 84, 488–500.
Allen, F., Faulhaber, G.R., 1989. Signaling by underpricing in the IPO market. Journal of Financial Economics 23, 303–323.
Amihud, Y., Hauser, S., Kirsh, A., 2003. Allocations, adverse selection and cascades in IPOs: Evidence from the Tel Aviv Stock Exchange. Journal of Financial Economics 68, 137–158.
Arugaslan, O., Cook, D.O., Kieschnick, R., 2004. Monitoring as a motivation for IPO underpricing. Journal of Finance 59, 2403–2420.
Asquith, D., Jones, J.D., Kieschnick, R., 1998. Evidence on price stabilization and underpricing in early IPO returns. Journal of Finance 53, 1759–1773.
Baron, D.P., 1982. A model of the demand for investment banking advising and distribution services for new issues. Journal of Finance 37, 955–976.
Baron, D.P., Holmström, B., 1980. The investment banking contract for new issues under asymmetric information: Delegation and the incentive problem. Journal of Finance 35, 1115–1138.
Beatty, R.P., Ritter, J.R., 1986. Investment banking, reputation, and the underpricing of initial public offerings. Journal of Financial Economics 15, 213–232.
Beatty, R.P., Welch, I., 1996. Issuer expenses and legal liability in initial public offerings. Journal of Law and Economics 39, 545–602.
Beller, A.L., Terai, T., Levine, R.M., 1992. Looks can be deceiving: A comparison of initial public offering procedures under Japanese and U.S. securities laws. Law and Contemporary Problems 55, 77–118.
Benveniste, L.M., Busaba, W.Y., Wilhelm Jr., W.J., 1996. Price stabilization as a bonding mechanism in new equity issues. Journal of Financial Economics 42, 223–255.
Benveniste, L.M., Busaba, W.Y., Wilhelm Jr., W.J., 2002. Information externalities and the role of underwriters in primary equity markets. Journal of Financial Intermediation 11, 61–86.
Benveniste, L.M., Erdal, S.M., Wilhelm Jr., W.J., 1998. Who benefits from secondary market stabilization of IPOs? Journal of Banking and Finance 22, 741–767.
Benveniste, L.M., Ljungqvist, A., Wilhelm Jr., W.J., Yu, X., 2003. Evidence of information spillovers in the production of investment banking services. Journal of Finance 58, 577–608.
Benveniste, L.M., Spindt, P.A., 1989. How investment bankers determine the offer price and allocation of new issues. Journal of Financial Economics 24, 343–361.

Benveniste, L.M., Wilhelm Jr., W.J., 1990. A comparative analysis of IPO proceeds under alternative regulatory environments. Journal of Financial Economics 28, 173–207.

Biais, B., Bossaerts, P., Rochet, J.-C., 2002. An optimal IPO mechanism. Review of Economic Studies 69, 117–146.

Booth, J.R., Chua, L., 1996. Ownership dispersion, costly information and IPO underpricing. Journal of Financial Economics 41, 291–310.

Booth, J.R., Smith, R., 1986. Capital raising, underwriting and the certification hypothesis. Journal of Financial Economics 15, 261–281.

Bradley, D.J., Jordan, B.D., 2002. Partial adjustment to public information and IPO underpricing. Journal of Financial and Quantitative Analysis 37, 595–616.

Brennan, M.J., Franks, J., 1997. Underpricing, ownership and control in initial public offerings of equity securities in the U.K. Journal of Financial Economics 45, 391–413.

Busaba, W.Y., Benveniste, L.M., Guo, R.-J., 2001. The option to withdraw IPOs during the premarket. Journal of Financial Economics 60, 73–102.

Carter, R.B., Manaster, S., 1990. Initial public offerings and underwriter reputation. Journal of Finance 45, 1045–1067.

Chowdhry, B., Nanda, V., 1996. Stabilization, syndication, and pricing of IPOs. Journal of Financial and Quantitative Analysis 31, 25–42.

Cook, D.O., Jarrell, S.L., Kieschnick, R., 2003. Investor sentiment and IPO cycles. Working Paper, University of Texas at Dallas (unpublished).

Cook, D.O., Kieschnick, R., Van Ness, R.A., 2006. On the marketing of IPOs, Journal of Financial Economics, in press.

Cornelli, F., Goldreich, D., 2001. Bookbuilding and strategic allocation. Journal of Finance 56, 2337–2369.

Cornelli, F., Goldreich, D., 2003. Bookbuilding: How informative is the order book? Journal of Finance 58, 1415–1443.

Cornelli, F., Goldreich, D., Ljungqvist, A., 2006. Investor sentiment and pre-IPO markets. Journal of Finance 61, 1187–1216.

Derrien, F., Womack, K.L., 2002. Auctions vs. bookbuilding and the control of underpricing in hot IPO markets. Review of Financial Studies 16, 31–61.

Dorn, D., 2002. Does sentiment drive the retail demand for IPOs? Working Paper, Columbia University (unpublished).

Drake, P.D., Vetsuypens, M.R., 1993. IPO underpricing and insurance against legal liability. Financial Management 22, 64–73.

Dunbar, C.G., 1995. The use of warrants as underwriter compensation in initial public offerings. Journal of Financial Economics 38, 59–78.

Dunbar, C.G., 2000. Factors affecting investment bank initial public offering market share. Journal of Financial Economics 55, 3–41.

Edelen, R.M., Kadlec, G.B., 2005. Issuer surplus and the partial adjustment of IPO prices to public information. Journal of Financial Economics 77, 347–373.

Ellis, K., Michaely, R., O'Hara, M., 2000. When the underwriter is the market maker: An examination of trading in the IPO after-market. Journal of Finance 55, 1039–1074.

Fernando, C.S., Gatchev, V.A., Spindt, P.A., 2005. Wanna dance? How firms and underwriters choose each other. Journal of Finance 60, 2437–2469.

Field, L.C., Karpoff, J.M., 2002. Takeover defenses at IPO firms. Journal of Finance 57, 1857–1889.

Field, L.C., Sheehan, D.P., 2004. IPO underpricing and outside blockholdings. Journal of Corporate Finance 10, 263–280.

Grinblatt, M., Hwang, C.Y., 1989. Signaling and the pricing of new issues. Journal of Finance 44, 393–420.

Grossman, S., Hart, O., 1980. Takeover bids, the free-rider problem and the theory of the corporation. Bell Journal of Economics 11, 42–64.

Habib, M.A., Ljungqvist, A., 1998. Underpricing and IPO proceeds: A note. Economics Letters 61, 381–383.

Habib, M.A., Ljungqvist, A., 2001. Underpricing and entrepreneurial wealth losses in IPOs: Theory and evidence. Review of Financial Studies 14, 433–458.

Hanley, K., 1993. The underpricing of initial public offerings and the partial adjustment phenomenon. Journal of Financial Economics 34, 231–250.
Hanley, K., Kumar, A.A., Seguin, P.J., 1993. Price stabilization in the market for new issues. Journal of Financial Economics 34, 177–197.
Hanley, K., Wilhelm, W.J., 1995. Evidence on the strategic allocation of initial public offerings. Journal of Financial Economics 37, 239–257.
Heckman, J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.
Hensler, D.A., 1995. Litigation costs and the underpricing of initial public offerings. Managerial and Decision Economics 16, 111–128.
Hughes, P.J., 1986. Signaling by direct disclosure under asymmetric information. Journal of Accounting and Economics 8, 119–142.
Hughes, P.J., Thakor, A.V., 1992. Litigation risk, intermediation, and the underpricing of initial public offerings. Review of Financial Studies 5, 709–742.
Ibbotson, R.G., 1975. Price performance of common stock new issues. Journal of Financial Economics 2, 235–272.
Ibbotson, R.G., Jaffe, J.F., 1975. "Hot issue" markets. Journal of Finance 30, 1027–1042.
Jagannathan, R., Sherman, A., 2006. Why do IPO auctions fail? Mimeo, University of Notre Dame.
James, C., Wier, P., 1990. Borrowing relationships, intermediation and the cost of issuing public securities. Journal of Financial Economics 28, 149–171.
Jegadeesh, N., Weinstein, M., Welch, I., 1993. An empirical investigation of IPO returns and subsequent equity offerings. Journal of Financial Economics 34, 153–175.
Jenkinson, T.J., 1990. Initial public offerings in the United Kingdom, the United States, and Japan. Journal of the Japanese and International Economies 4, 428–449.
Jenkinson, T.J., Jones, H., 2004. Bids and allocations in European IPO bookbuilding. Journal of Finance 59, 2013–2040.
Jensen, M., Meckling, W., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics 3, 306–360.
Keloharju, M., 1993. The winner's curse, legal liability, and the long-run price performance of initial public offerings in Finland. Journal of Financial Economics 34, 251–277.
Koh, F., Walter, T., 1989. A direct test of rock's model of the pricing of unseasoned issues. Journal of Financial Economics 23, 251–272.
Krigman, L., Shaw, W.H., Womack, K.L., 1999. The persistence of IPO mispricing and the predictive power of flipping. Journal of Finance 54, 1015–1044.
Kunz, R.M., Aggarwal, R., 1994. Why initial public offerings are underpriced: Evidence from Switzerland. Journal of Banking and Finance 18, 705–724.
Lee, P.J., Taylor, S.L., Walter, T.S., 1996. Australian IPO pricing in the short and long run. Journal of Banking and Finance 20, 1189–1210.
Lee, P.M., Wahal, S., 2004. Grandstanding, certification, and the underpricing of venture capital backed IPOs. Journal of Financial Economics 73, 375–407.
Levis, M., 1990. The winner's curse problem, interest costs, and the underpricing of initial public offerings. Economic Journal 100, 76–89.
Li, X., Masulis, R., 2003. Venture capital investments by IPO underwriters: Certification or conflict of interest? Working Paper, Vanderbilt University (unpublished).
Ljungqvist, A., 1997. Pricing initial public offerings: Further evidence from Germany. European Economic Review 41, 1309–1320.
Ljungqvist, A., 2003. Conflicts of interest and efficient contracting in IPOs. Working Paper, New York University (unpublished).
Ljungqvist, A., Jenkinson, T.J., Wilhelm, W.J., 2003. Global integration in primary equity markets: The role of U.S. banks and U.S. investors. Review of Financial Studies 16, 63–99.
Ljungqvist, A., Nanda, V., Singh, R., 2004. Hot markets, investor sentiment, and IPO pricing. Journal of Business, in press.

Ljungqvist, A., Wilhelm, W.J., 2002. IPO allocations: Discriminatory or discretionary? Journal of Financial Economics 65, 167–201.

Ljungqvist, A., Wilhelm, W.J., 2003. IPO pricing in the dot-com bubble. Journal of Finance 58, 723–752.

Ljungqvist, A., Wilhelm, W.J., 2005. Does prospect theory explain IPO market behavior? Journal of Finance 60, 1759–1790.

Logue, D., 1973. Premia on unseasoned equity issues, 1965–69. Journal of Economics and Business 25, 133–141.

Loughran, T., Ritter, J.R., 2002. Why don't issuers get upset about leaving money on the table in IPOs? Review of Financial Studies 15, 413–443.

Loughran, T., Ritter, J.R., 2004. Why has IPO underpricing increased over time? Financial Management 33, 5–37.

Lowry, M., Schwert, G.W., 2004. Is the IPO pricing process efficient? Journal of Financial Economics 71, 3–26.

Lowry, M., Shu, S., 2002. Litigation risk and IPO underpricing. Journal of Financial Economics 65, 309–335.

Mauer, E., Senbet, L., 1992. The effect of the secondary market on the pricing of initial public offerings: Theory and evidence. Journal of Financial and Quantitative Analysis 27, 55–79.

Megginson, W., Weiss, K.A., 1991. Venture capitalist certification in initial public offerings. Journal of Finance 46, 879–903.

Michaely, R., Shaw, W.H., 1994. The pricing of initial public offerings: Tests of adverse-selection and signaling theories. Review of Financial Studies 7, 279–319.

Mikkelson, W.H., Partch, M.M., Shah, K., 1997. Ownership and operating performance of firms that go public. Journal of Financial Economics 44, 281–307.

Miller, E.M., 1977. Risk, uncertainty, and divergence of opinion. Journal of Finance 32, 1151–1168.

Miller, R.E., Reilly, F.K., 1987. An examination of mispricing, returns, and uncertainty for initial public offerings. Financial Management 16, 33–38.

Muscarella, C.J., Vetsuypens, M.R., 1989. A simple test of Baron's model of IPO underpricing. Journal of Financial Economics 24, 125–135.

Nanda, V., Yun, Y., 1997. Reputation and financial intermediation: An empirical investigation of the impact of IPO mispricing on underwriter market value. Journal of Financial Intermediation 6, 39–63.

Ofek, E., Richardson, M., 2003. Dot-com mania: The rise and fall of internet stock prices. Journal of Finance 58, 1113–1138.

Pagano, M., Panetta, F., Zingales, L., 1998. Why do companies go public? An empirical analysis. Journal of Finance 53, 27–64.

Purnanandam, A.K., Swaminathan, B., 2004. Are IPOs really underpriced? Review of Financial Studies 17, 811–848.

Reuter, J., 2004. Are IPO allocations for sale? Evidence from the mutual fund industry, Journal of Finance, in press.

Ritter, J.R., 1984. The hot issue market of 1980. Journal of Business 57, 215–240.

Ritter, J.R., 1987. The costs of going public. Journal of Financial Economics 19, 269–282.

Ritter, J.R., 1991. The long-run performance of initial public offerings. Journal of Finance 46, 3–27.

Ritter, J.R., 2003. Investment banking and securities issuance. In: Constantinides, G.M., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. North-Holland, Amsterdam.

Ritter, J.R., Welch, I., 2002. A review of IPO activity, pricing, and allocations. Journal of Finance 57, 1795–1828.

Rock, K., 1986. Why new issues are underpriced. Journal of Financial Economics 15, 187–212.

Ruud, J.S., 1993. Underwriter price support and the IPO underpricing puzzle. Journal of Financial Economics 34, 135–151.

Rydqvist, K., 1994. Compensation, participation restrictions and the underpricing of initial public offerings. Mimeo, Stockholm School of Economics.

Rydqvist, K., 1997. IPO underpricing as tax-efficient compensation. Journal of Banking and Finance 21, 295–313.

Schultz, P.H., Zaman, M.A., 1994. After-market support and underpricing of initial public offerings. Journal of Financial Economics 35, 199–219.

Sherman, A., Titman, S., 2002. Building the IPO order book: Underpricing and participation limits with costly information. Journal of Financial Economics 65, 3–29.

Shleifer, A., Vishny, R., 1986. Large stakeholders and corporate control. Journal of Political Economy 94, 461–488.

Slovin, M.B., Sushka, M.E., Bendeck, Y.M., 1994. Seasoned common stock issuance following an IPO. Journal of Banking and Finance 18, 207–226.

Smart, S.B., Zutter, C.J., 2003. Control as a motivation for underpricing: A comparison of dual- and single-class IPOs. Journal of Financial Economics 69, 85–110.

Smith, C.W., 1986. Investment banking and the capital acquisition process. Journal of Financial Economics 15, 3–29.

Spatt, C., Srivastava, S., 1991. Preplay communication, participation restrictions, and efficiency in initial public offerings. Review of Financial Studies 4, 709–726.

Spiess, D.K., Pettway, R.H., 1997. The IPO and first seasoned equity sale: Issue proceeds, owner/managers' wealth, and the underpricing signal. Journal of Banking and Finance 21, 967–988.

Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557–586.

Stoughton, N.M., Zechner, J., 1998. IPO mechanisms, monitoring and ownership structure. Journal of Financial Economics 49, 45–78.

Stoughton, N.M., Wong, K.P., Zechner, J., 2001. IPOs and product quality. Journal of Business 74.

Taranto, M., 2003. Employee stock options and the underpricing of initial public offerings. Working Paper, University of Pennsylvania (unpublished).

Thaler, R., 1980. Toward a positive theory of consumer choice. Journal of Economic Behavior and Organization 1, 39–60.

Thaler, R., 1985. Mental accounting and consumer choice. Marketing Science 4, 199–214.

Tinic, S.M., 1988. Anatomy of initial public offerings of common stock. Journal of Finance 43, 789–822.

Titman, S., Trueman, B., 1986. Information quality and the valuation of new issues. Journal of Accounting and Economics 8, 159–172.

Welch, I., 1989. Seasoned offerings, imitation costs, and the underpricing of initial public offerings. Journal of Finance 44, 421–449.

Welch, I., 1992. Sequential sales, learning and cascades. Journal of Finance 47, 695–732.

Welch, I., 1996. Equity offerings following the IPO: Theory and evidence. Journal of Corporate Finance 2, 227–259.

Zingales, L., 1995. Insider ownership and the decision to go public. Review of Economic Studies 62, 425–448.

Chapter 8

# CONGLOMERATE FIRMS AND INTERNAL CAPITAL MARKETS*

VOJISLAV MAKSIMOVIC

*Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA*
*e-mail: vmax@rhsmith.umd.edu*

GORDON PHILLIPS

*Robert H. Smith School of Business and NBER, University of Maryland, College Park, MD 20742, USA*
*e-mail: gphillips@rhsmith.umd.edu*

## Contents



* We thank Espen Eckbo and N.R. Prabhala for helpful comments.

*Handbook of Corporate Finance, Volume 1*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01008-7*

## Abstract

Conglomerate firm production represents more than 50 percent of production in the United States. Given the size of production by conglomerate firms, understanding the costs and benefits of this form of organization has important implications. Several studies have shown that there exists a discount in stock market value of conglomerate firms relative to single-segment focused firms. This discount represents an economically important puzzle. Early literature came to the conclusion that the conglomerate discount was the result of problems with resource allocation and internal capital markets. Recent empirical literature has found that self-selection by firms with different investment opportunities can explain the conglomerate discount. Additional theoretical and empirical research has shown how a model of profit-maximizing firms with different abilities and investment opportunities across divisions can explain observed resource allocation by conglomerate firms.

## Keywords

conglomerates, multidivisional firms, firm organization, investment, internal capital markets.

## 1. Introduction

In this chapter we survey the large literature on corporate diversification in corporate finance. For corporate finance, the primary questions about diversification are: "When does corporate diversification affect firm value?" And, "When diversification adds value, how does it do so?" By a diversified firm in corporate finance, we usually mean a firm that operates in more than one industry, as classified by the Standard Industrial Code (SIC).[1]

Questions about the relation between diversification and value arise naturally from the larger problem of determining how the boundaries of firms should be set. Coase (1937) argues that boundaries are set at the point at which the costs of carrying out transactions within a firm equal those of carrying them out in the open market or in another firm. Thus, for corporate diversification to be of interest, it must be that the cost of carrying out transactions within the firm are affected if it contains more than one industry within its boundaries. Implicit in this belief is that the skills and resources which are required to operate efficiently differ materially across industries, and that the *diversity* of operating environments affects the cost of performing transactions within the firm. These cost differences could be due to financial externalities across industries, such as improved risk sharing within the firm, or real externalities that could arise due to the use of a shared factor of production, such as the attention of the firm's decision makers.

Diversification across industries is also of interest to researchers because data on most intra-firm decisions is in general hard to acquire. By contrast, some data on how firm revenues and capital expenditures are distributed across the industries is readily available, which makes the research on diversification a good starting point for studying the more general problem of setting firm boundaries.

A more pragmatic reason for studying corporate diversification is that corporate managers face decisions about diversifying and refocusing their firms. In addition, managers face decisions about investing across multiple businesses they operate. Companies such as Berkshire Hathaway and General Electric generate large amounts of cash that can be invested in different business or returns to shareholders via dividends. Empirical data about how such decisions worked out in the past may be useful in strategic planning. Estimates of specific of costs and benefits might also be useful to investors and to regulators.

The corporate finance literature on diversification took off with the discovery of the conglomerate discount by Lang and Stulz (1994) and Berger and Ofek (1995). Our review therefore begins with a discussion of these papers and of subsequent work that has extended and reinterpreted their results. We then briefly discuss the theoretical

[1] In practice, researchers usually define firms as diversified if they generate less then 90% of their revenues in a single SIC code industry. Industries are commonly defined at the 3-digit level, although some studies use the 2-digit or 4-digit levels. Scharfstein (1998) is an exception in using a more qualitative criterion for diversification.

approaches that have been developed to explain the conglomerate discount and its investment decisions in Section 3. The empirical research motivated by these studies is reviewed in Section 4. Section 5 concludes.[2]

## 2. The conglomerate discount

### 2.1. Documenting the discount: Early research

In contemporary corporate finance the seminal papers on conglomerates are Lang and Stulz (1994) and Berger and Ofek (1995). Essentially, these papers decomposed conglomerate firms into their constituent industry segments and then valued these segments using the "comparables" approach to valuation. These papers found that the typical conglomerate is undervalued and selling at a discount compared to a collection of comparable single-segment firms. The existence of this conglomerate discount presents a puzzle. While Lang and Stulz (1994) do not take a position on the provenance of the discount, the early literature on conglomerates sought to explain this puzzle by arguing that conglomerates are subject to greater agency problems than single-segment firms. As a result, managers of conglomerate firms destroy value. By implication stockholder value would be maximized if most firms were organized as a single segment firms.

Since Lang and Stulz (1994) and Berger and Ofek (1995) are the seminal papers in the study of conglomerates it is worth examining their methodology in some detail. Preceding work on conglomerates in the industrial organization and strategy literatures had examined differences in ex-post accounting performance between conglomerates and single-segment firms. By contrast, Lang and Stulz (1994) and Berger and Ofek (1995) start from the question: "When do shareholders gain from diversification?" where gain is measured by the relative value of the diversified firm compared to single-segment firms in the same industry. To adjust for scale, firm value is in the first instance proxied by Tobin's $q$, the market value of the firm (equity and debt) divided by an estimate of the replacement value of the firm's assets.[3] To obtain the comparables, for each division of a conglomerate Lang and Stulz (1994) compute mean Tobin's $q$ of single-segment firms operating in the same 3-digit SIC code. The conglomerate's comparable $q$ is then found by the weighed average of the divisional $q$s. While the weights used can be derived in several ways, Lang and Stulz show that to obtain an unbiased estimate of the comparable, a division's weight should be computed as the ratio of the replacement cost of a division's assets to the replacement cost of the whole conglomerate's assets. However, as replacement values are generally unavailable, Lang and Stulz use book values in their place. The conglomerate discount is defined to be the difference between

[2] By its nature, this type of review inevitably omits many significant papers. Interested readers may want to consult other summaries of the literature, such as Martin and Savrak (2003).

[3] In some of their tests Lang and Stulz (1994) use the ratio of market to book values of a firm. The results are very similar.

a conglomerate's Tobin's $q$ and its comparable $q$ computed in the manner described above.

Lang and Stulz measure diversification in two ways. As their principal measure they count the number of the business segments that each firm reports in the Business Information File of Compustat. They use segment information from the Business Information File to compute two Herfindahl indices of diversification for each firm: an index computed from by using segment sales data and a second index computed from data on assets per segment.

Lang and Stulz main statistical tests consist of annual cross-sectional regressions for the period 1978 to 1990. They first regress firms' Tobin's $q$s on a constant and four dummy variables, $D(j)$, $j = 2, \ldots, 5$. The $j$th dummy variable takes on the value 1 if the conglomerate has more than $j$ segments in different SIC codes. Thus, $D(j)$ can be interpreted as the marginal contribution to $q$ of diversifying from $j - 1$ to $j$ segments. In a second round of tests they replace Tobin's $q$ as the dependent variable by the conglomerate discount, computed using comparables as above.

Across the annual cross-sectional regressions, Lang and Stulz consistently find that the coefficient of $D(2)$ is negative and significant, indicating that a two-segment firm sells at a discount both to single-segment firms in general, and to "comparable" single-segment firms, as defined above. There is much less evidence for the existence of a marginal effect of diversification on the discount for a larger number of segments. Lang and Stulz also show that a substantial portion of the discount remains even after controlling for differences in size and in the extent to which the firm faces financial constraints, as proxied, following Fazzari, Hubbard and Peterson (1988), by whether or not it pays dividends.

In addition, Lang and Stulz investigate whether the discount can be explained by differences in the propensity of single-segment and diversified firms to invest in research and development. Since the firm's balance sheet does not fully capture investment in R&D, the Tobin's $q$s of firms that engage in a great deal of R&D are going to be overstated relative to those of firms that engage in less R&D. If it were the case that single-segment firms were relatively R&D intensive, this relative valuation effect could explain the conglomerate discount. Lang and Stulz find that this is not the case. Thus, Lang and Stulz conclude that the diversification discount that they find cannot be explained by "reporting biases or subtle advantages of diversified firms".

The existence of a conglomerate discount naturally leads to the question: Are multi-segment firms worth less than single-segment firms because they diversify, or do less valuable firms choose to diversify?[4] The evidence from summary statistics is not clear-cut. Lang and Stulz find that single-segment firms that diversify have lower $q$s than single-segment firms that do not choose to diversify. However, the industry-adjusted $q$

[4] It is also possible that the decision to diversify is not causally related to the discount. This possibility is discussed below.

of diversifiers prior to diversification is not lower than that of non-diversifiers. Thus, the conglomerate discount is not explained by the low performance of firms that choose to become diversifiers.[5] However, not all findings they report are statistically significant or point in the same direction.

Thus, Lang and Stulz show the existence of a conglomerate discount. However, they judge their evidence to be "less definitive on the question of the extent to which diversification hurts performance". They find that the evidence is consistent with notion that firms diversify because they face diminishing returns in their industries. Lang and Stulz argue that to establish whether this is the case requires a more detailed disaggregated analysis and an explicit model.

Berger and Ofek (1995) confirm the Stulz and Lang result that there exists a conglomerate discount in the range of 13–15% of firm value for the period 1986–1991. They also investigate further potential causes of the discount. They find that the discount is smaller when the firm is not too diversified and all the segments are in the same 2-digit SIC code. They also find evidence that cross-subsidization and overinvestment contribute to the discount, and more limited evidence that diversified firms obtain tax benefits.

Berger and Ofek compute the estimated value of each segment in three related ways using a valuation approach similar to the multiples approach of Lang and Stulz. Berger and Ofek multiply each segment's assets, sales or earnings, reported in the Compustat industry segment database, by the corresponding median valuation multiple. The industry median is obtained by matching the segment to all the single-segment firms with sales above $20m in the most refined SIC code that contains at least five such firms. The valuation multiples are the ratios of the single-segment firms' total value (as proxied by the market value of equity and book value of debt) to the its reported assets, sales or earnings.[6]

Berger and Ofek also investigate whether diversified firms destroy value by overinvesting in unprofitable industries. Their measure of over-investment is the ratio of the sum of a conglomerate's capital expenditures and depreciation in 3-digit SIC code industries whose median Tobin's $q$ in the bottom quartile, to the conglomerate's total sales. They find that overinvestment so defined is associated with a loss of excess value.

Next, Berger and Ofek investigate whether cross-subsidization can explain the conglomerate discount. They regress the firm's excess value on an indicator which takes a value of one if the firm has a segment with a negative cash flow and zero otherwise.[7] The coefficient of this negative cash flow dummy is negative for diversified firms and

[5] Graham, Lemmon and Wolf (2002) reach the opposite conclusion. Their study is discussed below.

[6] Berger and Ofek do not use the conglomerate discount directly as their dependent variable, but the natural logarithm of the ratio of the actual firm value to the imputed value obtained by multiplying the reported accounting value by the appropriate multiplier. This number they term excess value.

[7] To compute excess value they estimate separate multiples in each industry for segments that have positive cash flows and those that do not.

indistinguishable from zero for single-segment firms. They thus conclude that having a segment with negative cash flows reduces the value of diversified firms by a greater amount than it reduces the value of focused firms.

Berger and Ofek also compare the long-term debt of diversified firms with the total debt level that would be predicted by summing the debt levels of a collection of single-segment firms that match the diversified firm's segments in size, profitability and investment opportunities. They find that while diversified firms borrow more than predicted, this effect is minor.

In sum, Berger and Ofek argue that their results provide evidence of a "significant loss of value in corporations that followed a diversification strategy in the 1980s". They also supply potential explanations for this loss. First, they find that conglomerate firms invest more in low-$q$ industries. Thus high investment in low-$q$ industries by conglomerate firms is associated with lower value. Second, they find that having a negative cash flow division lowers the value of a conglomerate. They interpret this loss in value as arising from "the subsidization of poorly performing segments contributing to the value loss from diversification".

Using a different methodology, Comment and Jarrell (1995) provide complementary evidence about the valuation of conglomerate firms during the 1978–1989 period. They find that increases in focus, subsequent to asset sales, are associated with increases in value. Their results are summarized in Figure 1.

Figure 1 shows that, on average, increases (decreases) in focus are associated with positive (negative) abnormal stock returns in the year in which focus increases.[8] They also find that some of the presumed economies of scope, such as the ability to support more debt and the ability to reduce transactions in the capital markets, are not exploited more by diversified firms.

The early evidence in Lang and Stulz, and Berger and Ofek shows convincingly that conglomerates sell at a discount when compared to benchmark industry single-segment firms. It is also consistent with the notion that the discount is caused by inefficient operations and that, as Comment and Jarrell argue, the presumed economies of scope do not appear to be exploited. However, both Lang and Stulz and Berger and Ofek draw the reader's attention to potential deficiencies with the data. These potential problems raise several questions:

- To what extent are the well known difficulties with the data material to the estimates of the discount?
- Do the comparables used fully take into account the differences between single-segment and diversified firms? Clearly firms choose their organizational form and this choice may be related to firm and industry characteristics.
- Can the differences in valuation be explained? Do conglomerate firms and single segment firms invest differently?

[8] For a discussion of some of the difficulties in interpreting long-run event studies, see the chapter by Kothari and Warner (2006) (Chapter 1 in this volume).

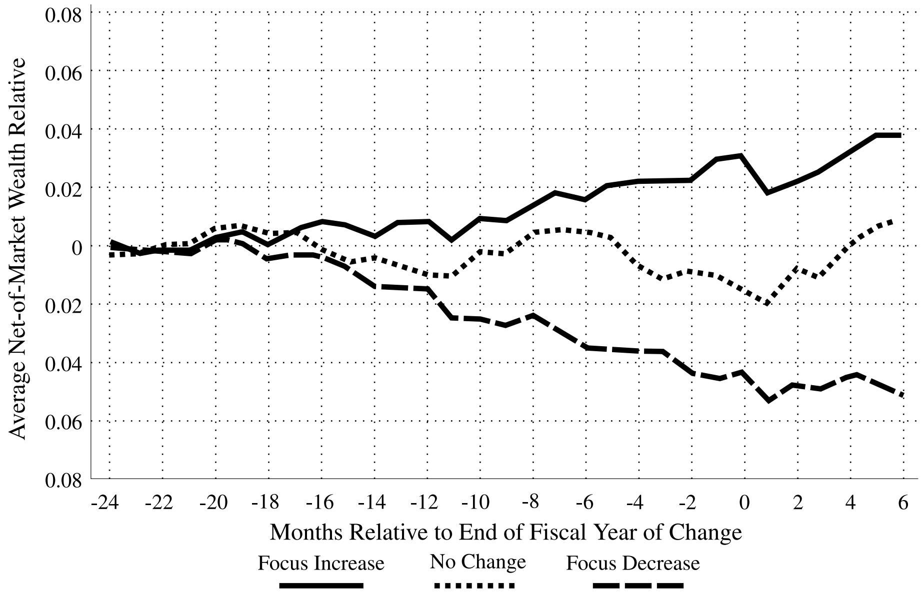



Fig. 1. Event-study showing the average wealth effect of focus changes for three groups of firms segmented by the direction of focus change. The sample consists of fiscal years in the period 1978–1989 for exchange-listed firms. Plotted points represent, for each group of firms, the average month-by-month value of a SI initial investment in the firm less the corresponding month-by-month value of a $1 initial investment in the CRSP equally-weighted market portfolio. Month zero is defined as the last month of the fiscal year of the change in focus, Focus is measured by a Herfindahl index defined on revenue. There are 5,088 fiscal-years with increases in the Herfindahl index, 4,469 with decreases, and 7,056 with no change. Source: Comment and Jarrell (1995), Journal of Financial Economics, Vol. 74, p. 74.

We will be reviewing how the literature has addressed these issues in the remainder of this chapter.

### 2.2. *Initial caveats: The data*

Research in firm organization is particularly tricky because researchers are required to look inside the corporation to assess the efficiency of resource allocation between various subunits. Such data is not readily available, and much of the data that is available is subject to potential manipulation and reporting biases. The data problems mean that researchers in this area must pay special attention to data issues and to the potential for measurement error.

The principal data source for the early research on conglomerates is the Compustat Industry Segment (CIS) database. Pursuant to the Statement of Financial Standards (SFAS) No. 14 and SEC Regulation S-K, after 1977 firms were required to report certain audited segment information on segments whose assets, sales or profits are deemed material by exceeding 10% of the firms' consolidated totals.[9] The CIS database contains information for such segments on net sales, earnings before interest and taxes (EBIT), depreciation, capital expenditures, and assets, as well as the total number of reported segments for the firm. This data is available for all active Compustat firms except utility subsidiaries and is easy for most researchers to access.

There are, however, several well-known problems with CIS data. Firms self-report segment data and changes in the number of reported segments may reflect changes in reporting practice. Hyland (1997) finds that up to a quarter of reported changes in the number of segments stem from changes in reporting policy, not changes in the level of diversification.[10] The reporting requirement also only applies to segments that meet a 10% materiality condition. Thus, segments reported by large firms may be span several industries.[11] Moreover, there is no presumption that a self-reported segment approximates a single industry. According to SFAS 14, a segment is distinguished by the fact that its constituents "are engaged in providing a product service or a group of related products and services ... to unaffiliated customers". Thus, segments may be vertically integrated. The 4-digit SIC in which they are classified by CIS are assigned by COMPUSTAT, not by the firms themselves. This last problem is quite severe: using Census data Villalonga (2004a, 2004b) shows that in over 80% of cases the SIC code assigned by COMPUSTAT is not the code of the segment's largest industry. Taken together, these problems raise the possibility that a substantial number of segments are misclassified into 4-digit SIC codes and that a substantial number of firms that report only one segment in fact operate in related or vertically integrated industries.[12]

Several researchers have used alternative data sources from the US Bureau of Census which do not rely on data which is aggregated up to segment level by firms. Maksimovic and Phillips (1998, 2001, 2002, 2007) and Schoar (2002) use the Longitudinal Research Database (LRD), maintained by the Center for Economic Studies at the Bureau of the Census.[13] The LRD database contains detailed plant-level data on the value of shipments produced by each plant, investments broken down by equipment and buildings, and the number of employees. The LRD tracks approximately 50,000 manufacturing plants every year in the Annual Survey of Manufactures (ASM) from 1974 to 2003. The ASM covers all plants with more than 250 employees. Smaller plants are randomly selected every fifth year to complete a rotating five-year panel. Note that while the annual

[9] Revised disclosure requirements, SFAS 131, superseded SFAS 14 in 1997. Most of the studies that use Compustat data discussed in this review rely on pre-1997 data. Under SFAS 131 firms do not have to report line of business data unless they are organized that way for performance evaluation (Berger and Hahn, 2003).

[10] See also Denis, Denis and Sarin (1997), Pacter (1993) and Hayes and Lundholm (1996).

[11] Villalonga (2004a, 2004b) notes that the maximum number of 4-digit segments belonging to a single firm for her sample of firms drawn from the BITS database of the U.S. Bureau of Census is 133.

[12] Note that the definition of relatedness according to SFAS 14 does not correspond to the SIC classification. Thus, divisions from different 2-digit SIC codes may be related according to SFAS 14.

[13] For a more detailed description of the Longitudinal Research Database (LRD) see McGuckin and Pascoe (1988).

data is called the Annual Survey of Manufactures, reporting is not voluntary for large plants and is not voluntary once a smaller firm is selected to participate in a rotating panel. All data has to be reported to the Censuc Bureau by law and fines can be levied for misreporting.

Annual Survey of Manufactures offers several advantages over Compustat: First, it is comprehensive and covers both public and private firms in manufacturing industries. Second, coverage is at the plant level, and output is assigned by plants at the four-digit SIC code level. Thus, firms that produce under multiple SIC codes are not assigned to just one industry. Third, plant-level coverage means that plants can be tracked even when they change owners.

Villalonga (2004a, 2004b) uses the Business Information Tracking Series (BITS) database, also from the Bureau of the Census. BITS provides data between 1989 and 1996 for all U.S. business establishments, private and public, in all some 50 million establishment-year observations.[14] For each establishment, the BITS database contains data on the number of employees, the payroll and on the identity and revenue of the firm that owns it. Each establishment is assigned to a 4-digit SIC code.

Because the BITS database covers all sectors of the economy and is not limited to the manufacturing sector like the LRD, it is more comprehensive. However, since the available data for each establishment is limited, BITS cannot be used to determine an establishment's productivity.

Villalonga (2004a, 2004b) links the BITS dataset with COMPUSTAT, enabling her to determine the composition of a Compustat firm without relying on SFAS 14 disclosures. She then recomputes the conglomerate discounts of the COMPUSTAT firms that she has linked, using as comparables those COMPUSTAT firms that BITS data identifies as being single-segment firms.

The results are startling. Villalonga finds that diversified firms trade at a significant *premium* over single-segment firms, as so classified using BITS. When COMPUSTAT segment data is used to classify firms, Villalonga obtains the standard conglomerate discount obtained in the earlier literature.

Villalonga explores several possible explanations for this discrepancy. A fundamental difference between BITS and COMPUSTAT is that former treats vertical integration as a form of corporate diversification, whereas the latter does not. However, when Villalonga reconstitutes BITS segments to group together vertically integrated businesses and recomputes the discount she still obtains a conglomerate premium.

These results highlight the fact that COMPUSTAT segments are related by construction, at least in the eyes of the firms. Thus, measures of diversification based on COMPUSTAT data may implicitly be measures of unrelated diversification. It is thus possible that diversification, measured by COMPUSTAT is a measure of inefficient diversification (hence the discount). Villalonga also raises the possibility that Compustat segments are lumped together to avoid disclosing to competitors which segments are most lucrative.

[14] An establishment is a location where a firm conducts business, such as a plant, a store or a warehouse.

The firms classified as single segment in BITS are smaller than the firms classified as single segment in COMPUSTAT. If as, suggested by Maksimovic and Phillips (2002) size is positively correlated with productivity, then the premium that Villalonga finds using BITS data may be occurring because she is implicitly comparing conglomerates, which are larger, with unproductive small single-segment firms.

Finally, several interesting results showing that alternative measures of diversification may affect the interpretation of current results are obtained by Denis, Denis and Yost (2002). They examine global diversification over time. These firms are not necessarily diversified industrially. They document that global diversification results in average valuation discounts of the same magnitude as those for industrial diversification. Analysis of the changes in excess value associated with changes in diversification status reveals that increases in global diversification reduce excess value. One possible implication of their results that is consistent with Maksimovic and Phillips (2002) is that as firms expand they take on less profitable projects but ones that still may have positive NPV, thus reducing *ratio* measures of excess value.

Denis, Denis and Yost (2002) also find that firms that are both globally and industrially diversified do not suffer a diversification discount on average, suggesting that global diversification may in this case benefit firm value. This result is driven by the latter half of the sample period, in which firms that are both globally and industrially diversified are valued at a premium relative to single segment, domestic firms. Their results imply that the value and costs of diversification may change over time.

### *2.3. Self-selection and the endogeneity of the decision to become a conglomerate*

The early research on the conglomerate discount relied on the comparison of conglomerates' divisions with a control sample of comparables using single-segment firms chosen using heuristic criteria described above. The implicit assumption was that conglomerate and single-segment firms faced the same investment opportunities and were of similar ability.

This way of selecting comparables raises issues on two grounds. First, it ignores potentially observable differences between the divisions and the matching single-segment firms that might affect valuation. Second, the heuristic matching procedures implicitly assume that firms become conglomerates randomly, and not as argued by Maksimovic and Phillips (2002), because they differ in material ways from firms that remain single-segment. If the decision to diversify is not random, and is instead based on information observed by the firm but not by the researcher, then the estimation procedure must take into account the endogeneity of the decision.[15]

The underlying hypothesis in the discount literature is that the value of firm $i$ at time $t$ relative to its comparables, $V_{it}$ is a linear function of a set of control variables $X_{it}$ and

[15] For early discussions of this endogeneity in the context of corporate finance decisions, see Eckbo, Maksimovic and Williams (1991) and Prabhala (1997). Chapter 2 in this volume (Li and Prabhala, 2007) contains a much more comprehensive discussion of selection issues in this type of research.

on whether the firm is a conglomerate, denoted by the indicator variables $D_{it}$ which takes on the value 1 if the firm is a conglomerate and 0 if it is not.

$$V_{it} = \beta_1 + \beta_2 X_{it} + \beta_3 D_{it} + e_{it}, \tag{1}$$

where $e_{it}$ is an error term.

A necessary condition for the OLS estimate of coefficient $\beta_3$ to be unbiased is for $D_{it}$ to be independent from the error term $e_{it}$ in equation (1). The earlier literature, such as Lang and Stulz, implicitly assume that this condition holds and that conglomerate status can be treated as being exogenous in the estimation. But suppose instead that the firm's decision to operate in more than one industry depends on a set of characteristics $W_{it}$ and a stochastic error term $u_{it}$. Specifically assume that $D_{it} = 1$ when $\lambda W_{it} + u_{it} > 0$ and $D_{it} = 0$ when $\lambda W_{it} + u_{it} < 0$. Then, the coefficient of in equation (1) will be biased if, as seems plausible, a common determinant of both the value $V_{it}$ and the decision to become conglomerate is omitted from estimated equation (1).

Several recent empirical papers on the conglomerate discount, by Campa and Kedia (2002), Graham, Lemmon and Wolf (2002), Lamont and Polk (2002) and Villalonga (2003, 2004a, 2004b) address these issues. Maksimovic and Phillips (2002) gives an equilibrium justification for the endogeneity of the discount and also empirically shows that there is a size-efficiency relationship that holds for conglomerate and single-segment firms.

The most direct evidence on the importance of self-selection in the determination of conglomerate discounts is provided by Graham, Lemmon and Wolf (2002).[16] They show directly that diversification through acquisitions creates a measured discount in the sense of Berger and Ofek (1995) even when the diversification is value increasing. Using a sample of 356 mergers that occurred between 1978 and 1995 and (i) which met the Berger and Ofek criteria of inclusion in the sample of diversifiers and (ii) for which they had data on both the bidder and the target, Graham et al. show that acquirers register a discount computed in the sense of Berger and Ofek in a two-year window surrounding the acquisition. However, the greater part of this discount can be explained by the fact that the targets are selling at a discount relative to single-segment firms prior to the merger. Thus, much of the discount associated with corporate diversification by acquisition cannot be attributed to the costs associated with operating more diversified firms but can be attributed to the fact that diversifying firms are on average acquiring assets already valued at a discount relative to the industry benchmarks. To the extent that conglomerate firms engage in more acquisition activity than single-segment firms (as shown in Maksimovic and Phillips, 2007), it is possible that their growth pattern might induce a discount even when it is value maximizing.

Campa and Kedia (2002) also argue that the documented discount of diversified firms is not by itself evidence that diversification destroys value. They use three alternative

[16] See Chevalier (2000) for a related argument.

econometric techniques in an attempt to control for the endogeneity of the diversification decision—firm fixed effects, simultaneous-equation estimation using instrumental variables and Heckman's two-step procedure. Their data is from COMPUSTAT and their sample and the measurement of excess value follow the earlier literature. Segments of multiple-segment firms are valued using median sales and asset multipliers of single-segment firms in that industry. The imputed value of a segment is obtained by multiplying segment sales (asset) with the median sales (asset) multiplier of all single-segment firm-years in that SIC. The imputed value of the firm is the sum of the segment values.

Campa and Kedia find a strong negative relation between a firm's choice to be diversified and its value. Firms that are diversified have a lower value than firms that do not. However, once the endogeneity between the decision to be diversified and firm value is taken into account, the diversification discount always drops, and sometimes turns into a premium.

The statistical modeling of the endogeneity of conglomerate status, in turn, raises questions about the nature of the decision to become conglomerate. In their statistical specification, Campa and Kedia implicitly assume that the decision to remain diversified is itself endogenous in each period. This is appropriate if the decision to diversify is easily reversible. However, if the decision is costly to reverse, then it is natural to focus attention on the endogeneity of the decision to diversify (as opposed to the endogeneity of the decision to maintain conglomerate status), or more generally on changes in the level of diversification.

Villalonga (2004a, 2004b) focuses on the decision to become diversified. Using a Compustat for the years 1978–1997 she identifies 167 firm years in which single-segment firms diversified. Her control sample consists of 40,757 single-segment firm years. She adopts a two-stage procedure. In the first stage, she uses a probit model to obtain the probability that a firm becomes diversified, which she terms the propensity to diversify. For the probits Villalonga tries several specifications, including one that uses the same explanatory variables as Campa and Kedia (2002).

In the second stage Villalonga controls for the estimated propensity to diversify in determining whether becoming diversification destroys value. She uses two types of matching estimators (the methods proposed by Dehejia and Wahba, 1999, and Abadie and Imbens, 2002) and Heckman's (1979) correction for selection bias. As in Campa and Kedia's (2002) tests, Heckman's method directly corrects for biases due to unobserved characteristics of firms that choose to diversify. The matching estimators use the estimate of the propensity to merge as one of the characteristics for finding matching non-diversifying single-segment firms that are comparable to the diversifying single-segment firms. Consistent with Campa and Kedia (2002), Villalonga finds that the decision to diversify did not affect the value of the 167 firms that she identifies as having diversified during her sample period.

Lamont and Polk (2002) adopt a different approach and a difference definition of the extent of diversification in their study of the relation between diversification and value. They argue that a key characteristic of an industry is the ratio of investment to capital stock. In their view a firm that operates in industries that have a greater disparity of investment to capital stock ratios is more diversified than a firm that operates in industries that have similar investment to capital stock ratios. Thus, for each 2-digit SIC code industry to which COMPUSTAT assigns the firm's segments, Lamont and Polk calculate the median investment to capital ratio among the single-segment firms. The measure of a conglomerate firm's diversity in year $t$ is then computed as $\sigma$, the weighted standard deviation of these median ratios for all segments.

Lamont and Polk argue that changes in $\sigma$ over time can be decomposed into endogenous and exogenous components. The exogenous change in diversity, $\Delta\sigma_{\mathrm{X}}$, is the change in diversity between $t-1$ that would have occurred if COMPUSTAT had assigned the firm in the current year to precisely the same 2-digit SIC codes as in the previous year. The endogenous change in diversity, $\Delta\sigma_{\mathrm{N}}$, is the change in diversity that occurs because the 2-digit SIC codes assigned to the firm have changed between years $t-1$ and $t$.

Lamont and Polk use COMPUSTAT data for 1,987 diversified firms during the period 1980–1997. They find that 80% of the variation in firms' diversity is due to exogenous industry shocks. In their regressions they regress the change in excess value on $\Delta\sigma_{\mathrm{X}}$ and $\Delta\sigma_{\mathrm{N}}$ alone and with control variables such as lagged $\sigma$. They find that increases in both $\Delta\sigma_{\mathrm{X}}$ and $\Delta\sigma_{\mathrm{N}}$ reduce firm's excess value. They interpret the negative coefficient of $\Delta\sigma_{\mathrm{X}}$ as evidence that diversification reduces firm value. This finding persists even when plausible measurement error is taken into account. Lamont and Polk (2002) also analyze similarly defined changes in diversity of leverage, cash flows and sales growth one at a time. They do not find that "exogenous" changes in diversity of these variables have a significant negative effect by themselves.

Lamont and Polk's interpretation of their results on investment diversity are in sharp contrast to Campa and Kedia (2002) and Villalonga (2004a, 2004b) and have not been fully reconciled with these studies. Villalonga (2003) argues that Lamont and Polk's measure does not pick up "diversification" as traditionally measured in the literature—the presence of the firm's operations in more than one industry—but "diversity" which is the within firm dispersion of some industry characteristics. Indeed, she reports tests that show that measures of diversification, such as the number of two digit industries that the firm operates in are uncorrelated with Lamont and Polk's measure of exogenous cash flow diversity $\Delta\sigma_{\mathrm{X}}$. However, this observation raises the question of which measure better captures economic differences between firms.

## 3. Theory explaining the conglomerate discount and organizational form

The early literature on the conglomerate discount leaves several questions unanswered. Perhaps the most fundamental of them is why there should be a conglomerate discount? Is the existence of a discount evidence of bad investment choices or is the discount an endogenous outcome of a process by which different types of firms optimally select different types of expansion paths, given different investment opportunities? If there is evidence of inefficient investment choices, why do they occur?

Conceptually, the conglomerate discount is an unlikely subject for academic research. In most introductory corporate finance classes MBA students are painstakingly taught that firms should maximize the net present value of their investments, not the ratio of market value to replacement cost. In fact, they are explicitly warned that maximizing the latter, which is equivalent to maximizing the profitability index, leads to inefficient investment choices in the presence of capital constraints.[17] Yet when we evaluate the performance of conglomerates, we do so using the conglomerate discount, which is equivalent to comparing the profitability indices of conglomerate and single-segment firms. We do this because of the practical difficulties of obtaining properly scaled measures of value, and not because the literature has shown that it is a measure of a relevant measure of performance.

In this section, we review the theoretical frameworks that have been used to motivate the recent empirical literature on diversification and the investment of diversified firms. We begin with the literature which assumes there is a "bright" side of conglomerates—that conglomerates internal allocation of financial capital has benefits. This literature assumes that firms would not become conglomerates unless there is some benefit of doing so in terms of allocating financial capital within the firm. However, this literature does not explain why there is a discount. Implicitly the literature on the bright side of conglomerates assumes that the discount would be larger if the conglomerate's segments were stand alone single-segment firms, which prompts questions about the appropriate comparables to use in determining the discount. We illustrate this line of research with Stein's (1997) model of how diversified firms' internal capital markets lead to a different selection of investment projects than when firms operate in a single industry. Second, we discuss Matsusaka's (2001) model of how organizational competencies may drive the diversification decision.

Third, we discuss the literature which takes the opposite perspective and models how conflicts of interest between the firm's managers and the firm's owners may lead to inefficient diversification. Fourth, several models taking the same perspective of inefficient diversification have argued that intra-firm bargaining in firms operating in several different environments leads to poor investment choices (Rajan, Servaes and Zingales, 2000; Scharfstein and Stein, 2000).

Finally, we end with discussion of equilibrium models of the conglomerate firm which show that the conglomerate discount can arise endogenously and that conglomerate investment is a profit-maximizing approach to differential investment opportunities.

The papers that we review are only a small portion of the theoretical literature on the conglomerate firm. The models are all highly stylized and rather informally presented. In part, this is because data constraints make it very hard to test complex structural models of intra-firm dynamics. They are nonetheless important for our purposes because

[17] See, for example, Brealey and Myers (2003) and Ross, Westerfield and Jordan (2006, p. 283).

they have motivated several of the empirical studies we examine below. In the interest of brevity, we do not discuss several models which deserve a separate review, including Berkovitch, Israel and Tolkowsky (2005), Faure-Grimaud and Inderest (2005), Fluck and Lynch (1999) and Inderst and Meuller (2003).

### 3.1. *Efficient internal capital markets*

Stein (1997) analyzes how internal capital markets create value and the optimal size and scope of such markets. In Stein's model firms consist of either a single stand-alone project or of several projects overseen by a headquarters. Stein assumes that each project's managers obtain private benefits from managing their project. These benefits are higher for better projects. The private benefits give managers an incentive to overstate their project's prospects. This is known to potential investors, who therefore supply less capital than the managers request. As a result, good projects are capital rationed if they operate as individual firms.

Stein assumes that a conglomerate's headquarters has the ability to monitor the projects it oversees. It uses its information in two ways. First, it can transfer capital from one project to another. Second, it can appropriate for itself some of the private benefits of the project managers, albeit at the cost of diluting the incentives of the managers.

Because the headquarters can extract private benefits from several projects simultaneously it has the incentive to allocate capital to the better projects. The ability to transfer funds across projects, allocating some more funds than they would be able to raise as stand-alone firms, and others less, makes better allocation possible.

A key assumption in Stein (1997) is that as the number of projects overseen by the headquarters increases, the quality of monitoring provided by the headquarters declines. However, as the number of projects the quarters oversees increases, the headquarters in Stein's model also gains in two ways. First, the value of its ability to transfer funds from the worst to the best projects increases. Second, if the project payoffs are not perfectly correlated the volatility of the firm's payoffs declines and it becomes able to raise more funds from the capital market, thereby reducing credit rationing and increasing value. The firm reaches its optimal size when the marginal decline in value due to declining monitoring ability is equal to the marginal increase resulting from the relaxation of financing constraints and the funding of good projects.

The theory also has implications for the optimal scope of the firm. Stein addresses two effects which work in opposite directions. To the extent that the returns of different divisions of a conglomerate are uncorrelated diversification increases the value of the headquarters' ability to direct investment funds and raise capital externally. However, there may be another effect at work. Because headquarters' allocation decisions are dependent on the *ranking* of investment projects rather than their absolute values, and to the extent that accurate rankings are more likely to be made if all projects are within the same industry (because valuation errors are likely to be correlated), diversification is costly.

Thus, Stein suggests that diversification is value increasing when valuation errors are small and when the returns of projects within an industry are highly correlated, Diversification is value reducing when valuation errors are likely to be large and when the payoffs of projects within industries are likely to have a low correlation.[18]

### *3.2. Conglomerates and organizational competencies*

Matsusaka (2001) develops a matching model to explain why conglomerate firms exist. In his model firms have different organizational competencies. The organizational competencies are somewhat transferable across industries. When sales decline in an industry it is not optimal for firm to go out of business. Instead it should diversify into new lines of business in order to find a good match between their organizational competence and the line of business. If they find a good match they may transit into the new industry and exit their original industry.

Matsusaka's (2001) elegant framework generates several predictions. Diversifying firms trade at a discount because on average the match between their organizational competence and their existing main divisions is bad. Because the match in the new industry may also turn out to be bad, many diversification attempts are in fact reversed. However, the announcement of a diversification is a signal that the firm is worth maintaining, resulting in a positive announcement effect. The theory also predicts that successful diversifiers quit their original industry. Thus the theory is quite consistent with the early evidence on the diversification discount (e.g., Lang and Stulz, 1994, and Berger and Ofek, 1995), as diversification results from a poor match between industries and firm's organizational competence, and on announcement returns (e.g., Schipper and Thompson, 1983; Hubbard and Palia, 1998) which document positive or non-negative returns to changes in the level of diversification.

### *3.3. Diversification and the failure of corporate governance*

Given the message from the early literature that diversification destroys value, the obvious question is why we observe so many diversified firms. One plausible answer is that while diversification destroys investor value it benefits the managers of corporations. Thus diversification might arise as a result of a failure of corporate governance which should be penalizing managers who diversify inappropriately.

Jensen (1986) and Stulz (1990) argue that managers may obtain increased status and perquisites when they diversify their firms. Diversification allows managers to act on a broader stage, and in particular may allow them to participate in "hot" and exciting industries. It may also be easier to skim from a diversified firm (Bertrand and Mullainathan, 2001).

[18] The model does not analyze the possibility that a focused firm may rank projects correctly but over- or under-invest in the aggregate because the valuation errors it makes are correlated across projects.

Diversification may also yield concrete career benefits, because experience running a complex diversified firm might provide experience that the increases the value of the manager's future employment prospects (Gibbons and Murphy, 1992). On the other side of the coin, diversification may entrench the manager because it may be harder to find a replacement who has a demonstrated ability in managing the firm's particular mix of businesses (Shleifer and Vishny, 1989).

Taken together, the literature on agency makes a powerful prima fasciae case that agency conflicts may drive unprofitable diversification. An issue in determining the extent to which this is the case is that most of the contributions are set in a partial equilibrium framework. Thus, it is not clear why the incentives are not set in ways that penalize unprofitable diversification. Moreover, it is not clear why diversification is inefficient. A rational empire-building CEO of a diversified firm can in principle decentralize its operations and provide incentives to the managers running the firm's divisions so that firm value is not destroyed. Thus, it must either be the case that increasing the firm's scope the firm itself destroys value or that managers of firms that diversify are irrational and have a hubristic belief in their ability to run acquired businesses (Roll, 1986).

There have been only a few attempts to analyze the manager's incentive to diversify in a more general model of the trade-offs. This is in part because the incentives of, and the constraints faced by, the board of directors, the party that formally employs the manager, are not well understood.[19]

Aggarwal and Samwick (2003) model the diversification process by assuming that the board maximizes the value of the firm. The key assumption is that diversification, which is assumed to be value destroying, is not contractible and cannot be forbidden by the board. The CEO benefits from diversification, because it enables him to diversify his risk and because he has private gains from diversification. The board can attempt to motivate the manager to work harder by tying his compensation to firm value. However, this type of compensation has a byproduct of increasing the manager's risk exposure, making value destroying diversification more attractive. In equilibrium, managerial compensation is set as a result of contracting in a standard principal agent problem where managerial effort is costly.

In Aggarwal and Samwick (2003) the manager's compensation $w$ is given by $w = w_0 + \alpha\pi + \gamma n$, where $\alpha$ and $\gamma$ are constants chosen by the firm, $n$ is the amount of diversification and $\pi$ is the firm value. Firm value is given by $\pi = x - n + \varepsilon(n)$, where $x$ is the costly managerial effort and $\varepsilon(n)$ is a normally distributed shock to firm value with zero mean and variance $\sigma^2/n$. For the manager, diversification has three consequences. First, it affects the value of the corporation and thereby the manager's compensation through the $\alpha$ term. Second, it enables the manager to diversify risks since it reduces the risk of the corporation $\sigma^2/n$. Third, diversification enters directly in to the

[19] See Hermalin and Weisbach (1998) for a theoretical model in which the relationship between the board of directors and the CEO evolves over time.

manager's utility function because it affects the value of his proprietary benefits. Given the assumed relation between diversification and value, and the assumed compensation contract, the firm's directors can affect the manager's actions by tying his rewards to performance.

The board of directors offer the CEO a linear contract based on $\pi$ and $n$. The CEO chooses the level of diversification $n$ and effort $x$. The $\pi$ is realized and the CEO is compensated on $\pi$ and $n$.

This framework leads to some interesting predictions, which differ from those that would be derived by intuition alone. For example, suppose that there is an exogenous increase in the amount of private benefits that the manager can gain from diversification that is value destroying for the firm. In equilibrium it would be optimal to increase his performance pay in order to reduce his incentive to diversify. However, in an interior equilibrium this increase will not be enough to totally negate the effect of the exogenous increase in private benefits from diversification. As a result, empirically we would observe contemporaneous increases in incentive based pay and in diversification. However, the positive correlation would not be an indication of a causal relation. More generally, the empirical relation between incentives and diversification shows that the interpretation of simple correlations between incentive based compensation and diversification is not straightforward.

Aggarwal and Samwick (2003) derive testable relations regarding changes in firm value, incentive compensation and level of diversification in response to changes in exogenous parameters, such as managerial risk aversion or the ability to gain private benefits from diversification. They estimate this relation on about 1600 firms in the 1990s using COMPUSTAT, CRSP and ExecuComp data. The pattern of relations they find is consistent with their model's predictions for the case in which diversification decisions are driven by increases in managers' private benefits from diversifying.The advantage of an explicit modeling approach as in Aggarwal and Samwick (2003) is that it yields a set of transparent predictions that can be taken to data and checked for consistency. For this clarity to be attained the researcher has to take a point of view about the underlying relation. Other initial structures, in which, for example, the board can monitor and approve diversification or where not all diversification reduces value—may yield different predictions on the value of diversification.

### *3.4. Diversification and the power within the firm*

Another strand in the literature argues that investment decisions within diversified firms are driven by the need to moderate conflicts of interests between different divisions and different levels of the hierarchy within the firm. The starting point for this research are the observations by Lamont (1997) and Shin and Stulz (1998) that diversified firms capital expenditures are not as sensitive to proxies of industry opportunities as focused firms. Such distortions would be unlikely to occur in the standard agency framework where the CEO has an incentive to maximize firm value so as to maximize his ability to expropriate investors. While such distortions might occur in more complex agency models, where top management diversifies out of career concerns or to reduce risk, it is also plausible that the distortions be caused by intra-firm conflicts.

In the classic influence cost model of intra-firm conflict, Meyer, Milgrom and Roberts (1992) model a resource process where lower-level managers of a firm attempt to lobby top management to increase the investment flows available to their firm. The lobbying is costly, but in equilibrium top managers infer the true value of investment opportunities by observing the costly lobbying. Thus, the lobbying leads to inefficiency but does not lead to misallocation of resources.

In Scharfstein and Stein (2000), managers of divisions which lack investment opportunities have a low opportunity cost of their time and therefore engage lobbying which is creates costs for the firm as a whole. An efficient response to such lobbying might be for the firm's owners to bribe the managers of weaker divisions to desist. However, the top managers of firms are themselves the agents of the firm's owners and this affects how they pay off the divisional managers. Scharfstein and Stein (2000) derive conditions under which top management finds it optimal to bribe troublesome divisional managers by giving them too large a share of the investment budget rather than with cash. This occurs because top managers cannot directly expropriate the firm's capital budget whereas they can extract benefits from any operating funds that they would have used to pay divisional managers. Thus, to reduce the cost of lobbying, top management overinvests in the divisions with poor growth opportunities.[20]

A central assumption of this approach is that the top management has limited power over the divisional managers. An alternative response by top managers who do have such power might be to change the reporting structure within these divisions or add extraneous task which can be easily monitored to the divisional managers' workload so as to increase the opportunity cost of their time and thereby reduce their propensity to lobby. Another possibility might be for the firm to sell or spin off its weaker divisions.

Rajan, Servaes and Zingales (2000) explore another implication of limited head office power over divisions. They argue that while top management can direct capital expenditures across divisions it cannot commit to a future distribution out of the value created by the investment. The distribution of the surplus is determined through negotiations between divisions after the surplus has been realized. The inability of top management to commit to a distribution means that a division's investment choices may be distorted.

A key assumption about the ex-post bargaining process between divisions is that the divisions' bargaining power is influenced by their initial investment decisions. As a result, it might be in the top management's interest to initially allocate initial investment capital in a way that will influence the outcomes of future bargaining between divisions over the distribution of the surplus rather than to maximize value. Given a distribution of capital across divisions, divisional managers will update their predictions about the

[20] See also Fulghieri and Hodrick (1997).

likely outcomes of bargaining over the surplus and make investment choices accordingly. It is in the top management's interest to allocate initial investment funds in ways that induce the divisional managers to choose projects that maximize the firm's value.

Hence, in RSZ top management uses the initial allocation of investment to divisions as a commitment device to substitute for its inability to commit to a distribution of surplus. This form commitment is clearly not as efficient as a first-best case in which top management can commit to the distribution of profits that the divisions realize. Empirically, the capital expenditures of conglomerates might seem, and would be, less efficient than those of single-segment firms. However, they are value maximizing given the constraints that top managers face.

RSZ make specific assumptions about the way the bargaining between divisions works and obtain predictions about the distortions that arise. Specifically they assume that each division can choose to invest in two types of investment projects. "Efficient" projects are value maximizing. "Defensive" projects produce less value, but the value generated can be better defended against redistribution to other divisions. The top management's problem is to allocate the right amount of capital to each division and to motivate the divisional management to invest the capital in the efficient project.

The divisional manager's incentive to choose a defensive project is higher when the surplus generated by the efficient project, which he has to partially give up in ex-post bargaining with other divisions, is high relative to the manager's share of the other divisions' surplus that he expects to gain in bargaining. Under plausible assumptions, this occurs when the manager's division has better investment opportunities than the other divisions. As a result, perverse investment incentives are more likely to occur in firms with divisions facing diverse investment opportunities.

The RSZ model predicts that the value of diversified firms is inversely related to the diversity in their investment opportunities. The model also predicts that capital transfers will occur from large high-value divisions to small low-value divisions. Both of these predictions are testable. We discuss these tests later.

A central feature of most theoretical models of the conglomerate firm is that they are partial equilibrium in the sense that they do not analyze the firm's internal allocation of capital in the context of the market for whole firms and partial-firm assets. As Maksimovic and Phillips (2001) show, there is a large market for assets in which conglomerates are important players. Thus, as an alternative to distorting the firm's investment expenditures, a firm facing the problems modeled by RSZ might trade divisions to obtain a diversified portfolio of assets that faces comparable investment opportunities. Thus, a generalized RSZ framework might suggest that the firm can operate on an alternative margin, yielding the prediction that at times when the market for firms' assets is active, firms are less likely to distort investment flows.

### *3.5. Neoclassical modelof conglomerates and resource allocation*

The case in which firms maximize value and there are no unresolved agency problems provides a benchmark for an analysis of conglomerate growth and diversification.

Arguably, if investment patterns in conglomerates can be predicted by a neoclassical model, then the effort in explaining misallocation of resources by managers may be better directed at examining other forms of shareholder expropriation. Maksimovic and Phillips (2002) consider a neoclassical model where firms differ because managerial and organizational talent or some other fixed resource varies across firms. Interestingly, the neoclassical model for conglomerate firms was introduced after the initial models of power within the firm. It has motivated empirical models of investment within the conglomerate firm and also endogeneity and sample selection models.

In Maksimovic and Phillips (2002) the firm decides endogenously whether to produce in one or in several industries. As in Coase (1937) and Lucas (1978), it is assumed that there are diseconomies of scale within firms. Firms exhibit neoclassical decreasing returns-to-scale, so that their marginal costs increase with output. Specifically, firms use the variable inputs of labor, and capacity units to produce output.

In each industry, firms with higher organizational ability or talent can produce more output with the same amount of input, and thus have higher productivity, than firms with lower ability or talent. Thus, differences in talent have greater economic significance when output prices are high. The productivity with which any given firm operates plants can differ across industries in which it operates. For a given output price and a given talent level, there are decreasing returns to scale in each industry in which the firm operates and at the level of the whole firm.

For concreteness, consider a population of firms that can operate in a maximum of two industries, which we denote as industry 1 and industry 2, respectively. The productivity of each firm can be modeled by a vector $(d_1, d_2)$, where the firm's talent in industry $i$ is $d_i$. Firms that have a higher productivity in industry $i$, $d_i$, produce more output for a given level of inputs if they choose to operate in industry $i$. All firms are assumed to be price-takers and to produce a homogeneous output. Firms use two inputs: industry-specific homogeneous production capacity $k$ and labor $l$. Further assume that firms can trade capacity with other firms in the same industry or build capacity at price $r$ per unit. For tractability, we assume that each unit of capacity produces one unit of output. For each firm, the profit function is

$$d_1 p_1 k_1 + d_2 p_2 k_2 - r_1 k_1 - r_2 k_2 - \alpha l_1^2 - \alpha l_2^2 - \beta(l_1 + l_2)^2, \tag{2}$$

where $p_i$ and $r_i$ are the prices of output and capacity in industry $i = 1$ or 2, $\alpha$ and $\beta$ are positive cost parameters, and $k_i$ is the capacity the firm maintains in industry $i$. The profit function embodies the assumption of neoclassical diminishing returns within each industry (the $\alpha l_i^2$ terms) and the assumption that when organizational talent is a scarce resource, costs depend on the firm's total size (the $\beta(l_1 + l_2)^2$ term). A firm is diversified if $k_1 > 0$ and $k_2 > 0$ and single segment if capacity in only one of the two industries is greater than 0.

The model can be solved at the firm level to give the firm's optimal capacity $(k_1, k_2)$ in each of the industries as a function of its own productivity vector $(d_1, d_2)$, and industry-level variables, demand $(p_1, p_2)$ and the cost of capacity $(r_1, r_2)$. Optimal outputs by the firms in each industry can be obtained by direct optimization. Dropping the firm

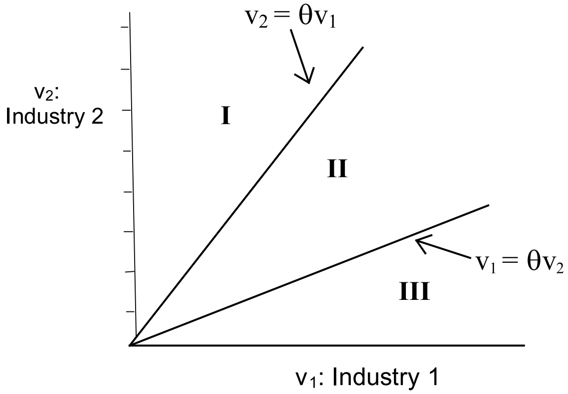



Fig. 2. Optimal production with differing ability across industries.

subscripts and defining $v_i = d_i p_i - r_i$, it can be shown that the optimum output for a firm, assuming conglomerate production, is given by

$$k_1 = \frac{(\alpha+\beta)(d_1 p_1 - r_1) - \beta(d_2 p_2 - r_2)}{2\alpha(\alpha+2\beta)} = \frac{(\alpha+\beta)v_1 - \beta v_2}{2\alpha(\alpha+2\beta)},$$
$$k_2 = \frac{(\alpha+\beta)(d_2 p_2 - r_2) - \beta(d_1 p_1 - r_1)}{2\alpha(\alpha+2\beta)} = \frac{(\alpha+\beta)v_2 - \beta v_1}{2\alpha(\alpha+2\beta)},$$

for $v_2 > \beta v_1/(\alpha+\beta)$ and $v_2 < (\alpha+\beta)v_1/\beta$. For values of $v_1, v_2$ outside of this range, a firm will choose to be a single-segment firm.

Figure 2 illustrates which firms choose to be either conglomerates or single-segment firms. Letting $\theta = (\alpha+\beta)/\beta$, we can illustrate optimal organizational form across industries.[21] If $v_2 > \theta v_1$, then the firm will produce only in industry 1, so that $k_2(v_1, v_2) = \frac{v_2}{2(\alpha+\beta)}$ and $k_1(v_1, v_2) = 0$. Similarly, if $v_1 > \theta v_2$, then $k_1(v_1, v_2) = \frac{v_1}{2(\alpha+\beta)}$ and $k_2(v_1, v_2) = 0$.

Firms in region II optimally choose to be conglomerates, whereas firms in regions I and III choose to produce in a single segment. Specialization is optimal if the firm is much more productive in one industry than the other; diversification is optimal if the productivities are similar. Thus, the decision to diversify depends in part on the firm's comparative productivity in the two industries. An implication of this result is that, all else being equal, a conglomerate's large segment is more productive than its small segment.

The relation between productivity and focus in a population of firms depends both on the distribution of ability within these firms and on the distribution of ability across firms. If organizational talent is industry-specific, firms that are highly productive in

[21] The figure assumes that $r_1 = r_2$. More general cases are discussed in Maksimovic and Phillips (2002).

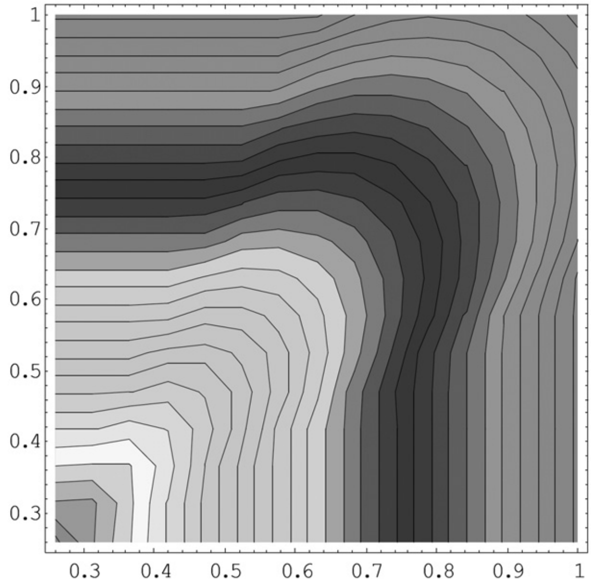



Fig. 3. Market-to-book value contour plot.

one industry are likely to be relatively less productive in the other industries and thus are more likely to operate in a single industry. Firms whose organizations are not highly adapted to any one industry are less focused. By contrast, if organizational talent is not industry-specific, so that $d_1 = d_2$, all firms divide their production equally between the industries. In this case, there is no relation between productivity and focus, and there are no differences in productivity across segments. Larger firms, however, are more productive than smaller firms across all segments.

We can show this relation between the productivity in industry 1 ($d_1$) and productivity in industry 2 ($d_2$) graphically. In Figure 3, we plot "iso-valuation" lines, plotting a firm's market-value-to-book-value (replacement cost of assets) ratio as a function of its productivity in industry 1 ($d_1$) and 2 ($d_2$).[22] We can define a firm's market over book as follows:

$$\frac{\text{MV}}{\text{BOOK}} = \frac{d_1 p_1 k_1 + d_2 p_2 k_2 - \alpha l_1^2 - \alpha l_2^2 - \beta\,(l_1 + l_2)^2}{r_1 k_1 + r_2 k_2}. \tag{3}$$

The axes of Figure 3 are a firm's productivity in productivity in industry 1 ($d_1$) and 2 ($d_2$). The band (the height if the graph were 3D) of the graph tells us the amount produced in each industry and equivalently the average market value to book value of

[22] In this simple context the market-to-book ratio is equivalent to Tobin's $Q$.

the firm. Each band in the figure represents firms with equal market value to book value ratios.

We can observe that for a firm to produce in two distinct industries near a 45 degree line in the center of the graph, it has to have *higher* productivity than firms with equivalent market value to book value ratios. Equivalently, if we match by productivity (or size) single segment firms in two industries to a conglomerate firm producing in both industries, the conglomerate firm will have a *lower* market value to book value ratio than the weighted average of the single segment firms. Thus one cannot in general conclude that multi-segment firms with lower market to book ratios are allocating resources inefficiently.

We now illustrate the effect when we generalize the model allowing firms to produce across ten different industries. We illustrate this using two numerical examples that show how differences in organizational talent across industries causes firms to choose to operate segments of different sizes and different observed productivities.

In each example we take the number of industries to be ten. We assume there are 25,000 potential firms, each of which is assigned firm-specific ability for each of the ten industries. In terms of the previous discussion and the empirical work, high ability is the same as high productivity. We draw the ability assignment d from a normal distribution with a mean ability of 1 and a standard deviation of 0.5. The output and input prices and the cost parameters in all industries are held constant (in this case we set the parameters from equation (1) as follows: $p = 200$, $r = 200$, $\alpha = 5$, $\beta = 2$). In the first example, firm ability is industry-specific. Firms' ability to manage in one industry is independent of their ability to manage in the other industries. Thus, the draws are independent and identically distributed both within firms and across firms. In the second example, there is a firm-specific effect: The draws within a firm for each of the ten industries are correlated. We draw the common ability from a normal distribution with a mean equal to 0 and standard deviation equal to 0.25. We add this common ability to the random industry ability drawn earlier. Thus, part of a firms' ability can be applied equally to all industries. In each case we determine the industries in which it is optimal for each firm to produce and also the amount of each firm production in each industry, given the price of output and the prices of inputs. We keep track separately of firms that choose to produce in one industry only, two industries only, etc., up to firms that choose to produce in all the industries (if such firms exist). Thus, we have simulated data on one-segment firms, two-segment firms, etc. For all firms with a given number of segments, we rank the segments by size, and we compute the mean firm ability d for that segment.

In Figure 4 we allow the draws of firm ability in each of the 10 industries to be independent. We call the industry in which the firm produces its "segments". We label the segment in which the firm produces the most its segment #1, the industry in which produces its second most, its "segment #2", increasing this for each of the firm's remaining segments. The height of the graph ($z$-axis) gives the managerial ability *and* equivalently the size of the firm in that industry in which the firm produces. Each row of the figure

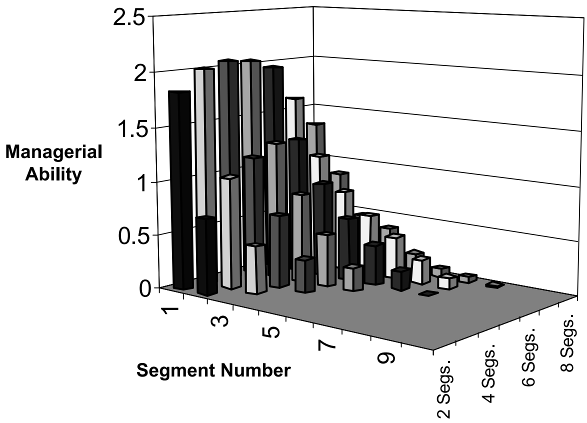



Fig. 4. Ability by segment. Model with no common managerial ability across industries.

thus contains the average of productivity by segment number ($x$-axis) for firms with a given number of segments ($y$-axis).

Figure 4 thus illustrates the case in which the assignment of firm ability is independent across industries in which the firm produces. The figure shows how average firm talent in the economy varies by the number of segments a firm operates in and by segment rank. As predicted, the figure shows that within firms the main segments of conglomerates have higher productivity than peripheral segments. As we go across the number of segments in which a firm operates equally ranked segments at first become more productive and then less productive. The drop-off in productivity occurs because it is very unlikely that any single firm is productive in all ten industries. Thus, firms that choose to produce in many industries are likely to have mediocre ability in all of them. In this simulated example, no firms in the sample produce in all the industries. A simple OLS regression on the simulated data shows that firms' mean productivity is positively and significantly related to their focus, measured by the Herfindahl index, and size. These relations between focus and productivity are obtained even without assuming the existence of agency costs.

In Figure 5, we allow ability in each segment to have a firm-specific component, so that a firm which highly productive in one industry is likely to be highly productive in other segments. As in Figure 4, the height of the graph ($z$-axis) gives the managerial ability *and* equivalently the size of the firm in that industry in which the firm produces. Each row of the figure thus contains the average of productivity by segment number ($x$-axis) for firms with a given number of segments ($y$-axis).

In Figure 5, we still see that the main segments are more productive than the peripherals. However, now equally ranked segments are more productive in firms that operate in more segments. Firms that choose to operate in many segments are on average more productive. Interestingly, a simple OLS regression shows that firms' mean productivity

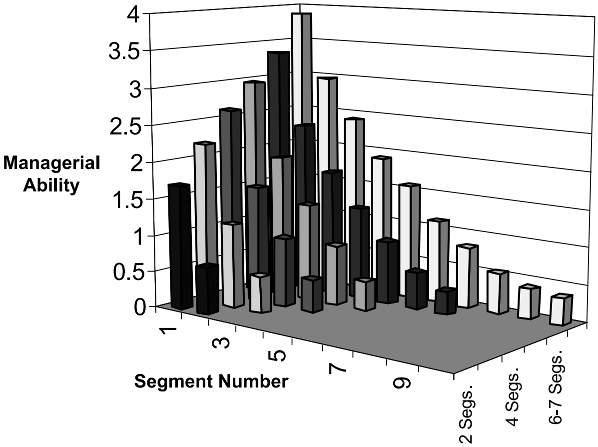



Fig. 5. Ability by segment. Model with common managerial ability across industries.

is again positively and significantly related to their focus, measured by the Herfindahl index, and size, albeit less so than with no common firm talent.

While the simple model makes predictions about the distribution of firms' production, this distribution of production across industries depends on the distribution of ability. However, ability is hard to measure. As a result, the predictions on the distribution of segment size and productivity industries do not directly differentiate the model from other models which predict that firms inefficiently expand into industries outside their core competence. To differentiate the neo-classical from other views, it is necessary to obtain predictions about the firm's responses to exogenous shocks to the industry environment. We discuss this below.

More recently, Gomes and Livdan (2004) embed and calibrate the model in Maksimovic and Phillips (2002) in a dynamic setting.[23] They show explicitly that for the parameter values they select the calibrated model is consistent with Lang and Stulz's (1994) findings on the diversification discount. They can also reproduce Schoar's (2002) finding that expanding focused firms are less productive after diversification than non-expanding focused firms.

[23] Gomes and Livdan (2004) argue that the models differ in certain respects. However, these differences do not affect any of the main intuitions. In essence, the differences come down to technical assumptions that ensure the existence of an equilibrium in which some firms specialize and others do not. Maksimovic and Phillips' implicitly assumes that a firm that chooses to produce in two industries has higher costs than would two identical firms that together produce the same output as the diversified firm but that are *constrained* to specialize in one industry each. By contrast, Gomes and Livdan assume that there is a fixed cost to producing in any industry. Both assumptions serve to counterbalance the assumption of diminishing returns to scale in each industry which both papers make, and which would otherwise make diversification more attractive.

## 4. Investment decisions of conglomerate firms

We next review the recent evidence on the conglomerate discount and conglomerate firms' investment decisions by examining first the investment and resource allocation decision of existing conglomerate firms. We then review the literature on spinoffs and divestitures of conglomerate firms.

There have been four major ways that the literature has addressed how conglomerate firms may invest differentially. First, there has been a branch that has examined whether conglomerate firms have differential investment—cash flow sensitivity. Second, there have been studies examining investment allocation across projects by firms within a single industry. The advantage of the single-industry studies is that in controls for differences investment opportunities that might be hard to measure. Third, several studies have examined how firms should invest when faced with differential opportunities based on the neoclassical investment model. Fourth, studies have examined divestitures and spin-offs for evidence of decreased agency costs after the divestiture. We review each of these areas in turn.

### *4.1. Investment–cash flow sensitivity*

The models of conglomerate investment relate the conglomerate firm's investment expenditures in each segment to the segment's investment opportunities and to the state of the firm's internal capital market.

Neoclassical theory suggests that the firm's level of investment should depend only on its perceived investment opportunities measured by the firm's marginal Tobin's $q$, where marginal Tobin's $q$ is the value of the investment opportunity divided by the cost of the required investment.[24]

Shin and Stulz (1998) and Scharfstein (1998) use this relation between Tobin's q and investment to examine how a firm's internal capital market allocates investment. If the internal capital market is as efficient as the public market for capital we would expect to see a similar relation between investment and Tobin's $q$ for the segments of conglomerates and for single-segment firms.

One set of tests estimates an investment equation on single-segment firms and conglomerates' segments. Consider equation (4)

$$i_j = z_j\gamma + q_j\beta + \zeta_j, \tag{4}$$

where $i$ is the firm's capital expenditures, $q$ is the marginal Tobin's $q$ and $z$ is a vector of exogenous explanatory variables. For single segment firms the marginal Tobin's $q$

[24] Tobin's $q$ is usually defined as the value of the firm (equity and debt claims) scaled by the replacement value of the firm's assets. In the corporate finance literature this quantity is often approximated by the ratio of the market value of a firm's assets (market value of equity + book value of assets − book value of equity-deferred taxes) to the book value of assets. See Whited (2001).

is usually proxied by the firm's average Tobin's $q$.[25] For conglomerate segments we cannot observe the segment's average $q$ directly, but must use a proxy. The usual proxies in the diversification literature are based on the average or median Tobin's $q$s of single-segment firms operating in segment $j$'s industry.

When equation (4) is run via OLS, the coefficient $\beta$ is higher in single-segment firms than in conglomerates, suggesting that conglomerates' segments are insufficiently responsive to differences in investment opportunities. This implies that conglomerates overinvest when opportunities are low and underinvest when they are better (Scharfstein, 1998).

A second set of tests recognizes that in an imperfect financial market the firm's investment expenditures may depend on its cash flow as well as on its marginal Tobin's $q$. For a conglomerate, a segment's investment may depend both on its own cash flows and on the cash flows of the whole firm. Thus, we can augment the investment equation by putting in the cash flows of the segment and that of the whole firm in the investment equation,

$$i_j = z_j\gamma + q_j\beta + \delta\,\mathrm{CF}_j + \phi\,\mathrm{CF}_{-j} + \zeta_j, \tag{5}$$

where $\mathrm{CF}_j$ is the cash flow of segment $j$ and $\mathrm{CF}_{-j}$ is the cash flow of entire conglomerate less segment $j$.

Shin and Stulz (1998) argue that if the internal capital market is working efficiently investment will not depend on a segment's cash flow but on that of the firm as a whole and $\phi \gg \delta$.

It is reasonable to suppose that in an efficient internal capital market the level of investment in one segment will be affected by the level of investment opportunities in other segments. Thus, as further test of the efficiency of the internal capital market equation (5) can be augmented by estimates of Tobin's $q$ for the firm's other segments $-j$.

Using COMPUSTAT Shin and Stulz (1998) examine the workings of internal capital markets of about 14,000 conglomerates for the period 1980 to 1992, paying careful attention to data issues (see the Appendix to their paper). They find that (a) the investment of a conglomerate segment depends more on its own cash flows than on the cash flows of the firm's other segments ($\delta$ exceeds $\phi$); (b) in highly diversified firms, a segment's cash flow is less sensitive to its cash flow than in comparable single-segment firms, (c) a segment's investment increases with its $q$ but is not related to the other segments' $q$s, and (d) the segments with the highest $q$s have the same cash flow sensitivity $\delta$ as other segments.

In sum, Shin and Stulz (1998) find that the internal capital market does not equalize the effect of cash shortfalls across segments. At the same time, a segment's investment is affected by the cash flows of the other segments, notwithstanding differences in Tobin's $q$ across segments. They conclude that conglomerates internal capital markets do not meet their standard of efficiency.

[25] See Hayashi (1982) and Abel and Eberly (1994) for the conditions under which the marginal Tobin's $q$ is well proxied by the average Tobin's $q$.

Shin and Stulz's results suggest that conglomerates may invest less efficiently than single-segment firms, and that, while firm's internal financial markets are integrated, the integration is partial so that the markets are not allocatively efficient. These studies, based on COMPUSTAT data, stand in marked contrast to the findings of MP (2002) using LRD data, who find that conglomerate investment is, on the whole, efficient.

More recent work has tried to reconcile the findings of these papers. As is often the case in research on conglomerates, the issues center on the thorny issue of measurement of the within firm quantities, in this case investment and Tobin's $q$.

A key variable which is difficult to measure at the conglomerate-segment level is Tobin's $q$. As discussed above, the COMPUSTAT based literature attempts to proxy Tobin's $q$ for a segment by using observed $q$s of "comparable" firms. Whited (2001) directly tests whether the findings of the COMPUSTAT based literature can be attributed to measurement error caused by the use of segments' $q$s based on estimated derived from "comparable" single-segment firms.

Whited's arguments can be illustrated with equation (1). As noted above, we cannot observe $q$ directly, but must use a proxy, perhaps based on the average Tobin's $q$s of single segment firms operating in segment $j$'s industry. Whited (2001) models the consequences of the use of a noisy proxy on the estimates of coefficients of $\beta$ in equation (1) and $\beta$, $\delta$ and $\phi$ in equation (2) above. Suppose that the relation between the proxy, $p$ and the Tobin's $q$ takes the following form:

$$p_j = \alpha + q_j + \varepsilon_j.$$

We can eliminate $z$ from this system by regressing all the variables on $z$ and using the residuals. For simplicity we can also initially fold the variables $\mathrm{CF}_j$ and $\mathrm{CF}_{-j}$ with the other exogenous variables into $z$. Doing so we obtain

$$\tilde{i}_j = \tilde{q}_j \beta + \tilde{\zeta}_j, \tag{6}$$

$$\tilde{p}_j = \tilde{q}_j + \tilde{\varepsilon}_j. \tag{7}$$

These equations can be used to generate a set of eight moments such as

$$E\left(\tilde{i}_j^2\right) = \beta^2 E\left(\tilde{q}_j^2\right) + E\left(\tilde{\zeta}_j^2\right), \qquad E(\tilde{i}_j \tilde{p}_j) = \beta E\left(\tilde{q}_j^2\right),$$
$$E\left(\tilde{p}_j^2\right) = E\left(\tilde{q}_j^2\right) + E\left(\tilde{\varepsilon}_j^2\right),$$
$$E\left(\tilde{i}_j \tilde{p}_j^2\right) = \beta E\left(\tilde{q}_j^3\right), \qquad E\left(\tilde{i}_j^2 \tilde{p}_j\right) = \beta^2 E\left(\tilde{q}_j^3\right), \quad \text{etc.}$$

The estimation technique consists of replacing the eight left-hand side moments with their sample estimates and then using GMM to find a vector of six right-hand side unobservable quantities ($\beta$, $E(\tilde{q}_j^2)$, $E(\varepsilon_j^2)$, $E(\zeta_j^2)$, $E(\tilde{q}_j^3)$, $E(\tilde{q}_j^4)$). This vector is one that comes closest to minimizing the distance between the left-hand and right-hand sides of equations, when evaluated using the minimum variance GMM weighting matrix derived by Erickson and Whited (2000).

The estimate of sensitivity of investment $\beta$ is obtained from $\beta = E(\tilde{i}_j^2 \tilde{p}_j)/E(\tilde{i}_j \tilde{p}_j^2)$. Given the estimate of $\beta$, the remaining moment conditions can then be solved to give the other unknowns. Because the estimator provides estimates of $E(\tilde{q}_j^2)$, $E(\varepsilon_j^2)$ and $E(\zeta_j^2)$, Whited (2001) also obtains estimates of the $R^2$ of the first equation, that is the proportion of the variation of capital investment explained by the true Tobin's $q$, as well as the $R^2$ of the second equation, the proportion of the variation of $p$'s (the proxy for Tobin's $q$) variation explained by the true $q$.

Whited (2001) reestimates equations (4) and (5) correcting for the possible error measurement error in the estimates of Tobin's $q$. She finds that the corrected estimate of $\beta$ in equation (4) when estimated over conglomerate segments is insignificantly different from the estimate of $\beta$ for single-segment firms. Thus, she finds that the previous findings of inefficient investment by conglomerates segments may be due to measurement error. She also finds that the corrected estimates of $\phi$ and $\delta$ in equation (5) are insignificantly different from zero, suggesting that the previous finding that the firm's internal capital market is at least partially inefficient might also have been caused by measurement error.

While the formal tests in Whited (2001) are specific to the model she investigates, they raise a serious concern about the use of segment Tobin's $q$s derived from COMPUSTAT data in all studies of intra-firm investment efficiency.

Maksimovic and Phillips (2007) argue that previous studies of investment using Compustat data are subject to another form of measurement error: They exclude a major type of investment expenditure by conglomerates. MP show that single-segment and conglomerate firms differ both in the level of total investment and the *type* of investment. The overall level of capital expenditures on existing plants by conglomerates and single-segment firms in U.S. manufacturing industries is similar. However, conglomerates and single-segment firms differ markedly in their rates of purchases of new plants, even when controlling for segment size. Thus, the COMPUSTAT based studies which use segment capital expenditures as a proxy for investment and do not include acquisitions exclude a major category of investment by conglomerates. Using LRD data, for each single-segment firm and conglomerate segment MP predict $\overline{\mathrm{FD}}_j$ the probability that the segment will be run a financial deficit if it invests at the level predicted by its productivity and industry conditions not taking account whether it is a conglomerate segment or not. They then run the regressions of the following form:[26]

$$\mathrm{acq}_j(\text{or } i_j) = z_j\gamma + \overline{\mathrm{FD}}_j\beta + \delta\,\mathrm{cong} \times \overline{\mathrm{FD}}_j + \phi\,\mathrm{cong}_j + \varphi\,\mathrm{TFP}_j + \zeta_j,$$

where $\mathrm{acq}_j$ is a measure of segment $j$'s acquisition activity, $i_j$ is a measure of segment $j$'s capital expenditures, $\mathrm{TFP}_j$ is the segment's industry standardized productivity and $\mathrm{cong}_j$ is a dummy that takes a value of 1 if the segment belongs to a conglomerate and 0 otherwise. Maksimovic and Phillips (2007) finds that $\beta < 0$, so that a predicted financing deficit leads to a reduction of acquisition and capital expenditure. However,

[26] The regressions in Maksimovic and Phillips (2007) allow for differences across types of industries, but these differences suppressed in this exposition.

$\delta > 0$, indicating that belonging to conglomerate segments reduces a segment's financing constraints. The effect is particularly striking for the rate of acquisitions, which is considerably higher for conglomerates segments, even the ones predicted to run a financing deficit. In further analysis MP show that this effect is greater for the more efficient conglomerate segments and that subsequent to acquisition the acquired plants either maintain or improve their productivity on average. Thus, using LRD data and using TFP together with industry conditions as a measure of investment opportunities MP find no evidence for a negative effect of the internal financial market on resource allocation.[27]

More generally, Maksimovic and Phillips (2002) argue that specifications such as (5) above may be problematical since (a) the decision to become a conglomerate is endogenous and there is likely to be selection bias and (b) the investment of a conglomerate segment does not depend in the same way on investment opportunities as that of a conglomerate firm which maximizes value across different segments. Thus, the estimate of growth opportunities derived from a single-firm Tobin's $q$s may be an inappropriate for the study of investment by conglomerate segments.

### *4.2. Industry studies*

Four case-studies exploring the workings of internal capital markets in specific industries provide another form of evidence on the workings of internal capital markets is provided by. Lamont (1997) studies investment decisions of diversified oil companies following the oil price shock of 1986 when oil prices plunged by over 50%. Khanna and Tice (2001) study the responses of diversified in response to Wal-Mart's entry into their market. Campello (2002) studies banking. Guedj and Scharfstein (2004) analyze the effect of organizational scope on the development strategies and performance of biopharmaceutical firms.

The oil price drop of 1986 provides a natural experiment for the effect of external demand shocks on a conglomerates internal capital market. Lamont identifies approximately 40 non-oil segments owned by 26 oil companies. He tests whether the investment of these non-oil segments of oil firms segments depends on the firm's internal capital market by comparing their capital expenditures with the capital expenditures of similar segments owned by firms less-dependent on the price of oil. Lamont shows that following a significant negative oil price shock, non-oil segments owned by oil companies significantly cut their investment in 1986 compared to the control group of segments not owned by oil companies. Thus, firm-level adverse shock in the oil segment was transmitted to the other segments. Moreover, Lamont finds evidence that the oil companies overinvested in their non-oil segments in prior to the oil price drop.

[27] Maksimovic and Phillips (2007) do not have data on prices paid for the acquisitions. Thus, they cannot determine if the observed increases in productivity are enough to compensate the acquiring firms for the costs of the acquisitions.

Lamont's (1997) interpretation has been queried by Schnure (1997). Schnure examines the cash positions of the 26 oil companies over the period 1985–1986 and finds little evidence that they faced cash constraints. For example, more than half the oil companies in the sample repurchased stock in 1986, many increased dividends and the cash holdings of the sample increased substantially in 1986. This suggests that the relation between the oil price shock and the investment by non-oil segments of oil companies is more complex than the simple transmission of a negative shock via internal capital markets.[28]

Khanna and Tice (2001) examine the responses of discount retailers in response to Wal-Mart's entry into their local markets in the period between 1975 and 1996. Prior to Wal-Mart's entry most markets had several incumbent discount retailers. Khanna and Tice identify 24 stand-alone incumbent discount retailers and 25 incumbent discount divisions of diversified firms. They examine the effect of organizational form by studying the incumbents' responses to Wal-Mart's entry while controlling for factors such as productivity and size.

Khanna and Tice find that conditional on staying in a market following Wal-Mart's entry, diversified firms invest more than focused firm and their investment in more sensitive to their own productivity levels than that of focused firms. They find evidence that diversified firms transfer funds away from failing discount divisions. Moreover, diversified firms appear to be quicker in deciding whether to stay and compete with Wal-Mart or to exit the market.

Some caveats are in order. The diversified firms in Khanna and Tice are for the most part retailers, albeit with non-discount divisions. Thus, their study addresses the effect of capital markets in related diversification. The discount retailing divisions of diversified firms tend to be more productive than the stand-alone firms with which they are compared, raising the possibility of self-selection in the decision to become diversified. However, overall Khanna and Tice conclude that internal capital markets work well for these firms and that the competitive responses of diversified firms are more efficient than those of focused retailers.

Campello (2002) examines the internal capital markets in financial conglomerates (bank holding companies) by comparing the responses of small subsidiary and independent banks to monetary policy. These conglomerates are not diversified across different industries. The advantage of examining diversification within an industry is that it is easier to control for differences in their investment opportunities. Campello finds that internal capital markets in financial conglomerates relax the credit constraints faced by smaller bank affiliates and that internal capital markets lessen the impact of Fed policies on bank lending activity.

Guedj and Scharfstein (2004) contrast the research and development strategies and subsequent performance of small biopharmaceutical firms with those of more mature

[28] The model in Rajan, Servaes and Zingales (2000) suggests that the relative decline in the investment opportunities in oil have made the oil segment less willing to acquiesce to uneconomic transfers to other segments.

firms. The former have no history of successful drug development and are typically focused on one stand-alone project, such as the development of a specific drug, whereas the latter usually have the option of picking among several projects to develop. To the extent that the projects are discrete, the large firms closely resemble the theoretical model of internal capital markets in Stein (1997).

Guedj and Scharfstein analyze a sample of 235 cancer drugs that entered clinical trials in the period 1990–2002. In order to be marketed in the U.S. a drug has to undergo three separate phases of clinical trials. In each phase more information is revealed about the drug's prospects. These trials are expensive, and after each phase is completed the sponsoring firm must determine whether to proceed onto the next stage or whether to curtail the development of the particular drug.

Guedj and Scharfstein find that standalone firms are more likely to push drugs that have completed Phase I trials into Phase II trials. However, standalone firms also have much worse results at Phase II. This pattern especially evident for those standalone firms that have large cash reserves. Thus, as in Stein (1997), single-product firms do not abandon projects optimally, whereas managers of multi-project firms shift resources in response to new information. In that light firm diversification can be viewed as a response to an agency conflict between the managers of single-product firms and shareholders.

Khanna and Tice (2001), Campello (2002), and Guedj and Scharfstein (2004) identify several specific advantages of internal capital markets. Lamont (1997) identifies a potentially countervailing disadvantage: a tendency to transmit investment shocks to the firm's main division to unrelated projects.[29] We next look at attempts to analyze the effect of internal capital markets on a broader scale.

### *4.3. Efficient internal capital markets*

Stein's (1997) model suggests that there is a positive relation between the internal market's efficiency and the amount of external capital a diversified firm raises. Moreover, the efficiency of external capital markets is greater when a firm has more divisions and when the investment opportunities across divisions are not correlated.

Peyer (2001) and also Billet and Mauer (2003) test predictions on how conglomerate firms allocate firms across divisions. For diversified firms, Payer estimates the firm's excess external capital raised as the difference between the firm's use of external capital compares and an estimate of how much a matching portfolio of single-segment firms would have used.

For each diversified firm Peyer obtains the amount of external capital used as the difference between the external capital raised from outside investors and the external capital returned to outside investors.

[29] Maksimovic and Phillips (2002) also find that the operations of peripheral units of conglomerates are cut back much more severely in recessions than their main units. It is unclear whether these cuts occur because of a reduction of the resources available to the firm's internal capital market or because a shock triggers off a re-evaluation of the firm's long-term strategy. Schnure's (1997) results suggests that it might be the latter.

He also computes the firm's imputed use of external capital: for each of the diversified firms' divisions he computes the external capital that would have been raised by the been the median single-segment firm in the same 3-digit SIC code as the division. These estimates are then weighted by divisional sales to obtain the firm's imputed net external capital need. The firm's excess net external capital (EEC) raised by the diversified firm is then computed as

$$\text{EEC} = \frac{\text{Net external capital used} - \text{Imputed net external capital used}}{\text{Lagged book value of assets}}.$$

Peyer estimates the following regression:

$$\begin{aligned}\text{EEC} = {} & \alpha_i + \beta_i + \gamma_1(\text{ICM size})_{i,t} + \gamma_2(\text{ICM efficiency})_{i,t-1} \\ & + \gamma_3(\text{Informational asymmetry})_{i,t-1} \\ & + \gamma_4(\text{Informational asymmetry} \times \text{ICM efficiency})_{i,t-1} \\ & + \gamma_5(\text{Capital need})_{i,t} + \gamma_6(\text{Relative value})_{i,t-1} + \gamma_7(\text{Firm size})_{i,t-1}.\end{aligned}$$

Motivated by Stein (1997), Peyer uses the inverse of the Herfindahl index and the coefficient of variation in $q$ across the firm's divisions as measures of Internal Capital Market (ICM) size.[30]

As a measure of ICM efficiency use RSZ's Relative Value Added by Allocation (RVA), where RVA is defined as

$$\text{RVA}_j = \sum_{k=1}^{n} \frac{\text{BA}_{jk}}{\text{BA}_j}(q_{jk} - \bar{q}_j) \times \text{IAI}_{jk},$$

where $\text{BA}_j$ is the book value of assets of firm $j$, $\text{BA}_{jk}$ is the book value of assets of segment $k$ *and* IAI is a measure of the excess investment in segment $k$.[31]

RVA has the following interpretation: IAI is given a positive weight when the division has relatively good investment opportunities ($q_j - \bar{q} > 0$) and a negative weight when the firm has relatively bad investment opportunities ($q_j - \bar{q} < 0$). Thus a positive RVA indicates that the ICM is efficient because additional investment in being channeled into segments with better than average (for the firm) investment opportunities.

Peyer uses several measures of informational asymmetry: the ratio of intangible to tangible assets, residual variance of daily stock returns and the dispersion in analysts' forecasts. He also computes two additional variables. Excess capital need is measured by Excess internal cash flow = (internal cash flow − imputed internal cash flow)/lagged book value of assets. Relative firm valuation (to control for the propensity of firms to

[30] Peyer also uses diversity as a measure in one of his runs. Following Rajan, Servaes and Zingales (2000) diversity is defined as the standard deviation of the segment asset-weighted imputed $q$ divided by the equally weighted average imputed segment q. As noted above, RSZ predict a negative relation between diversity and ICM efficiency.

[31] We discuss the IAI below. Other measures of excess investment used by Peyer perform similarly.

issue equities after a run-up) is measures using the Lang and Stulz (1994) and Berger and Ofek (1995) measures. Firm size is measured using market valuations.

Peyer finds that firms with efficient ICMs and diversified firms use more net external capital than comparable standalone firms. Measures of information asymmetry are negatively correlated with the use of external capital. The relation is attenuated for firms with efficient ICMs.

EEC is positively related to excess value, especially for firms that have efficient ICMs and firms with larger ICMs. Peyer interprets the positive correlation between the use of external capital and firm value supports the notion that diversified firms are raising external capital to invest in a firm-value-increasing manner.

For robustness, Peyer examines changes in EEC in response to changes in the explanatory variables. He finds that increases in ICM efficiency and increases in the size of the ICM are positively related to changes in EEC. Increases in information asymmetry have a smaller negative effect on EEC if the firm has an efficient ICM. Moreover, there exists an association between the increased use of external capital and firm valuations, measured as in Berger and Ofek (1995).

In all, the Peyer (2001) findings that more efficient ICM firms and firms with larger ICMs use more external capital makes and have a higher firm provides empirical support for Stein (1997).

Billet and Mauer (2003) construct an index of the diversified firm's internal capital market that includes the amount of subsidies and transfers and the efficiency of these flows. Subsidies to division $i$ of firm $j$ are calculated as:

$$\text{Subsidy}_{ij} = \text{Max}(\text{Capital expenditures}_{ij} - \text{After tax cash flow}_{ij}, 0).$$

They calculate the potential transfer from division $i$ to other divisions as:

$$\begin{aligned}&\text{Potential transfer}_{ij}\\&\quad = \text{Max}(\text{After tax cash flow}_{ij} - w_{ij} * \text{dividends}_j - \text{CAPX}_{ij}, 0).\end{aligned}$$

Dividends are determined at the firm level. The firm-level dividends are weighted by $w_{ij}$, the share of assets division $i$ represents of the firm $j$'s assets in the calculation of potential transfers.

Billet and Mauer demonstrate that funds flow toward financially constrained efficient divisions of conglomerates and that these types of transfers to constrained segments with good investment opportunities increase firm value. They show that the higher the transfers to financially constrained segments with good investment opportunities, the higher the overall valuation of the conglomerate.

### *4.4. Bargaining power within the firm and differential investment opportunities*

Rajan, Servaes and Zingales (2000) (RSZ) examine how differential investment opportunities within the firm affect investment efficiency. The empirical tests in RSZ are of two kinds. First, they test whether conglomerates distort their investment expenditures by underinvesting in divisions with better growth opportunities and overinvesting in divisions with worse opportunities. Second, they test their model's predictions about the relation been distortions and the diversity of the firm's operations.

RSZ find that diversified firms invest more in segments with good opportunities than in segments with poor opportunities. However, conglomerates might still misallocate investment flows relative to comparable single-segment firms. Specifically, their theoretical model predicts that segments with good investment opportunities and above average resources will transfer assets to segments with poorer investment opportunities and below average resources.[32] The purpose of the transfer is to reduce the threat that segment with poorer investment opportunities and resources will expropriate the better segments ex-post, thereby improving the better segments' investment incentives.

RSZ cannot directly observe resource transfers between a diversified firm's segments. Instead, they have to infer those transfers for each segment by comparing the segment's investment to the investment of comparable single-segment firms. They attribute differences between the actual investment and the investment of comparable single-segment firms to transfers across divisions. However, RSZ also allow for the possibility that conglomerates may systematically over-invest relative to single-segment firms because they have better access to capital. Thus RSZ measure of the extent to which a segment deviates from its benchmark, the Industry-Adjusted Investment (IAI), subtracts out the weighted average industry-adjusted investment across all the segments of a firm. Thus,

$$\mathrm{IAI}_{jkt} = \frac{I_k}{\mathrm{BA}_k} - \frac{I_k^{ss}}{\mathrm{BA}_k^{ss}} - \sum_{k=1}^{n} w_k \left( \frac{I_k}{\mathrm{BA}_k} - \frac{I_k^{ss}}{\mathrm{BA}_k^{ss}} \right),$$

where $I_k$ is the investment in segment $k$, $\mathrm{BA}_k$ is the book value of assets in segment $k$, is the (asset-weighted) ratio of the capital expenditures to assets of comparable single-segment firms, and $w_k$ is the ratio of segment $k$'s assets to the firm's assets.

In the econometric model they take to data, RSZ predict that a segment's investment depends on the magnitude of its asset-weighted investment opportunities relative to those of the rest. In particular, their model predicts that an increase in diversity should decrease investment in segments that have asset-weighted investment opportunities above the firm average, and increase investment in segments below the firm average.

To test their model, RSZ divide up the segments of each diversified firm in each year along two dimensions (above vs. below average investment opportunities, above vs. below average resources) to obtain a $2 \times 2$ classification matrix of all the segments in their sample. Then for each firm in each year they sum up the IAIs for the firm's segments that fall into each cell receive (thus, in each year each firm will have four observations for the transfers, one for each cell, although some may be missing).

[32] The RSZ model is discussed above in Section 3.4.

RSZ run the following regression equation separately for the segments in each cell of the classification matrix:

$$\sum_{k=1}^{m(j,t)} \mathrm{IAI}_{jkt} = \alpha + \beta \frac{1}{q_{jt}} + \gamma(\mathrm{Diversity})_{jt} + \delta(\mathrm{Firm\ sales})_{jt} + \mathrm{controls} + \varepsilon_{jt},$$

where $\sum_{k=1}^{m(j,t)} \mathrm{IAI}_{jkt}$ is the sum of the IAI across the $m(j, t)$ segments belonging to firm $j$ at time $t$ in the cell, and $q_{jt}$ is the equally weighted average $q$s of firm $j$ segments at time $t$ and the firm's diversity is measured as the standard deviation of the firm's weighted segment $q$s divided by the mean $q$, or

$$\mathrm{Diversity}_{jt} = \sqrt{\sum_{k=1}^{m(j,t)} \frac{1}{m(j,t)-1}\left(w_{jkt}q_{jkt} - \frac{\sum_{k=1}^{m(j,t)} w_{jkt}q_{jkt}}{m(j,t)}\right)} \Bigg/ \frac{\sum_{k=1}^{m(j,t)} q_{jkt}}{m(j,t)}.$$

The control variables include the firm fixed effects and calendar year dummies.

The predictions of the RSZ model are summarized in Table 1.

Investment falls in high opportunity segments with high resources as the firm's diversity increases (cell (1)). Investment increases in low opportunity segments with low resources as diversity increases (cell (4)). Investment increases with diversity in high opportunity resource segments (cell (2)). Investment falls with diversity in large unprofitable segments (cell (1)).

These predictions contrast this with Efficient Internal Market models that emphasize the positive aspects of internal capital markets: top management has the option to reallocate resources from divisions with low investment opportunities to divisions with high investment opportunities. An increase in the diversity increases the value of this option and, thus, should increase the amount of resources transferred to segments with better investment opportunities. Thus, if firms' internal capital markets are efficient, we would observe $\gamma > 0$ in cells (1) and (2) and $\gamma < 0$ in cells (3) and (4) as increases in diversity make transfers between segments more valuable.

Table 1
Predictions of the RSZ Model

| | Segments with resources > firm avg. resources | Segments with resources < firm avg. resources |
|---|---|---|
| Segments with $Q >$ firm average $Q$ | (1) $\gamma < 0$ | (2) $\gamma > 0$ |
| Segments with $Q <$ firm average $Q$ | (3) $\gamma < 0$ | (4) $\gamma > 0$ |

By contrast, Scharfstein and Stein's (2000) model of intra-firm bargaining would imply that the least productive divisions receive transfers from the most productive divisions. Again, an increase in diversity will lead to an increase in this transfer. That model would predict $\gamma < 0$ in cells (1) and (2) and $\gamma > 0$ in cells (3) and (4).

RSZ test their model on 13,947 firm-years in the sample data is obtained from COMPUSTAT for the period 1980–1993. They separate regressions for each cell of Table 1 and obtain parameter estimates that accord with the predictions in the table.

RSZ perform extensive robustness checks. They also verify that (a) investment deviations that they that they classify as value increasing actually are positively related with to diversified firms' value and (b) that diversity itself is negatively related to firm value.

To summarize, even though some transfers in the right direction increase with diversity (cells (2) and (3) in the table above), RSZ find that on average as diversity increases, investment in segments with above-average opportunities becomes too small and investment in segments with below average opportunities becomes too large. This leads reduces the value of the firm of diverse firms.

### *4.5. Investment under a profit—maximizing neoclassical model*

The Maksimovic and Phillips (2002) model differs form the preceding literature in several regards. First, the tests are motivated by the neoclassical profit-maximizing model. The model assumes that each firm has a corporate ability or talent, a fixed resource. It chooses the industries in which it operates so as to extract the maximum value from its ability, diversifying and focusing in response to demand shocks, and the consequent changes in the opportunity cost of assets, across industries. Thus, the focus of the model is not specifically on how well the internal capital market works, but on whether the diversified firms expand in segments in which they have a comparative advantage.

An implication of the MP model is that the decision to diversify is endogenous and depends on segment productivity and industry demand shocks. This implies that the use of single-segment firms as benchmarks for the values of conglomerates' segments is subject to selection bias.

Second, in empirical tests MP use plant-level Survey of Manufactures LRD data to classify each firm's plants into 3-digit SIC code industries. Thus, their classification of firms' assets in not subject to the same discretion that characterizes COMPUSTAT segment data. Moreover, their sample is larger than that of comparable studies. However, the Survey of Manufactures only covers manufacturing industries, so MP, cannot separately identify manufacturers who also operate outside manufacturing.

Third, instead of analyzing capital expenditures at the segment level, MP analyzes the growth in value added. Thus, their measure takes into account growth through whole and partial firm acquisitions as well as through direct capital expenditures.[33]

[33] Maksimovic and Phillips (2007) show that diversified firms are more likely to grow through acquisitions than single-segment firms.

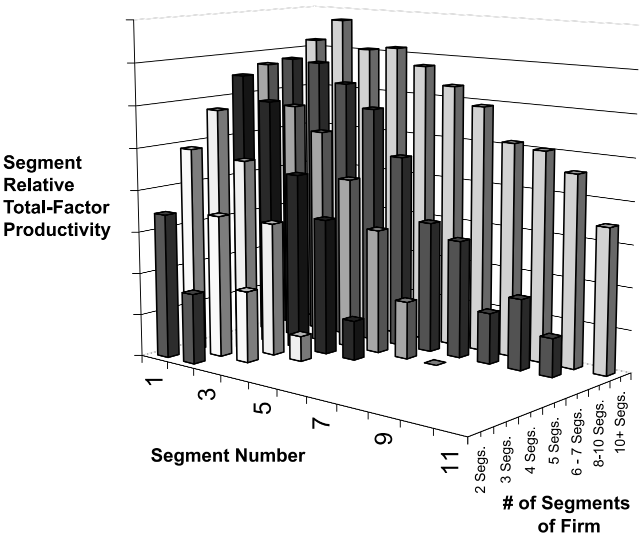



Fig. 6. Productivity ordered by segment size. Segment productivity estimated from LRD plant-level data.

Fourth, MP do not use Tobin's $q$s of single-segment firms to proxy for a segment's growth options. Instead, they use the industry growth in real total value added to obtain industry level measures of investment opportunities. To provide a measure of segment productivity at the micro level, they benchmark each plant in an industry against every other plant in the industry to obtain each plant's predicted real value added in each year given the inputs (capital, energy, labor and materials) the plant used in that year. They use the difference between the plant's actual value added and the predicted value as a measure of the plant's relative productivity. They aggregate up their measure plant productivity to derive the segment-level Total Factor Productivity (TFP) for each year.[34]

Using LRD data, MP can directly observe how the productivity of diversified firms' segments varies by the number of segments and the segment's relative size within the firm. Figure 6 summarizes the data. Controlling for the number of segments in a diversified firm, TFP decreases as the segment's relative size within the firm falls.

The pattern of productivity in Figure 6 is consistent with a neoclassical model in which firms spread their operations across a range of industries in which they have a comparative advantage and in which they have decreasing returns to scale.[35] Consistent with the model, larger firms are more productive on average than smaller firms. However, Figure 6 can also be given an agency interpretation: large productive firms may

[34] Schoar (2002) shows that TFP at the firm level predicts the conglomerate discount for diversified firms.
[35] Decreasing returns to scale may arise from the production technology. More broadly, a firm may perceive itself as having decreasing returns to scale because expansion may provoke a competitive response by rival firms.

waste resources by diversifying into industries in which they do not have a comparative advantage.

To distinguish between these interpretations, Maksimovic and Phillips (2002) examine how firms respond to industry demand shocks. According to their model, firms should grow the segments in which they are particularly productive and that have received a positive demand shock. They should reduce the growth of segments in which they are not efficient and which have received a negative demand shock.

To see this, recall from Section 3.5 that the output of a conglomerate firm $i$ in industry 1 depends on $v_{1i} = d_{1i} p_1 - r_1$. Suppose that there is positive demand shock in industry 1. First, output prices increase, $\Delta p_1 > 0$. Second, the price of capacity in the industry also increases, $\Delta r_1 > 0$. Productive firms in the industry increase output in industry 1. More formally, they experience an increase in their $v_{1i}$ because the marginal positive effect of a price rise, $d_{1i} \times \Delta p_1$, outweighs the effect of an increase in the cost of capacity $\Delta r_1$. The higher the ability $d_{1i}$, the more capacity a firm adds in response to a positive price shock.

The effect of a price shock in industry 1 on the marginal producers not is more complex. If the effect of the expansion by the productive firms on $\Delta r_1$ is minor, then the marginal firms may also expand, although at a slower rate than the more productive producers. However, if the price of capacity is bid up sufficiently high so that for some firms with small $d_{1i}$, $d_{1i} \times \Delta p_1 - \Delta r_1 < 0$, then these marginal producers will sell some capacity to more productive producers and focus instead in industry 2. These firms' operations in industry 1 decline not only relative to those of more productive firms, but in absolute size as well.[36]

The MP model also generates another testable prediction regarding cross-segment effects. When managerial capacity is a fixed factor of production, investment decisions by conglomerate firms in one segment create opportunity costs for investments in other industries in which they operate. Thus, segments' investment decisions will depend on the relative demand growth across all the industries in which the conglomerate operates. Specifically, suppose that there is a large positive demand shock in one of the industries in which a conglomerate operates. If the conglomerate's segment in that industry is highly productive, it will grow relatively fast. This growth increases the conglomerate's costs in other segments, thus decreasing the other segments' optimal size. Thus, a conglomerate which has a productive segment in an industry that has received a positive demand shock will grow more slowly than it otherwise would have in its *other* segments.

Suppose, instead, that the conglomerate had an unproductive segment in a fast growing industry. Then the conglomerate may find it optimal to divest or reduce operations in the high growth industry for two reasons. The positive demand shock in the industry will have increased the value of its capacity, increasing a low productivity conglomerate's

[36] Here we assume that the effect of a demand increase in industry 1 affects $v_1$ only and does not affect $v_2$. Appendix A discusses both effects.

opportunity cost of staying in the industry rather than selling out to a high productivity producer. This effect is amplified because any reductions in growth or divestitures in a segment in which the conglomerate is less productive will produce positive externalities in its segments. Hence, a conglomerate which is a relatively unproductive producer and therefore divests or grows more slowly in an industry that has received a positive demand shock will grow faster than it otherwise would have in its *other* segments.

The predictions concerning cross-segment effects are derived in Maksimovic and Phillips (1999, 2002) and briefly reviewed in Appendix A. They differ from predictions of models that stress influence costs, which suggest that resources are transferred to unproductive segments, and empire-building models, such as Lamont (1997) that suggest that wealth generated by positive shocks is dispersed throughout the firm.

To test their model Maksimovic and Phillips (2002) run the following regression on a sample of 270 thousand segment years over the period 1977–1992.

$$\begin{aligned}\text{GROWTH} = \alpha &+ \beta(\text{Industry shock}) + \gamma(\text{Segment TFP})\\ &+ \delta(\text{Industry shock}) \times (\text{Segment TFP})\\ &+ \phi(\text{Other segments' TFP})\\ &+ \theta(\text{Relative demand}) \times (\text{Other segments' TFP}) + \text{controls}.\end{aligned}$$

MP use TFP as measure of each segment's productivity. The TFP takes the actual amount of output produced for a given amount of inputs and compares it to a predicted amount of output. The measure is computed at the plant level and aggregated up to segment level. "Predicted output" is what a plant should have produced, given the amount of inputs it used. A plant that produces more than the predicted amount of output has a greater-than-average TFP. This measure is more flexible than a cash flow measure, and does not impose the restrictions of constant returns to scale and constant elasticity of scale that a "dollar in, dollar out" cash flow measure requires. Demand shocks are measured by changes in the industry real shipments.[37]

Consistent with the model, MP find that productive segments grow faster ($\gamma > 0$), especially in industries which have experienced a positive demand shock ($\delta > 0$). Most importantly, a segment's growth rate is lower if the firm has more productive operations in other industries ($\phi < 0$). The segment's growth is further reduced if these more productive operations are in industries which have received a positive demand shock ($\theta < 0$). The last two finding are consistent with the cross-segment predictions of MP's neoclassical model but difficult to reconcile with an agency model in which the firm invests inefficiently.

As a robustness check MP identify a subsample of "failed conglomerates" (diversified firms which restructure by decreasing the number of segments by at least a quarter)

[37] MP show that their results also hold for several other measures of productivity and demand shocks.

and a control subsample of “regular” conglomerates that do not reduce the number of segments so substantially. They run their regression separately on the two subsamples. For the failed conglomerates the coefficients $\phi$ and $\theta$ are not significantly different from zero for the period prior to restructuring. The subsample of “regular” conglomerates these coefficients are negative and significant, as predicted by the model. Thus, MP find evidence that there is subset of “failed” conglomerates that grow inefficiently, and are subsequently broken up. However, even for these failed conglomerates MP do not find a positive significant relation the segments’ growth rates and other segments productivity. Thus, they find no evidence that even these failed conglomerates systematically grow their unproductive segments at the expense of productive segments.[38]

MP also find that a segment’s relative size in the firm does affect its growth, even controlling for productivity. Main segments of firms (i.e., segments that produce at least a quarter of its value added) grow faster in response to positive demand shocks than peripheral segments. In part this is because main segments are on average more productive. However, a substantial growth differential remains even after controlling for productivity.

The growth differential is especially pronounced in recessions. Rather than being cushioned in recessions as predicted by models that stress bargaining within the firm, peripheral segments of conglomerates are cut sharply in response to negative demand shocks. These cuts are greater than predicted by MP’s simple neoclassical model. They suggest that a more complex mechanism is at work. Thus, negative demand shocks may cause diversified firms to reassess the prospects of their peripheral segments and to shift resources into more promising ventures, as modeled by Stein (1997).

The decline in peripheral divisions is also reflected in aggregate Census data. In the beginning of the 1980s main divisions of diversified firms produced about half of the value added by U.S. manufacturing and this share was maintained through the end of the 1990ies. By contrast, the share of peripheral segments of diversified firms fell from 27.5% to 23.5% over that period.

In sum, Maksimovic and Phillips (2002) find that a simple profit maximizing neoclassical model of firm growth across segments is consistent with plant-level data and that there is little evidence of systematic resource misallocation by diversified firms. There is some evidence that failed conglomerates that are subsequently broken up do not allocate resources model efficiently. However, even these firms do not systematically grow unproductive segments at the expense of productive segments. Instead, there is evidence that smaller, less productive units of conglomerates grow more slowly than their main divisions or similarly productive stand alone firms.

[38] MP perform robustness tests using several alternative measures of productivity and investment. Their model predicts, for example, predicts that segment size is a proxy for segment productivity. The results using segments size yield the qualitative results.

### *4.6. Mergers and acquisitions, divestitures and spinoffs*

#### *4.6.1. Diversified firms and the market for assets*

The early theoretical literature on internal capital markets, such as Stein (1997), explicitly recognizes the importance of the size of the internal market for its efficiency. Thus, while the importance setting the firm's boundaries for the quality of the internal capital market was recognized early, much of the literature takes a partial equilibrium approach and assumes that the firm's boundaries as given. This is potentially important since many of the hypothesized conflicts within the firm can be solved or mitigated by selling assets that do not fit well with the firm's total portfolio. Thus, if the market for firms' assets is efficient, the magnitude of the investment distortions that can be created by conflicts within the firm is likely to be tightly bounded. Of course, there may still be conflicts of interest between top management and shareholders. However, top managers have clear incentives to maximize firm value.[39]

An objection to this might be that the market for corporate assets is insufficiently liquid so that firms which attempt to readjust their portfolio by selling segments face a capital loss. This is unlikely. Tables 2A and 2B, from Maksimovic and Phillips (2001), shows that there exists a large, procyclical market for segments and individual plants.

Using Census, data Maksimovic and Phillips (2001) show that in the period 1974–1992, 1.94% of all manufacturing plants change ownership annually in partial-firm transactions.[40] This is comparable to the total rate at which manufacturing plants change ownership in all-firm mergers and takeovers over this period, 1.95% annually. Similar rates of partial firm sales occur in both growing and declining industries. The market for divisions and plants is a market dominated by conglomerates. MP report that the sellers operate in an average of 10 4-digit SIC industries and the buyers in an average of 8 such industries.

MP test whether diversified firm's decision to sell a manufacturing plant can be explained by their neoclassical model. They run a probit regression on a panel of plants 1979–1992 from the LRD, where the dependent variable, PLANT SALE, takes on the value of 1 if the plant is sold and the value of 0 if the plant is not sold in a given year.

$$\begin{aligned}\text{PLANT SALE} = {} & \alpha + \beta(\text{Industry shock}) + \gamma(\text{Segment TFP}) \\ & + \delta(\text{Industry shock}) \times (\text{Segment TFP}) \\ & + \phi(\text{Other segments' TFP}) \\ & + \theta(\text{Relative demand}) \times (\text{Other segments' TFP}) + \text{controls}.\end{aligned}$$

Consistent with the profit maximizing model, MP find that plants in productive segments are less likely to be sold ($\gamma < 0$), especially in industries which have experienced

[39] For a contrary view, see Aggarwal and Samwick (2003).

[40] See Maksimovic and Phillips (2001) for a detailed description of the sample. See also Schlingemann, Stulz and Walking (2002) for a discussion of liquidity in the market for assets on the rate of sales.

Table 2A

| | Total | Mergers and acquisitions | Asset sales | |
|---|---|---|---|---|
| | | | Full segment | Partial segment |
| Reallocation rates across and within industries<br>Full period: 1974–1992 | | | | |
| Number of plants reallocated | 35,291 | 17,720 | 8,556 | 9,015 |
| Average annual % of plants reallocated | 3.89% | 1.95% | 0.95% | 0.99% |
| % Plants sold to buyer inside industry | | | | |
| Same three-digit SIC code | 56.8% | 54.1% | 55.5% | 63.1% |
| Same four-digit SIC code | 47.7% | 44.9% | 47.9% | 53.0% |
| Average plant size (Real $ in thousands, value of shipments) | $30,332 | $28,435 | $30,916 | $33,506 |
| Average industry plant size (Real $ in thousands, value of shipments) | $33,790 | $34,569 | $36,440 | $37,574 |

a positive demand shock ($\delta > 0$). Most importantly, a plant's probability of being sold is higher if the firm has more productive operations in other industries ($\phi > 0$). The probability of being sold further increased if these more productive operations are in industries which have received a positive demand shock ($\theta < 0$). The last two finding are consistent with the simple neoclassical model but and do not suggest an agency model in which the firm retains and subsidizes inefficient plants using resources generated by more successful divisions.

MP also find that there is negative relation between the probability that a plant is sold and the share of the firm's output produced by the segment to which the plant belongs. The finding is consistent with the notion that diversified firms divest from their smallest and least productive divisions and redeploy their assets.

MP also examine who purchases plants and firms and find that the probability of a purchase goes up with the buyer's productivity. When they examine the productivity of the plants after the purchase, MP find that the change in productivity increases with difference between the buyer's productivity and purchased plant's productivity. In sum, the evidence is consistent with transfers of assets going from less to more productive firms—especially when industries receive positive demand shocks.

More recently, Maksimovic, Phillips and Prabhala (2006) show that acquirers sell about 40% of the target's plants in the four years after the acquisition. The sold plants tend to be those in the target's peripheral divisions. The plants that are kept increase in productivity after the acquisition, the plants that are sold do not. This pattern is consistent with the hypothesis that acquirers keep the assets which they can exploit efficiently and that they economize on managerial attention by selling or closing the assets that they cannot exploit efficiently.

Taken together, plant-level evidence suggests that the direction and timing of sales of corporate assets is consistent with an efficient allocation of resources within the firm.

Table 2B

| | Sample of firms | | | |
|---|---|---|---|---|
| | Total | Mergers and takeovers | Asset sales | |
| | | | Full division | Partial division |
| Transactions by aggregate economy conditions | | | | |
| Recession years (1981, 1982, 1991) | | | | |
| Average % reallocated (total number) | 3.57% (5,148) | 2.16% (3,112) | 0.70% (1,003) | 0.72% (1,033) |
| Expansion years (1986, 1987, 1988) | | | | |
| Average % reallocated (total number) | 6.19% (8,989) | 2.69% (3,904) | 1.73% (2,509) | 1.77% (2,576) |
| Indeterminate years | 3.21% | 1.73% | 0.70% | 0.78% |
| Transactions by industry capacity utilization | | | | |
| Low industry capacity utilization (bottom quartile) | | | | |
| Average % reallocated (total number) | 3.86% (8,618) | 1.90% (4,244) | 0.99% (2,210) | 0.97% (2,164) |
| High industry capacity utilization (top quartile) | | | | |
| Average % reallocated (total number) | 3.69% (8,413) | 1.92% (4,375) | 0.87% (1,977) | 0.90% (2,061) |
| Transactions by long-run industry growth/decline | | | | |
| Quartile 1: Declining industry growth | | | | |
| Average % reallocated (total number) | 4.01% (6,290) | 1.95% (3,058) | 1.09% (1,707) | 0.97% (1,525) |
| Quartile 2 | | | | |
| Average % reallocated (total number) | 3.86% (5,250) | 1.96% (2,666) | 1.05% (1,425) | 0.85% (1,160) |
| Quartile 3 | | | | |
| Average % reallocated (total number) | 3.52% (10,008) | 1.80% (5,131) | 0.88% (2,505) | 0.83% (2,372) |
| Quartile 4: High industry growth | | | | |
| Average % reallocated (total number) | 4.03% (15,746) | 2.01% (7,870) | 0.87% (3,405) | 1.14% (4,471) |

Source: Maksimovic and Phillips (2001), Journal of Finance.

On average, the good assets are kept and the assets that cannot be exploited efficiently are sold. When the opportunity cost of retaining marginal assets is higher because other segments are more productive and growing faster, the rate at which marginal assets are disposed off is higher.

Schoar (2002) also used the LRD plant-level data to examine productivity of conglomerate firms and changes in productivity following plant acquisitions. Schoar (2002) establishes that market valuations of single-segment and conglomerate firms track estimates of productivity derived from LRD data. The tracking is equally strong for single-segment and conglomerate firms. This suggests that the conglomerate discount, if it exists, is unlikely to be caused by investors' inability to evaluate diversified firms' operations as efficiently as those of single-segment firms.

Schoar also finds no evidence that conglomerates' plants are less efficient than those of single-segment firms. Specifically, using plant-level data she runs the following regression

$$\text{TFP} = a + b * \text{DIV} + c * (\text{plant size}) + d * (\text{plant age}),$$

where TFP is total factor productivity and DIV is a dummy that takes on a value of 1 if the plant belongs to a diversified firm and zero otherwise. The coefficient of DIV is positive and significant and remains so when the equation is augmented by segment-level control variables.

Like Maksimovic and Phillips (2001), Schoar finds that acquired plants on average increase in productivity while the acquirer's own plants decline in productivity. She calls this the "new toy" effect, and argues that post-acquisition productivity of the acquirer is on balance negative. However, as Schoar points out this time-series effect does not cancel out the cross-sectional finding that diversified firms' plants have a higher TFP.

An intriguing possibility raised by Schoar's work is that a diversification discount may arise because conglomerates pay out a higher proportion of their revenues in salaries and benefits than standalone firms. She finds that diversified firms pay higher hourly wage rates than similar standalone firms. Assuming that these differences do not reflect differences in the educational level or quality of their respective workforces, the wage difference is enough to explain a 2–3% discount for diversified firms.

### *4.6.2. Spinoffs*

Several studies, including Gertner, Powers and Scharfstein (2002), Dittmar and Shivasani (2003), Burch and Nanda (2003), Anh and Denis (2004), and Colak and Whited (2005), examine spinoff and divestiture decisions that reduce the number of divisions that a conglomerate firm operates. These papers examine the investment efficiency of firms before and after the refocusing decision. This approach has potential advantages over studies that examine a sample of firms, some of which refocus and some which do not. If it can be assumed that the severity of measurement error does not change over time, measurement error bias that in the comparison of before and after refocusing performance, is mitigated. These papers further argue that they have reduced omitted variables bias by focusing on changes in value and efficiency in a single sample of firms.

Gertner, Powers and Scharfstein (2002) examine sensitivity of segment investment to the median Tobin's $q$ of the single-segment firms in that segment's industry.

The sensitivity of Tobin's $q$ captures the idea that the more efficient a firm is, the more it should respond to changes in investment opportunities by altering its investment policy. In order to get around the problem that the median industry Tobin's $q$ is an imperfectly measure of investment opportunities for an individual firm, Gertner, Powers and Scharfstein (2002) paper examines the same firm's sensitivity of investment to Tobin's $q$ before and after the spinoff. They find that segment sensitivity to industry Tobin's $q$ increases after the segment spinoff and that changes are related to the stock market's reaction to the spinoff decision.

Dittmar and Shivasani (2003) find that the announcement returns for divestitures are significantly correlated with the change in the diversification discount. Larger decreases in diversification are associated with higher announcement returns. Dittmar and Shivdasani also find that RSZ measures of the efficiency of segment investment increase substantially following the divestiture and that this improvement is associated with a decrease in the diversification discount. One can interpret this evidence in several ways. The evidence is consistent with the firm divesting divisions will now be run more efficiently. Alternatively, the evidence is also consistent with changes in investment opportunities for the divesting firm or its divisions and thus the market responds positively as firms change their investment.

Burch and Nanda (2003) examine whether changes in value following spinoffs are related to measures of investment diversity by reconstructing the diversified firm after the spinoff. They construct changes in value using both industry multiples and also using firm-specific measures. To avoid the measurement error problem of assessing opportunities using industry measures, they also use an ex post, direct measure of excess value based on the post-spinoff market-to-book values of the divested division(s) and remaining parent firm. As they note, using ex-post data implicitly assumes that diversity in post-spinoff investment opportunities is a reasonable proxy for the diversity prior to the spinoff. Using these measures, they find that improvements in aggregate excess value (changes in the implicit discount less the actual pre-spinoff discount) depends significantly on direct measures of diversity and changes in measures of diversity based on industry proxies.

Anh and Denis (2004) also examine the changes in measure of investment efficiency from RSZ pre- and post-spinoff. They find that post-spinoff, measures of investment efficiency increase for the hypothetical combined firm—combining the post-spinoff divisions with the parent in order to examine the total impact of the spinoff decision. They also find that the measures of investment efficiency increase the most for firms with the highest dispersion in the segment Tobin's $q$s from single-segment firms. They do note two caveats to their analysis. First, they note that by focusing just on firms that choose to spinoff divisions, they may be focusing on the set of firms with more severe investment inefficiencies. Second, they note that other changes in the investment opportunity set may be driving firms to spinoff and also contributing to the observed changes in investment efficiency.

Colak and Whited (2005) show the caveats noted in these papers are important. Their results challenge the view that these spinoffs and divestitures provide evidence that firms were misallocating resources prior to the spinoff. Using three-different approaches to control for endogeneity they show that refocusing decisions does not necessarily cause improvements in efficiency. In particular, firms that choose to spin-off and divest divisions are larger, more diversified, and subject to more serious problems of asymmetric information. Further, the spun-off segments tend to be in fast growing industries with a great deal of IPO and corporate control activity. Finally, they appear to have experienced recent unanticipated shocks to profit. They find that although spin-offs and divestitures may be associated with improvements in investment efficiency, they do not cause these improvements. When they control for measurement error, they also show that the sensitivity of investment to both industry Tobin's $q$ does not significantly change following the refocusing decision.

## 5. Conclusions: What have we learned?

There have been a substantial number of careful empirical papers on internal financial markets in the last few years. Any summary of what has been learned is bound to be subjective and reflect the interests of the authors. With that caveat in mind, we can summarize the existing evidence about internal capital markets.

- The early work established clearly that, using single-segment firms as benchmarks, there exists a conglomerate discount.
- Initial attempts to explanation the discount focused on agency conflicts and conflicts among divisions that led to overinvestment in divisions with poor prospects and underinvestment in divisions with high $q$s.
- Conclusions drawn from econometric studies of segment capital expenditures, which use the Tobin's $q$s of single-segment firms to proxy for segment investment opportunities, are subject to measurement error and may not be valid.
- Diversified firms rely more on acquisitions than single-segment firms. Thus, studies that focus on capital expenditures may miss important components of investment by diversified firms.
- A conglomerate discount is not, by itself, evidence of agency or inefficiency—it may be due to the fact that single segment and diversified firms operate on different regions of the production function.
- A simple neoclassical model that recognizes that the decision to diversify is endogenous and that firms grow fastest in industries where they have a comparative advantage in response to positive demand shocks in those industries is consistent with the growth patterns of diversified firms.
- The sales of plants by firms are also consistent with a simple neoclassical profit-maximizing model.
- Much of conglomerate discount can be explained by sample selection. Firms that choose to diversify, or to stay diversified or to be acquired by diversifiers inherently differ from single-segment firms.

- On balance, industry case studies and econometric analyses of firm growth suggest that internal capital markets are efficient in reallocating resources.
- Even controlling for productivity, main and peripheral segments of diversified firms are treated differently. Main divisions grow faster, are less likely to be cut back in recessions, and less likely to be sold.

In our review of the evidence and econometric results, we have come to the conclusion that diversified firms predominantly behave like value maximizers given their productivity and internal capital markets facilitate the efficient transfer of resources. The evidence is broadly consistent with firms making endogenous value-maximizing choice of organizational form and allocating resources across industries consistent with a neoclassical model of resource allocation.

However, there is a large part of the literature that reaches different a conclusion, that conglomerate firms usually misallocate resources. Given the latest evidence, we are unable to reach this conclusion *for the majority* of conglomerate firms. However, there is some evidence that conglomerate firms that are busted up had investment patterns that varied from the neoclassical model. In addition, other puzzles do remain. In particular, the differences in growth patterns of main and peripheral divisions of diversified firms still have to be explained.

The conclusion that internal capital markets do not, on average, promote resource misallocation does not imply that firms are not subject to agency problems. Managers may allocate resources efficiently, but then expropriate the shareholder value created using those resources. Similarly, diversified firms may overpay for acquisitions that increase the firm's total value added from manufacturing activities.

More generally, the empirical literature on internal capital markets is an excellent case study of the importance of specifying the underlying benchmark model, paying attention to strengths and weaknesses of alternative data sources, and addressing econometric issues such as sample selection and measurement error. Seemingly reasonable choices at any of these steps are fully capable of leading to different results. As a result, the area remains of active interest to researchers.

## Appendix A. Neoclassical model of resource allocation across industries

In this appendix we illustrate how demand shocks affect the relative resource allocation and output of efficient and inefficient producers in an industry and also efficient and inefficient segments within a multi-industry setting. The exposition is based on the working paper version of MP (2002) and complements the discussion in Section 3.5.

We begin by analyzing how firms change capacity in response to demand shocks in a single industry and then generalize the model to multiple industries. We also discuss how these predictions differ from those of agency models in the literature.

### *A.1. Shocks and growth in a single industry*

We first analyze the relative growth rates and the flow of assets between differing productivity over the business cycle in a single industry. Accordingly, in this subsection we assume that all firms in the industry are single-segment firms that produce only in one industry.

We start by simplifying the firm's profit function given in equation (2) in the text to the one industry case

$$pd^j k^j - rk^j - \beta\big(k^j\big)^2,$$

where, for simplicity, we have abstracted from labor costs (so that $\alpha = 0$ in equation (2) in the text). The subscript $j$ refers to firm $j$. Recall that $r$ is the market price of a unit of capacity and $\beta$ is the standard neoclassical diseconomy of scale. To reduce notation, we further assume without loss of generality that $d^j$ can take one of only two values. Let high productivity, or H firms, produce one unit of that industry's output per unit of capacity so that for those firms $d^j = d^{\mathrm{H}} = 1$. Let low productivity, or L firms, produce only $d^j = d < 1$ units of output per unit of capacity. Thus, the profit functions specialize to $pk^{\mathrm{H}} - rk^{\mathrm{H}} - \beta(k^{\mathrm{H}})^2$ for H firms and $pdk^{\mathrm{L}} - rk^{\mathrm{L}} - \beta(k^{\mathrm{L}})^2$ for L firms, after adjusting the notation to reflect the fact that all the H (L) firms are identical, and where and the number of capacity units operated by H and L firms is $k^{\mathrm{H}}$ and $k^{\mathrm{L}}$, respectively. Assume that total amount of capacity available to the industry is $K = \sigma + \rho r$, $\sigma, \rho > 0$. Thus, we assume that the supply of capacity is not perfectly elastic, reflecting the addition of new capacity (for high levels of $r$) and sales for scrap (for low levels of $r$).

Assume that there is an exogenously determined number, $n$, of entrepreneurs and that the proportion of entrepreneurs that can operate H firms is $\lambda$. To avoid discussion of firm entry and exit, which would require more notation, also assume that the opportunity cost of capacity outside the industry is sufficiently low so that it is optimal for all high- and low-quality firms to operate at the level of demand we are considering.

The time sequence is as follows. There is one period and two dates: $t = 1, 2$. At time $t = 1$, the entrepreneurs learn the actual realization of the next period's level of demand in the industry. A market for capacity opens in which firms can purchase capacity units at a price $r$. The price of capacity, $r$, adjusts so that supply equals demand for capacity. At time $t = 2$, the firms realize the cash flows. For simplicity, we assume that capacity has no salvage value at $t = 2$.

To make explicit the role of demand shocks and the distribution of capacity units on firm growth, we describe the equilibrium in the market for output. The market price that the customers pay in industry for the output is determined as $p = a - bn(\lambda k^{\mathrm{H}} + (1 - \lambda)k^{\mathrm{L}})$, where $n(\lambda k^{\mathrm{H}} + (1 - \lambda)k^{\mathrm{L}})$ is the aggregate output and $a, b$ are positive constants.

**Remark 1.** A positive demand shock causes, productive profit maximizing firms increase in size relative to less productive profit maximizing firms.

**Proof of Remark 1.**

We obtain the output of type H firms by maximizing the firm's operating profit, $pk^{\mathrm{H}} - rk^{\mathrm{H}} - \beta(k^{\mathrm{H}})^2$. Solving for $k^{\mathrm{H}}$, we obtain $\frac{p-r}{2\beta}$ as the optimal capacity that type H firms operate at the given opportunity cost, $r$. The capacity at which the low-quality firms operate is similarly obtained as $k^{\mathrm{L}} = \frac{pd-r}{2\beta}$. Notice that $k^{\mathrm{H}} > k^{\mathrm{L}}$, so that a type H firm uses more capacity than the low-quality firm at every price level.

If both H and L firms are active in the industry and the price of capacity exceeds its salvage value, the market price of the output is $p = a - bn(\lambda k^{\mathrm{H}} + d(1-\lambda)k^{\mathrm{L}})$. We determine the price of capacity by equating the demand for capacity by each type of firm to the total number of capacity units available, either on the secondary market or as supplied by manufacturers, so that

$$\sigma + \rho r = \lambda n \frac{p-r}{2\beta} + (1-\lambda) n \frac{pd-r}{2\beta}, \tag{A.1}$$

where the total amount of capacity employed by the industry is $K = \sigma + \rho r$. The first term on the right hand side of the equation is the demand for capacity by the $\lambda n$ high-quality firms. The second term is the demand for capacity by the $(1-\lambda)n$ low-quality firms. Solving equation (1) for the opportunity cost of capacity yields

$$r = \frac{p(\lambda + d(1-\lambda))}{n + 2\beta\rho} - \frac{2\beta\sigma}{n + 2\beta\rho}. \tag{A.2}$$

Substituting the expression for the rental cost of capital (A.2) into the expressions for the desired capacity by high- and low-quality firms, we obtain

$$k^{\mathrm{H}} = \frac{\sigma}{n + 2\beta\rho} + \frac{(1-d)(1-\lambda)n + 2\beta\rho}{2w(n + 2\beta\rho)} p,$$
$$k^{\mathrm{L}} = \frac{\sigma}{n + \beta\rho} - \frac{(1-d)\lambda n - 2\beta\rho d}{2w(n + 2\beta\rho)} p.$$

The derivative of the ratio $(k^{\mathrm{H}}/k^{\mathrm{L}})$ with respect to the output price, $p$, is

$$\frac{2(1-d)(n+2\rho)\beta K}{(2w\sigma + (2\beta d\rho - (1-d)\lambda n)p)^2} > 0.$$

The last expression shows that a positive price shock (increase in $p$) increases the ratio $k^{\mathrm{H}}/k^{\mathrm{L}}$. Thus, positive price shocks are associated with higher growth of high-quality firms relative to low-quality firms. Since positive demand shocks to $a$ at time $t = 1$ translate into increases in $p$, it is straightforward, but messy, to show that the same relation obtains for the ratio $k^{\mathrm{H}}/k^{\mathrm{L}}$ and $a$.[41] □

**Remark 2.** Consider a multiperiod generalization of the above industry equilibrium in which the model is repeated over a sequence of dates, with the demand intercept $a$ changing over time. Positive (negative) innovations in a will cause more productive firms to engage in purchases of new capacity and purchases from other firms (divest) and less productive firms to divest (acquire) capacity.

[41] The analysis presented here assumes an interior equilibrium. A full analysis would take into account the exit and entry of entrepreneurs.

In a multi-period setting firms don't need to acquire all their capacity in each period. After the first period, they have an endowment of capacity form the previous period. Thus, they need only make marginal adjustments to capacity in response to changes in $a$. Firms can choose to use all their capacity to produce, to sell some capacity and use the remainder to produce, or to buy more capacity and produce. Capacity may be purchased from and sold to other firms operating in the same industry, or from sources outside the industry. The net capacity adjustments they make follow from Remark 1.

### *A.2. Cross-segment effects and the growth of conglomerates*

As discussed above, when a positive demand shock occurs in industry 1 more productive producers increase their market share. When the productive producer is a conglomerate which operates both in industry 1 and industry 2 this increase in production in industry 1 creates a negative externality for this producer in industry 2. Thus, the conglomerate producer becomes a relatively less aggressive competitor in industry 2. By contrast, producers in industry 1 that are sufficiently less productive reduce capacity in industry 1 by selling capacity to the more productive firms.[42] This reduction in capacity reduces their control costs and creates a positive externality for the producers in industry 2. As a result, the less productive producers in industry 1 that also operate in industry 2 become more aggressive competitors in industry 2 and grow faster than they otherwise would in that industry. Thus, we can observe that:

**Remark 3.** Given a distribution of managerial talent, a positive price shock in industry 1 provides incentives for: (a) Conglomerates that are more productive producers in industry 1 relative to industry competitors to reduce their focus on industry 2 and increase their focus on industry 1 (b) Conglomerates that are marginally productive producers in industry 1 to reduce their focus on industry 1 and increase their focus on industry 2.

We illustrate case (b). This is easiest to show if we assume that there exist some firms in each industry which are single-segment. We use the suffix $ss$ to indicate that the firm is single-segment. For simplicity, all single-segment firms in both industries have the same technology.

We assume that of the total number of firms $n$ a fraction, $\lambda_c$, are conglomerates and operate in both industries. Assume that all conglomerates have identical abilities $d_1^c$ and $d_2^c$. An equal number of single-segment firms operates in both industries, so that the fraction of the $n$ firms operating in each industry as single-segment firms is $\lambda_{ss}$,

[42] Note that "sufficiently" depends on the elasticity of supply of capacity into the industry. If supply is fixed ($\beta = 0$), then it is sufficient that $d < 1$. In a more general model it would not be necessary for the sale of capacity to occur in industry 1 in order for the less productive firms to become more aggressive competitors in industry 2. It would be sufficient for the less productive producers to grow more slowly in industry 1 than the more productive producers following a positive price shock.

where $\lambda_{ss} = (1 - \lambda_c)/2$. We assume that the capacity in each industry is fixed at $K_i$ for $i = 1, 2$.

The profit function of a single-segment firm that operates only in industry $i$ is, as before

$$p_i k_i^{ss} - r_i k^{ss} - \beta_i \left(k_i^{ss}\right)^2.$$

Maximizing profits yield an expression for optimal output analogous to that in the single industry case above, so that $k_i^{ss} = \frac{p_i - r_i}{2\beta}$.

A conglomerate's profit function is given in equation (2) in the text. For the special case discussed here it can be rewritten as

$$d_1^c p_1 k_1^c + d_2^c p_2 k_2^c - r_1 k_1^c - r_2 k_2^c - \beta\left(k_1^c + k_2^c\right)^2.$$

We want to show that following a positive price shock in industry 2, conglomerate segments that are less efficient than the competing single segment firms in industry 1 become smaller relative to the single segment firms in industry 1, so that the ratio $(k_1^c / k_1^{ss})$ declines with increases in $p_2$. We thus assume that $d_1^c < 1$ and, without loss of generality, $d_2^c = 1$.

By solving for $k_i^c$ and $k_i^{ss}$ and substituting into the industry equilibrium conditions $\sigma + \rho r_i = (\lambda_c k_i^c + \lambda_{ss} k_i^{ss})n$ where $i = 1, 2$, we can solve for the price of capital in each industry $r_1$ and $r_2$. Substituting $r_1$ and $r_2$ back into the expressions for $k_i^c$ and $k_i^{ss}$, we obtain

$$\frac{\delta(k_2^c / k_1^{ss})}{\delta p_1} = -A\left(\frac{n\lambda_{ss}}{n\lambda_{ss} + 2\beta(v + w_j)} - d_1^c\right).$$

It can be shown that $A$ is positive for feasible $\lambda_{ss}$ ($\lambda_{ss} < 0.5$). Thus, for all sufficiently low $d_1^c$ ($d_1^c < \frac{n\lambda_{ss}}{n\lambda_{ss} + 2\beta\rho}$) the result follows. Note that if the supply of capacity is fixed in each industry so that $\rho = 0$, it is sufficient that $d_1^c < 1$.

Case (a) can be shown similarly. We can also show that:

**Remark 4.** The greater the productivity of a conglomerate's operations in an industry, the greater the effect of price shocks in that industry on the optimal size of operations of the conglomerate in other industries.

Thus, we would expect that shocks in a conglomerate's main segment (which, all else being equal, has a higher relative productivity) would produce greater effects on the industries in which it has its peripheral segments than if the opposite were true.

Note that we do not predict this pattern of growth across conglomerates business units because the conglomerate firms have an internal capital markets that are superior to those of single-industry firms. Rather, they result from the comparative advantage of conglomerates and single-segment firms over different ranges of demand. Moreover, the predictions of model differ from the agency or empire building models in the literature. The agency and empire building models predict that if a conglomerate receives a positive shock in industry 1 it grows faster in industry 2. By contrast, by Remark 3 the model here predicts that if a conglomerate receives a positive shock in industry 1 and is very productive in industry 1 it grows more slowly in industry 2. Only when a conglomerate that receives a positive shock in industry 1 and is relatively unproductive in industry 1 does it grow faster in industry 2. Thus these predictions differ from agency and empire building models.

## References


Abadie, A., Imbens, G., 2002. Simple and bias-corrected matching estimators for average treatment effects. Working Paper, University of California, Berkeley, CA.

Abel, A., Eberly, J.C., 1994. A unified model of investment under uncertainty. American Economic Review 84 (5), 1369–1384.

Aggarwal, R., Samwick, A., 2003. Why do managers diversify their firms? Agency reconsidered. Journal of Finance 58, 71–188.

Anh, S., Denis, D.J., 2004. Internal capital markets and investment policy. Journal of Financial Economics 71, 489–516.

Berger, P., Hahn, R., 2003. The impact of SFAS no. 131 on information and monitoring. Journal of Accounting Research 41, 163–223.

Berger, P., Ofek, E., 1995. Diversification's effect on firm value. Journal of Financial Economics 37, 39–66.

Berkovitch, E., Israel, R., Tolkowsky, E., 2005. The boundaries of the firm: The choice between stand-alone and conglomerate firms. Working Paper.

Bertrand, M., Mullainathan, S., 2001. How much should we trust difference-in-difference estimates? NBER Working Paper 8841.

Billet, M., Mauer, D., 2003. Cross-subsidies, external financing constraints, and the contribution of the internal capital market to firm value. Review of Financial Studies 16, 1167–1202.

Brealey, R., Myers, S., 2003. Principles of Corporate Finance, 7th edition. McGraw-Hill, Irwin.

Burch, T.R., Nanda, V., 2003. Divisional diversity and the conglomerate discount: The evidence from spin-offs. Journal of Economic Finance 70, 69–98.

Campa, J.M., Kedia, S., 2002. Explaining the diversification discount. Journal of Finance 52, 1731–1762.

Campello, M., 2002. Internal capital markets in financial conglomerates: Evidence from small bank responses to monetary policy. Journal of Finance, 2773–2804.

Chevalier, J., 2000. What do we know about cross-fertilization? Evidence from the investment policies of merging firms. Working Paper, University of Chicago.

Coase, R., 1937. The nature of the firm. Economica 4, 386–405.

Colak, G., Whited, T., 2005. Spin-offs, divestitures and conglomerate investment. Working Paper, University of Iowa and University of Michigan.

Comment, R., Jarrell, G., 1995. Corporate focus and stock return. Journal of Financial Economics 37, 67–87.

Dehejia, R.H., Wahba, S., 1999. Causal effects in non-experimental studies: Re-evaluating the evaluation of training programs. Journal of the American Statistical Association 94, 1053–1062.

Denis, D.J., Denis, D.K., Sarin, A., 1997. Agency problems, equity ownership, and corporate diversification. Journal of Finance 52, 135–160.

Denis, D., Denis, D., Yost, K., 2002. Global diversification, industrial diversification, and firm value. Journal of Finance 57, 1951–1979.

Dittmar, A., Shivasani, A., 2003. Divestitures and divisional investment policies. Journal of Finance 58, 2711–2744.

Eckbo, E., Maksimovic, V., Williams, J., 1991. Truncation bias and cross-sectional estimates in event studies: Econometric issues with an application to horizontal mergers. Review of Financial Studies 1990, 343–367.

Erickson, T., Whited, T., 2000. Measurement error and the relationship between investment and $q$. Journal of Political Economy 108, 1027–1057.

Faure-Grimaud, A., Inderest, R., 2005. Conglomerate entrenchment under optimal financial contracting. American Economic Review, in press.

Fazzari, S., Hubbard, R., Peterson, B., 1988. Financing constraints and corporate investment. Brookings Papers on Economic Activity 1988, 141–195.

Fluck, Z., Lynch, A., 1999. Why firms merge and then divest: A theory of financial synergy. The Journal of Business 72 (July), 319–346.

Fulghieri, P., Hodrick, L., 1997. Synergies and internal agency conflicts: The double-edged sword of mergers. Working Paper, Columbia University.

Gertner, R., Powers, E., Scharfstein, D.S., 2002. Learning about internal capital markets from corporate spinoffs. Journal of Finance, 2479–2506.

Gibbons, R., Murphy, K.J., 1992. Optimal incentive contracts in the presence of career concerns: Theory and evidence. Journal of Political Economy 100, 468–505.

Gomes, J., Livdan, D., 2004. Optimal diversification: Reconciling theory and evidence. Journal of Finance 59, 507–535.

Graham, J., Lemmon, M., Wolf, J., 2002. Does corporate diversification destroy value? Journal of Finance 57, 695–720.

Guedj, I., Scharfstein, D., 2004. Organizational scope and investment: Evidence from the drug development strategies and performance. NBER Working Paper 10933.

Hayashi, F., 1982. Tobin's marginal $q$ and average $q$: A neoclassical interpretation. Econometrica 50, 213–224.

Hayes, R., Lundholm, R., 1996. Segment reporting to the capital market in the presence of a competitor. Journal of Accounting Research 34, 261–280.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.

Hermalin, B., Weisbach, M.S., 1998. Endogenously chosen boards of directors and their monitoring of management. American Economic Review 88, 96–118.

Hubbard, R.G., Palia, D., 1998. A re-examination of the conglomerate merger wave in the 1960s: An internal capital markets view. NBER Working Paper 6539.

Hyland, D.C., 1997. Why firms diversify. An empirical examination. Unpublished doctoral dissertation, Ohio State University, Columbus, OH.

Inderst, Meuller, H., 2003. Internal vs. external financing: An optimal contracting approach. Journal of Finance.

Jensen, M.C., 1986. Agency costs of free cash flow, corporate finance and takeovers. American Economic Review 76, 323–329.

Khanna, N., Tice, S., 2001. The bright side of internal capital market. Journal of Finance 56, 1489–1528.

Kothari, S., Warner, J., 2006. Econometrics of event studies. In: Eckbo, E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. Elsevier, Amsterdam, Chapter 1, this volume.

Lamont, O., 1997. Cash flow and investment: Evidence from internal capital markets. Journal of Finance 52, 83–109.

Lamont, O.A., Polk, C., 2002. Does diversification destroy value? Evidence from the industry shocks. Journal of Financial Economics 63, 51–77.

Lang, L., Stulz, R., 1994. Tobin's $q$, corporate diversification, and firm performance. Journal of Political Economy 102, 1248–1280.

Li, K., Prabhala, N.R., 2007. Self-selection models in corporate finance. In: Eckbo, B.E. (Ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1. Elsevier, Amsterdam, Chapter 2, this volume.

Lucas, R., 1978. On the size distribution of business firms. Bell Journal of Economics 9, 508–523.

Maksimovic, V., Phillips, G., 1998. Asset efficiency and reallocation decisions of bankrupt firms. Journal of Finance 53, 1495–1532.

Maksimovic, V., Phillips, G.M., 1999. Do conglomerate firms allocate resources inefficiently? Available at SSRN: http://ssrn.com/abstract=108068 or DOI: 10.2139/ssrn.108068.

Maksimovic, V., Phillips, G., 2001. The market for corporate assets: Who engages in mergers and assets sales and are there efficiency gains? Journal of Finance 56, 2019–2065.
Maksimovic, V., Phillips, G., 2002. Do conglomerate firms allocate resources inefficiently across industries? Theory and evidence. Journal of Finance 57, 721–767.
Maksimovic, V., Phillips, G., 2007. The industry life cycle and acquisitions and investment: Does firm organization matter? Journal of Finance, in press.
Maksimovic, V., Phillips, G., Prabhala, N.R., 2006. Mergers, acquisitions and the boundaries of the firm. Working Paper, University of Maryland.
Martin, J., Savrak, A., 2003. Corporate diversification and shareholder value: A survey of recent literature. Journal of Corporate Finance 9, 37–57.
Matsusaka, J.G., 2001. Corporate diversification, value maximization and organizational capabilities. Journal of Business 74, 409–431.
McGuckin, R., Pascoe, G., 1988. The longitudinal research database (LRD): Status and research possibilities. Survey of Current Business 68 (11), 31–37.
Meyer, M., Milgrom, P., Roberts, J., 1992. Organizational prospects, influence costs, and ownership changes. Journal of Economics & Management Strategy 1, 9–35.
Pacter, P., 1993. Reporting disaggregated data. Financial Accounting Standards Board Research Report.
Peyer, U., 2001. Internal and external capital markets. Working Paper, University of North Carolina, Chapel Hill, NC.
Prabhala, N.R., 1997. Conditional methods in even studies and an equilibrium justification for standard event-study procedures. Review of Financial Studies 10, 1–38.
Rajan, R., Servaes, H., Zingales, L., 2000. The cost of diversity: The diversification discount and inefficient investment. Journal of Finance 55, 35–80.
Roll, R., 1986. The Hubris hypothesis of corporate takeovers. Journal of Business 59, 197–216.
Ross, S., Westerfield, R., Jordan, B., 2006. Fundamentals of Corporate Finance, 7th edition. McGraw-Hill, Irwin.
Scharfstein, D.S., 1998. The dark side of internal capital markets II: Evidence from diversified conglomerates. NBER Working Paper 6352.
Scharfstein, D.S., Stein, J., 2000. The dark side of internal capital markets: Division rent-seeking and inefficient investment. Journal of Finance 55, 2537–2564.
Schipper, K., Thompson, R., 1983. The impact of merger-related regulations on the shareholders of bidding firms. Journal of Accounting Research 21, 184–221.
Schlingemann, F.P., Stulz, R.M., Walking, R., 2002. Divestitures and the liquidity of the market for corporate assets. Journal of Financial Economics 64, 117–144.
Schnure, C., 1997. Internal capital markets and investments: Do the cash flow constraints really bind? SSRN Working Paper.
Schoar, A., 2002. Effects of diversification on productivity. Journal of Finance 57, 2379–2403.
Shin, H., Stulz, R.M., 1998. Are internal capital markets efficient? Quarterly Journal of Economics 113, 531–552.
Shleifer, A., Vishny, R., 1989. Management entrenchment: The case of manager-specific investments. Journal of Financial Economics 25, 123–140.
Stein, J., 1997. Internal capital markets and the competition for corporate resources. Journal of Finance 52, 111–133.
Stulz, R.M., 1990. Managerial discretion and optimal financing policies. Journal of Financial Economics 26, 3–27.
Villalonga, B., 2003. Research roundtable discussion: The diversification discount. SSRN Working Paper.
Villalonga, B., 2004a. Diversification discount or premium? New evidence from the business information tracing series. Journal of Finance 59, 479–506.
Villalonga, B., 2004b. Does diversification cause the diversification discount. Financial Management 33, 5–23.
Whited, T., 2001. Is it efficient investment that causes the diversification discount? Journal of Finance 56, 1667–1691.

*Chapter 9*

# VENTURE CAPITAL*

PAUL GOMPERS

*Graduate School of Business, Harvard University, Morgan 483, Soldiers Field, Boston, MA 02163, USA*
*e-mail: pgompers@hbs.edu*

## Contents



* This chapter is based in part on Gompers and Lerner (1999b, 2001). All errors are our own.

*Handbook of Corporate Finance, Volume 1*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01009-9*

**Abstract**

Venture capital has emerged as an important intermediary in financial markets, providing capital to young high-technology firms that might have otherwise gone unfunded. Venture capitalists have developed a variety of mechanisms to overcome the problems that emerge at each stage of the investment process. At the same time, the venture capital process is also subject to various pathologies from time to time, which can create problems for investors or entrepreneurs. This handbook chapter reviews the recent empirical literature on these organizations and points out area where further research is needed.

**Keywords**

venture capital, agency costs, optimal contracting, capital constraints monitoring

## 1. Introduction

Venture capital has attracted increasing attention in both the popular press and academic literature. It is alternately described as the engine fueling innovation in the US economy and as the industry that fueled the boom and bust of the Internet era. The recent dramatic growth and subsequent decline in the venture capital industry during the past decade has been accompanied by new academic research that explores its form and function. This research has increasingly shown that far from being a destabilizing factor in the economy, the venture capital industry, while relatively small compared to the public markets, has had a disproportionately positive impact on the economic landscape. There are several critical research questions, however, that still need to be addressed. This includes the extent to which the US venture capital model will be transferred outside of the US and measuring risk and return in the venture capital sector. Thus, this chapter has a two-fold role: to summarize and synthesize what is known about the nature of venture capital investing from recent research and to raise several areas that have yet to be fully answered.

The current view from the existing research is that venture capital has developed as an important intermediary in financial markets, providing capital to firms that might otherwise have difficulty attracting financing. These young firms are plagued by high levels of uncertainty and large differences in what entrepreneurs and investors know, possess few tangible assets, and operate in markets that can and do change very rapidly. The venture capital process can be seen as having evolved useful mechanisms to overcome potential conflicts of interest at each stage of the investment process. At the same time, the venture capital process is also subject to various pathologies from time to time. Various researchers have documented periods of time and settings in which these imbalances have created problems for investors or entrepreneurs.

A natural first question is what constitutes venture capital. Venture capital is often interpreted as many different kinds of investors. Many start-up firms require substantial capital. A firm's founder may not have sufficient funds to finance these projects alone and therefore must seek outside financing. Entrepreneurial firms that are characterized by significant intangible assets, expect years of negative earnings, and have uncertain prospects are unlikely to receive bank loans or other debt financing. Venture capital organizations finance these high-risk, potentially high-reward projects, purchasing equity or equity-linked stakes while the firms are still privately held. At the same time, not everyone who finances these types of firms is a venture capitalist. Banks, individual investors (or "angels"), and corporations are among the other providers of capital for these firms. Venture capital is defined as independent and professionally managed, dedicated pools of capital that focus on equity or equity-linked investments in privately held, high growth companies.

The primary focus of this chapter is on reviewing the empirical academic research on venture capital and highlighting the critical role that venture capital has played in filling an important financing gap. Our empirical understanding of venture capital has grown dramatically over the past decade as large scale databases on venture investing have become widely available to researchers. The theoretical literature on venture capital has likewise exploded during the past decade. The improvement in efficiency might be due to the active monitoring and advice that is provided (Cornelli and Yosha, 1997; Marx, 1994; Hellmann, 1998), the screening mechanisms employed (Amit, Glosten, and Muller, 1990a, 1990b; Chan, 1983), the incentives to exit (Berglöf, 1994), the proper syndication of the investment (Admati and Pfleiderer, 1994), or the staging of the investment (Bergmann and Hege, 1998). This work has improved our understanding of the factors that affect the relationship between venture capitalists and entrepreneurs.

## 2. The development of the venture capital industry

The venture capital industry was a predominantly American phenomenon in its initial decades. It had its origins in the family offices that managed the wealth of high net worth individuals in the last decades of the nineteenth century and the first decades of this century. Wealthy families such as the Phippes, Rockefellers, Vanderbilts, and Whitneys invested in and advised a variety of business enterprises, including the predecessor entities to AT&T, Eastern Airlines, and McDonald-Douglas. Gradually, these families began involving outside professional managers to select and oversee these investments.

The first venture capital firm satisfying the criteria delineated above, however, was not established until after World War II. MIT President Karl Compton, Harvard Business School Professor Georges F. Doriot, and local Boston business leaders formed American Research and Development (ARD) in 1946. This small group of venture capitalists made high-risk investments into emerging companies that were based on technology developed for World War II. The success of the investments ranged widely: almost half of ARD's profits during its 26-year existence as an independent entity came from its \$70,000 investment in Digital Equipment Company (DEC) in 1957, which grew in value to \$355 million. Because institutional investors were reluctant to invest, ARD was structured as a publicly traded closed-end fund and marketed mostly to individuals (Liles, 1977). The few other venture organizations begun in the decade after ARD's formation were also structured as closed-end funds.

The closed-end fund structure employed by these funds had some significant advantages that made them more suited to venture capital investing than the more familiar open-end mutual funds. While the funds raised their initial capital by selling shares to the public, the funds did not need to repay investors if they wished to no longer hold the fund. Instead, the investors simply sold the shares on a public exchange to other investors. This provision allowed the fund to invest in illiquid assets, secure in the knowledge that they would not need to return investors' capital in an uncertain time frame. Most importantly, because it was a liquid investment that could be freely bought or sold, Security and Exchange Commission regulations did not preclude any class of investors from holding the shares.

The publicly traded structure, however, was soon found to have some significant drawbacks as well. In a number of cases, brokers sold the funds to inappropriate investors: i.e., elderly investors who had a need for high current income rather than long-term capital gains. When the immediate profits promised by unscrupulous brokers did not materialize, these investors vented their frustration at the venture capitalists themselves. For instance, much of General Doriot's time during the mid-1950s was spent addressing investors who had lost substantial sums on their shares of American Research and Development.

The first venture capital limited partnership, Draper, Gaither, and Anderson, was formed in 1958. Unlike the closed-end funds, partnerships were exempt from securities regulations, including the exacting disclosure requirements of the Investment Company Act of 1940. The set of the investors from which the funds could raise capital, however, was much more restricted. The interests in a given partnership could only be held by a limited number of institutions and high net-worth individual investors.

The Draper partnership and its followers applied the template of other limited partnerships: e.g., to develop real estate projects and explore oil fields. The partnerships had pre-determined, finite lifetimes (usually ten years, though extensions were often allowed). Thus, unlike closed-end funds, which often had indefinite lives, the partnerships were required to return the assets to investors within a set period. From the days of the first limited partnerships, these distributions were typically made in stock. Rather than selling successful investments after they went public and returning cash to their investors, the venture capitalists would simply give them their allocation of shares in the company in which the venture firm had invested. In this way, the investors could choose when to realize the capital gains associated with the investment. This feature was particular important for individuals and corporate investors, as they could arrange the sales in a manner that would minimize their capital gains tax obligation.

While imitators soon followed, limited partnerships accounted for a minority of the venture pool during the 1960s and 1970s. Most venture organizations raised money either through closed-end funds or small business investment companies (SBICs), federally guaranteed risk capital pools that proliferated during the 1960s. While the market for SBICs in the late 1960s and early 1970s was strong, the sector ultimately collapsed in the 1970s. The combination of federal guarantees and limited scrutiny of applicants led to scenario that foreshadowed the savings and loan crisis of the 1980s. Unscrupulous and naïve operators were frequently granted SBIC licenses. Frequently, their investments proved to be either in firms with poor prospects or in outright fraudulent enterprises.

Activity in the venture industry increased dramatically in late 1970s and early 1980s. Tables 1A, 1B and Figure 1 provide an overview of fundraising by venture partnerships, highlighting the changing volume of investments over the years, as well as the shifting mixture of investors. Industry observers attributed much of the shift to the U.S. Department of Labor's clarification of the "prudent man" rule in 1979. Prior to this year, the Employee Retirement Income Security Act (ERISA) limited pension funds from investing substantial amounts of money into venture capital or other high-risk asset classes. The Department of Labor's clarification of the rule explicitly allowed pension managers to invest in high-risk assets, including venture capital. In 1978, when $424 million was

Table 1A
Summary statistics for venture capital fund-raising by independent venture partnerships. All dollar figures are in millions of 2004 dollars

| | 1978 | 1979 | 1980 | 1981 | 1982 | 1983 | 1984 | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *First closing of funds* | | | | | | | | | | | | | |
| Number of funds | 23 | 27 | 57 | 81 | 98 | 147 | 150 | 99 | 86 | 112 | 78 | 88 | 50 |
| Size (millions of 1992) | 414 | 469 | 1,208 | 1,661 | 2,026 | 5,289 | 4,694 | 4,065 | 4,295 | 5,217 | 3,606 | 3,354 | 2,431 |
| *Sources of funds* | | | | | | | | | | | | | |
| Private pension funds | 15% | 31% | 30% | 23% | 33% | 26% | 25% | 23% | 39% | 27% | 27% | 22% | 31% |
| Public pension funds | [a] | [a] | [a] | [a] | [a] | 5% | 9% | 10% | 12% | 12% | 20% | 14% | 22% |
| Corporations | 10% | 17% | 19% | 17% | 12% | 12% | 14% | 12% | 11% | 10% | 12% | 20% | 7% |
| Individuals | 32% | 23% | 16% | 23% | 21% | 21% | 15% | 13% | 12% | 12% | 8% | 6% | 11% |
| Endowments | 9% | 10% | 14% | 12% | 7% | 8% | 6% | 8% | 6% | 10% | 11% | 12% | 13% |
| Insurance companies/banks | 16% | 4% | 13% | 15% | 14% | 12% | 13% | 11% | 10% | 15% | 9% | 13% | 9% |
| Foreign investors/other | 18% | 15% | 8% | 10% | 13% | 16% | 18% | 23% | 11% | 14% | 13% | 13% | 7% |
| *Independent venture partnerships as a share of the total venture pool*[b] | | | 40% | 44% | 58% | 68% | 72% | 73% | 75% | 78% | 80% | 79% | 80% |

Source: Compiled from the unpublished Venture Economics funds database and various issues of the *Venture Capital Journal*. The numbers differ slightly from Lerner and Gompers (1996) due to continuing emendations to the funds database.

[a]Public pension funds are included with private pension funds in these years.

[b]This series is defined differently in different years. In some years, the *Venture Capital Journal* states that non-bank SBICs and publicly traded venture funds are included with independent venture partnerships. In other years, these funds are counted in other categories. It is not available after 1994.

Table 1B
Summary statistics for venture capital fund-raising by independent venture partnerships. All dollar figures are in millions of 1992 dollars

| | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 |
|---|---|---|---|---|---|---|---|---|---|
| *First closing of funds* | | | | | | | | | |
| Number of funds | 34 | 31 | 46 | 80 | 84 | 80 | 103 | 161 | 186 |
| Size (millions of 1992 $) | 1,483 | 1,950 | 2,480 | 3,582 | 4,045 | 6,805 | 8,060 | 16,933 | 31,299 |
| *Sources of funds* | | | | | | | | | |
| Private pension funds | 25% | 22% | 59% | 47% | 38% | 43% | 40% | 37% | 9% |
| Public pension funds | 17% | 20% | [a] | [a] | [a] | [a] | [a] | 10% | 9% |
| Corporations | 4% | 3% | 8% | 9% | 2% | 13% | 30% | 18% | 16% |
| Individuals | 12% | 11% | 7% | 12% | 17% | 9% | 13% | 11% | 19% |
| Endowments | 24% | 18% | 11% | 21% | 22% | 21% | 9% | 8% | 15% |
| Insurance companies/banks | 6% | 14% | 11% | 9% | 18% | 5% | 1% | 3% | 11% |
| Foreign investors/other | 12% | 11% | 4% | 2% | 3% | 8% | 7% | 13% | 22% |
| *Independent venture partnerships as a share of the total venture pool*[b] | | | | | | | | | |
| | 80% | 81% | 78% | 78% | | | | | |

Source: Compiled from the unpublished Venture Economics funds database and various issues of the *Venture Capital Journal*. The numbers differ slightly from Lerner and Gompers (1996) due to continuing emendations to the funds database.

[a]Public pension funds are included with private pension funds in these years.

[b]This series is defined differently in different years. In some years, the *Venture Capital Journal* states that non-bank SBICs and publicly traded venture funds are included with independent venture partnerships. In other years, these funds are counted in other categories. It is not available after 1994.

invested in new venture capital funds, individuals accounted for the largest share (32 percent). Pension funds supplied just 15 percent. Eight years later, when more than $4 billion was invested, pension funds accounted for more than half of all contributions.[1]

The subsequent years saw both very good and very trying times for venture capitalists. On the one hand, venture capitalists backed many of the most successful high-technology companies during the 1980s and 1990s, including Apple Computer, Cisco Systems, Genentech, Microsoft, Netscape, and Sun Microsystems. A substantial number of service firms (including Staples, Starbucks, and TCBY) also received venture financing.

At the same time, commitments to the venture capital industry were very uneven. As Figure 1 and Tables 1A, 1B depict, the annual flow of money into venture funds increased by a factor of ten during the early 1980s, peaking at around six billion (in 2004 dollars). From 1987 through 1991, however, fundraising steadily declined. This

[1] The annual commitments represent pledges of capital to venture funds raised in a given year. This money is typically invested over three to five years starting in the year the fund is formed.

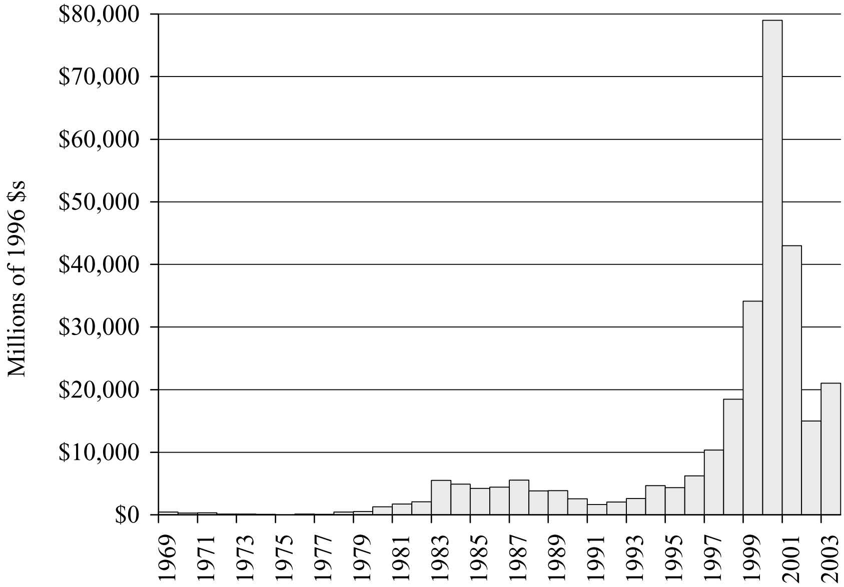



Fig. 1. Commitments to the venture capital industry. Commitments are defined as the amount of money that is pledged to venture capital funds in that year. Amounts are in millions of 1996 dollars. Source: Venture Economics and Asset Alternatives.

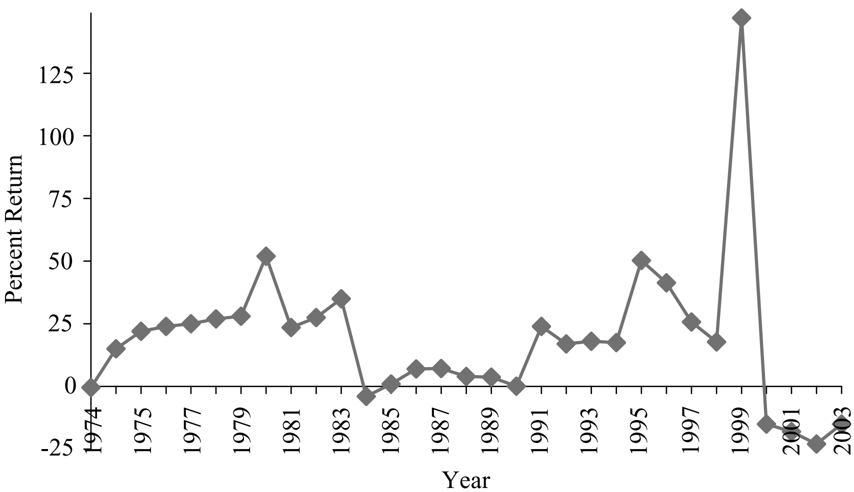



Fig. 2. Return on venture capital. The average annual internal rate of return on venture capital funds, net of fees and profit-sharing, is plotted by year. Source: Compiled from Venture Economics data.

fall-back reflected the disappointment that many investors encountered with their investments. As Figure 2 shows, returns on venture capital funds declined in the mid-1980s, apparently because of overinvestment in various industries and the entry of inexperienced venture capitalists. As investors became disappointed with returns, they committed less capital to the industry.

This pattern reversed dramatically in the 1990s, which saw rapid growth in venture fundraising. The explosion of activity in the IPO market and the exit of many inexperienced venture capitalists led to increasing venture capital returns. New capital commitments rose in response, increasing by more than twenty times between 1991 and 2000. While previous investment surges have been associated with falling venture capital returns, this expansion in fundraising saw a rise in the returns to venture funds. Much of the growth in fundraising was fueled by public pension funds, many of which entered venture investing for the first time in a significant way.

The explosion in venture capital investing was also driven by two other classes of investors: corporations and individuals. While the late 1960s and mid 1980s had seen extensive corporate experimentation with venture funds, the late 1990s saw an unprecedented surge of activity. The determinants of this increase were various. Some were similar to those in earlier waves of corporate venturing activity. For instance, the high degree of publicity associated with the successful venture investments of the period, such as Amazon.com, eBay, and Yahoo! triggered the interest of many CEOs, who sought to harness some of the same energy in their organization

This rapid rise in venture capital investing, however, gave way to just as rapid a deflation in venture capital investment activity. The causes of the decline are myriad. Some have commented on the overshooting of the venture industry and how the level of investment activity in 1999 and 2000 was driven up by irrational sentiment towards technology stocks. This sentiment fueled the rise in public equity values and the IPO market. When the business model for many of the startup companies, especially Internet-related firms, failed to deliver profits, investors began to realize that valuation levels assigned to these companies did not make rational sense.

In addition, corporations which had fueled much of the purchasing of new technology suddenly found themselves with excess capacity and slow end user demand. Technology spending by these companies quickly dried up and startups no longer had markets for their products. This decline in spending was protracted and many venture capital-backed startups could not recover.

Finally, the venture capital industry itself contributed to the overshooting and subsequent decline. Many venture capital firms played "follow the leader" strategies and invested in companies that were too similar to one another. This meant that even in attractive markets, product prices were driven down to unprofitable levels. Good ideas and good companies failed because the size of the markets addressed could not support the level of investment activity that took place in 1999 and 2000.

These factors led to a rise in venture capital-backed company failures and a rapid write-down in investment values. As fund portfolio values declined, interim internal rates of return became negative and investment levels declined. In the aftermath of the

Table 2
Number and dollar amount of venture capital disbursements in the U.S. in the first three quarters of 1999, by VentureOne industry classification. All dollar figures are in millions of current dollars

| Industry | Number of transactions | Total $ invested |
|---|---|---|
| Communications and networking | 255 | $4,498 |
| Electronics and computer hardware | 59 | $423 |
| Information services | 296 | $3,053 |
| Semiconductors and components | 58 | $518 |
| Software | 489 | $4,233 |
| *Total of information technology* | 1157 | $12,726 |
| Healthcare services | 47 | $411 |
| Medical compounds | 84 | $649 |
| Medical devices and equipment | 114 | $827 |
| Medical information systems | 44 | $336 |
| *Total of life sciences* | 289 | $2,233 |
| Retail and consumer products | 30 | $227 |
| Other companies | 454 | $5,580 |
| *Total of non-technology or other* | 484 | $5,807 |
| *Grand total* | 1979 | $20,957 |

Source: Compiled from unpublished VentureOne databases.

retrenchment, many venture capital firms decided to reduce the amount of capital that they had raised, essentially foregoing commitments that their investors had made to their funds. As the investment pace slowed, the level of fundraising declined even more dramatically. While fundraising in the past few years has begun to recover, how far it rises and whether it reaches the speculative levels of 1999 and 2000 is an open question.

## 3. The venture capital investment process

Venture capitalists typically invest the money in young firms that may be little more than in the head of a talented engineer or scientist. Most of the firms that venture capitalists finance have few other sources of cash and many are subject to severe credit rationing. In order to overcome this capital rationing, however, the control and monitoring aspects of venture capitalists' investment process are paramount. Researchers have explored how the types of contracts utilized, the timing of investment, and the active involvement of the venture capital investor play important roles in improving the likelihood of success for the startup company

Tables 2–4 present historical information on the mixture of investments. Table 2 provides a detailed summary of investments in 1998; Table 3 presents a more aggregated summary of investments (in manufacturing firms only) over the past three decades; and

Table 4 provides a summary of investments in the ten states with the most venture capital activity over the past three decades.

Before considering the mechanisms employed by venture capitalists, it is worth highlighting that a lengthy literature has discussed the financing of young firms. Uncertainty and informational asymmetries often characterize young firms, particularly in high-technology industries. These information problems make it difficult to assess these firms, and permit opportunistic behavior by entrepreneurs after financing is received. This literature has also highlighted the role of financial intermediaries in alleviating these information problems.

To briefly review the types of conflicts that can emerge in these settings, Jensen and Meckling (1976) demonstrate that conflicts between managers and investors ("agency problems") can affect the willingness of both debt and equity holders to provide capital. If the firm raises equity from outside investors, the manager has an incentive to engage in wasteful expenditures (e.g., lavish offices) because he may benefit disproportionately from these but does not bear their entire cost. Similarly, if the firm raises debt, the manager may increase risk to undesirable levels. Because providers of capital recognize these problems, outside investors demand a higher rate of return than would be the case if the funds were internally generated.

More generally, the inability to verify outcomes makes it difficult to write contracts that are contingent upon particular events. This inability makes external financing costly. Many of the models of ownership (e.g., Grossman and Hart, 1986, and Hart and Moore, 1990) and financing choice (e.g., Hart and Moore, 1998) depend on the inability of investors to verify that certain actions have been taken or certain outcomes have occurred. While actions or outcomes might be observable, meaning that investors know what the entrepreneur did, they are assumed not to be verifiable: i.e., investors could not convince a court of the action or outcome. Start-up firms are likely to face exactly these types of problems, making external financing costly or difficult to obtain.

If the information asymmetries could be eliminated, financing constraints would disappear. Financial economists argue that specialized financial intermediaries, such as venture capital organizations, can address these problems. By intensively scrutinizing firms before providing capital and then monitoring them afterwards, they can alleviate some of the information gaps and reduce capital constraints. Thus, it is important to understand the tools employed by venture investors discussed below as responses to this difficult environment, which enable firms to ultimately receive the financing that they cannot raise from other sources. It is the nonmonetary aspects of venture capital that are critical to its success.

One of the most common features of venture capital is the meting out of financing in discrete stages over time. Sahlman (1990) notes that staged capital infusion is the most potent control mechanism a venture capitalist can employ. Prospects for the firm are periodically reevaluated. The shorter the duration of an individual round of financing, the more frequently the venture capitalist monitors the entrepreneur's progress and the greater the need to gather information. Staged capital infusion keeps the owner/manager

Table 3

Number and dollar amount of venture capital disbursements for U.S. manufacturing industries, by industry and five-year period. The count of venture capital investments in each five-year period is the sum of the number of firms receiving investments in each year. All dollar figures are in millions of 1992 dollars

| # | Industry | 1965–69 | 1970–74 | 1975–79 | 1980–84 | 1985–89 | 1990–96 |
|---|---|---|---|---|---|---|---|
| Panel A: Venture capital investments (#s) | | | | | | | |
| 1 | Food and kindred | 1 | 9 | 6 | 23 | 80 | 93 |
| 2 | Textile and apparel | 4 | 12 | 9 | 19 | 27 | 70 |
| 3 | Lumber and furniture | 2 | 8 | 6 | 24 | 62 | 37 |
| 4 | Paper | 2 | 2 | 2 | 2 | 12 | 14 |
| 5 | Industrial chemicals | 1 | 1 | 1 | 6 | 18 | 23 |
| 6 | Drugs | 1 | 12 | 34 | 245 | 554 | 746 |
| 7 | Other chemicals | 1 | 7 | 8 | 10 | 52 | 46 |
| 8 | Petroleum refining and extraction | 3 | 3 | 26 | 92 | 27 | 14 |
| 9 | Rubber products | 1 | 5 | 6 | 19 | 11 | 7 |
| 10 | Stone, clay and glass products | 0 | 1 | 3 | 14 | 48 | 31 |
| 11 | Primary metals | 0 | 3 | 5 | 20 | 44 | 33 |
| 12 | Fabricated metal products | 0 | 0 | 0 | 2 | 1 | 2 |
| 13 | Office and computing machines | 39 | 84 | 108 | 744 | 641 | 442 |
| 14 | Other non-electrical machinery | 12 | 12 | 32 | 254 | 280 | 162 |
| 15 | Communication and electronic | 23 | 65 | 60 | 497 | 736 | 709 |
| 16 | Other electrical equipment | 0 | 6 | 16 | 36 | 52 | 50 |
| 17 | Transportation equipment | 1 | 7 | 5 | 6 | 24 | 25 |
| 18 | Aircraft and missiles | 0 | 0 | 0 | 12 | 20 | 4 |
| 19 | Professional and scientific instruments | 13 | 37 | 70 | 383 | 549 | 544 |
| 20 | Other machinery | 7 | 14 | 16 | 62 | 89 | 98 |
| | *Total* | 111 | 288 | 413 | 2,470 | 3,327 | 3,150 |
| Panel B: Venture capital disbursements (millions of 1992 $s) | | | | | | | |
| 1 | Food and kindred | 4 | 19 | 7 | 25 | 212 | 258 |
| 2 | Textile and apparel | 6 | 15 | 14 | 27 | 45 | 186 |
| 3 | Lumber and furniture | 4 | 17 | 9 | 26 | 200 | 354 |
| 4 | Paper | 1 | 8 | 3 | 3 | 22 | 46 |
| 5 | Industrial chemicals | 0 | 1 | 1 | 41 | 34 | 33 |
| 6 | Drugs | 0 | 15 | 136 | 623 | 1,869 | 3,017 |
| 7 | Other chemicals | 1 | 40 | 4 | 9 | 155 | 87 |
| 8 | Petroleum refining and extraction | 12 | 6 | 92 | 359 | 110 | 29 |
| 9 | Rubber products | 1 | 3 | 15 | 28 | 8 | 18 |
| 10 | Stone, clay and glass products | 0 | 1 | 5 | 34 | 99 | 45 |
| 11 | Primary metals | 0 | 8 | 11 | 25 | 67 | 166 |
| 12 | Fabricated metal products | 0 | 0 | 0 | 1 | 0 | 1 |
| 13 | Office and computing machines | 67 | 404 | 288 | 3,253 | 2,491 | 1,426 |
| 14 | Other non-electrical machinery | 64 | 17 | 37 | 677 | 669 | 323 |
| 15 | Communication and electronic | 44 | 189 | 82 | 1,746 | 2,646 | 2,627 |
| 16 | Other electrical equipment | 0 | 8 | 53 | 78 | 107 | 104 |
| 17 | Transportation equipment | 0 | 10 | 4 | 9 | 47 | 96 |
| 18 | Aircraft and missiles | 0 | 0 | 0 | 19 | 19 | 8 |
| 19 | Professional and scientific instruments | 13 | 86 | 114 | 811 | 1,449 | 1,509 |
| 20 | Other machinery | 7 | 28 | 22 | 113 | 176 | 350 |
| | *Total* | $225 | $874 | $895 | $7,907 | $10,423 | $10,685 |

Source: Based on Kortum and Lerner (1999) and supplemented with tabulations of unpublished Venture Economics databases.

Table 4
Number and dollar amount of venture capital disbursements for all industries in the ten states with the most venture capital activity, by state and five-year period. The count of venture capital investments in each five-year period is the sum of the number of firms receiving investments in each year. All dollar figures are in millions of 1992 dollars

| State | 1965–69 | 1970–74 | 1975–79 | 1980–84 | 1985–89 | 1990–96 |
|---|---|---|---|---|---|---|
| Panel A: Venture capital investments (#s) | | | | | | |
| California | 65 | 179 | 310 | 1,863 | 2,645 | 3,380 |
| Massachusetts | 45 | 93 | 155 | 708 | 1,014 | 1,028 |
| Texas | 18 | 71 | 84 | 373 | 584 | 489 |
| New York | 28 | 90 | 73 | 311 | 324 | 276 |
| New Jersey | 15 | 35 | 47 | 171 | 291 | 336 |
| Colorado | 5 | 22 | 31 | 194 | 258 | 298 |
| Pennsylvania | 8 | 21 | 32 | 120 | 290 | 311 |
| Illinois | 16 | 29 | 31 | 133 | 214 | 312 |
| Minnesota | 12 | 34 | 42 | 170 | 186 | 194 |
| Connecticut | 3 | 20 | 37 | 136 | 217 | 210 |
| *Total, all states* | 302 | 847 | 1,253 | 5,365 | 8,154 | 9,406 |
| Panel B: Venture capital disbursements (millions of 1992 $s) | | | | | | |
| California | 218 | 546 | 691 | 6,711 | 9,670 | 13,603 |
| Massachusetts | 61 | 155 | 197 | 1,943 | 2,829 | 3,386 |
| Texas | 37 | 140 | 148 | 1,161 | 2,171 | 2,010 |
| New York | 32 | 154 | 162 | 688 | 1,404 | 1,394 |
| New Jersey | 33 | 82 | 77 | 370 | 1,214 | 1,711 |
| Colorado | 12 | 50 | 46 | 493 | 805 | 951 |
| Pennsylvania | 18 | 41 | 116 | 370 | 1,530 | 1,109 |
| Illinois | 59 | 134 | 117 | 287 | 1,208 | 1,413 |
| Minnesota | 6 | 90 | 44 | 270 | 406 | 522 |
| Connecticut | 1 | 32 | 85 | 319 | 1,463 | 724 |
| *Total, all states* | $687 | $1,935 | $2,259 | $15,261 | $30,742 | $37,162 |

Source: Based on tabulations of unpublished Venture Economics databases.

on a "tight leash" and reduces potential losses from bad decisions.[2] Venture capitalists

[2] Two related types of agency costs exist in entrepreneurial firms. Both agency costs result from the large information asymmetries that affect young, growth companies in need of financing. First, entrepreneurs might invest in strategies, research, or projects that have high personal returns but low expected monetary payoffs to shareholders. For example, a biotechnology company founder may choose to invest in a certain type of research that brings him/her great recognition in the scientific community but provides little return for the venture capitalist. Similarly, entrepreneurs may receive initial results from market trials indicating little demand for a new product, but may want to keep the company going because they receive significant private benefits from managing their own firm. Second, because entrepreneurs' equity stakes are essentially call options, they have incentives to pursue highly volatile strategies, such as rushing a product to market when further testing may be warranted.

should weigh potential agency and monitoring costs when determining how frequently they should reevaluate projects and supply capital. The duration of funding should decline and the frequency of reevaluation should increase when the venture capitalist expects conflicts with the entrepreneur are more likely.

If monitoring and information gathering are important, venture capitalists should invest in firms in which asymmetric information is likely to be a problem. The value of oversight will be greater for these firms. The capital constraints faced by these companies will be very large and the information gathered will help alleviate the constraint. Early-stage companies have short or no histories to examine and are difficult to evaluate. Similarly, high-technology companies are likely to require close monitoring. A significant fraction of venture investment should therefore be directed towards early-stage and high-technology companies.

In practice, venture capitalists incur costs when they monitor and infuse capital. Monitoring costs include the opportunity cost of generating reports for both the venture capitalist and entrepreneur. If venture capitalists need to "kick the tires" of the plant, read reports, and take time away from other activities, these costs can be substantial. Contracting costs (e.g., legal fees) and the lost time and resources of the entrepreneur must be imputed as well. These costs lead to funding being provided in discrete stages.

The nature of the firm's assets also has important implications for expected agency costs and the structure of staged venture capital investments. Intangible assets should be associated with greater agency problems. As assets become more tangible, venture capitalists can recover more of their investment in liquidation. This reduces the need to monitor tightly and should increase the time between refinancings. Industries with high levels of R&D should also have more frequent agency problems, and venture capitalists should shorten funding duration. Finally, a substantial finance literature (e.g., Myers, 1977) argues that firms with high market-to-book ratios are more susceptible to these agency costs, thus venture capitalists should increase the intensity of monitoring of these firms.

Gompers (1995) tests these predictions using a random sample of 794 venture capital-financed companies. The results confirm the predictions of agency theory. Venture capitalists concentrate investments in early stage companies and high technology industries where informational asymmetries are significant and monitoring is valuable. Venture capitalists monitor the firm's progress. If they learn negative information about future returns, the project is cut off from new financing. Firms that go public (these firms yield the highest return for venture capitalists on average) receive more total financing and a greater number of rounds than other firms (which may go bankrupt, be acquired, or remain private). Gompers also finds that early stage firms receive significantly less money per round. Increases in asset tangibility increase financing duration and reduce monitoring intensity. As the role of future investment opportunities in firm value increases (higher market-to-book ratios or R&D intensities), firms are refinanced more frequently. These results suggest the important monitoring and information generating roles played by venture capitalists.

Consistent evidence regarding the strength of contractual terms in these agreements is found in Kaplan and Stromberg's (2003) analysis of 130 venture partnership agreements. The overall use of contracts to control potential adverse behavior on the part of entrepreneurs has been modeled in a in a number of settings. Kaplan and Stromberg test a variety of these theories to determine whether factors like information asymmetries are critical to the types of contracts that are signed between venture capitalists and entrepreneurs. They find that venture contracts are effective at separating cash flow ownership from board rights, liquidation rights, voting rights and other control rights. Similarly, future financing and allocation of ownership in the firm is often based on reaching contingent milestones. The results support the contracting view of Aghion and Bolton (1992) and Dewatripont and Tirole (1994).

In addition to the staged capital infusions, venture capitalists will usually make investments with other investors. One venture firm will originate the deal and look to bring in other venture capital firms. This syndication serves multiple purposes. First, it allows the venture capital firm to diversify. If the venture capitalist had to invest alone into all the companies in his portfolio, then he could make many fewer investments. By syndicating investments, the venture capitalist can invest in more projects and largely diversify away firm-specific risk.

A second potential explanation for syndication patterns is that involving other venture firms provides as a second opinion on the investment opportunity. There is usually no clear-cut answer as to whether any of the investments that a venture organization undertakes will yield attractive returns. Having other investors approve the deal limits the danger that bad deals will get funded. This is particularly true when the company is early-stage or technology-based.

Lerner (1994a) tests this "second opinion" hypothesis in a sample of biotechnology venture capital investments. In a sample of 271 firms, Lerner finds that in the early rounds of investing, experienced venture capitalists tend to syndicate only with venture capital firms that have similar experience. Lerner argues that if a venture capitalist were looking for a second opinion, then he would want to get a second opinion from someone of similar or better ability, certainly not from someone of lesser ability.

A related topic is explored by Hochberg, Ljungqvist, and Lu (2006) who examine the relationship among various venture capital investors in syndicate networks and the performance of the companies in which they invest. Hochberg et al. create a measure of centrality based on syndicate patterns in the network. This measure, the Bonacich (1987) measure, controls for how central a venture capital firm is to the entire industry. Firms with greater Bonacich measures are more central to the industry based upon their syndicate patters. Hochberg et al. find that this measure is a strong predictor of performance for the underlying portfolio companies. Portfolio companies that receive an investment by a venture firm that is more central to the industry are more likely to be successful (as measured by the probability of exiting through an IPO or acquisition). In addition, they are more likely to survive to a subsequent financing round than are similar firms financed by venture capitalists that are less central based on their syndication patterns. These patterns support the results found by Lerner in his earlier work.

The advice and support provided by venture capitalists is often embodied by their role on the firm's board of directors. Lerner (1995) examines the decision of venture capitalists to provide this oversight. He examines whether venture capitalists' representation on the boards of the private firms in their portfolios is greater when the need for oversight is larger. This approach is suggested by Fama and Jensen (1983) and Williamson (1983), who hypothesize that the composition of the board should be shaped by the need for oversight. These authors argue that the board will bear greater responsibility for oversight—and consequently that outsiders should have greater representation—when the danger of managerial deviations from value maximization is high. If venture capitalists are especially important providers of managerial oversight, their representation on boards should be more extensive at times when the need for oversight is greater.

Lerner examines changes in board membership around the time that a firm's chief executive officer (CEO) is replaced, an approach suggested by Hermalin and Weisbach's (1988) study of outside directors of public firms. The replacement of the top manager at an entrepreneurial firm is likely to coincide with an organizational crisis and to heighten the need for monitoring. He finds that an average of 1.75 venture capitalists are added to the board between financing rounds when the firm's CEO is replaced in the interval; between other rounds, 0.24 venture directors are added. No differences are found in the addition of other outside directors. This oversight of new firms involves substantial costs. The transaction costs associated with frequent visits and intensive involvement are likely to be reduced if the venture capitalist is proximate to the firms in his portfolio. Consistent with these suggestions, he find that geographic proximity is an important determinant of venture board membership: organizations with offices within five miles of the firm's headquarters are twice as likely to be board members as those more than 500 miles distant. Over half the firms in the sample have a venture director with an office within sixty miles of their headquarters.

The role that venture capitalists play in shaping the overall board of directors at the time of the IPO is also explored in Baker and Gompers (2004a). In particular, they examine the determinants of board structures and the effects that these board structures play in determining the success of the firm. With data from 1,116 IPO prospectuses, they describe board size and composition for a set of firms with a median age of less than six years and a median equity capitalization of $42 million. This analysis gives insights on the role that venture capitalists play—beyond providing money—and the bargaining process between the CEO and outside shareholders.

The venture capital-backed board has fewer insiders and quasi-outsiders and more independent outside directors. These results hold when we control for ownership structure and the endogeneity of venture financing, suggesting a causal relationship where venture capitalists, in addition to monitoring management and providing capital, give advice and value-added services that otherwise might be performed by instrumental board members. The evidence is consistent with the Hermalin and Weisbach (1988) notion that board structure is the outcome of a bargain between the CEO and the outside investors. First, the fraction of outsiders on the board of directors falls with CEO tenure and voting control. Venture capitalists appear to be a counterweight to CEO control. Venture capitalists not only reduce inside representation indirectly by reducing the control of the CEO with their concentrated outside ownership stakes, but also reputable venture firms are directly associated with greater outsider representation on the board. Second, a possible interpretation of the venture reputation effect is that reputable venture firms gain power by having access to adequate replacements for the founder. Consistent with this notion, the probability that a founder remains on as CEO at the time of the IPO falls with venture firm reputation. Baker and Gompers also explore the performance implications of better boards and find that the better board structure of venture capital backing improves long-term firm outcomes.

Hellmann and Puri (2002) examine the value that is added by venture capitalists, i.e., the role that they play in the professionalization of start-up companies. They examine a sample of 170 Silicon Valley start-ups and find that venture capitalists play a role at the top of the organization, in terms of replacing the original founders with an outside CEO. Moreover, they seem to influence developments further down the organization, in terms of playing a role for the introduction of stock option plans, the hiring of a VP of sales and marketing, and the formulation of human resource policies.

There are several specific questions that Hellmann and Puri address. First, they explore whether venture capitalists provide support in building up the internal organization. They look at several measures including the recruitment processes, the overall human resource policies, the adoption of stock option plans, and the hiring of a vice president of marketing and sales. When they compare similar companies that did and did not receive venture capital financing, they find that companies that obtain venture capital are more likely and are faster to professionalize along these various dimensions.

In work similar to Baker and Gompers (2004a, 2004b), Hellmann and Puri look at the position of the CEO and ask whether a founder is more likely to be replaced by an outsider as CEO when a venture capitalist invests in the firm. Not surprisingly, venture capitalists are more likely to replace a founder as CEO. To attract a new CEO, venture capital is particularly important for early stage companies that do not have any signs of success, still important for companies with a product on the market, and no longer important by the time companies have gone public.

Another mechanism utilized by venture capitalists to avoid conflicts is the widespread use of stock grants and stock options. Managers and critical employees within a firm receive a substantial fraction of their compensation in the form of equity or options. This tends to align the incentives of managers and investors. Baker and Gompers (2004b) examine the role that venture capitalists play in setting compensation and incentives of entrepreneurs. They find that venture capitalists increases the sensitivity of management's compensation to the firm's performance relative to similar nonventure capital-financed companies. Fixed salaries are lower and the size of the equity stake held is higher for venture capital-backed CEOs.

The venture capitalist also employs additional controls on compensation to reduce potential gaming by the entrepreneur. First, venture capitalists usually require vesting of the stock or options over a multi-year period. In this way, the entrepreneur cannot leave the firm and take his shares. Similarly, the venture capitalist can significantly dilute the entrepreneur's stake in subsequent financings if the firm fails to realize its targets. This provides additional incentives for the entrepreneur. In order to maintain his stake, the entrepreneur will need to meet his stated targets.

Until this point, this section has highlighted the ways in which venture capitalists can successfully address agency problems in portfolio firms. The argument is often made by venture capital practitioners, however, that the industry has gone through periods of disequilibrium. During periods when the amount of money flowing into the industry has dramatically grown, they argue, the valuations at which investments are made or the likelihood that certain transactions get funded can shift dramatically. If there are only a certain number of worthy projects to finance, then a substantial increase in the amount of venture fundraising may increase the prices that are paid to invest in these companies. These higher prices may ultimately affect the returns on investment in the industry.

Sahlman and Stevenson (1987) chronicle the exploits of venture capitalists in the Winchester disk drive industry during the early 1980s. Sahlman and Stevenson believe that a type of "market myopia" affected venture capital investing in the industry. During the late 1970s and early 1980s, nineteen disk drive companies received venture capital financing. Two-thirds of these investments came between 1982 and 1984, the period of rapid expansion of the venture industry. Many disk drive companies also went public during this period. While industry growth was rapid during this period of time (sales increased from $27 million in 1978 to $1.3 billion in 1983), Sahlman and Stevenson question whether the scale of investment was rational given any reasonable expectations of industry growth and future economic trends.[3] Similar stories are often told concerning investments in software, biotechnology, and the Internet. The phrase "too much money chasing too few deals" is a common refrain in the venture capital market during periods of rapid growth.

Gompers and Lerner (2000) examine these claims through a dataset of over 4000 venture investments between 1987 and 1995 developed by the consulting firm VentureOne. They construct a hedonic price index that controls for various firm attributes that might affect firm valuation, including firm age, stage of development, and industry, as well as macroeconomic variables such as inflow of funds into the venture capital industry. In addition, they control for public market valuations through indexes of public market values for firms in the same industries and average book-to-market and earnings-to-price ratios.

The results support contentions that a strong relation exists between the valuation of venture capital investments and capital inflows. While other variables also have significant explanatory power—for instance, the marginal impact of a doubling in public market values was between a 15% and 35% increase in the valuation of private equity

[3] Lerner (1997) suggests, however, that these firms may have displayed behavior consistent with strategic models of "technology races" in the economics literature. Because firms had the option to exit the competition to develop a new disk drive, it may have indeed been rational for venture capitalists to fund a substantial number of disk drive manufacturers.

transactions—the inflows variable is significantly positive. A doubling of inflows into venture funds leads to between a 7% and 21% increase in valuation levels.

While prices rose somewhat in 1987, they declined and remained quite flat through the 1990s. Starting in 1994, however, prices steadily increased. This increase coincided with the recent rise in venture fundraising. The regression results show that this rise in fundraising is an important source of the increase in prices. The results are particularly strong for specific types of funds and funds in particular regions. Because funds have become larger in real dollar terms, with more capital per partner, many venture capital organizations have invested larger amounts of money in each portfolio company. Firms have attempted to do this in two ways. First, there has been a movement to finance later-stage companies that can accept larger blocks of financing. Second, venture firms are syndicating less. This leads to greater competition for making later-stage investments. Similarly, because the majority of money is raised in California and Massachusetts, competition for deals in these regions should be particularly intense and venture capital inflows may have a more dramatic effect on prices in those regions. The results support these contentions. The effect of venture capital inflows is significantly more dramatic on later-stage investments and investments in California and Massachusetts.

### *3.1. Exiting venture capital investments*

In order to make money on their investments, venture capitalists need to turn illiquid stakes in private companies into realized return. Typically, as was discussed above, the most profitable exit opportunity is an initial public offering (IPO). In an IPO, the venture capitalist assists the company in issuing shares to the public for the first time. Table 5 summarizes the exiting of venture capital investments through initial public offerings as well as comparable data on non-venture capital offerings.

Initial empirical research into the role of venture capitalists in exiting investments focused on the structure of IPOs. Barry et al. (1990) focus on establishing a broad array of facts about the role of venture capitalists in IPOs, using a sample of 433 venture-backed and 1123 non-venture IPOs between 1978 and 1987.

Barry et al. (1990) document that venture capitalists hold significant equity stakes in the firms they take public (on average, the lead venture capitalist holds a 19% stake immediately prior to the IPO, and all venture investors hold 34%), and hold about one-third of the board seats. They continue to hold their equity positions in the year after the IPO. Finally, venture-backed IPOs have less of a positive return on their first trading day. The authors suggest that this implies that investors need less of a discount in order to purchase these shares (i.e., the offerings are less “underpriced”), because the venture capitalist has monitored the quality of the offering.

Megginson and Weiss (1991) argue that because venture capitalists repeatedly bring firms to the public market, they can credibly stake their reputation. Put another way, they can certify to investors that the firms they bring to market are not overvalued. Certification requires that venture capitalists possess reputational capital, that the acquisition

Table 5
The distribution of venture-backed and non-venture IPOs for the period 1978–1999. This table compares the distribution of IPOs in this sample versus all IPOs recorded over this period of time. All dollar figures are in millions of 1992 dollars

| Year | Number of venture-backed IPOs | Amount raised in venture-backed IPOs | Total number of IPOs | Total amount raised in all IPOs | Venture-backed IPOs as percent of all IPOs (number) | Venture-backed IPOs as percent of all IPOs (amount) |
|---|---|---|---|---|---|---|
| 1978 | 6 | $134 | 42 | $485 | 12.50% | 21.59% |
| 1979 | 4 | $62 | 103 | $777 | 3.74% | 7.34% |
| 1980 | 24 | $670 | 259 | $2,327 | 8.48% | 22.35% |
| 1981 | 50 | $783 | 438 | $4,848 | 10.25% | 13.91% |
| 1982 | 21 | $738 | 198 | $1,901 | 9.59% | 27.97% |
| 1983 | 101 | $3,451 | 848 | $17,999 | 10.64% | 16.09% |
| 1984 | 44 | $731 | 516 | $5,179 | 7.86% | 12.37% |
| 1985 | 35 | $819 | 507 | $13,307 | 6.46% | 5.80% |
| 1986 | 79 | $2,003 | 953 | $23,902 | 7.66% | 7.73% |
| 1987 | 69 | $1,602 | 630 | $19,721 | 9.87% | 7.52% |
| 1988 | 36 | $915 | 435 | $6,679 | 8.28% | 13.70% |
| 1989 | 39 | $1,110 | 371 | $6,763 | 10.51% | 16.41% |
| 1990 | 43 | $1,269 | 276 | $4,828 | 15.58% | 16.29% |
| 1991 | 119 | $3,835 | 367 | $16,872 | 32.43% | 22.73% |
| 1992 | 157 | $4,317 | 509 | $23,990 | 30.84% | 17.99% |
| 1993 | 193 | $4,905 | 707 | $40,456 | 27.30% | 12.12% |
| 1994 | 159 | $3,408 | 564 | $27,786 | 28.19% | 12.26% |
| 1995 | 205 | $6,251 | 566 | $36,219 | 36.22% | 17.26% |
| 1996 | 284 | $10,976 | 845 | $38,245 | 33.61% | 28.70% |
| 1997 | 138 | $4,419 | 628 | $40,278 | 21.34% | 10.60% |
| 1998 | 78 | $3,388 | 319 | $31,075 | 24.45% | 10.90% |
| 1999 | 271 | $20,757 | 485 | $56,952 | 55.87% | 36.45% |

Sources: Barry et al. (1990), Ritter (2006), and various issues of the *Going Public: The IPO Reporter* and the *Venture Capital Journal*.

of such a reputation is costly, and that the present value of lost reputational capital by cheating is greater than the one-time gain from behaving in a duplicitous manner.

Megginson and Weiss test these ideas using a matched set of 640 venture-backed and non-venture IPOs between 1983 and 1987. First, they examine the quality of the underwriters who bring firms to market. They show that the underwriters of venture-backed firms are significantly more experienced than the underwriters of comparable non-venture offerings. Megginson and Weiss also find that institutional holdings of venture-backed firms after the IPO are larger than comparable non-venture companies. Third, Megginson and Weiss gather evidence on expenses associated with going public. Venture-backed IPOs have significantly lower fees than non-venture IPOs. Fourth, Megginson and Weiss demonstrate that venture capitalists retain a majority of their equity after the IPO. Megginson and Weiss argue that this is a commitment device. Finally, Megginson and Weiss present evidence that the underpricing of venture capital-backed IPOs is significantly less than the underpricing of non-venture IPOs.

More recent research has examined the timing of the decision to take firms public and to liquidate the venture capitalists' holdings (which frequently occurs well after the IPO). Several potential factors affect when venture capitalists choose to bring firms public. One of these is the relative valuation level of publicly traded securities. Lerner (1994b) examines when venture capitalists choose to finance a sample of biotechnology companies in another private round versus taking the firm public in. Using a sample of 350 privately held venture-backed firms, he shows take firms public at market peaks, relying on private financings when valuations are lower. Seasoned venture capitalists appear more proficient at timing IPOs. The results are robust to the use of alternative criteria to separate firms and controls for firms' quality. The results are not caused by differences in the speed of executing the IPOs, or in the willingness to withdraw the proposed IPOs.

Another consideration may be the reputation of the venture capital firm. Gompers (1996) argues that young venture capital firms have incentives to "grandstand": i.e., they take actions that signal their ability to potential investors. Specifically, young venture capital firms bring companies public earlier than older venture capital firms in an effort to establish a reputation and successfully raise capital for new funds. He examines a sample of 433 venture-backed initial public offerings (IPOs) between 1978 and 1987, as well as a second sample consisting of the first IPOs brought to market by 62 venture capital funds. The results support predictions of the grandstanding hypothesis. For example, the effect of recent performance in the IPO market on the amount of capital raised is stronger for young venture capital firms, providing them with a greater incentive to bring companies public earlier. Young venture capital firms have been on the IPO company's board of directors 14 months less and hold smaller percentage equity stakes at the time of IPO than the more established venture firms. The IPO companies that they finance are nearly two years younger and more underpriced when they go public than companies backed by older venture capital firms. Much of the difference in underpricing and the venture capitalists' percentage equity stake is associated with a shorter duration of board representation, indicating that rushing companies to the IPO market imposes costs on the venture firm. The results suggest that the relation between performance and capital raising affects the incentives and actions of venture capitalists.

The typical venture capital firm, however, does not sell their equity at the time of the IPO. The negative signal that would be sent to the market by an insider "cashing out" would prevent a successful offering. In additional, most investment banks require that all insiders, including the venture capitalists, do not sell any of their equity after the offering for a pre-specified period (usually six months) as noted in Brav and Gompers (2003). Once that lock-up period is over, however, venture capitalists can return money to investors in one of two ways. They can liquidate their position in a portfolio company by selling shares on the open market after it has gone public and then paying those proceeds to investors in cash. More frequently, however, venture capitalists make distributions of shares to investors in the venture capital fund. Many institutional investors have received a flood of these distributions during the past several years and have grown increasingly concerned about the incentives of the venture capitalists when they declare these transfers.

Gompers and Lerner (1998a) examine how investors might be affected by distributions. These distributions have several features that make them an interesting testing ground for an examination of the impact of transactions by informed insiders on securities prices. Because they are not considered to be "sales", the distributions are exempt from the anti-fraud and anti-manipulation provisions of the securities laws. The legality of distributions provides an important advantage. Comprehensive records of these transactions are compiled by the institutional investors and the intermediaries who invest in venture funds, addressing concerns about sample selection bias. Like trades by corporate insiders, transactions are not revealed at the time of the transaction. Venture capitalists can immediately declare a distribution, send investors their shares, and need not register with the SEC or file a report under Rule 16(a). Rather, the occurrence of such distributions can only be discovered from corporate filings with a lag, and even then the distribution date cannot be precisely identified. To identify the time of these transactions, one needs to rely on the records of the partners in the fund. They characterize the features of the venture funds making the distributions, the firms whose shares are being distributed, and the changes associated with the transactions in a way that can discriminate between the various alternative explanations for these patterns.

From the records of four institutions, Gompers and Lerner construct a representative set of over 700 transactions by 135 funds over a decade-long period. The results are consistent with venture capitalists possessing inside information and of the (partial) adjustment of the market to that information. After significant increases in stock prices prior to distribution, abnormal returns around the distribution are a negative and significant −2.0 percent, comparable to the market reaction to publicly announced secondary stock sales. The sign and significance of the cumulative excess returns for the twelve months following the distribution appear to be negative in most specifications, but are sensitive to the benchmark used.

Significant differences appear in the returns for some sub-samples. Distributions that occur in settings where information asymmetries may be greatest—especially where the firm has been taken public by a lower-tier underwriter and the distribution is soon after the IPO—have larger immediate price declines. Post-distribution price performance is related to factors that predict event window returns.

Finally, Brav and Gompers (1997) explore the long-run performance implications of venture capital backing after they perform an IPO. In particular, they examine whether the pre-IPO performance differences noted by Hellmann and Puri (2002) or Gompers and Lerner (1998b) carry over to when the companies go public, long after they received venture financing. Brav and Gompers find that venture capital-backed companies do indeed outperform comparable nonventure-capital-backed companies, with venture capital backed companies earning 40% more over five years after the IPO.

## 4. Venture investing and innovation

In this section, I explore the issue of venture capital impact on innovation. I begin by reviewing the evidence regarding the overall impact of venture capital on innovation. I then turn to exploring the impact of the historic boom-and bust pattern on these shifts. I highlight that while the overall relationship between venture capital and innovation is positive, the relationships across the cycles of venture activity may be quite different.

A lengthy theoretical literature has been developed in recent years, as financial economists have sought to understand the mechanisms employed by venture capitalists. These works suggest that these financial intermediaries are particularly well suited for nurturing innovative new firms.

It might be thought that it would be not difficult to address the question of the impact of venture capital on innovation. For instance, one could look in regressions across industries and time whether, controlling for R&D spending, venture capital funding has an impact on various measures of innovation. But even a simple model of the relationship between venture capital, R&D, and innovation suggests that this approach is likely to give misleading estimates.

Both venture funding and innovation could be positively related to a third unobserved factor, the arrival of technological opportunities. Thus, there could be more innovation at times that there was more venture capital, not because the venture capital caused the innovation, but rather because the venture capitalists reacted to some fundamental technological shock which was sure to lead to more innovation. To date, only two papers have attempted to address these challenging issues.

The first of these papers, Hellmann and Puri (2000), examines a sample of 170 recently formed firms in Silicon Valley, including both venture-backed and non-venture firms. Using questionnaire responses, they find empirical evidence that venture capital financing is related to product market strategies and outcomes of startups. They find that firms that are pursuing what they term an innovator strategy (a classification based on the content analysis of survey responses) are significantly more likely and faster to obtain venture capital. The presence of a venture capitalist is also associated with a significant reduction in the time taken to bring a product to market, especially for innovators. Furthermore, firms are more likely to list obtaining venture capital as a significant milestone in the lifecycle of the company as compared to other financing events.

The results suggest significant interrelations between investor type and product market dimensions, and a role of venture capital in encouraging innovative companies. Given the small size of the sample and the limited data, they can only modestly address concerns about causality. Unfortunately, the possibility remains that more innovative firms select venture capital for financing, rather than venture capital causing firms to be more innovative.

Kortum and Lerner (2000), by way of contrast, examine these patterns can be discerned on an aggregate industry level, rather than on the firm level. They address concerns about causality in two ways. First, they exploit the major discontinuity in the recent history of the venture capital industry: as discussed above, in the late 1970s, the U.S. Department of Labor clarified the Employee Retirement Income Security Act, a policy shift that freed pensions to invest in venture capital. This shift led to a sharp increase in the funds committed to venture capital. This type of exogenous change should identify the role of venture capital, because it is unlikely to be related to the arrival of entrepreneurial opportunities. They exploit this shift in instrumental variable regressions. Second, they use R&D expenditures to control for the arrival of technological opportunities that are anticipated by economic actors at the time, but that are unobserved to econometricians. In the framework of a simple model, they show that the causality problem disappears if they estimate the impact of venture capital on the patent–R&D ratio, rather than on patenting itself.

Even after addressing these causality concerns, the results suggest that venture funding does have a strong positive impact on innovation. The estimated coefficients vary according to the techniques employed, but on average a dollar of venture capital appears to be three to four times more potent in stimulating patenting than a dollar of traditional corporate R&D. The estimates therefore suggest that venture capital, even though it averaged less than three percent of corporate R&D from 1983 to 1992, is responsible for a much greater share—perhaps ten percent—of U.S. industrial innovations in this decade.

The evidence that venture capital has a powerful impact on innovation might lead us to be especially worried about market downturns. A dramatic fall in venture capital financing, it is natural to conclude, would lead to a sharp decline in innovation.

But this reasoning, while initially plausible, is somewhat misleading. For the impact of venture capital on innovation does not appear to be uniform. Rather, during periods when the intensity of investment is greatest, the impact of venture financing appears to decline. The uneven impact of venture on innovation can be illustrated by examining the experience during two "boom" periods in the industry.

One example was the peak period of biotechnology investing in the early 1990s. While the potential of biotechnology to address human disease was doubtless substantial, the extent and nature of financing seemed to many observers at the time hard to justify. In some cases, dozens of firms pursuing similar approaches to the same disease target were funded. Moreover, the valuations of these firms often were exorbitant: for instance, between May and December 1992, the average valuation of the privately held biotechnology firms financed by venture capitalists was $70 million. These doubts were validated when biotechnology valuations fell precipitously in early 1993: by December 1993, only 42 of 262 publicly traded biotechnology firms had a valuation over $70 million.

Most of the biotechnology firms financed during this period ultimately yielded very disappointing returns for their venture financiers and modest gains for society as a whole. In many cases, the firms were liquidated after further financing could not be arranged. In others, the firms shifted their efforts into other, less competitive areas, largely abandoning the initial research efforts. In yet others, the companies remained mired with their peers for years in costly patent litigation.

The boom of 1998–2000 provides many additional illustrations. Funding during these years was concentrated in two areas: Internet and telecommunication investments, which, for instance, accounted for 39% and 17% of all venture disbursements in 1999. Once again, considerable sums were devoted to supporting highly similar firms—e.g., the nine dueling Internet pet food suppliers—or else efforts that seemed fundamentally uneconomical and doomed to failure, such as companies which undertook the extremely capital-intensive process of building a second cable network in residential communities. Meanwhile, many apparently promising areas—e.g., advanced materials, energy technologies, and micro manufacturing—languished unfunded as venture capitalists raced to focus on the most visible and popular investment areas. It is difficult to believe that the impact of a dollar of venture financing was as powerful in spurring innovation during these periods as in others.

## 5. What we don't know about venture capital

While financial economists know much more about venture capital than they did a decade ago, there are many unresolved issues that would reward future research. In this final section, I highlight three areas for further research that I consider particularly promising.

### *5.1. Understanding risk and return*

One critical, but unanswered area, is the assessment of venture capital as a financial asset. Many institutions, primarily public and private pension funds, have increased their allocation to venture capital and private equity in the belief that the returns of these funds are largely uncorrelated with the public markets.

It is natural to see how they come to this conclusion. Firms receiving capital from private equity funds very often remain privately held for a number of years after the initial investment. These firms have no observable market price. In order to present a conservative assessment of the portfolio valuation, private equity managers often refrain from marking portfolio firm values to market, preferring to maintain the investments at book value.

But as discussed throughout this analysis, there appear to be many linkages between the public and private equity market values. Thus, the stated returns of private equity funds may not accurately reflect the true evolution of value, and the correlations reported by Venture Economics (1997) and other industry observers may be deceptively low. To ignore the true correlation is fraught with potential dangers.[4]

[4] In a preliminary analysis using data from one venture group, Gompers and Lerner (1997) find that the correlation between venture capital and public market prices increases substantially when the underlying venture portfolio is "marked-to-market". An alternative approach is to examine the relatively modest number of publicly traded venture capital funds, as is done by Martin and Petty (1983).

Recent work by Kaplan and Schoar (2004) and Cochrane (2004) has attempted to deal with this stale price problem. Kaplan and Schoar use the change in the level of the S&P 500 as a benchmark from the time of investment while Cochrane uses econometric corrections for stale prices and selection biases in the data. While the results of each are somewhat contradictory, they are important first steps in addressing a problem that is clearly central to the asset allocation decision of many investors.

### *5.2. The internationalization of venture capital*

The rapid growth in the U.S. venture capital market has led institutional investors to look increasingly at venture capital alternatives abroad. Until very recently, outside of the United Kingdom (where performance of funds has been quite poor) and Israel, there has been little venture capital activity abroad.[5] (Table 6 provides an international comparison of venture capital activity.) Black and Gilson (1998) argue that the key source of the U.S. competitive advantage in venture capital is the existence of a robust IPO market. Venture capitalists can commit to transfer control back to the entrepreneur when a public equity market for new issues exists. This commitment device is unavailable in economies dominated by banks, such as Germany and Japan.

These arguments, however, have less credibility in light of the events of the past twelve months. There has been a surge in venture capital investment, particularly relating to the Internet, in a wide variety of nations across Asia, Europe, and Latin America. While local groups (many recently established) have made some of these investments, much of the activities have been driven by U.S.-based organizations.

In a pioneering study, Jeng and Wells (1999) examine the factors that influence venture capital fundraising in 21 countries. They find that the strength of the IPO market is an important factor in the determinant of venture capital commitments, echoing the conclusions of Black and Gilson. Jeng and Wells find, however, that the IPO market does not seem to influence commitments to early-stage funds as much as later-stage ones. While this work represents an important initial step, much more remains to be explored regarding the internationalization of venture capital.

One provocative finding from the Jeng and Wells analysis is that government policy can have a dramatic impact on the current and long-term viability of the venture capital sector. In many countries, especially those in Continental Europe, policymakers face a dilemma. The relatively few entrepreneurs active in these markets face numerous daunting regulatory restrictions, a paucity of venture funds focusing on investing in high-growth firms, and illiquid markets where investors do not welcome IPOs by

[5] One potential source of confusion is that the term venture capital is used differently different in Europe and Asia. Abroad, venture capital often refers to all private equity, including buyout, late stage, and mezzanine financing (which represent the vast majority of the private equity pool in most overseas markets). In the U.S., these are separate classes. I confine our discussion of international trends—as the rest of the paper—to venture capital using the restrictive, U.S. definition.

Table 6
The size of the venture capital pool in 21 nations in 1995. I use Jeng and Wells' figures for early-stage funds in each country outside the U.S. because I believe it to be most comparable to venture capital funds as defined in the U.S. Figures for Australia and New Zealand are 1994 estimated levels; figures for Israel are a 1995 estimate; and figures for Portugal are the actual level in 1994. All dollar figures are in millions of current U.S. dollars

| Country | Total venture capital under management |
|---|---|
| Australia | 54 |
| Austria | 0.4 |
| Belgium | 8 |
| Canada | 182 |
| Denmark | 4 |
| Finland | 1 |
| France | 35 |
| Germany | 116 |
| Ireland | 1 |
| Israel | 550 |
| Italy | 60 |
| Japan | 11 |
| Netherlands | 100 |
| New Zealand | 1 |
| Norway | 7 |
| Portugal | 9 |
| Spain | 24 |
| Sweden | 9 |
| Switzerland | 1 |
| United Kingdom | 36 |
| United States | 3,651 |

Source: Compiled from Jeng and Wells (1999), as slightly amended by the author.

young firms without long histories of positive earnings. It is often unclear where to being the process of duplicating the success of the United States. Only very recently have researchers begun to examine the ways in which policymakers can catalyze the growth of venture capital and the companies in which they invest. (Three recent exceptions are Irwin and Klenow (1996), Lerner (1999), and Wallsten (1996).) Given the size of recent initiatives undertaken both in the United States and abroad (summarized in Lerner, 1999, and Gompers and Lerner, 1999a), much more needs to be done in this arena.

## References

Admati, A., Pfleiderer, P., 1994. Robust financial contracting and the role for venture capitalists. Journal of Finance 49, 371–402.

Aghion, P., Bolton, P., 1992. In incomplete contracts approach to financial contracting. Review of Economic Studies 59, 473–494.

Amit, R., Glosten, L., Muller, E., 1990a. Does venture capital foster the most promising entrepreneurial firms? California Management Review 32, 102–111.

Amit, R., Glosten, L., Muller, E., 1990b. Entrepreneurial ability, venture investments, and risk sharing. Management Science 36, 1232–1245.

Baker, M., Gompers, P.A., 2004a. The determinants of board structure and function in entrepreneurial firms. Journal of Law and Economics, in press.

Baker, M., Gompers, P.A., 2004b. An analysis of executive compensation, ownership, and control in entrepreneurial firms. Working Paper, Harvard University.

Barry, C.B., Muscarella, C.J., Peavy III, J.W., Vetsuypens, M.R., 1990. The role of venture capital in the creation of public companies: Evidence from the going public process. Journal of Financial Economics 27, 447–471.

Berglöf, E., 1994. A control theory of venture capital finance. Journal of Law, Economics, and Organizations 10, 247–267.

Bergmann, D., Hege, U., 1998. Dynamic venture capital financing, learning, and moral hazard. Journal of Banking and Finance, in press.

Black, B., Gilson, R., 1998. Venture capital and the structure of capital markets: Banks versus stock markets. Journal of Financial Economics 47, 243–277.

Bonacich, P., 1987. Power and centrality: A family of measures. American Journal of Sociology 92, 1170–1182.

Brav, A., Gompers, P.A., 1997. Myth or reality? Long-run underperformance of initial public offerings; Evidence from venture capital and nonventure capital-backed IPOs. Journal of Finance 52, 1791–1821.

Brav, A., Gompers, P.A., 2003. The role of lock-ups in initial public offerings. Review of Financial Studies 16, 1–29.

Chan, Y., 1983. On the positive role of financial intermediation in allocation of venture capital in a market with imperfect information. Journal of Finance 38, 1543–1568.

Cochrane, J., 2004. The risk and return of venture capital. Journal of Financial Economics, in press.

Cornelli, F., Yosha, O., 1997. Stage financing and the role of convertible debt. Working Paper, London Business School and Tel Aviv University.

Dewatripont, M., Tirole, J., 1994. A theory of debt and equity: Diversity of securities and manager–shareholder congruence. Quarterly Journal of Economics 109, 1027–1054.

Fama, E.F., Jensen, M.C., 1983. Separation of ownership and control. Journal of Law and Economics 26, 301–325.

Gompers, P., 1995. Optimal investment, monitoring, and the staging of venture capital. Journal of Finance 50, 1461–1489.

Gompers, P., 1996. Grandstanding in the venture capital industry. Journal of Financial Economics 42, 133–156.

Gompers, P., Lerner, J., 1997. Risk and reward in private equity investments: The challenge of performance assessment. Journal of Private Equity 1 (Winter), 5–12.

Gompers, P., Lerner, J., 1998a. Venture capital distributions: Short- and long-run reactions. Journal of Finance 53, 2161–2183.

Gompers, P., Lerner, J., 1998b. What drives venture fundraising? Brookings Papers on Economic Activity—Microeconomics, 149–192.

Gompers, P., Lerner, J., 1999a. Capital market imperfections in venture markets: A report to the Advanced Technology Program. Advanced Technology Program, U.S. Department of Commerce, Washington.

Gompers, P., Lerner, J., 1999b. The Venture Capital Cycle. MIT Press, Cambridge.

Gompers, P., Lerner, J., 2000. Money chasing deals? The impact of fund inflows on private equity valuations. Journal of Financial Economics 55, 239–279.

Gompers, P., Lerner, J., 2001. The Venture Capital Revolution. Harvard Business School Press, Boston.

Grossman, S., Hart, O., 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. Journal of Political Economy 94, 691–719.

Hart, O., Moore, J., 1990. Property rights and the nature of the firm. Journal of Political Economy 98, 1119–1158.

Hart, O., Moore, J., 1998. Default and renegotiation: A dynamic model of debt. Quarterly Journal of Economics 113, 1–41.
Hellmann, T., 1998. The allocation of control rights in venture capital contracts. Rand Journal of Economics 29, 57–76.
Hellmann, T., Puri, M., 2000. The interaction between product market and financing strategy: The role of venture capital. Review of Financial Studies 13, 959–984.
Hellmann, T., Puri, M., 2002. Venture capital and the professionalization of start-up firms: Empirical evidence. Journal of Finance 57, 169–197.
Hermalin, B.E., Weisbach, M.S., 1988. The determinants of board composition. Rand Journal of Economics 19, 589–606.
Hochberg, I., Ljungqvist, A., Lu, Y., 2006. Networking as entry deterrence and the competitive supply of venture capital (March 2006). AFA 2007 Chicago Meeting Paper. Available at SSRN: http://ssm.com/abstract=891468.
Irwin, D.A., Klenow, P.J., 1996. Hightech R&D subsidies: Estimating the effects of Sematech. Journal of International Economics 40, 323–344.
Jeng, L., Wells, P., 1999. The determinants of venture capital funding: An empirical analysis. Working Paper, Harvard University.
Jensen, M.C., Meckling, W.H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics 3, 305–360.
Kaplan, S., Schoar, A., 2004. Private equity returns: Persistence and capital flows. Journal of Finance, in press.
Kaplan, S., Stromberg, P., 2003. Financial contracting theory meets the real world: Evidence from venture capital contracts. Review of Economic Studies 70, 281–315.
Kortum, S., Lerner, J., 1999. Does venture capital spur innovation? Working Paper No. 6846, National Bureau of Economic Research.
Kortum, S., Lerner, J., 2000. Assessing the impact of venture capital on innovation. Rand Journal of Economics 31, 674–692.
Lerner, J., 1994a. The syndication of venture capital investments. Financial Management 23, 16–27.
Lerner, J., 1994b. Venture capitalists and the decision to go public. Journal of Financial Economics 35, 293–316.
Lerner, J., 1995. Venture capitalists and the oversight of private firms. Journal of Finance 50, 301–318.
Lerner, J., 1997. An empirical exploration of a technology race. Rand Journal of Economics 28, 228–247.
Lerner, J., 1999. The government as venture capitalist: The long-run effects of the SBIR program. Journal of Business 72, 285–318.
Liles, P., 1977. Sustaining the Venture Capital Firm. Management Analysis Center, Cambridge.
Martin, J.D., Petty, J.W., 1983. An analysis of the performance of publicly traded venture capital companies. Journal of Financial and Quantitative Analysis 18, 401–410.
Marx, L., 1994. Negotiation and renegotiation of venture capital contracts. Working Paper, University of Rochester.
Megginson, W.C., Weiss, K.A., 1991. Venture capital certification in initial public offerings. Journal of Finance 46, 879–893.
Myers, S., 1977. Determinants of corporate borrowing. Journal of Financial Economics 5, 147–175.
Ritter, J.R., 2006. Jay R. Ritter's home page: http://bear.cba.ufl.edu/ritter.
Sahlman, W.A., 1990. The structure and governance of venture capital organizations. Journal of Financial Economics 27, 473–524.
Sahlman, W.A., Stevenson, H., 1987. Capital market myopia. Journal of Business Venturing 1, 7–30.
Venture Economics, 1997. Investment Benchmark Reports Venture Capital. Venture Economics, New York.
Wallsten, S.J., 1996. The Small Business Innovation Research program: Encouraging technological innovation and commercialization in small firms? Working Paper, Stanford University (unpublished).
Williamson, O.E., 1983. Organization form, residual claimants, and corporate control. Journal of Law and Economics 26, 351–366.

# AUTHOR INDEX

n indicates citation in footnote.

# SUBJECT INDEX

# INTRODUCTION TO THE SERIES



The Handbooks in Finance are intended to be a definitive source for comprehensive and accessible information in the field of finance. Each individual volume in the series presents an accurate self-contained survey of a sub-field of finance, suitable for use by finance and economics professors and lecturers, professional researchers, graduate students and as a teaching supplement. The goal is to have a broad group of outstanding volumes in various areas of finance.

William T. Ziemba
University of British Columbia

Publisher's Note
For a complete overview of the Handbooks in Finance Series, please refer to the listing at the end of this volume.

# PREFACE: EMPIRICAL CORPORATE FINANCE

Judging by the sheer number of papers reviewed in this Handbook, the empirical analysis of firms' financing and investment decisions—empirical corporate finance—has become a dominant field in financial economics. The growing interest in everything "corporate" is fueled by a healthy combination of fundamental theoretical developments and recent widespread access to large transactional data bases. A less scientific—but nevertheless important—source of inspiration is a growing awareness of the important social implications of corporate behavior and governance. This Handbook takes stock of the main empirical findings to date across an unprecedented of corporate finance issues, ranging from econometric methodology, to raising capital and capital structure choice, and to managerial incentives and corporate investment behavior. The surveys are written by leading empirical researchers that remain active in their respective areas of interest. With few exceptions, the writing style makes the chapters accessible to industry practitioners. For doctoral students and seasoned academics, the surveys offer dense roadmaps into the empirical research landscape and provide suggestions for future work.

## Part 1 (Volume 1): Econometric Issues and Methodological Trends

The empirical corporate finance literature is progressing through a combination of large-sample data descriptions, informal hypotheses testing, as well as structural tests of theory. Researchers are employing a wide spectrum of econometric techniques, institutional settings, and markets structures in order to distill the central message in the data. Part 1 of Volume 1 begins by reviewing key econometric issues surrounding event studies, and proceeds to explain the econometrics of self-selection. It then explains and illustrates methodological issues associated with the growing use of auction theory, and it ends with a discussion of key elements of the corporate finance evidence from a behavioral perspective.

In Chapter 1, "Econometrics of Event Studies," S.P. Kothari and Jerold Warner review the power of the event-study method; the most successful empirical technique to date for isolating the price impact of the information content of corporate actions. The usefulness of event studies arises from the fact that the magnitude of abnormal performance at the time of an event provides a measure of the (unanticipated) impact of this type of event on the wealth of the firms' claimholders. Thus, event studies focusing on announcement effects for a short horizons around an event provide evidence relevant for understanding corporate policy decisions. Long-horizon event studies also serve an important purpose in capital market research as a way of examining market efficiency. The survey discusses sampling distributions and test statistics typically used in event studies, as well as criteria for reliability, specification and power. While much is known about the statistical properties of short-horizon event studies, the survey provides a critical review of potential pitfalls of long-horizon abnormal return estimates. Serious challenges related to model specification, skewness and cross-correlation remain. As they also point out, events are likely to be associated with return-variance increases, which are equivalent to abnormal returns varying across sample securities. Misspecification induced by variance increases can cause the null hypothesis to be rejected too often unless the test statistic is adjusted to reflect the variance shift. Moreover, the authors emphasize the importance of paying close attention to specification issues for nonrandom samples of corporate events.

Self-selection is endemic to voluntary corporate events. In Chapter 2, "Self-Selection Models in Corporate Finance," Kai Li and Nagpurnanand Prabhala review the relevant econometric issues with applications in corporate finance. The statistical issue raised by self-selection is the wedge between the population distribution and the distribution within a selected sample, which renders standard linear (OLS/GLS) estimators biased and inconsistent. This issue is particularly relevant when drawing inferences about the determinants of event-induced abnormal stock returns from multivariate regressions, a technique used by most event studies today. These regressions are typically run using samples that exclude non-event firms. The standard solution is to include a scaled estimate of the event probability—the inverse Mills ratio (the expected value of the true but unobservable regression error term)—as an additional variable in the regression. Interestingly, as the author spoint out, testing for the significance of the inverse Mills ratio is equivalent to testing whether the sample firms use private information when they self-select to undertake the event. Conversely, if one believes that the particular event being studied is induced by or reflect private information (market overpricing of equity, arrival of new investment projects, merger opportunities, etc.), then consistent estimation of the parameters in the cross-sectional regression requires the appropriate control for self-selection. What is "appropriate" generally depends on the specific application and should ideally be guided by economic theory. The survey also provides a highly useful overview of related econometric techniques—including matching (treatment effect) models, panel data with fixed effects, and Bayesian self-selection models—with specific applications.

In Chapter 3, "Auctions in Corporate Finance," Sudipto Dasgupta and Robert Hansen introduce auction theory and discuss applications in corporate finance. The authors explain theoretical issues relating to pricing, efficiency of allocation (the conditions under which the asset is transferred to the most efficient buyer), differential information, collusion among buyers, risk aversion, and the effects of alternative auctions designs (sealed-bids versus open auction, seller reserve price, entry fees, etc.). It is important for empirical research in corporate finance to be informed of auction theory for at least two reasons. First, when sampling a certain transaction type that in fact takes place across a variety of transactional settings, auction theory help identify observable characteristics that are likely to help explain the cross-sectional distribution of things like transaction/bid prices, expected seller revenues, valuation effects, and economic efficiency. This is perhaps most obvious in studies of corporate takeovers (negotiation versus auction, strategic bidding behavior, etc.) and in public security offerings (role of intermediaries, degree and role of initial underpricing, long-run pricing effects, etc.). Second, auction theory provides solutions to the problem of optimal selling mechanism design. This is highly relevant in debates over the efficiency of the market for corporate control (negotiations versus auction, desirability of target defensive mechanisms, the role of the board), the optimality of the bankruptcy system (auctions versus court-supervised negotiations, allocation of control during bankruptcy, prospects for fire-sales, risk-shifting incentives, etc.), and the choice of selling mechanism when floating new securities (rights offer, underwritten offering, fixed-price, auction, etc.).

In Chapter 4, "Behavioral Corporate Finance," Malcolm Baker, Richard Ruback and Jeffery Wurgler survey several aspects of corporate finance and discuss the scope for competing behavioral and rational interpretations of the evidence. The idea that inherent behavioral biases of CEOs—and their perception of investor bias—may affect corporate decisions is both intuitive and compelling. A key methodological concern is how to structure tests with the requisite power to discriminate between behavioral explanations and classical hypotheses based on rationality. The "bad model" problem—the absence of clearly empirically testable predictions—is a challenge for *both* rational and behavioral models. For example, this is evident when using a scaled-price ratio such as the market-to-book ratio (B/M), and where the book value is treated as a fundamental asset value. A high value of B/M may be interpreted as "overvaluation" (behavioral) or, alternatively, as B poorly reflecting economic fundamentals (rational). Both points of view are consistent with the observed inverse relation between B/M and expected returns (possibly with the exception of situations with severe short-selling constraints). Also, measures of "abnormal" performance following some corporate event necessarily condition on the model generating expected return. The authors carefully discuss these issues and how researchers have tried to reduce the joint model problem, e.g., by considering cross-sectional interactions with firm-characteristics such as measures of firm-specific financing constraints. The survey concludes that behavioral approaches help explain a number of important financing and investment patterns, and it offers a number of open questions for future research.

## Part 2 (Volume 1): Banking, Public Offerings, and Private Sources of Capital

In Part 2, the Handbook turns to investment banking and the capital acquisition process. Raising capital is the lifeline of any corporation, and the efficiency of various sources of capital, including banks, private equity and various primary markets for new securities is an important determinant of the firm's cost of capital.

In Chapter 5, "Banks in Capital Markets," Steven Drucker and Manju Puri review empirical work on the dual role of banks as lenders and as collectors of firm-specific private information through the screening and monitoring of loans. Until the late 1990s, U.S. commercial banks were prohibited from underwriting public security offerings for fear that these banks might misuse their private information about issuers (underwriting a low quality issuer and market it as high quality). Following the repeal of the Glass–Steagall Act in the late 1990s, researchers have examined the effect on underwriter fees of the emerging competition between commercial and investment banks. Commercial banks have emerged as strong competitors: in both debt and equity offerings, borrowers receive lower underwriting fees when they use their lending bank as underwriter. The evidence also shows that having a lending relationship constitutes a significant competitive advantage for the commercial banks in terms of winning underwriting mandates. In response, investment banks have started to develop lending units, prompting renewed concern with conflicts of interest in underwriting. Overall, the survey concludes that there are positive effects from the interaction between commercial banks' lending activities and the capital markets, in part because the existence of a bank lending relationship reduces the costs of information acquisition for capital market participants.

In Chapter 6, "Security Offerings," Espen Eckbo, Ronald Masulis and Øyvind Norli review studies of primary markets for new issues, and they extend and update evidence on issue frequencies and long-run stock return performance. This survey covers all of the key security types (straight and convertible debt, common stock, preferred stock, ADR) and the most frequently observed flotation methods (IPO, private placement, rights offering with or without standby underwriting, firm commitment underwritten offering). The authors review relevant aspects of securities regulations, empirical determinants of underwriter fees and the choice of flotation method, market reaction to security issue announcements internationally, and long-run performance of U.S. issuers. They confirm that the relative frequency of *public* offerings of seasoned equity (SEOs) is low and thus consistent with a financial pecking order based on adverse selection costs. They also report that the strongly negative announcement effect of SEOs in the U.S. is somewhat unique to U.S. issuers. Equity issues in other countries are often met with a significantly positive market reaction, possibly reflecting a combination of the greater ownership concentration and different selling mechanisms in smaller stock markets. They conclude from this evidence that information asymmetries have a first-order effect on the choice of which security to issue as well as by which method. Their large-sample estimates of post-issue long-run abnormal performance, which covers a wide range of security types, overwhelmingly reject the hypothesis that the performance is 'abnormal.' Rather, the long-run performance is commensurable with issuing firms' exposures to commonly accepted definitions of pervasive risk factors. They conclude that the long-run evidence fails to support hypotheses which hold that issuers systematically time the market, or hypotheses which maintain that the market systematically over- or under-reacts to the information in the issue announcement.

The cost of going public is an important determinant of financial development and growth of the corporate sector. In Chapter 7, "IPO Underpricing," Alexander Ljungqvist surveys the evidence on one significant component of this cost: IPO underpricing, commonly defined as the closing price on the IPO day relative to the IPO price. He classifies theories of underpricing under four broad headings: 'asymmetric information' (between the issuing firm, the underwriter, and outside investors), 'institutional' (focusing on litigation risk, effects of price stabilization, and taxes), 'control' (how the IPO affects ownership structure, agency costs and monitoring), and 'behavioral' (where irrational investors bid up the price of IPO shares beyond true value). From an empirical perspective, these theories are not necessarily mutually exclusive, and several may work to successfully explain the relatively modest level of underpricing (averaging about 15%) observed before the height of the technology-sector offerings in 1999–2000. Greater controversy surrounds the level of underpricing observed in 1999–2000, where the dollar value of issuers' underpricing cost ('money left on the table') averaged more than four times the typical 7% investment banking fee. Two interesting—and mutually exclusive—candidate explanations for this unusual period focus on inefficient selling method design (failure of the fix-priced book-building procedure to properly account for the expected rise in retail investor demand) and investor irrationality (post-offering pricing 'bubble'). Additional work on the use and effect of IPO auctions, and on the uniquely identifying characteristics of a pricing 'bubble,' is needed to resolve this issue.

Multidivisional (conglomerate) firms may exist in part to take advantage of internal capital markets. However, in apparent contradiction of this argument, the early literature on conglomerate firms identified a 'conglomerate discount' relative to pure-play (single-plant) firms. In Chapter 8, "Conglomerate Firms and Internal Capital Markets," Vojislav Maksimovic and Gordon Phillips present a comprehensive review of how the literature on the conglomerate discount has evolved to produce a deeper economic understanding of the early discount evidence. They argue that issues raised by the data sources used to define the proper equivalent 'pure-play' firm, econometric issues arising from firms self-selecting the conglomerate form, and explicit model-based tests derived from classical profit-maximizing behavior, combine to explain the discount without invoking agency costs and investment inefficiencies. As they explain, a firm that chooses to diversify is a different type of firm than one which stays with a single segment—but either type may be value-maximizing. They conclude that, on balance, internal capital markets in conglomerate firms appear to be efficient in reallocating resources.

After reviewing internal capital markets, bank financing, and public securities markets, Volume 1 ends with the survey "Venture Capital" in Chapter 9. Here, Paul Gompers defines venture capital as "independent and professionally managed, dedicated pools of capital that focus on equity or equity-linked investments in privately held, high-growth companies." The venture capital industry fuels innovation by channeling funds to start-up firms and, while relatively small compared to the public markets, has likely had a disproportionately positive impact on economic growth in the United States where the industry is most developed. The empirical literature on venture capital describes key features of the financial contract (typically convertible preferred stock), staging of the investment, active monitoring and advice, exit strategies, etc., all of which affect the relationship between the venture capitalist and the entrepreneur. While data sources are relatively scarce, there is also growing evidence on the risk and return of venture capital investments. Paul Gompers highlights the need for further research on assessing venture capital as a financial asset, and on the internationalization of venture capital.

### Part 3 (Volume 2): Dividends, Capital Structure, and Financial Distress

The first half of Volume 2 is devoted to the classical issue of capital structure choice. This includes the effect of taxes, expected bankruptcy costs, agency costs, and the costs of adverse selection in issue markets on the firm's choice of financial leverage and dividend policy. More recent empirical work also links debt policy to competition in product markets and to the firm's interaction with its customers and suppliers. There is also substantial empirical work on the effect on expected bankruptcy and distress costs of the design of the bankruptcy code, where claim renegotiation under court supervision (such as under Chapter 11 of the U.S. code) and auctions in bankruptcy (such as in Sweden) are major alternatives being studied.

In Chapter 10, "Payout Policy," Avner Kalay and Michael Lemmon refer to payout policy as "the ways in which firms return capital to their equity investors." Classical dividend puzzles include why firms keep paying cash dividends in the presence of a tax-disadvantage relative to capital gains, and why dividend changes have information contents. In contrast to increases in debt interest payments, dividend increases are not contractually binding and therefore easily reversible. So, where is the commitment to maintain the increased level of dividends? While there is strong evidence of a positive information effect of unanticipated dividend increases, they argue that available signaling models are unlikely to capture this empirical phenomenon. Moreover, there is little evidence that dividend yields help explain the cross-section of expected stock returns—which fails to reveal a tax effect of dividend policy. Recent surveys indicate that managers today appear to consider dividends as a second order concern after investment and liquidity needs are met, and to an increased reliance on stock repurchase as an alternative to cash payouts.

In Chapter 11, "Taxes and Corporate Finance," John Graham reviews research specifically relating corporate and personal taxes to firms' choice of payout policy, capital structure, compensation policy, pensions, corporate forms, and a host of other financing arrangements. This research often finds that taxes do appear to affect corporate decisions, but the economic magnitude of the tax effect is often uncertain. There is cross-sectional evidence that high-tax rate firms use debt more intensively than do low-tax rate firms, but time-series evidence concerning whether firm-specific changes in tax status affect debt policy is sparse. Many firms appear to be "underleveraged" in the sense that they could capture additional tax-related benefits of debt at a low cost—but refrain from doing so. Conclusions concerning "underleverage" are, however, contingent on a model of the equilibrium pricing implications of the personal tax-disadvantage of interest over equity income, a topic that has been relatively little researched. Graham also points to the need for a total tax-planning view (as opposed to studying tax issues one by one) to increase the power of tests designed to detect overall tax effects on firm value.

In Chapter 12, "Tradeoff and Pecking Order Theories of Debt," Murray Frank and Vidhan Goyal review the empirical evidence on firms capital structure choice more generally. Under the classical tradeoff theory, the firm finds the optimal debt level at the point where the marginal tax benefit of another dollar of debt equals the marginal increase in expected bankruptcy costs. This theory is somewhat challenged by the evidence of underleverage surveyed by Graham. However, corporate leverage ratios appears to be mean-reverting over long time horizons, which is consistent with firms trying to maintain target leverage ratios. This target may reflect transaction costs of issuing securities, agency costs, and information asymmetries as well as taxes and bankruptcy costs, and the available evidence does not indicate which factors are the dominant ones. They report several stylized facts about firms leverage policies. In the aggregate for large firms (but not for small firms), capital expenditures track closely internal funds, and the "financing deficit" (the difference between investments and internal funds) track closely debt issues. This is as predicted by the "pecking order" hypothesis, under which debt is preferred over equity as a source of external finance. For small firms, however, the deficit tracks closely equity issues, which reverses the prediction of the pecking order. The authors conclude that "no currently available model appears capable of simultaneously accounting for the stylized facts."

In Chapter 13, "Capital Structure and Corporate Strategy," Chris Parsons and Sheridan Titman survey arguments and evidence that link firms' leverage policies to structural characteristics of product markets. Capital structure may affect how the firm chooses to interact with its non-financial stakeholders (customers, workers, and suppliers concerned with the firm's survival) as well as with competitors. To account for endogeneity problems that commonly arise in this setting, most papers in this survey analyze firms' responses to a "shock," whether it be a sharp (and hopefully unanticipated) leverage change, an unexpected realization of a macroeconomic variable, or a surprising regulatory change. This approach often allows the researcher to isolate the effect of leverage on a firm's corporate strategy, and in some cases, makes it possible to pinpoint the specific channel (for example, whether a financially distressed firm lowers prices in response to predation by competitors or by making concessions to its customers). There is evidence that debt increases a firm's employment sensitivity to demand shocks (perhaps perpetuating recessions), but can also protect shareholder wealth by moderating union wage demands. Excessive leverage can also inhibit a firm's ability to compete in the product market, as measured by prices and market shares. Firms that depend crucially on non-fungible investments from stakeholders are most sensitive to these losses, and choose more conservative capital structures as a result.

To avoid formal bankruptcy, financially distressed firms engage in asset sales, equity issues and debt renegotiations. In Chapter 14, "Bankruptcy and the Resolution of Financial Distress," Edith Hotchkiss, Kose John, Robert Mooradian and Karin Thorburn survey empirical work on the costs, benefits, and effectiveness of out-of-court debt workouts and of formal "one size fits all" bankruptcy procedures. Failing to renegotiate their debt claims out of court, the firm files for bankruptcy, where it is either liquidated piecemeal or restructured as a going concern under court protection. For reasons that are poorly understood, different bankruptcy systems have evolved in different countries, with a trend toward the structured bargaining process characterizing Chapter 11 of the U.S. code. The U.S. code substantially restricts the liquidation rights of creditors as filing triggers automatic stay of debt payments, prevents repossession of collateral, and allows the bankrupt firm to raise new debt with super-priority (debtor-in-possession financing). In contrast, UK bankruptcy is akin to a contract-driven receivership system where creditor rights are enforced almost to the letter. Here, assets pledged as collateral can be repossessed even if they are vital for the firm, and there is no stay of debt claims. This makes it difficult to continue to operate the distressed firm under receivership, even if the bankrupt firm is economically viable. A third system is found in Sweden where the filing firm is automatically turned over to a court-appointed trustee who arranges an open auction (while all debt claims are stayed). The authors survey the international evidence on bankruptcies (which also includes France, Germany, and Japan). They conclude that it remains an open question whether Chapter 11 in the U.S.—with its uniquely strong protection of the incumbent management team—represents an optimal bankruptcy reorganization procedure.

## Part 4 (Volume 2): Takeovers, Restructurings, and Managerial Incentives

Modern corporate finance theory holds that in a world with incomplete contracting, financial structure affects corporate investment behavior and therefore firm value. The Handbook ends with comprehensive discussions of the value-implications of major corporate investment and restructuring decisions (outside of bankruptcy) and of the role of pay-for-performance type of executive compensation contracts on managerial incentives and risk taking behavior.

In Chapter 15, "Corporate Takeovers," Sandra Betton, Espen Eckbo and Karin Thorburn review and extend the evidence on mergers and tender offers. They focus in particular on the bidding process as it evolves sequentially from the first bid through bid revision(s) and towards the final bid outcome. Central issues include bid financing, strategic bidding, agency issues and the impact of statutory and regulatory restrictions. The strategic arsenal of the initial bidder includes approaching the target with a tender offer or a merger bid, acquiring a toehold to gain an advantage over potential competitors, offering a payment method (cash or stock) which signals a high bidder valuation of the target, and/or simply bid high (a preemptive strike). The survey provides new evidence on the magnitude of successive bid jumps, and on the speed of rival firm entry and the time between the first and the final bids in multi-bidder contests. The survey confirms that the average abnormal return to bidders is insignificantly different from zero, and that the sum of the abnormal returns to targets and bidders is positive, suggesting that takeovers improve the overall efficiency of resource allocation. Takeover bids also tend to generate positive abnormal returns throughout the industry of the target, in part because they increase the likelihood that industry rivals may become targets themselves (industry "in-play" effect). The evidence strongly rejects the hypothesis that horizontal mergers reduce consumer welfare through increased market power—even when the merger-induced change in industry concentration is non-trivial. However, some input suppliers suffer losses following downstream mergers that increase the downstream industry's bargaining power.

In Chapter 16, "Corporate Restructuring: Breakups and LBOs," Espen Eckbo and Karin Thorburn review a number of financial and asset restructuring techniques—other than corporate takeovers and bankruptcy reorganizations. They distinguish between transactions that securitize corporate divisions from those that recapitalize the entire firm. Forms of divisional securitization include spinoff, splitoff, divestiture, equity carveout and tracking stock. Forms of recapitalizations of the entire firm include leveraged recapitalization, leveraged buyout (LBO), demutualization, going-private transactions, and state privatizations. They show transaction frequency, describe the financing technique, discuss regulatory and tax issues, and review evidence on the associated valuation effects. Announcement-induced abnormal stock returns are generally reported to be positive. Potential sources of this wealth creation include improved alignment of management and shareholder incentives through post-transaction compensation contracts that include divisional stock grants, the elimination of negative synergies, improved governance systems through the disciplinary effect of leverage, the avoidance of underinvestment costs, wealth transfers from old bondholders experiencing claim dilution and risk increase following new debt issues, and an "in-play" effect as divisional securitization increases the probability that the division will become a future acquisition target. Unbundling corporate assets and allowing public trade of securities issued by individual divisions also leads to a general welfare increase from increased market completeness and analyst following. The evidence indicates improved operating performance following spinoffs and LBOs, and increased takeover activity after spinoffs and carveouts, and that a minority of LBO firms goes public within five years of the going-private transaction.

Delegation of corporate control to managers gives rise to costly agency conflicts as the personal interests of managers and owners diverge. The literature on executive compensation seeks to identify the form of the employment contract that minimizes agency costs. In Chapter 17, "Executive Compensation and Incentives," Rajesh Aggarwal surveys the empirical findings of this literature over the past two decades, focusing in particular on evidence concerning stock options and restricted stock grants. The optimal provision of incentives in managerial compensation contracts depends on factors such as executive risk and effort aversion, managerial productivity, and information asymmetries. A key limitation on incentive provision appears to be the need to share risk between managers and shareholders. Also, while optimal contracting theory implies that firm performance should be evaluated relative to an industry or market wide benchmark, relative performance provisions (e.g., by indexing the exercise price of a stock option to the market) are rarely observed. This puzzle may be explained in part by accounting and tax rules, and in part by the cost to shareholders of indexed options (relative to other forms of compensation) when managers are risk averse. Observed compensation practices may also reflect a governance problem if the CEO has undue influence over the determination of her own level of pay. Some researchers argue that rent extraction by the CEO is a major issue of concern for shareholders, an issue that remains controversial.

For a given compensation contract, risk-averse managers have a personal incentive to limit risk exposure by lowering the volatility of the firm's cash flow ex post. If unchecked, this incentive may lead to value-reducing overinvestment in risk-reducing technologies and projects. However, as reviewed by Clifford Smith in Chapter 18, "Managing Corporate Risk," it is widely accepted that active cash flow risk management can also lead to increased shareholder value. For example, if hedging alters the timing of taxable cash flows, there may be a net tax benefit. Hedging may also reduce expected costs of financial distress which in turn may allow the firm to capture additional benefits from leverage. Hedging opportunities (using various forms of derivatives and hybrid instruments) have increased substantially over the past decade, and their costs have decreased. As a result, today some form of hedging activity is common among large publicly traded firms. The evidence indicates that smaller firms—with greater default risk—tend to hedge a larger percentage of their exposures than larger firms. However, Smith points to several data problems that limit the power of the empirical research in this area.

I would like to thank all the contributors for their hard work and patience in seeing this Handbook to fruition. A special thank goes to the Series Editor William T. Ziemba for his enthusiasm for this project.

B. Espen Eckbo
Dartmouth College, 2008

*Chapter 10*

# PAYOUT POLICY*

AVNER KALAY

MICHAEL LEMMON

*University of Utah, David Eccles School of Business*

## Contents



* Avner Kalay is from the David Eccles School of Business, The University of Utah and Tel Aviv University and Mike Lemmon is from the David Eccles School of Business, The University of Utah. The authors acknowledge helpful comments from Espen Eckbo (the editor).

*Handbook of Empirical Corporate Finance, Volume 2*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

## Abstract

This chapter provides a survey of payout policy—the return of capital by firms to their equity investors through dividends and share repurchases. The modern study of payout policy is rooted in the irrelevance propositions developed by Nobel Laureates Merton Miller and Franco Modigliani. Payout policy is irrelevant when capital markets are perfect, when there is no asymmetric information, and when the firm's investment policy is fixed. Relaxing these assumptions leads to a role for payout policy to control agency problems and convey information to investors. Although changes in dividend policy are associated with changes in firm value, there is mixed evidence regarding tax effects and little evidence that payout decisions are driven by motives to signal true firm value to investors. The evidence does support a link between payout decisions and conflicts of interest between the firm's various claimholders. This chapter also surveys the evidence relating to share repurchases as an alternative form of payout and describes recent behavioral theories of payout policy.

**Keywords**

payout policy, dividends, stock repurchases, asymmetric information, agency problems, taxes

## 1. Introduction

Payout policy refers to the ways in which firms return capital to their equity investors. Payouts to equity investors take the form of either dividends or share repurchases. The modern study of payout policy is rooted in the irrelevance propositions developed by Nobel Laureates Merton Miller and Franco Modigliani.[1] The irrelevance propositions clearly delineate the conditions under which the method and pattern of the firm's payouts are irrelevant in the sense that the firm's payout decisions do not alter firm value. Miller and Modigliani show that payout policy is irrelevant when capital markets are perfect, when there is no asymmetric information, and when the firm's investment policy is fixed. In practice, however, it appears that payout policy follows systematic patterns and that firm value responds to changes in payout policy in predictable ways.

For example, in a classic study, Lintner (1956) surveyed the managers of 28 firms regarding their dividend policies. Based on the interviews, Lintner established several stylized facts about dividend policy. First, dividends are sticky in the sense that they do not change dollar for dollar with earnings. Specifically, managers exhibited a reluctance either to cut or to raise existing dividends unless they were confident that the new dividend level could be sustained in the future. Second, the level of dividends was tied to sustainable long-term earnings. Third, dividends were smoothed from year to year in order to move toward a long-term target payout ratio. Finally, based on the survey evidence, Lintner developed a simple partial adjustment model of dividend changes. Lintner's model was able to explain 85% of the year-to-year changes in dividends of his sample firms.

Understanding payout policy is important because firms return significant amounts of capital to shareholders in the form of dividends and share repurchases. Table 1 shows summary statistics on the payout policies of U.S. companies via dividend payments and share repurchases for each year from 1972 to 2004. As seen in the figure, the aggregate total payout (TP) has generally been between 40 and 70% of aggregate firm earnings and between 2 and 5% of the aggregate market value of equity. The figure also shows that repurchases have become a more important form of payout over time, particularly since 1983. In addition, the incidence of dividend increases and decreases is seen in the figure to have declined over time, although this decline is largely driven by the fact that the fraction of firms paying dividends has also declined over time.

In this chapter, we survey the academic literature on payout policy and offer some guidance on directions for future research. Following the study by Lintner, a large literature in finance, both theoretical and empirical, has emerged that attempts to understand these systematic patterns in payout policy. Our discussion of the literature is organized around the assumptions underlying the irrelevance propositions of Miller and Modigliani, and around what effect relaxing these various assumptions might have on the firm's payout choices. Because of the scope of this task and limitations on space, our review will undoubtedly be incomplete. We apologize in advance to authors whose work we do not cite.

[1] See Miller and Modigliani (1961).

Table 1

Summary statistics on payout policy of firms, 1972–2003

The table presents summary information on payout policies of firms in the Compustat database.

| Year | Number of Firms | Dividends as a % of Earnings | Repurchases as a % of Earnings | Total Payout as a % of Earnings | Dividends as as a % of Market Value | Repurchases as a % of Market Value | Total Payout as a % of Market Value | % of Firms with Dividends > 0 | % of Firms Decreasing Dividends | % of Firms Increasing Dividends |
|---|---|---|---|---|---|---|---|---|---|---|
| 1972 | 2794 | 43.0% | 3.5% | 46.5% | 2.2% | 0.2% | 2.4% | 58.84% | 15.50% | 22.87% |
| 1973 | 3000 | 35.3% | 5.5% | 40.8% | 3.0% | 0.5% | 3.5% | 60.60% | 9.17% | 63.15% |
| 1974 | 3096 | 34.9% | 2.3% | 37.2% | 4.8% | 0.3% | 5.1% | 63.15% | 8.85% | 63.40% |
| 1975 | 3292 | 39.3% | 1.3% | 40.6% | 3.6% | 0.1% | 3.7% | 63.40% | 13.79% | 37.30% |
| 1976 | 3329 | 35.9% | 2.1% | 38.0% | 3.4% | 0.2% | 3.6% | 66.84% | 12.86% | 43.29% |
| 1977 | 3283 | 38.4% | 4.0% | 42.4% | 4.4% | 0.4% | 4.8% | 69.54% | 10.47% | 50.17% |
| 1978 | 3318 | 35.7% | 4.1% | 39.7% | 4.4% | 0.5% | 4.9% | 68.99% | 12.31% | 47.43% |
| 1979 | 3495 | 31.9% | 4.1% | 36.1% | 4.2% | 0.5% | 4.8% | 63.75% | 11.94% | 42.78% |
| 1980 | 3557 | 34.6% | 4.7% | 39.3% | 3.5% | 0.5% | 3.9% | 60.28% | 15.21% | 35.90% |
| 1981 | 4090 | 36.6% | 5.2% | 41.8% | 4.0% | 0.6% | 4.6% | 51.20% | 15.90% | 25.05% |
| 1982 | 4118 | 50.2% | 11.2% | 61.4% | 3.7% | 0.8% | 4.5% | 47.16% | 13.79% | 22.92% |
| 1983 | 4354 | 44.0% | 8.5% | 52.6% | 3.1% | 0.6% | 3.6% | 42.28% | 16.56% | 15.20% |
| 1984 | 4437 | 37.9% | 22.4% | 60.3% | 3.4% | 2.0% | 5.4% | 39.78% | 8.93% | 20.90% |
| 1985 | 4354 | 48.5% | 37.4% | 85.9% | 3.1% | 2.4% | 5.4% | 38.24% | 9.94% | 18.70% |
| 1986 | 4596 | 61.6% | 35.3% | 96.9% | 2.9% | 1.7% | 4.6% | 34.12% | 10.95% | 14.41% |
| 1987 | 4772 | 47.0% | 31.3% | 78.4% | 3.1% | 2.1% | 5.2% | 31.71% | 9.44% | 15.44% |

(*Continued*)

Table 1 *(Continued)*

| Year | Number of Firms | Dividends as a % of Earnings | Repurchases as a % of Earnings | Total Payout as a % of Earnings | Dividends as as a % of Market Value | Repurchases as a % of Market Value | Total Payout as a % of Market Value | % of Firms with Dividends > 0 | % of Firms Decreasing Dividends | % of Firms Increasing Dividends |
|---|---|---|---|---|---|---|---|---|---|---|
| 1988 | 4586 | 47.8% | 26.5% | 74.3% | 3.5% | 1.9% | 5.5% | 31.86% | 6.17% | 19.76% |
| 1989 | 4407 | 45.2% | 27.7% | 72.9% | 2.8% | 1.7% | 4.5% | 32.77% | 7.29% | 19.86% |
| 1990 | 4385 | 51.5% | 21.8% | 73.3% | 3.1% | 1.3% | 4.4% | 32.16% | 7.62% | 18.55% |
| 1991 | 4458 | 68.0% | 16.4% | 84.4% | 2.4% | 0.6% | 3.0% | 31.16% | 8.87% | 14.35% |
| 1992 | 4680 | 67.4% | 19.1% | 86.5% | 2.4% | 0.7% | 3.0% | 30.77% | 9.09% | 13.22% |
| 1993 | 5011 | 59.0% | 17.2% | 76.2% | 2.1% | 0.6% | 2.7% | 29.44% | 7.73% | 13.55% |
| 1994 | 5328 | 43.1% | 14.9% | 58.0% | 2.1% | 0.7% | 2.9% | 28.10% | 5.62% | 14.25% |
| 1995 | 5936 | 43.7% | 20.9% | 64.6% | 2.0% | 0.9% | 2.9% | 26.45% | 4.90% | 14.39% |
| 1996 | 6135 | 40.1% | 24.7% | 64.8% | 1.8% | 1.1% | 2.8% | 26.00% | 5.78% | 12.75% |
| 1997 | 6193 | 38.0% | 33.8% | 71.8% | 1.4% | 1.2% | 2.6% | 25.22% | 6.54% | 11.53% |
| 1998 | 6002 | 50.2% | 48.1% | 98.2% | 1.4% | 1.3% | 2.7% | 24.43% | 6.63% | 11.47% |
| 1999 | 5955 | 39.3% | 40.5% | 79.8% | 0.9% | 0.9% | 1.8% | 22.15% | 6.20% | 11.24% |
| 2000 | 6052 | 36.8% | 41.9% | 78.7% | 1.0% | 1.2% | 2.2% | 20.42% | 6.63% | 9.63% |
| 2001 | 5770 | −284.2% | −263.7% | −547.9% | 1.3% | 1.2% | 2.5% | 19.79% | 6.51% | 8.86% |
| 2002 | 5438 | 169.1% | 156.1% | 325.2% | 1.6% | 1.5% | 3.1% | 19.24% | 5.78% | 9.02% |
| 2003 | 3824 | 36.6% | 30.9% | 67.4% | 1.5% | 1.2% | 2.7% | 23.77% | 3.09% | 15.24% |

The remainder of the chapter presents a review of the Miller and Modigliani arguments regarding the irrelevance of payout policy; a summary of the literature on the interaction between both corporate and personal taxes and the firm's payout choices; a discussion of how conflicts of interest and agency problems among the firm's various claimants affect payout choices; an examination of the role of asymmetric information in determining the firm's payout decisions; a review of the literature on share repurchase; a study of some alternative theories and new stylized facts regarding payout policy; and a summary of the state of knowledge on payout policy.

## 2. The Miller and Modigliani irrelevance propositions

Miller and Modigliani (1961) show that in perfect and complete capital markets, payout policy is irrelevant to firm value. Their basic thesis is that investment policy determines firm value and that payout is simply the residual between earnings and investment. Payout policy is irrelevant from the investor's perspective because any desired temporal pattern of payments can be replicated by appropriate purchases and sales of equity. Because investors can create "homemade" dividends, they will not pay a premium for a firm with a particular dividend policy.

In perfect capital markets, the following conditions are assumed to hold:

1. Information is costless and equally available to everyone.
2. There are no taxes.
3. There are no transactions costs associated with purchasing or selling securities.
4. There are no contracting or agency costs.
5. No investor or firm individually can influence the price of securities.

Given the perfect capital markets assumptions noted earlier and the assumption that the firm's investment policy is fixed, it is relatively straightforward to show that dividend policy does not affect firm value.

### 2.1. *Dividend policy irrelevance*

Assume that a firm financed completely by equity is established at time $t = 0$. The value of the all-equity firm is the present value of future dividends received by the investors, given by

$$S_0 = \sum_{t=0}^{\infty} \frac{E_0\left[D_t\right]}{(1+r)^t} \tag{1}$$

where $S_0$ is the stock price at time $t = 0$, $E_0[D_t]$ is the expected value of the dividend to be paid at time $t$ conditional on information available at $t = 0$, and $r$ is the risk-adjusted rate of return that investors require to hold the stock.

The sources and uses of funds identity dictate that in each period

$$CF_t + F_t = D_t + I_t + (1+r)F_{t-1} \tag{2}$$

where $CF$ is the firm's operating cash flows, $F_t$ is new financing raised at time $t$, $D_t$ is the dividend paid, $I_t$ is investment, and $(1+r)F_{t-1}$ is repayment of financing raised at time $t-1$.

Solving the sources and uses identity in Equation (2) for dividend payments and substituting the result in Equation (1), we can rewrite the value of the firm as

$$S_0 = \sum_{t=0}^{\infty} \frac{E_0(CF_t - I_t)}{(1+r)^t} \tag{3}$$

Note that dividend payments do not appear in Equation (3). The value of the firm depends only on the residual of operating cash flows net of investment. This "free cash flow" is available to be paid out as a dividend. If investment needs exceed current cash flows, then the firm must sell additional securities. Because both cash flows and investment outlays are not a function of dividend policy, dividend policy is irrelevant to firm value.

Paying out a dividend that exceeds the difference between current cash flow and investment does not increase owners' wealth; instead, it requires the firm to sell additional securities to fund the optimal investment plan. Because any new financing is done on fair terms (i.e., new financing is zero net present value [NPV]), an increase in today's dividend by a dollar requires the firm to raise additional financing worth a dollar in present value. Thus, dividend policy is irrelevant to the value of the firm under the perfect capital market assumptions used by Miller and Modigliani.

The Miller and Modigliani arguments clearly delineate the conditions under which dividend policy is irrelevant to firm value. If dividend policy is to have an effect on shareholder wealth, then it must be that one or more of the perfect capital markets assumptions are violated. The remainder of this chapter examines the implications of relaxing the various assumptions underlying the Miller and Modigliani irrelevance propositions in order to study the ways in which dividend policy can affect firm value.

## 3. Dividends and taxes

In the United States and many other countries, dividend income is taxed at a higher rate than is capital gains. Assuming that investors act rationally, the preferential tax treatment of capital gains should have significant effects on the corporate and personal dividend decisions. Yet, as detailed in this section, even after several decades of research, many questions remain unanswered. Our theories tell us that taxes should matter, but the empirical evidence is still difficult to interpret.

For most individuals, capital gains are not taxed until they are realized, and the tax rate applied to realized long-term capital gains of individuals has generally been lower than the tax rate applied to dividend income.[2] Consequently, by choosing when and what securities to trade, investors can affect the timing and amount of their tax payments. Rational investors can, for example, liquidate mostly losing parts of their

[2] The definition of "long term" for tax purposes has varied over the years between six months and a year. Both short-term capital gains and dividends are taxed as ordinary income.

portfolio, indefinitely deferring the payment of taxes on their capital gains.[3] The savings associated with postponing the payment of taxes can substantially reduce the effective tax rate. For example, deferring tax payments for 20 years when the appropriate annual discount rate is 10% reduces the effective tax rate by 85%. Investors can defer the realization of capital gains while keeping their preferred consumption path. They can do it by borrowing against their portfolio to finance current consumption. Alternatively, they can fund consumption by liquidating losing parts of their portfolio. Finally, investors can finance their current consumption by taking opposite (short and long) positions in similar financial instruments realizing only the losing component of the package.[4]

In the presence of preferential tax treatment of capital gains, rational investors should have a tax-related dividend aversion.[5] Other things being equal, investors should prefer low-dividend yield stocks.[6] In equilibrium, dividend aversion results in larger pretax risk-adjusted returns for stocks with larger dividend yields. Tests of this hypothesis—a tax-induced positive correlation between dividend yield and risk-adjusted returns—can be divided into two groups. The first set of tests examines the relationship between dividend yield and risk-adjusted return within a static equilibrium model (most notably Brennan, 1970). The second set examines the dynamic behavior of stock prices around the ex-dividend period.

Our review and analysis starts with the first set of tests. We survey the conflicting empirical evidence of these tests and then relate it to the literature on the ex-dividend period. We show that combining these two strands of research helps resolve the apparent inconsistent empirical results obtained by Black and Scholes (1974) on the one hand and by Litzenberger and Ramaswamy (1979) on the other.

### 3.1. Tests of the Brennan model

Brennan's (1970) capital asset pricing model (CAPM) states that a security's pretax excess return is linearly and positively related to its systematic risk and dividend yield. Formally,

$$E(r_{it} - r_{ft}) = a_1 + a_2\beta_{it} + a_3(d_{it} - r_{ft}) \tag{4}$$

[3] Constantinides (1983, 1984) modeled this feature of the tax code and called it the tax timing option. Financial theory tells us that investors should be willing to pay for this option. The market value of this option captures the tax advantage of the long-term capital gains associated with the option to choose when to realize these gains.

[4] The IRS imposes some limitations on such strategies. For a strategy to be feasible, the financial instruments should be sufficiently different that the strategy involves business risk. Buying long and selling short IBM, for example, is not a feasible strategy.

[5] Miller and Scholes (1978) suggest a scheme whereby investors can convert dividend income to tax-deferred capital gains. If it can be done costlessly, investors should not have a dividend aversion. However, the scheme is costly, and the evidence indicates that investors hardly use it.

[6] Faced with investors' dividend aversion, corporations should avoid paying dividends to the extent possible. Why then do companies continue to pay dividends? The next section presents possible motivations for corporate dividend payments.

where $r_{it}$ is the rate of return on stock $i$ during period $t$, $\beta_{it}$ is its systematic risk, $d_{it}$ is the dividend yield, and $r_{ft}$ is the risk-free rate of interest during period $t$. A significantly positive $a_3$ is interpreted as evidence of a tax effect. The two most influential tests of the Brennan model—Black and Scholes (1974, hereafter BS) and Litzenberger and Ramaswamy (1979, hereafter LR) present seemingly conflicting results. BS find no evidence of a tax effect, whereas LR find evidence consistent with the tax hypothesis.[7]

### *3.1.1. The Black and Scholes experiment*

To test the Brennan model, BS form portfolios of stocks using a long-run estimate of the dividend yield—the dividends paid in the preceding year divided by the end-of-year share price. They classify stocks with a high estimated dividend yield as having a high expected yield over the following year. They find no difference in pretax risk-adjusted returns across stocks with high- and low-dividend yields. They also find no difference in after-tax risk-adjusted returns as a function of the dividend yield. Based on this evidence, they advise investors to ignore dividends when forming portfolios.

### *3.1.2. The Litzenberger and Ramaswamy experiment*

In contrast to the way that BS estimate the expected dividend yield, LR estimate a short-run measure of the expected dividend yield, computed as follows. If a dividend announcement is made in month $t-1$ and the stock goes ex-dividend during month $t$, the estimate of dividend yield is simply $d_t/p_{t-1}$. In this case, the end of month $t-1$ stock price, $p_{t-1}$, contains the information associated with the dividend announcement during the month. When the announcement and the ex day occur in the same month, $t$, LR estimate the market's time $t$ expected dividend as of the end of month $t-1$ as the last dividend paid during the previous 12 months. For months in which no dividends are paid, LR assume that the expected dividend yield is zero.

LR use a three-step procedure to test for tax effects. The first step of the LR experiment is the estimation of the systematic risk of each stock for each of the test months. Formally, the following regression is estimated for each month, $t$

$$R_{ij} - R_{fj} = a_{it} + \beta_{it}(R_{mj} - R_{fj}) + \varepsilon_{ij} j = t-60, \ldots, t-1 \tag{5}$$

where $R_{mj}$ is the return on a proxy for the market portfolio, $R_{ij}$ is the rate of return on stock $i$, $R_{fj}$ is the risk-free rate of interest during period $j$, and $\varepsilon_{ij}$ is a noise term. The coefficient $\beta_{it}$ is the estimated beta for stock $i$ for month $t$.

The second step uses the estimated beta for stock $i$ during month $t$, $\beta_{it}$, and an estimate of stock $i'$s expected dividend yield for month $t$, $d_{it}$, as independent variables in the following cross-sectional regression for month $t$:

$$R_{it} - R_{ft} = a_{1t} + a_{2t}\beta_{it} + a_{3t}(d_{it} - R_{ft}) + \varepsilon_{it} \tag{6}$$

[7] Other studies include Blume (1980), Gordon and Bradford (1980), Morgan (1982), Poterba and Summers (1984), and Rosenberg and Marathe (1979).

The cross-sectional regression is estimated separately for each month during the period from 1936 through 1977, resulting in a time series of estimates of $a_{3t}$. The third step computes an estimate of $a_3$ in Equation (4) as the mean of this time series of estimates. LR find $a_3$ to be significantly positive and interpret this as evidence of a dividend tax effect.

#### *3.1.3. Litzenberger and Ramaswamy's estimate of dividend yield and potential information-induced biases*

In order to minimize the potential for information-induced biases to affect their inferences, the estimate of the expected short-term dividend yield for month $t$ uses only information available at the end of month $t-1$. Nevertheless, Miller and Scholes (1982) point out that some information-induced bias can still remain. The LR experiment uses the Center for Research in Security Prices (CRSP) tapes, which do not report announcements of dividend omissions. A dividend omission, when contrasted with a positive expected dividend, is equivalent to an announcement of a drastic dividend reduction to which the market responds negatively. By ignoring omissions, LR's experiment erroneously assumes that the months corresponding to dividend omissions have zero expected dividends. Consequently, the experiment relates the resulting negative excess return to a zero expected dividend yield. Classifying months with dividend omissions as zero expected dividend months can result in a positive cross-sectional relationship between LR's estimate of expected dividend yield and measured stock returns.[8]

Kalay and Michaely (2000) investigate the potential information-induced biases by performing a modified LR experiment using weekly returns. They limit the sample to cases in which the announcement week precedes the ex-dividend week (96.6% of the sample), excluding weeks containing announcements of dividend omissions. The modified experiment results in a significantly positive dividend yield coefficient. Interestingly, the point-estimate of this coefficient is almost identical to the one reported by LR (obtained using monthly returns). Based on this result, they conclude that the positive dividend yield coefficient documented in the LR experiment is not driven by information-induced biases. At this juncture, it seems that the two major tests of the Brennan (1970) model (LR and BS) lead to conflicting results. Later, we will present additional analysis and a possible resolution of this conflict. Before we do so, we examine the other set of tests; namely, the ex-dividend day studies.

### *3.2. The ex-dividend day studies*

Studying the ex-dividend period enables a direct comparison of the market valuation of a dollar paid in dividends to the valuation of a dollar of realized capital gains. There are

[8] LR (1982) address this potential problem by constructing alternative measures of expected short-term dividend yields that are based only on current and past information. These experiments also result in statistically significant and positive dividend yield coefficients.

three important dates in every dividend period: the announcement day, the ex-dividend day, and the payment day. On the announcement day, the firm declares the dividend per share to be paid on the payment date to its stockholders of record at the closing of trade on the last cum-dividend day. The announcement day precedes the ex-dividend day by about two weeks and the payment day by about four weeks. A stock purchased on the last cum- (with) dividend day includes a claim to the dividend declared (to be paid two weeks later), while a stock purchased on the ex-dividend day does not. The ex-dividend price should therefore be lower to reflect the lost dividend.

#### *3.2.1. The ex-dividend day studies—the theory*

The theoretical analysis of stock price behavior around the ex-dividend day compares the expected price drop to the dividend per share.[9] In perfect capital markets, assuming complete certainty, the stock price drop should equal the dividend per share. Any other stock price behavior provides potential arbitrage opportunities. A smaller (larger) price drop provides arbitrage profits by buying (selling short) on the cum-dividend day and selling (covering) on the ex-dividend day. A similar analysis can be conducted in the presence of uncertainty if we assume that any excessive ex-dividend period risk is not priced. This is the case if the risk is diversifiable and/or investors are risk-neutral.[10] We will continue our analysis assuming that the ex-dividend period required rate of return is not different from that of any other day.

Elton and Gruber (1970) model the conditions for no profit opportunities around the ex-dividend day in the presence of differential taxation of realized capital gains and dividend income. Denote the realized long-term capital gains tax rate as $t_g < t_d$, where $t_d$ is the tax rate on dividend income. Let $D$ be the dividend per share, $P_b$ the last cum-dividend stock price, and $E(P_a)$ the expected ex-dividend stock price. Equating the after-tax returns from these two sources of income results in

$$(1 - t_g)(P_b - E[P_a]) = (1 - t_d)D \tag{7}$$

[9] The earlier papers on this issue are Campbell and Beranek (1955) and Barker (1959).

[10] The Introduction of uncertainty requires some modifications. Market participants can form an estimate of the expected ex-day price drop based on past realizations. In general, financial economists expect these estimates to be unbiased. Nevertheless, taking a position (long or short) to exploit profit opportunities around the ex-dividend day involves risk. Thus, a difference between the expected ex-day price drop and the dividend per share can provide profits but not arbitrage opportunities. Indeed, empirically, the ex-dividend period is a time of excessive volatility (see Lakonishok and Vermaelen, 1986). The possible effects of risk on ex-day trading are pointed out in Kalay (1984) and modeled by Heath and Jarrow (1988) and Michaely and Vila (1995). With unusually large and priced risk, the ex-day price drop should be smaller than the dividend per share, giving the stockholders a larger required rate of return. One has to remember, however, that there are several thousand ex-dividend events in a given calendar year. The risk associated with these events should be at least temporally independent, thereby presenting substantial diversification possibilities. Investors can also hedge part of the risk by using options. Given the risk-reduction technology and the short time interval between closing on cum day and opening on ex, one can still expect the ex-dividend price drop to be "almost equal" to the dividend per share.

and

$$\frac{P_b - E[P_a]}{D}\frac{1 - t_d}{1 - t_g} \tag{8}$$

A larger tax rate on dividend income (i.e., $t_d > t_g$) results in an ex-dividend price drop smaller than the dividend per share. In such an economy, one can infer the tax rates from the ex-day relative price drop.[11]

Elton and Gruber present empirical evidence documenting an ex-dividend price drop smaller than the respective dividend per share. This evidence seems consistent with the hypothesis that investors have a tax-induced preference for capital gains. The tax code, however, is a bit more complex. Short-term capital gains are taxed as ordinary income. Thus, as Kalay (1982a) points out, short-term traders can profit from a difference between the drop in the ex-dividend day stock price and the respective dividend per share. For example, assume the cum-dividend stock price is $50, the dividend per share is $2, and the expected ex-day price drop is 70% of the dividend per share—$1.4. A short-term investor can buy the stock cum-dividend and sell it on the ex day. She would have a capital loss of $1.4 but would gain $2 of cash dividends, netting a before-tax gain of 60 cents per share. This corresponds to a before-tax daily percentage excess return of 1.2%, corresponding to an annual excess return of 1,873% (assuming 250 trading days per year).

Kalay (1982a) argues that, without transaction costs, elimination of profit opportunities implies an expected ex-dividend price drop equal to the dividend per share.[12] Although there are limitations on the amount of short-term capital losses individuals can write off to offset dividend income (about $3000 a year), dealers are not subject to these restrictions. Hence, in the absence of transaction costs, short-term traders are expected to trade as long as there is a difference between the expected ex-day stock price drop and the dividend per share.

Nevertheless, we should not necessarily observe equality between the expected ex-day stock price drop and the dividend per share. Consider the role of corporations investing in other corporations. As stockholders, corporations are taxed only on 30% (up from 15%) of the cash dividends they receive, while realized capital gains are taxed at the corporate income tax rate. Thus, corporations have a preference for cash dividends. For a corporation, equality between the expected ex-day price drop and the dividend per share provides profit opportunities. Assume in the preceding example that the stock price drops by $2 on the ex day. The firm pays tax only on 30% of the $2

[11] Typically, the long-term capital gains were taxed at a fraction of the respective individual income tax. Hence, the marginal relative and absolute tax rates can be calculated.

[12] This statement assumes that the required rate of return during the cum-ex period is arbitrarily close to zero and thus can be ignored. Also ignored are the trivial effects of the delayed payments of the dividends. The actual payments are received in about two weeks following the ex day; thus, the realized capital losses should be compared to the present value of the dividends. In this case, the appropriate discount rate is the risk-free rate since we know of no default on a promised dividend. The potential effects of such modification are indeed trivial. A $2 dividend, for example, has a present value of 1.998$.

dividends it receives, while it can deduct the full $2 capital loss. If the corporate tax rate is 34%, the per-share after-tax dividend the corporation receives is $1.796, while its per-share after-tax capital losses are only $1.32. The net gain is therefore 47.6 cents, or a before-tax return on investment of about 1% per day—equivalent to a 1100% annual return![13]

What relationship between the dividend per share and the expected price drop would amount to no profit opportunities around the ex-dividend day for a corporate investor? The ex-day price drop should exceed the dividend per share. In our example, a $2 dividend should correspond to a $2.721 expected stock price drop. Yet recall that such ex-day stock price behavior provides profits to short-term traders. So is there an ex-dividend stock price drop that will provide no profit opportunities to all the traders? Interestingly, in the absence of transaction costs, the answer is no. The differential tax treatment of major economic players creates a large variety of relative valuations of dividends and capital gains. Any market-relative valuation of these cash flows results in profit opportunities to some groups.[14] One can say that the ex-dividend period provides unavoidable profit opportunities.

Transaction costs enable the existence of an ex-dividend day equilibrium. Transaction costs allow for a variety of relationships between the expected stock price drop and the dividend per share. The only requirement is that the profits from the different relative valuations are smaller than the cost of a round-trip transaction for the inframarginal traders. The relative ex-dividend price drop can be anywhere within the bounds that provide no profit opportunities to all traders. Consequently, as Kalay (1982a) points out, one cannot infer the marginal tax rates of the marginal investors from the relative ex-day price drop.

### *3.2.2. The ex-dividend day studies—the evidence*

The existing empirical evidence documents a stock price drop that is significantly smaller than the dividend per share—an unusually large cum-dividend rate of return (see Campbell and Beranek (1955), Durand and May (1960), Elton and Gruber (1970), Kalay (1982a), Lakonishock and Vermaelen (1983), and Eades, Hess, and Kim (1984), among others). Although theory does not predict a specific relationship between the relative ex-day price drop and the preferential long-term capital gains tax rate, this evidence is consistent with the hypothesis that a dollar of capital gains is worth more than a dollar of dividends. However, further investigations of this behavior cast serious doubt on this explanation.

Indeed, the more empirical evidence on stock price ex-day behavior we obtain, the harder it is to interpret. Consider the evidence presented in Eades, Hess, and Kim (1984;

[13] This common corporate practice is termed *dividend capture*. Dividend capture programs became so widespread in the United States that the government imposed limitations on their use. As of 1984, a corporation has to own the stock for at least 45 days to qualify for the 70% exclusion on its dividend income.
[14] See Dammon and Green (1987) and Dybvig and Ross (1986).

hereafter EHK). They find positive excess returns before and including the ex-dividend day and abnormally negative returns following it. In fact, the abnormal returns on the ex-dividend day are smaller than the excess return on the last cum-dividend day.

Based on Table 2 in EHK, we see that the cumulative excess return from day −5 to (and including) day 0 (the ex day) is 0.43%, while the ex-day excess return is only 0.142%. The cumulative negative excess return from day 1 to day 5 following the ex day is −0.24%. How can theory explain the relatively large and systematic price changes before and after the ex day? What is the reason for cumulative excess returns of 0.288% during the five days prior to the ex day? Why are professional investors with trivial transaction costs unable to time their trades to exploit this phenomenon?

To help quantify the relationship between the excess return and the relative ex-day price drop, note that the mean quarterly dividend yield during EHK's sample period

Table 2

Test of the null hypothesis of zero excess returns for the ex-dividend period with a sample of all taxable distributions by N.Y.S.E. common stocks. Average daily excess and standardized daily excess returns of equally weighed ex-day portfolios for each day in the ex-dividend period for the period July 2, 1962, to December 31, 1980. The number of ex-dividend day portfolios is 4.471; the number of trading days is 4,640; and the average number of stocks in each ex-day portfolio is 18.6.

| Trading day relative to ex-day | Average percent excess return[a] | Average standardized excess return[b] | $t$-statistic | Significance level | Posterior odds ratios[c] Uniform | Normal |
|---|---|---|---|---|---|---|
| −5 | 0.067 | 0.0631 | 4.218 | $<10^{-4}$ | 0.0073 | 0.0005 |
| −4 | 0.046 | 0.0621 | 4.155 | $<10^{-4}$ | 0.0095 | 0.0006 |
| −3 | 0.061 | 0.0832 | 5.561 | $<10^{-4}$ | $<10^{-4}$ | $<10^{-4}$ |
| −2 | 0.066 | 0.0892 | 5.968 | $<10^{-4}$ | $<10^{-4}$ | $<10^{-4}$ |
| −1 | 0.188 | 0.2340 | 15.647 | $<10^{-4}$ | $<10^{-4}$ | $<10^{-4}$ |
| Ex-day | 0.142 | 0.1756 | 11.741 | $<10^{-4}$ | $<10^{-4}$ | $<10^{-4}$ |
| +1 | −0.053 | −0.0651 | −4.355 | $<10^{-4}$ | 0.0041 | 0.0003 |
| +2 | −0.058 | −0.0734 | −4.911 | $<10^{-4}$ | 0.0003 | $<10^{-4}$ |
| +3 | −0.036 | −0.0405 | −2.707 | 0.0068 | 1.366 | 0.0824 |
| +4 | −0.046 | −0.0627 | −4.195 | $<10^{-4}$ | 0.0080 | 0.0005 |
| +5 | −0.043 | −0.0553 | −3.700 | 0.0002 | 0.0569 | 0.0037 |

[a]Excess return equals the difference between the ex-day portfolio return day $t$ and $\overline{RP_t}$ (the mean portfolio return for day $t$ estimated during the 60 day period surrounding the ex-day).

[b]Standardized excess return equals the excess return for the ex-day portfolio divided by the ex-day portfolio standard deviation estimated during the 60 day period surrounding the ex-day (30 days on each side of the ex-day).

[c]Both cases assume that the null hypothesis of no tax premium is true with probability 0.5. The prior beliefs about the alternative hypotheses are represented as a 0.5 probability that (1) the mean ex-day *SER* is between −1 and +1 with uniform probability, and (2) the mean *SER* is distributed as normal with a mean of zero and a standard deviation of 0.316.

is about 1%. A relative ex-day price drop of 0.85 corresponds to an excess return of 0.15% for the average stock. A strategy of owning the stock from −5 to 0 gives an excess return of 0.43% (equivalent to a relative price drop of 0.57 for the average stock). A portfolio of stocks held between days −5 to +5 gives an excess return of 0.24% (equivalent to a relative price drop of 0.76 for the average stock). This evidence indicates that indeed, as the theory suggests, there are profit opportunities around the ex-dividend day.

EHK's additional tests are even more puzzling. Table 3 of their paper describes the behavior of stock prices around the ex day of stock dividends and of nontaxable cash dividends. Surprisingly, a similar pattern of stock returns emerges. A strategy of buying stocks five days before the ex-dividend day and selling them on day +5 yields excess returns of 1.061%. A strategy of buying stocks five days before the ex day of a nontaxable dividend and selling on the first day after ex yields excess returns of 0.52%. Selling short these stocks on the ex day (day 0) and covering five days later (day +5) yields almost identical returns. Taxes should be unrelated to this stock price behavior. In addition to documenting the existence of profit opportunities to a short-term trader, this evidence casts doubt on the tax-related explanation of the ex-day empirical regularities.

Michaely (1991) provides additional evidence by investigating the ex-dividend day behavior of stock prices around the 1986 tax reform. He finds no evidence of excess returns around the ex day *before* and *after* the tax reform. It seems that during the latter part of the 1980s, the ex-dividend day price drop was equal to the dividend per share. The change occurred, however, before the Tax Reform Act of 1986, thus providing no evidence of tax effects. A more detailed investigation of the time-series behavior of the ex-dividend day excess return reveals a similar puzzle. Eades, Hess, and Kim (1994) find substantial time-series variation in stock price ex-day behavior. The variation does not correspond to changes in the tax code.

Some studies have found ex-dividend day evidence that seems consistent with the tax hypothesis. Barclay (1987) describes different ex-day stock price behavior before federal income taxes were introduced to the U.S. economy. The ex-dividend stock price drop appears to equal the dividend per share in this time period. One must remember, however, that before 1910 the New York Stock Exchange (NYSE) was a far less liquid market. In such a market, the mechanical reduction (which is equal to the dividend per share) in the ex-day opening stock price can result in such a finding.

In summary, in the absence of transaction costs there is no ex-dividend day relative price drop that provides no profit opportunities to all traders. Consequently, in such an economy taxes cannot be inferred from the size of the ex-day price drop relative to the dividend. The existence of transaction costs enables an equilibrium where the ex-day relative price drop is within the bounds of no profit opportunities. Theory, however, cannot help us in determining the relative ex-day price drop within the bounds of no profit opportunities. Nevertheless, financial economists document excess returns during the ex-dividend period. Can this empirical regularity help explain the LR results? We turn now to integrate the theory and evidence obtained by investigating the ex-dividend period with evidence documented in tests of the Brennan (1970) model.

Table 3

Tests of the null hypothesis of zero excess returns for the ex-dividend period with a sample of non-taxable distributions by N.Y.S.E. common stocks. Average daily excess and standardized daily excess returns of equally ex-day portfolios for each day in the ex-dividend period for the period July 2, 1962 to December 31, 1980.

| Trading day relative to ex-day | Average percent excess return[a] | Average standardized excess return[b] | $t$-statistic | Significance level | Posterior odds ratios[c] | |
|---|---|---|---|---|---|---|
| | | | | | Uniform | Normal |
| Panel A | | | | | | |
| Stock dividends and stock splits. The number of ex-dividend day portfolios is 1,550; the number of trading days is 4,640; and average number of stocks in each ex-day portfolio is 1.4. | | | | | | |
| −5 | −0.016 | −0.0258 | −1.017 | 0.3092 | 18.753 | 1.1822 |
| −4 | 0.070 | 0.0159 | 0.626 | 0.5312 | 25.824 | 1.6281 |
| −3 | 0.001 | 0.0037 | 0.147 | 0.8829 | 31.081 | 1.9588 |
| −2 | 0.059 | 0.0314 | 1.238 | 0.2157 | 14.630 | 0.9222 |
| −1 | 0.194 | 0.0969 | 3.815 | $<10^{-4}$ | 0.0217 | 0.0014 |
| Ex-day | 0.387 | 0.1998 | 7.866 | $<10^{-4}$ | $<10^{-4}$ | $<10^{-4}$ |
| +1 | 0.128 | 0.0666 | 2.624 | 0.0088 | 1.010 | 0.0637 |
| +2 | 0.151 | 0.0748 | 2.947 | 0.0032 | 0.411 | 0.0259 |
| +3 | 0.112 | 0.0632 | 2.489 | 0.0128 | 1.421 | 0.0896 |
| +4 | −0.025 | 0.0058 | 0.229 | 0.8328 | 30.604 | 1.9290 |
| +5 | −0.004 | −0.0029 | −0.113 | 0.9100 | 31.209 | 1.9668 |
| Panel B | | | | | | |
| Non-taxable cash distributions. The number of ex-dividend day portfolios is 765; the number of trading days is 4,460; and the average number of stocks in each ex-day portfolio is 1.2. | | | | | | |
| −5 | 0.198 | 0.1296 | 3.585 | 0.0030 | 0.0358 | 0.0023 |
| −4 | 0.119 | 0.0374 | 1.033 | 0.3016 | 12.924 | 0.8148 |
| −3 | 0.122 | 0.1122 | 3.104 | 0.0019 | 0.1789 | 0.0113 |
| −2 | 0.042 | 0.0799 | 2.209 | 0.0271 | 1.920 | 0.1211 |
| −1 | 0.232 | 0.1633 | 4.517 | $<10^{-4}$ | 0.0008 | $<10^{-4}$ |
| Ex-day | −0.139 | −0.1417 | −3.918 | $<10^{-4}$ | 0.0102 | 0.0006 |
| +1 | −0.275 | −0.1496 | −4.137 | $<10^{-4}$ | 0.0042 | 0.0003 |
| +2 | −0.047 | −0.0653 | −1.807 | 0.0708 | 4.319 | 0.2723 |
| +3 | −0.022 | −0.0254 | −0.703 | 0.4821 | 17.242 | 1.087 |
| +4 | −0.031 | −0.0260 | −0.720 | 0.4716 | 17.040 | 1.074 |
| +5 | −0.221 | −0.1113 | −3.078 | 0.0021 | 0.1932 | 0.0122 |

[a] Excess return equals the difference between the ex-day portfolio return on day $t$ and $\overline{RP}_t$ (the mean portfolio return for day $t$ estimated during the 60 day period surrounding the ex-day).

[b] Standardized excess return equals the excess return for the ex-day portfolio divided by the ex-day portfolio standard deviation estimated during the 60 day period surrounding the ex-day (30 days on each side of the ex-day).

[c] Both cases assume that the null hypothesis of no tax premium is true with probability 0.5. The prior beliefs about the alternative hypotheses are represented as a 0.5 probability that (1) the mean ex-day *SER* is between −1 and +1 with uniform probability, and (2) the mean *SER* is distributed as normal with a mean of zero and a standard deviation of 0.316.

### 3.3. Ex-day and cross-sectional studies

In this section, we provide a reconciliation of the seemingly conflicting results documented by BS (1974) and LR (1979, 1982).[15] BS estimate cross-sectional differences of before-tax long-term returns associated with differences in the respective dividend yields. In other words, if an investor owns a stock for a year, would she earn a higher pretax return if the stock held had a higher expected dividend yield? In contrast, the LR experiment documents, for a given stock, time-series differences in pretax rates of returns earned during the ex-dividend period compared to those received during other periods. Because of these differences in the experimental designs, the results of these two types of studies can differ. Investors can receive higher pretax returns during ex-dividend periods even if the stocks' annual returns are not related to their respective dividend yields. More importantly, however, we also argue that the evidence presented in both the LR and BS studies is inconsistent with the tax hypothesis.

As detailed above, Brennan's (1970) capital asset pricing model states that stocks with higher dividend yields should offer larger risk-adjusted pretax returns throughout the year. In contrast, the LR test of the Brennan model is inadvertently designed to discover whether the ex-dividend period offers unusually large risk-adjusted returns (we refer to this seasonal effect as time-series return variation). Time-series return variation per se is not evidence of a tax effect. As detailed later in this chapter, it seems nearly impossible to provide a tax-based explanation for time-series return variation in an economy that shows no cross-sectional return variation.[16]

#### 3.3.1. Tax effects and time-series return variation

If investors could avoid the dividend tax penalty during non-ex-day periods, they would require a tax premium only during ex periods. This would create time-series return variation. Under U.S. tax law, however, investors attempting to own the stock only during non-ex periods must realize short-term capital gains, which are taxed as ordinary (dividend) income.

To illustrate, consider an investor who is attempting to own stock XYZ without receiving its dividends. Suppose the stock pays quarterly dividends, with the ex-dividend days being the last business day of March, June, September, and December. A possible strategy

[15] This section is based exclusively on Kalay and Michaely (2000).

[16] With unlimited short selling possibilities, one can suggest a tax arbitrage in the multiperiod version of Brennan. Sell a well-diversified low-yield portfolio short and buy a well-diversified high-yield portfolio. Liquidate the positions within six months. In this case, both capital gains and dividend income are treated as ordinary income. The difference between the returns on the high-yield and low-yield portfolios constitutes a profit opportunity. If such trading is allowed, there is no equilibrium. Short-term traders are at equilibrium only when no expected risk-adjusted return differential between high- and low-yield portfolios exists. But, in such a case, the long-term investor benefits from a shift to low-yield stocks. To reach equilibrium, restrictions on the economy must be imposed. The restrictions can be no (or limited) short sales, wealth limitations, and less than perfect diversification possibilities. With such restrictions there is an equilibrium in which the risk-adjusted pretax return is correlated with the dividend yield.

involves buying the stock, say, on January 1 and selling it cum-dividend after the next dividend announcement, thereby realizing only capital gains. On April 1, our investor can buy the stock ex-dividend, keep it until the end of June, and so on. The dividends are paid to the investor's trading partners. But because the attempt to avoid dividend income involves realization of short-term capital gains, the investor pays the same taxes he or she would pay on dividend income and therefore requires an identical tax premium. Thus, even though a long-term investor could prefer capital gains to dividend income, he or she does not require a larger pretax return during the ex-dividend period only.[17,18]

The economic incentives of the long-term investor should not lead to excess returns during the ex-dividend period. If many long-term investors prefer to sell the stock before the last ex-dividend day, the cum-dividend stock price could be depressed, creating larger returns during the ex period.[19] But if these investors time their trades to economize on the taxes they could pay on the last dividend, they will surely require compensation for the dividends distributed during their holding periods. Thus, if we observe tax-based price pressure that results in excess returns during the ex period, we ought to observe a "tax premium" for stocks with higher dividend yields. It is difficult to provide a tax-based explanation for time-series return variation, but almost *impossible* to explain time-series return variation in an economy that shows no cross-sectional return variation.

### *3.3.2. The Litzenberger and Ramaswamy experiment—time-series or cross-sectional return variation*

Based on the preceding discussion, it is important to determine whether LR's documented dividend effect is evidence of time-series or cross-sectional return variation. By its very nature, the LR experiment is likely to uncover time-series return variation and is inefficient in detecting cross-sectional return variation. The LR experiment defines

[17] One can argue that a constant tax premium per unit of time is the preferred compensation. A different premium structure can force the long-term investor (hereafter LTI) to own the stock longer or sell it sooner than his consumption investment decisions dictate. Also, note that the LTI is almost indifferent as to the timing of his *purchase* around the ex day. With quarterly dividends, the investor has to own the stock for at least two ex periods to qualify as an LTI. He can avoid the third ex period just as easily if he buys the stock before the current ex. Finally, Constantinides (1983, 1984) points out that investors have incentives to realize short-term losses and to defer capital gains for as long as they can. Therefore, the long-term buyers and the short-term traders constitute, almost by definition, a larger fraction of the market than the long-term sellers. One would expect them to offset any temporary price pressure resulting from the population of long-term sellers.

[18] The tax-related considerations of short-term traders and tax-exempt institutions cannot lead to time-series return variations. Corporations, on the other hand, have a tax-related preference for cash dividends over short-term capital gains. Seventy percent of the dividends they receive are tax-exempt. Corporations are willing to *pay a tax-related premium* to own the stock during an ex-dividend month.

[19] The empirical evidence is inconsistent with this conjecture. A positive excess return prior to the ex day has been documented by Eades, Hess, and Kim (1984) and Lakonishok and Vermaelen (1986).

a stock as having a positive dividend yield only during its ex-dividend period. Hence, firms that pay quarterly dividends are classified as offering a zero dividend in two-thirds of the months. This experimental design makes it difficult to relate the dividend yield coefficient to taxes.

Consider the following possibility. For reasons that may not be unrelated to taxes, a stock's expected rate of return is higher during the ex-dividend month than in other months. Suppose this difference in returns is unrelated to the dividend yield. An example is presented in Figure 1. All of the stocks in the economy are assumed to have the same risk-adjusted expected returns, but the expected returns in the ex-dividend month are assumed to be higher.

As is evident from Figure 1, an LR experiment performed on this data would result in a positive dividend yield coefficient. In this case, however, the yield coefficient would indicate only that stock returns show seasonal variation. Interestingly, as demonstrated in Figure 2, an LR clientele test (as in their 1980 study) in our assumed economy leads to evidence consistent with a tax-induced clientele effect. Divide the stocks into five subsamples based on their expected dividend yield, as in LR. Group 1 contains the lowest-yield stocks, and group 5 the highest. An LR experiment in our economy would result in a smaller dividend yield coefficient for the higher-yield groups. This result,

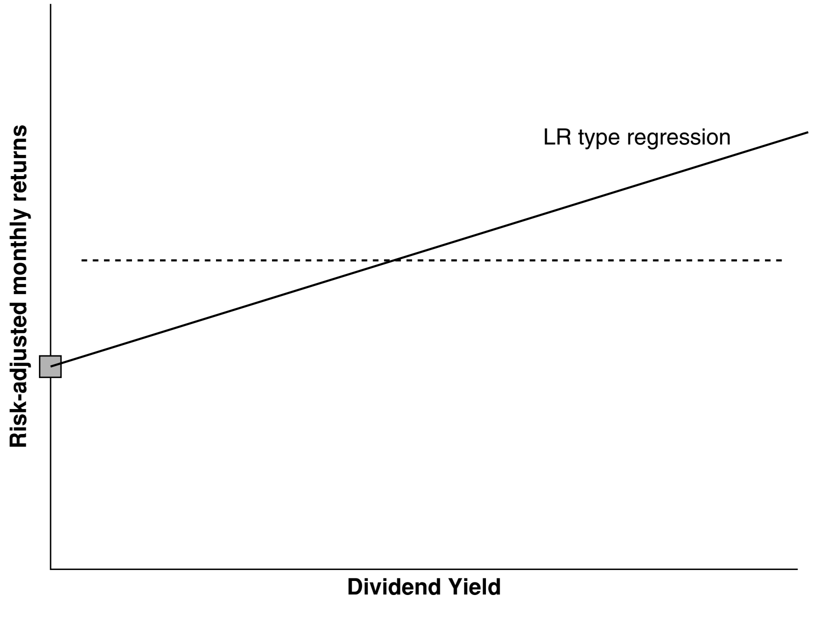



Fig. 1. Plots of risk-adjusted returns and the corresponding dividend yield along the horizontal broken line. The assumption made is that risk-adjusted returns during ex-dividend months are higher but unrelated to the respective dividend yield. The vertical axis intercept, denoted by a square, contains approximately two-thirds of the observations.

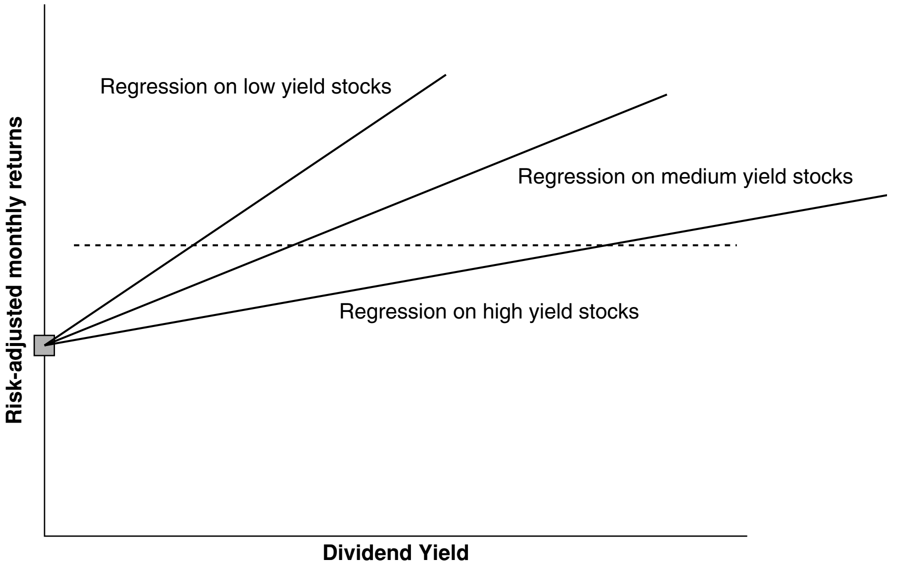



Fig. 2. Plots of risk-adjusted returns and the corresponding dividend yield along the horizontal broken line. The assumption made is that risk-adjusted returns during ex-dividend months are higher but unrelated to the respective dividend yield. The vertical axis intercept, denoted by a square, contains approximately two-thirds of the observations.

however, is not evidence of tax-induced clienteles. The explanation for this empirical regularity is quite simple. In the cross-sectional regression, the same return differential (between the ex and non-ex period) is related to a larger number for the higher-yield groups.

Although the LR experiment will uncover time-series return variation, it is not designed to find cross-sectional return variation. The example detailed in Figure 2 helps to illustrate this point. Assume a Brennan-type economy in which stocks with larger dividend yields are associated with larger pretax returns. Panel A of Figure 2 presents such an economy. Suppose that all firms in this economy pay quarterly dividends. Following the LR methodology, a dividend-paying stock is assumed to have a zero-dividend yield in two out of three months and its own yield during the ex month. As is evident in Panel B of the figure, the expected risk-adjusted return of the zero-yield category includes a "tax premium" that biases the experiment toward the null hypothesis of no tax effect.

### *3.3.3. The empirical evidence*

Kalay and Michaely (2000) replicate the LR experiment using weekly and monthly data. They document a positive and statistically significant dividend yield coefficient for both weekly and monthly data. As discussed earlier, this evidence by itself might very well indicate only time-series return variation that is unrelated to taxes. The ordinary-least-square (OLS) point estimate of the dividend yield coefficient in Kalay and Michaely's experiment is 0.246 for the weekly experiment and 0.226 for the monthly. The two estimates are almost identical. The difference between the ex-dividend week risk-adjusted returns and the returns in other weeks is similar to the difference between ex-dividend-month returns and the returns during other months. One can conclude that almost all of the excess returns occur within the ex-dividend week. This is strong evidence of time-series return variation.

To further test for cross-sectional return variation, Kalay and Michaely (2000) repeat the LR experiment using quarterly data. Expected quarterly dividends are assumed to be equal to the mean quarterly dividend yield of the previous calendar year. This is a direct test of cross-sectional return variation. The outcome is *an insignificant dividend yield coefficient*.[20]

In summary, during the ex-dividend period, stock returns are unusually high but are unrelated to the dividend yield. Thus, Black and Scholes (1974), who examine whether returns of high-dividend yield stocks are higher throughout the year, find no dividend effect. LR examine whether stocks experience higher risk-adjusted returns during the ex-dividend period and find that returns are higher during the ex-dividend period. The evidence presented in both studies is inconsistent with a tax effect.

### *3.3.4. Risk and the ex-day returns*

If taxes do not explain the higher ex-dividend period returns, what does? Changes in risk or risk premiums are one possible explanation. Indeed, the volume of trade in these stocks during this period is unusually high. One also finds a higher variance of stock returns during the ex-dividend period (see Lakonishok and Vermaelen, 1986). The incremental risk around the ex-dividend period could be priced. Namely, investors could require a larger ex period return to compensate them for the larger per-period risk around the ex day. On average, stocks experience an excess return of 0.24% during the 11 business days surrounding the ex day. The corresponding rate of return of a typical stock is about 0.54%. Therefore, the excess ex-day return constitutes an increase of 44% in the respective holding period returns. A similar increase in risk could explain these differences. To test the hypothesis that changes in risk explain the ex-day stock price behavior, we need a theory to guide us. Lacking a satisfactory theory, we are left with a conjecture.

## *3.4. The case of citizen utilities*

In 1956, Citizen Utilities created two classes of identical shares that differed only in their dividend payout. Series A shares paid cash dividends, and series B shares paid

[20] For additional evidence indicating that LR's findings are evidence of time-series return variation, see Kalay and Michaely (2000).

stock dividends. Based on a 1969 IRS ruling, the stock dividends paid on series B stocks were not subject to taxation. In his investigation of the case, Long (1978) compared the payouts and found that the firm paid consistently 8 to 10% more stock dividends than cash dividends. The relationships between these two streams of payouts were found to be extremely stable and predictable. Investors could have easily predicted that this relationship would continue in the future. Contrary to the tax hypothesis, however, Long (1978) found that, if anything, series A stocks receiving cash dividends commanded a slight premium over series B stocks. More recently, Hubbard and Michaely (1997) reexamined the relative valuation of series A and B stocks after the 1986 Tax Reform Act. They found that the reduction in the preferential tax treatment of capital gains did not change the relative valuation of the two series. Although these results are inconsistent with the tax hypothesis, Poterba (1986) has documented different ex-dividend day price behavior for the two series. He finds a smaller stock price drop than the dividends for the series that pays cash dividends and close to an equal price drop for the series that pays stock dividends. It is surprising to find ex-day evidence that, taken at face value, seems consistent with a preference for capital gains when the overall valuation indicates a preference for cash dividends. This evidence should not be taken as definitive, however, inasmuch as we are dealing with only one firm. That this firm is a utility further limits inference, for anecdotal evidence suggests that stockholders of utilities typically prefer cash dividends.

### *3.5. Recent evidence on dividends and taxes*

In two recent papers, Sialm (2005, 2006) provides some new evidence on the relationship between asset valuations, stock returns, dividends, and taxes. Rather than relying on the dividend yield as a proxy for the tax consequences of owning a security, Sialm computes a direct estimate of the effective tax burden associated with owning a security that accounts for time-series changes in tax rates on capital gains and dividend income and cross-sectional differences in the proportion of total returns that come in the form of dividends.

Sialm (2005) exploits the time-series variation in this direct measure of the tax burden to test whether expected tax burdens are capitalized into asset prices. Consistent with tax capitalization, he finds a negative relationship between asset valuations and effective tax rates after controlling for various macroeconomic factors. In addition, he also finds a positive association between taxes and aggregate asset returns.

Sialm (2006) provides new tests of the Brennan (1970) model based on the estimated tax yield of the security, rather than relying on the estimated dividend yield. The paper finds an economically and statistically significant relationship between risk-adjusted stock returns and effective personal tax rates using data on a cross section of equity securities between 1927 and 2004. Consistent with tax capitalization, stocks facing higher effective tax rates tend to compensate taxable investors by generating higher before-tax returns.

## 4. Agency relationships and dividend policy

Miller and Modigliani (1961) demonstrate that in perfect capital markets, assuming a given investment policy, the firm's market value is independent of its dividend decision. The assumption that investment policy is fixed is helpful, for it separates the effects on firm value of investment decisions from those of dividend decisions. In general, however, investment decisions and dividend decisions are often interrelated. For example, by deciding to reduce its dividend payment (holding other forms of payout constant), the firm increases its investment. Similarly, the firm can finance its payment of cash dividends by selling assets. Hence, a more realistic investigation of the dividend decision requires exploration of the effects that alternative dividend decisions have on the firm's investment policy. The motivation to investigate this issue further is enhanced by the modern view of the corporation as a complex structure of contractual arrangements among different parties. This section highlights the important role the dividend decision plays in the complicated relationships among the various parties.

### *4.1. The main claimholders of the firm*

There are different types of suppliers of capital, but all share the consequences of the economic activities of the corporation. Stockholders, bondholders, convertible bondholders, and owners of warrants supply capital and receive payouts based on the cash flows generated by the firm. Management and employees supply labor and receive payments. Suppliers of other factors of production and subcontractors also have a stake in the firm's success. Even past customers who continue to purchase maintenance services for equipment bought from the firm rely on the corporation's continued operation. These parties have conflicting interests, but two conflicts stand out as most important in the discussion of payout policy—the conflict between stockholders and bondholders and the conflict between management and stockholders.

### *4.2. Stockholder–Bondholder conflict and dividends*

Our investigation of the conflicts of interest between bondholders and stockholders assumes that management's interests are aligned with those of stockholders. Stockholders can hire and fire management, and managers often have significant equity positions in the firm. Moreover, the compensation of managers often contains a significant component that depends on the firm's stock price.[21] We therefore assume that management decisions are consistent with those the stockholders would make themselves, and we refer to managers and stockholders interchangeably in our discussion. We relax this assumption when we discuss the conflicts of interest between managers and stockholders in Section 4.4.

[21] Murphy (1999) provides a comprehensive summary of the literature examining managerial compensation.

Stockholders and bondholders share the cash flows generated by the firm.[22] Bondholders are entitled to periodic interest (coupon) payments and the payout of the face value (the promised amount) at the maturity of the debt. They are paid first, but the payoffs are limited to the promised coupons and face value. Stockholders, in contrast, receive all the remaining cash flows after the obligations to the bondholders have been met. If the value of the firm exceeds the value of the obligations to the bondholders, the stockholders will choose to pay them off and take the residual value. If, however, the value of the firm's assets is lower than the amount promised to the bondholders, then stockholders, having limited liability, can default. In this case, the bondholders have the right to take over the assets of the firm as partial compensation for their debt. There is a clear asymmetry in how the payoffs are split between the two parties. Stockholders, who receive the residual value after full payment to the bondholders, are the sole beneficiaries of the firm's upside potential; bondholders who will lose a portion of their principal in bad times bear the downside risk.

The asymmetry in the payoffs to the bondholders and the stockholders creates conflicting objective functions. The bondholders will attempt to maximize the likelihood that they will be paid in full. Hence, they will choose to minimize the downside risk and thereby increase the value of their claims. The stockholders will choose an investment policy with as large an upside potential as possible even if by doing so they increase the downside risk. Increasing both the upside potential and the downside risk would increase the market value of the equity with a corresponding reduction in the market value of the debt. The conflicting relationship between the bondholders and the stockholders is illustrated algebraically as follows:

$$B + S = F \tag{10}$$

where

$B =$ the market value of the bonds
$S =$ the market value of the stocks
$F =$ the market value of the firm

Consider an action that changes the market value of the firm by $\Delta F$

$$\Delta B + \Delta S = \Delta F \tag{11}$$

$$\Delta S = \Delta F - \Delta B \tag{12}$$

Stockholders gain from an action that leaves the market value of the firm unchanged if it reduces the market value of the debt. In such a case, stockholders' gains would equal bondholders' losses. Assuming perfect capital markets, the ex-dividend drop in firm value is equal to the dividend paid; $\Delta F = -D$.[23] In this case,

$$\Delta S = -D - \Delta B \tag{13}$$

[22] For simplicity we assume only two types of claims—straight debt and equity. The analysis is easily extended to convertible debt, warrants, and preferred stocks.

[23] In this section, cash dividends and payouts to stockholders via share repurchases are treated interchangeably.

If debt value is unchanged by the payment of the dividend, then the market value of the firm's stock falls by the full amount of the dividend, leaving stockholders unaffected by the dividend payment. However, the payment of dividends reduces the pool of assets supporting the debt, potentially increasing the risk of the bonds and reducing the value of the debt. If $\Delta B < 0$, stockholders are better off by paying a dividend.

An unexpected increase in risk will reduce the market value of the bonds. Hence, bondholders prefer smaller (or no) dividends. Stockholders have the opposite incentives. A reduction in the market value of the debt, holding other things constant, implies an increase in stockholders' wealth. Payment of dividends ensures stockholders some payoffs even if the firm eventually defaults on its obligations to the bondholders. Consequently, other things held constant, stockholders would like dividend payments to be as large as possible.

### *4.2.1. A partial solution to the conflict—dividend Constraints*

The incentives of stockholders are well known to market participants. Once debt is raised, the stockholders would choose to pay themselves more dividends than the amount that maximizes the market value of the firm. The potential purchasers of the bonds would assume a larger payout level and price the debt accordingly. Hence, a larger dividend payment does not hurt the bondholders as long as it is anticipated and bond prices reflect it. To the extent that there is a loss of value associated with the suboptimal dividend payment, it is the stockholders who bear the loss. Thus, it is in stockholders' best interests to restrict their future ability to pay dividends if they want to raise debt. Theory tells us that debt indentures should include restrictions on stockholders' ability to pay dividends.

The legal system imposes some restriction on the ability to pay. Dividends can be paid in an amount that does not exceed earned surplus. Uninformed investors are protected by imposing a limitation on stockholders' ability to empty the firm. Yet stockholders themselves offer bondholders additional protection in the debt indentures. Kalay (1982b) examines a large sample of debt indentures and finds all of them to contain covenants restricting dividend payments. As expected, the restrictions apply to all forms of payouts to the stockholders—cash dividends and share repurchases. Kalay finds two types of dividend restrictions, direct and indirect.

### *4.2.2. The direct dividend constraint*

The direct dividend constraint places a maximum on the amount of cash dividends and share repurchases at any given point in time during the life of the bond. The constraint specifies an initial reservoir of payable funds—the amount stockholders can pay out as dividends at the time of the debt issuance. To compute the funds available for dividend payments at any point in time during the life of the bond, one has to add to the initial reservoir an agreed upon fraction of the accumulated net earnings and funds raised by the sale of new stock, and subtract all the dividends paid up to that point in time. Note that if accumulated earnings are nonnegative, stockholders can pay out all the funds raised by the sale of new equity. Thus, equity-financed dividends are not restricted. If net earnings are negative, stockholders have to defer dividend payments until net earnings turn positive and cover the existing deficit. Finally, the constraint is cumulative—stockholders can forgo payments of dividends without necessarily losing their legal right to pay them in the future. The restriction applies to two types of dividends; those financed by reducing investment (or sale of assets) (hereafter investment-financed dividends) and those financed by the sale of new bonds (hereafter debt-financed dividends). The direct dividend constraint is soft in the sense that it can only stop payment of dividends in bad times, but it does not require any other action.

### *4.2.3. The indirect dividend constraint*

Indirect dividend constraints are implied by bond covenants such as stockholders' commitments to maintain a minimum level of net worth, a minimum level of working capital, or a maximum ratio of debt to assets. These restrictions are similar to the direct dividend constraint in that they are cumulative and allow the deferral of dividend payments, they are placed on the sum of dividend payments and share repurchases, and they most often apply only to debt- and investment-financed dividends. They differ from the direct dividend constraints in one important way. If violated, they force the stockholders to contribute new equity capital or give up the firm.

### *4.2.4. Stockholders pay less than they are allowed to—the reservoir of payable funds*

Based on the preceding analysis, once the set of dividend constraints is in place, we expect stockholders to pay themselves as much in dividends as they can. The payment of debt- or investment-financed dividends raises the risk of the outstanding bonds, thereby reducing their value. This reduction in debt value is a net benefit to the stockholders. Yet, Kalay (1982b) finds that stockholders do not choose to pay themselves as much dividends as they legally can. Kalay (1982b) reports the funds available for distribution to the stockholders under the most restrictive dividend constraint (defined as the reservoir of payable funds) that stockholders choose to maintain. He finds that for most firms this reservoir is positive and of nontrivial magnitude. It is, on average, 11.7% of firm value. At face value, this finding is very surprising. Stockholders forgo payments of investment- and debt-financed dividends that would substantially increase the risk of the outstanding bonds. Presumably, stockholders can get better prices for their bonds if they commit not to pay the amount maintained as reservoirs of payable funds. Pushing this logic to the extreme, stockholders can pre-commit to pay no dividends and thereby ensure that no reservoir would need to be maintained and eliminate any possible tax-related costs associated with payouts. Yet, stockholders choose a constraint that allows for payouts and pay less than they are allowed to legally.

### 4.2.5. Potential explanations

Stockholders incur costs associated with the forgone wealth transfer from bondholders by maintaining positive reservoirs. Hence, an economic rationale for their existence should point to some benefits. Kalay (1979, 1982b) suggests that one such benefit arises if the supply of investment projects with nonnegative NPV facing the corporation is limited.[24] In such a case, a total prohibition on payouts could force the firm to take negative NPV projects (i.e., to overinvest). Alternatively, the firm would thereby be forced to buy back its bonds, deviating from its optimal capital structure. Taking a project with negative NPV or maintaining suboptimal capital structure is costly. Thus, stockholders maintain reservoirs of payable funds to reduce the potential costs of overinvestment. Kalay (1979) documents evidence consistent with this hypothesis—firms with better future investment opportunities and higher leverage ratios maintain smaller reservoirs.[25]

Stockholders could also choose to maintain reservoirs of payable funds if they derive some benefits from receiving a smooth stream of payouts. The reservoir would allow the stockholders to pay similar dividends to the norm in periods of low or even negative earnings. The particular benefits stockholders derive from the smoothing of dividends are not clear, however. Investors can form portfolios with smooth dividends even if each firm provides a volatile series of payments.

### 4.2.6. Additional empirical evidence

Further examination of the role dividends play in the conflicts of interest between bondholders and stockholders involves the study of stock and bond prices around the announcement of an unexpected change in the dividends that firms pay. Other things held constant, an unexpected increase in dividends should be associated with a reduction in bond prices and an increase in stock prices. However, other things are not held constant. The empirical evidence indicates that unexpected large (small) dividends are associated with better (worse) performance. Even if current earnings are not improved (deteriorated), an increase (decrease) in dividends may reveal to the market management's expectations of future performance. If dividend increases reveal positive information about future performance, then an unexpected dividend payout could also be good news for bondholders—a prediction opposite to that from the conflicts of interest hypothesis. Handjinicolaou and Kalay (1984) examine this issue and find that both stock and bond prices decline around dividend decreases and that stock price rises and bond prices are unaffected by the announcement of dividend increases. This suggests that the information content of dividends is the stronger market force. Dhillon and Johnson (1994) reexamine the issue, focusing only on significant dividend changes. Their sample includes large

[24] Corporate investments in financial markets generally provide stockholders negative net present value because of double taxation.

[25] John and Kalay (1982) derive the optimal dividend constraint and find that stockholders have incentives to pay less investment-financed dividends than they are permitted. Their model cannot explain, however, stockholders' reluctance to pay as much debt-financed dividends as they are permitted.

dividend changes, dividend initiations, and omissions. They find that bond and stock price behavior around significant dividend changes is consistent with the agency hypothesis. Following announcements of dividend initiations, stock prices increase by an average of 0.72% and bond prices fall by an average of 0.7%. Similarly, following announcements of large dividend increases (of at least 30%), stock prices rise by 1.82% and bond prices fall by 0.5%.

Jayaraman and Shastri (1988) examine the behavior of bond and stock prices around specially designated dividends (SDDs). SDDs are less likely to convey information about the long-run prospects of the firm. A sample with limited information content seems appropriate to search for potential dividend-related wealth transfers between stockholders and bondholders. They find point estimates of average bond and stock price response to dividend announcements consistent with the agency hypothesis, yet the estimates are statistically insignificant.

### *4.3. Conflicts of interest between stockholders and other senior claimholders*

The conflicts of interest between stockholders and other senior claimholders are very similar to those they have with bondholders. Consider the case of holders of convertible bonds. Convertible bond consists of a straight bond and a call option. The holder of a convertible bond has the right to give up his or her bond (whose market value at the time of conversion is the exercise price) in exchange for a fixed number of shares of stock. By paying dividends, stockholders reduce the pool of assets supporting the payment of the straight bond and thus increase the bond's risk. The implied reduction in the market value of the straight bond is the wealth transferred to the stockholders. In addition, by paying dividends, stockholders reduce the ex-dividend price of the stock, thereby reducing the value of the call option. The decrease in the market value of the call option attached to the bond is also wealth transferred to the stockholders. Solutions to this conflict are similar to those used to ease the tension between the bondholders and the stockholders.[26]

### *4.4. Ownership versus control and the dividend decision*

The modern corporation is run by professional managers who are not the main contributors of capital. Stockholders own the firm, but professional managers exert significant control. Management makes the daily investment and financial decisions subject to oversight by the board of directors. The division of ownership and control is perhaps natural given the economies of scale associated with investment policy and the special skills required to manage a complex organization. Similar to the stockholder–bondholder conflicts described earlier, however, the incentives of management likely differ from those of the suppliers of capital.[27] The dividend decision is no exception.

[26] For example, convertible bonds typically contain antidilution clauses that limit dividend payments.

[27] In this section, we treat the suppliers of capital in all forms (stocks, bonds, convertibles, warrants, etc.) as one group that we call the stockholders.

Managers are often compensated through bonus plans that are proportional to profits. Larger corporations are likely to report bigger earnings. Holding the profitability of the firm constant, the more capital that is contributed to the corporation the bigger it is. In addition, managers' social status is typically tied to the amount of assets they control. Consequently, managers would like to retain as much earnings as they can to avoid reductions in the size of the asset base under their control. Managers are expected to be averse to paying dividends for additional reasons. Their human capital is associated with the success of the firm they run. By retaining a larger fraction of net earnings, they reduce the risk of financial distress. In larger firms, management may also be better able to hide perquisite consumption. Finally, in many firms, executives are granted stock options. To the extent that the exercise prices of these options are not adjusted downward on ex-dividend days, managers have incentives to avoid (or reduce) the payment of dividends. In summary, managers appear to have incentives to pay dividends in an amount that is lower than the dividends that would maximize stockholders' wealth. Paying less than is optimal allows managers to maximize the resources under their control, resulting in overinvestment and reducing the value of the firm's stock. It is therefore in the stockholders' best interest to mitigate these conflicts.

The board of directors is charged with representing the interests of the stockholders and appoints top management. An important part of their responsibilities is to monitor the performance of management and to ensure that the decisions taken are consistent with the maximization of stockholders' wealth. This is not an easy task. Management is privy to better information concerning the firm they run than are members of the board. The board may have a difficult time assessing to what extent management's decisions deviate from the maximization of firm value. The difficulty stems from the public good aspects of the board members' jobs and their limited skills and information. Faced with a significant measurement problem, boards attempt to design compensation packages that align the interests of management with those of the stockholders. Nevertheless, to the extent that imperfect monitoring and contracting by the board fails to eliminate completely the tensions between management and stockholders, conflicts of interest between these two groups will influence the dividend decision.[28]

### *4.4.1. Easterbrook's model*

Easterbrook (1984) conjectures that dividends play an important role in easing the tensions among the various claimholders of the firm. Dividends, according to Easterbrook (1984), are part of the solution to the agency relationships between management and stockholders. As the optimal size of the firm grows, entrepreneurs need more external funds to finance investment. The result is a larger and more diverse group of stockholders,

[28] Analyses by Zweibel (1996), Fluck (1999), and Myers (2000) examine managerial entrenchment in the context of capital structure, but their arguments can be extended to dividend payouts in a straightforward way. In these models, managers voluntarily commit to pay out cash because of the constant threat of discipline. An overview of issues in executive compensation is given by Aggarwal (2007).

most of whom own only a small fraction of the firm. A small stockholder is unlikely to monitor the actions of management. He or she incurs the full costs of monitoring but receives only a fraction of the benefits. Yet, if stockholders can organize and spread the monitoring costs proportionally, they stand to gain from improved decisions. Easterbrook (1984) argues that dividend payments serve such a role.

By paying more dividends, the firm increases the likelihood that it will have to raise external funds. Hence, paying dividends is associated with more frequent scrutiny by professionals such as lawyers, investment bankers, money managers, and public accountants. These professionals have strong incentives to scrutinize the firm and evaluate the management. These professionals can lose their reputation if they manage an unsuccessful stock or bond offering. They would be hurt by mispricing an issue, for market participants would be reluctant to buy their future security offerings. Consequently, management of dividend-paying firms is scrutinized more frequently and can extract less wealth from their stockholders. Recognizing the important role of dividends, stockholders insist that dividends be paid.

Adding to the conflicts of interest between the suppliers of capital and management, tensions between bondholders and stockholders highlight an additional role for dividends. Subject to existing constraints, the stockholders would like more dividends to be paid. As already pointed out, however, stockholders value maintaining a reservoir of payable funds to the extent that it serves to reduce the likelihood of overinvestment. But management has incentives to pay even less. By retaining more earnings, management reduces the firm's debt/equity ratio. Managers can benefit from a lower debt/equity ratio, other things held constant, because this reduces the probability of bankruptcy and thus reduces the likelihood that management will suffer a loss of reputation associated with controlling a bankrupt firm. Furthermore, managers invest a larger fraction of their wealth (including human capital) in the firm they manage than do other stockholders. Hence, reducing firm risk, even at the expense of profitability, can benefit them.

Easterbrook's (1984) analysis leads to the following empirical implications. First, closely held firms, other things held constant, should pay lower dividends. Similarly, firms having a large stockholder do not have to rely on the capital markets to monitor their managers. The large stockholder has more incentives to monitor management, and there is less need for dividends. Second, firms with lower optimal debt/equity ratios, other things constant, should pay less in dividends. Thus, based on Easterbrook's analysis, we expect a positive correlation between dividends and the debt/equity ratio.

#### *4.4.2. Jensen's model*

The starting point of Jensen's (1986) analysis (which is similar in spirit to Easterbrook) is the limited ability of stockholders to control management. Managers can choose actions that maximize their own utility but are not necessarily consistent with the maximization of shareholders' wealth. The important asset enabling management to depart from stockholders' interests is the corporation's free cash flows. Jensen defines free cash flows as cash flow in excess of that needed to fund positive NPV investment projects. The free cash flows are available to be used by management in pursuit of their own objectives. Paying out these funds reduces management's ability to shift resources away from the stockholders. Indeed, stockholders should insist that these funds be paid. Increasing leverage is another mechanism for reducing free cash flow as debt involves a commitment to periodic interest payments. Note that Jensen's theory predicts a negative relationship between dividends and the debt/equity ratio.

### *4.4.3. Empirical evidence*

Several studies have investigated the effects of the conflicts of interest between management and stockholders on the dividend decision. Lang and Litzenberger (1989) conjecture that the overinvestment problem associated with free cash flow is likely to be greater in stable, profitable companies with few growth opportunities. They examine the market response to unexpected dividend changes for firms with different investment opportunities, using Tobin's Q (market value of assets/book value of assets) as a measure of corporate investment opportunities. According to Lang and Litzenberger (1989), higher Tobin's Q implies better investment opportunities. Conversely, firms with low Q have poor investment opportunities, and market participants would want them to pay dividends. Consistent with the idea that dividend payments can limit agency problems in firms with poor investment opportunities, Lang and Litzenberger document a much stronger market response to dividend changes for firms with a Q less than one than for those with a Q greater than one. Yoon and Starks (1995) repeat the Lang and Litzenberger experiment using a longer time period. After controlling for additional factors, they find that the market reactions to dividend changes are similar for both high and low Q firms. Their results are not consistent with the free cash flow hypothesis.

Lie (2000) investigates the relation between excess funds and the market reaction to changes in payout policy and finds that firms that increase dividends or repurchase shares have excess cash relative to peer firms. He also finds that the market reaction to the announcement of special dividends is positively related to the firm's excess cash and negatively related to Tobin's Q. The results are consistent with the idea that distributing cash can limit potential overinvestment and increase shareholder wealth.

Along similar lines, DeAngelo, DeAngelo, and Stulz (2005) argue for a life-cycle theory of dividends in which the firm balances the benefits (e.g., reduced flotation costs) and costs (e.g., agency costs of free cash flow) associated with earnings retention. Under this theory, the trade-off between retention and distribution (i.e., payout of earnings) evolves over time as profits accumulate and investment opportunities decline. Consistent with this view, they find that the probability that a firm pays dividends is positively related to its mix of earned and contributed capital. They also find that the earned/contributed capital mix has a substantial effect on the likelihood of dividend initiations and omissions. Finally, they estimate that had the 25 largest long-standing dividend-paying firms in 2002 not paid dividends, their cash balances would total

\$1.8 trillion (51% of assets) compared to their actual cash balances of \$160 billion. In other words, had these firms not paid dividends, they would have huge cash balances and/or little or no leverage; dramatically increasing the potential for managers to pursue policies that benefit themselves at stockholders' expense.

Eckbo and Verma (1994) provide cross-sectional evidence on the relationship between voting rights and dividend payments based on data from Canada, where shareholders have the right to vote on the dividend payments proposed by management. They find that firms pay lower cash dividends when managers have greater voting control of the firm, which is consistent with both agency and tax arguments.

Using international data, LaPorta, Lopez-de Silanes, Shleifer, and Vishny (2000) examine two competing agency-based explanations for the payment of dividends. Under the bonding hypothesis, managers voluntarily pay dividends to commit not to expropriate shareholders. Under what they call the outcome hypothesis, shareholders are able to force managers to pay dividends in order to limit the resources under managers' control. LaPorta et al. use variation in the legal protection of shareholders across countries to examine these two hypotheses. Under the bonding hypothesis, firms in countries with weak legal protection of shareholders should be most likely to pay dividends. Under the outcome hypothesis, the opposite is true; firms in countries with strong legal protection for shareholders should be more likely to pay dividends. Their results support the outcome hypothesis over the bonding hypothesis. They document a strong positive relationship between the level of legal protection and the dividend payout, and they show that low-growth firms have higher payout ratios in countries with strong legal protection. Their findings suggest that investors use their legal power to force companies to pay dividends, particularly when growth prospects are poor. Their evidence does not support the idea that managers voluntarily pay dividends in order to reduce free cash flow problems.

Agrawal and Narayanan (1994) investigate the dividend decisions of a sample of all-equity firms as compared to the decisions of firms with debt in their capital structure. Debt involves commitment to payments of interest, and hence, other things held constant, leveraged firms have less free cash flow. Jensen's theory therefore predicts that firms with higher debt/equity ratios pay lower dividends. Indeed, Agrawal and Narayanan report that the dividend payout ratios of leveraged firms are significantly lower than those of all-equity firms. They examine the issue further by dividing their sample of all-equity firms into two groups. The first consists of firms in which managers have a significant ownership stake that aligns their interests more closely with those of the stockholders. The second contains firms in which managers have small ownership stakes. Consistent with the theory, they find that firms in the first group pay less dividends compared to firms in the second group.

Jensen, Solberg, and Zorn (1992) investigate the joint determination of dividend payouts, debt/equity ratio, and insiders' holdings. Controlling for profitability and investment opportunities, they found that dividend payouts are negatively related to the leverage ratio and insider holdings. The evidence seems to indicate that debt service obligations and significant insider ownership are substitutes for dividends in controlling the manager–stockholder conflict. Finally, Lambert, Lanen, and Larcker (1989) investigate the dividend policies of firms that granted their executives stock options. In general, exercise prices of executive stock options are not dividend protected, thus providing managers with significant option holdings and incentives to reduce dividend payments. Consistent with this view, they find a reduction in dividend payouts following the adoption of a stock option plan. Overall, the evidence suggests that agency problems between the various claimants of the firm have an effect on payout policy.

## 5. Asymmetric information and payout policy

Under the perfect capital markets conditions of the Miller and Modigliani dividend irrelevance proposition, all interested market participants have the same information about the firm. In perfect capital markets, the level and pattern of a firm's dividend payments will have no effect on the value of the firm's stock. The irrelevance of dividend policy for a firm's market value seems to be at odds with existing empirical research that documents the significant effects of dividend distributions on stock prices. To study the nature of the information conveyed through dividends, a number of empirical experiments examine whether unanticipated changes in dividends cause share prices to change in the same direction as the dividend change. Pettit (1972) finds that share prices tend to rise following announcements of dividend increases and to fall following announcements of dividend decreases. Aharony and Swary (1980) show that this relation between dividends and stock prices holds even after controlling for contemporaneous earnings announcements. Thus, dividends appear to contain incremental information about firm value beyond the value-relevant information contained in earnings. The variety of studies on announcement effects suggest that share prices tend to rise by about 0.4% following announcements of dividend increases and to fall by over 1% following announcements of dividend decreases.

Studies by Asquith and Mullins (1983), Healy and Palepu (1988), and Michaely, Thaler, and Womack (1995) examine large changes in dividend policy by focusing on the initiations and omissions of dividends. The analyses show that the market reactions to these events are dramatic. Excess returns following dividend initiations are about 3% and those following dividend omissions are over $-7\%$. Kalay and Lowenstein (1986) analyze whether the timing of dividend announcements conveys information to investors. Early announcements of dividends tend to be associated with good news, while delayed announcements are generally associated with bad news.

In summary, unanticipated announcements of dividend changes tend to be associated with revisions in share prices in the same direction as the dividend change. Share prices increase on average following dividend increases and initiations, and they fall on average following dividend decreases and omissions. The evidence clearly shows that dividends convey information that is relevant to investors.

If one relaxes the assumption of symmetric information, then dividend payments might convey information to the market. In this context, dividends might convey information not previously known to market participants, or dividends might arise as a mechanism for "signaling" the true value of the firm. It is possible that dividend decisions can convey value-relevant information to investors by resolving future uncertainty. Miller and Rock (1985) provide one illustration of this possibility. They note that dividends, investment, and earnings are intrinsically related through the accounting identity that describes the firm's sources and uses of funds as shown in the following equation.

$$\text{Dividend}_t = \text{Earnings}_t - \text{Investment}_t \tag{14}$$

In this accounting identity, dividends represent any dividends paid net of new financing raised. To illustrate how dividends might convey information, assume that the firm's investment policy is fixed and known. Further assume that the firm's earnings are serially correlated through time according to the following:

$$\text{Earnings}_{t+1} = \rho\text{Earnings}_t + \varepsilon_{t+1} \tag{15}$$

In other words, the level of the firm's current earnings contains information about the future level of earnings. As seen from the sources and uses identity, dividends are the residual of earnings over investment. Thus, larger than expected dividends imply higher earnings both today and in the future if earnings are serially correlated through time. Since the market does not know the current level of earnings, higher than expected dividends, which imply higher earnings, will lead to a positive stock price reaction to the announcement of the dividend.

As shown by Miller and Rock, however, if dividends can be used to convey information to market participants, then the usual rule of investing in all positive NPV projects is no longer consistent because managers have incentives to manipulate investment policy and pay higher dividends today in order to raise the current stock price. To restore consistency, a number of signaling models have been developed that treat dividends as a "costly" mechanism for signaling the firm's future prospects.

### *5.1. Dividend-signaling models*

Dividend-signaling models provide a logical framework for understanding the role of dividends in communicating relevant new information to the market. In this section, we briefly describe the essential features of several prominent dividend signaling models, such as those developed by Bhattacharya (1979), John and Williams (1985), and Miller and Rock (1985).

The basic feature of all these models is that managers possess private information about the firm's future earnings prospects and that they use dividend payments to communicate this private information to the market. While some information may be easy to communicate to investors through audited financial statements and other announcements, other crucial information may be more difficult to disseminate to a firm's investors. For example, a firm whose management is highly confident about the outcome of its ongoing research and development (R&D) may not easily convey this information to investors.

If the firm simply issues a public announcement regarding the likely success of its efforts, there is a likelihood that other firms, whose R&D is not progressing as well, will issue similar statements. Alternatively, if the firm announces too much detail in an effort to support its claims, then it may undermine its competitive advantage.

A second feature present in signaling models is that the firm has incentives to immediately establish its true market value. There are several reasons why this might be the case. For example, the firm may need to issue new equity securities to fund its investments, may have current investors who desire to sell shares for liquidity needs, or may be facing the threat of a takeover. If the private information held by managers is favorable but not reflected in the current market price, then any share sales or issues of new shares will transfer wealth from existing shareholders to new shareholders. Firms with more favorable private information have greater incentives to communicate this information to the market to eliminate this underpricing.

The payment of a dividend may proxy for this favorable information and lead to an upward adjustment of the firm's share price if investors believe that firms paying higher dividends have better future prospects. The signal is credible if other firms, with less optimistic prospects, cannot mimic the dividend policies of the firms with better future prospects. To be a credible signal, the dividend decision must be "costly" in the sense defined by Spence (1977). The "cost" of the dividend signal varies from model to model, and we describe these differences in more detail when we discuss specific signaling models. In the remainder of this section, we discribe several different dividend signaling models and survey relevant empirical evidence. The models are discussed in roughly chronological order.

### *5.1.1. The Bhattacharya model*

In Bhattacharya's signaling model (1979), the manager signals private information about the prospects of the firm's investment projects by committing ex ante to a dividend policy. This private information concerns the expected profitability of the project. The model covers two periods. At time 1, the project generates cash flows that are used to pay the dividends committed to at time 0. A crucial assumption of the model is that if the cash flows generated by the project are insufficient to cover the announced dividend payments, then the firm must resort to costly outside financing. After the dividends are paid at time 1, the firm is sold to a new group of shareholders who receive the payoffs generated by the firm at time 2. The payoffs from the project are independent and identically distributed across periods. The price that the new investors are willing to pay depends on their beliefs regarding the profitability of the project. At time 0, the manager can signal the profitability of the project by committing to pay a large dividend at time 1. Because issuing new securities is assumed to be costly, firms with less favorable investment projects will face higher expected financing costs for the same level of dividend payments. This ensures that the low-quality firm will find it unprofitable to mimic the dividend policy adopted by the high-quality firm, implying that the firm's dividend policy can serve as a credible vehicle for conveying management's private information to the market.

Despite its insights, the Bhattacharya model is subject to some criticisms. For one, Bhattacharya was not specific as to how management would commit to a specific dividend policy. Dividends do not represent a contractual obligation, and the firm is not obligated to resort to costly external financing should a cash flow shortfall occur. If market participants recognize this lack of commitment, they will not attach any importance to the existence of dividend payments. A second criticism of the model is that a commitment to a policy of share repurchases could also serve as a valid signaling mechanism. Given that share repurchases generally have more favorable tax consequences for investors compared to dividends, it is not clear why the firm would choose dividends rather than share repurchases to signal favorable information to the market.

### *5.1.2. The Miller and Rock model*

Miller and Rock (1985) also developed a signaling explanation for dividends. In their model, firms invest in a project at time 0. At time 1 the project produces earnings, which are used to pay dividends and finance the firm's new investments. Managers possess private information about the firm's realized earnings. Market participants do not directly observe either the level of earnings or the level of new investment. Miller and Rock assume that some shareholders desire to sell their holdings in the firm at time 1, which provides a motivation to establish the true valuation of the firm. Earnings are correlated through time, which implies that the firm has an incentive to convince the market that time 1 earnings are high so that selling shareholders receive a high price for their shares.

The relation between earnings, dividend payments, and investment is governed by the accounting identity that ensures that sources of cash flow are equal to uses of cash flow. Because of the sources and uses identity, all else being equal, a firm that increases its dividend must reduce its current investment. In the Miller and Rock model, "better" firms distinguish themselves by cutting investment to pay higher dividends. In equilibrium, dividend payments are sufficiently high such that lower quality firms will not find it in their interest to forgo profitable investment in order to mimic the dividend policy of higher quality firms. On the one hand, the idea that firms might reduce investment to pay higher dividends is a significant insight of their model. In addition, the Miller and Rock model generates predictions about the announcement effects of dividends. On the other hand, because of the sources and uses identity, the "dividend" payment in the Miller and Rock model is actually the sum of dividend payments and share repurchases net of any new financing. Thus, their model also cannot explain why firms would choose to signal using dividends (which are tax disadvantaged) rather than repurchases.

### *5.1.3. The John and Williams model*

John and Williams (1985) have overcome the criticism that share repurchases serve the same signaling role as dividends by developing a model in which the personal tax disadvantage of dividends represents the "cost" of signaling the firm's future prospects to the market.[29] The model can therefore explain why firms pay dividends, even when there are alternative methods of distributing cash to shareholders, such as share repurchases.

In the John and Williams model, shareholders have liquidity needs that must be met by selling shares. Managers act in the interest of existing shareholders and have information that outside investors do not have regarding the true value of the firm. If the firm is undervalued at the time that the existing shareholders need to sell their shares, they will sell at a price that is below fair value. John and Williams show that a "good" firm can signal its true value by paying a taxable dividend. If outside investors interpret the dividend as a positive signal, the share price will rise and stockholders will sell fewer shares in order to meet their liquidity needs. Dividend payments are costly to shareholders, who must pay tax on them. However, there are two benefits: (1) the shareholders sell their shares at a higher price, and more importantly (2), the shareholders maintain a larger fractional share of the firm's equity. If the firm is undervalued, the gain on the higher fractional share of the firm is valuable. A "bad" firm will not find it profitable to mimic the actions of the "good" firm because shareholders will lose on the fractional share retained when the overvaluation is corrected. Only shareholders in firms that are sufficiently undervalued will benefit enough from their higher fractional ownership to make it worthwhile to bear the tax cost of the dividend payment. The model suggests that firms expecting higher future operating cash flows optimally pay higher dividends, and that the optimal dividend is larger when the tax disadvantage of dividends relative to capital gains is smaller. Finally, the model can also explain why firms sometimes pay dividends and issue new equity securities in the same period. In this case, dividends are used to reduce the underpricing of new securities issued to raise outside financing.

### 5.2. *Dividend smoothing and dividend clienteles*

Another stylized fact is that corporations smooth dividends relative to cash flows. A firm often will not change its dividend payment over a substantial period of time, even though earnings might change dramatically over this same time period. John and Nachman (1986), using a dynamic version of the John and Wiliams (1985) model, and Kumar (1988) have developed signaling theories in which dividends are smoothed relative to underlying cash flows.

#### 5.2.1. *The John and Nachman model*

In the John and Nachman model, the equilibrium dividend paid is the product of two terms: the total extent of financing done at the firm and shareholder level, and the degree of optimism in the private information of the firm's managers. When the firm's managers possess a high level of optimism about future earnings then, because securities are mispriced, the firm will desire to raise only the amount needed to finance its profitable

[29] Bernheim (1991) also provides a signaling theory of dividends based on taxes.

investments. Alternatively, when the level of optimism is relatively low, management will desire to raise a large amount of funds and hold some in reserve to finance future investments. In these two cases, dividend payments can be roughly the same even though the cash flows are quite different.

### 5.2.2. *The Kumar model*

In the Kumar model, dividend smoothing arises as a way to separate different types of firms when there is a continuum of types ranging from low-quality to high-quality firms. Firm quality is broken into a finite number of discrete intervals, and the firms in each interval pay the same dividend even though they have different earnings.

### 5.2.3. *The Allen, Bernardo, and Welch model*

Allen, Bernardo, and Welch (2000) have developed a signaling model that can potentially explain both the payment of dividends and the smoothing of dividends based on tax clienteles. In their model, there are two types of investors with different tax rates on dividends: untaxed institutional investors and taxed individual investors. Moreover, they assume that institutional investors, because of their size, have greater incentives to invest resources to become informed about the quality of the firm. They are also more likely to facilitate mechanisms through which shortcomings in management are corrected. They further assume that dividends are a way of attracting institutional investors. In their model, the equilibrium market prices of dividend-paying firms make these securities a relatively better purchase for institutions compared to individual investors because of the relative tax advantage that institutional investors enjoy. This comparative advantage results in institutional investors holding an endogenously determined fraction of dividend-paying stocks in equilibrium.

In the signaling equilibrium derived by Allen, Bernardo, and Welch, taxable dividends exist to signal high-quality firm management because paying dividends increases the chance that the institutions holding the stock will detect firm quality. Poor-quality firms dislike attracting institutional ownership because it increases the probability that firm quality will be revealed. These poor quality firms will not find it worthwhile to incur the dividend tax costs to mimic high-quality firms. In contrast, high-quality firms do not fear detection and are willing to have their shareholders incur dividend taxes in order to signal firm quality.

Another issue surrounding signaling explanations of dividends is why firms use dividends to signal their quality even though share repurchases impose a lower tax burden on investors. As we described earlier, the John and Williams (1985) model can account for why dividends may be favored over share repurchases, because in this model it is the tax consequence of dividends that represents the cost of false signaling by lower quality firms. Firms do not use share repurchases to avoid taxes because it is precisely the tax consequences of dividend payments that support the signaling equilibrium. In contrast, in many other signaling models there is no reason dividends would be favored as a signaling mechanism relative to share repurchases. In fact, as long as share repurchases are taxed favorably relative to cash dividends, share purchases should be favored as a mechanism for signaling firm quality relative to dividends.

Nevertheless, dividends can sometimes arise as a viable signaling mechanism, even with adverse tax consequences for investors. One important distinction is that dividend payments represent a pro rata distribution of cash to all shareholders, but shares are repurchased only from investors who choose to sell. If some investors are more informed relative both to other investors and management, share repurchases can suffer from an adverse selection problem of their own. Specifically, if some investors possess superior information about the firm's future prospects, only those shareholders who believe that the firm is currently overvalued will tender their shares into the repurchase. Ambarish, John, and Williams (1987), Ofer and Thakor (1987), and Barclay and Smith (1988) develop models that predict the conditions under which share repurchases will be preferred to dividend payments when shareholders are differentially informed. In general, a variety of potential signals are available to management to signal the firm's "true" value to investors. Aside from dividends and repurchases, leverage, insider buying and selling, and capital expenditures, among others, have been suggested as possible mechanisms for communicating management's expectations to investors.

### *5.3. Empirical evidence on signaling*

Signaling models are clearly consistent with the empirical regularity that stock prices change in the same direction as the change in dividends. A second prediction of the dividend-signaling models is that future earnings changes should also be positively correlated with the current change in dividends. A number of studies have examined the relation between changes in dividends and the firm's subsequent earnings. If dividends convey private information about earnings to the market, forecasts of future earnings that include dividend information should be superior to those without dividend information.

Watts (1973) examines whether information in current dividends improves forecasts of future earnings based on current and past earnings alone. Working from a sample of 310 firms for the period 1946–1967, Watts finds little evidence that information about current dividends improve forecasts of future earnings. Similarly, Gonedes (1978) also found only weak evidence that current dividend information improves forecasts of subsequent earnings. Penman (1983) also finds little evidence that dividend changes help to forecast future earnings changes. Penman also finds that many firms with improved future earnings did not adjust their dividends.

A more recent study by Benartzi, Michaely, and Thaler (1997) confirms these findings. Using a large sample of firms over the period 1979–1991, these authors measure earnings changes following dividend announcements relative to the industry averages after adjusting for momentum and mean reversion in earnings. They find a strong positive association between lagged and contemporaneous earnings changes and dividend changes, but no systematic association between current dividend changes and future changes in earnings over the next two years. Following dividend decreases, Benartzi et al. actually report that earnings increase in the following two years.

DeAngelo, DeAngelo, and Skinner (1996) examine 145 firms with declines in earnings growth that occur after at least nine years of consecutive growth in earnings. The year of the earnings decline is defined as year 0, and DeAngelo et al. examine the year 0 dividend decision. They hypothesize that the year 0 dividend decision should contain significant information regarding the permanent or transitory nature of the current decline in earnings growth. In contrast to their predictions, they find no evidence that increases in dividends are associated with positive future earnings performance.

Ofer and Siegel (1987) further examine the relationship between dividend and earnings changes by focusing on whether unanticipated changes in dividends are associated with changes in the market's expectations of future earnings. In contrast to the studies that examine changes in earnings following dividend changes, Ofer and Siegel do find evidence that analysts change their forecasts about current-year earnings in the same direction as the dividend changes.

Other empirical studies have focused on testing the sufficient conditions for dividends to act as a costly signal. Tax-based dividend-signaling theories are based on the idea that dividends are a more costly payout mechanism compared to repurchases and that managers intentionally use dividend policy to "signal" the quality of the firm. Bernheim and Wantz (1995) examine whether the information contained in the market reaction to dividend changes varies in a systematic manner across different tax regimes. Tax-based dividend signaling models predict that, all else being equal, a dividend change of a given size will convey more information in periods when the relative taxes on dividends compared to capital gains are higher. Consistent with the dividend-signaling hypothesis, Bernheim and Wantz find that the market reaction to dividend changes is larger during periods of high relative taxes on dividends.

In contrast to the findings of Bernheim and Wantz, however, Bernhardt, Robertson, and Farrow (1994) use nonparametric techniques to account for the nonlinear nature of most signaling models and find little evidence in support of dividend signaling. Along similar lines, based on data surrounding the Tax Reform Act of 1986, Grullon and Michaely (2001) have found that market reactions to dividend increases were larger in the post-1986 period when dividends were taxed less heavily than were capital gains. Finally, Amihud and Murgia (1997) study the market reaction to dividend changes in Germany where dividends are favorably taxed relative to capital gains for most classes of investors. In this case, tax-based signaling models predict that dividend changes will have no signaling value. In contrast to the predictions of the signaling models, Amihud and Murgia find a market reaction to dividend changes in Germany similar to that documented in the United States.

In short, there is little evidence that changes in dividends predict future changes in earnings, which is one of the main predictions of dividend-signaling models. If anything, dividend changes tend to lag rather than lead earnings changes. In addition, there is at best only weak evidence in favor of the sufficient conditions of tax-based signaling models. Recent survey data based on the responses of 384 executives from 256 U.S. companies gathered by Graham, Harvey, and Michaely (2004) also provides little support for the idea that managers view dividends as a costly mechanism for signaling the true value of their firm. Although over 80% of the executives surveyed believe that dividend policy conveys information to investors, only 25% of the executives suggest that they use dividend policy to make their firm look better than their competitors, and only 4.4% of executives state that dividends are used to show that their firms can bear self-imposed costs (as would be required in the costly signaling models).

Given the weak evidence that dividends convey information about future earnings, the question remains as to what information dividend changes convey to the market. One possibility is that dividend changes convey information only about current earnings via the sources and uses identity. Another possibility is that dividend policy might convey information regarding the riskiness of the firm's cash flows. Consistent with this notion, nearly 40% of executives in the Graham et al. survey believe that dividends make the stock less risky. Some recent empirical evidence also supports this view. For example, Grullon, Michaely, and Swaminathan (2002) have shown that systematic risk decreases following increases in dividends and increases following a dividend decrease. They also find that profitability declines following dividend increases, but that the decline in profitability is more than offset by the decline in risk, which is consistent with the positive market reaction to dividend increases. Moreover, they also report that the magnitude of the market reaction is positively related to the reduction in risk, all else equal.

Yet another possibility is that dividend changes convey information about the persistence of past earnings changes. Consistent with this view, Koch and Sun (2004) have found that changes in dividend policy alter investors' assessments of the valuation consequences of past earnings changes, namely, the permanence of past earnings changes. These findings can potentially reconcile the fact that changes in dividends tend to lag past earnings changes (e.g., Benartzi, Michaely, and Thaler, 1997) but still appear to convey valuable information to market participants.

## 6. Share repurchases

As an alternative to paying cash dividends, a firm can return cash to shareholders by repurchasing some of its shares. There are four main ways to repurchase stock. First, in a repurchase tender offer, the firm offers to buy back a stated number of shares at a stated price—typically, about 20% above the current market price. Individual shareholders then decide whether to tender their shares at this price. Second, in a Dutch auction, the firm states a series of prices at which it is willing to buy back shares. Shareholders submit offers for the quantity of shares that they will sell at each price. The firm aggregates these orders and chooses the lowest price at which it can repurchase the desired number of shares. All tendering shareholders receive this price for their shares. Third, share repurchases can also result from direct negotiation with major blockholders, sometimes in conjunction with a takeover attempt.[30] Finally, an open market repurchase, in which

[30] For the purposes of this chapter, we ignore privately negotiated repurchases.

the company announces that it will buy back shares in the open market like any other investor, is by far the most common form of repurchase activity.

As shown in Table 1, prior to 1984, repurchases were relatively rare, but repurchase activity has accelerated dramatically since then. Grullon and Michaely (2002) suggest that one reason for the rise in repurchase activity is that the Securities and Exchange Commission (SEC) adopted Rule 10b-18 in 1982. Prior to the adoption of this rule, firms that repurchased shares ran the risk of being prosecuted for manipulating their share price. The rule laid out provisions that provided a safe harbor for firms engaged in repurchase activity. These authors also report that young firms are more likely than they were in the past to initiate cash payouts through repurchases, and that more established firms, while continuing to pay dividends, also exhibit a higher propensity to repurchase shares. They argue that firms have gradually substituted repurchases for dividends over time.

Although both dividends and repurchases return cash to shareholders, there are a number of relevant differences between the two. Perhaps the largest difference is in the tax treatment between dividends and repurchases. In the United States, dividends are taxed as ordinary income, whereas repurchases are taxed at the historically lower capital gains rate. To the extent that shareholders are not able to avoid the higher taxes on cash dividends, the differential taxation of dividends versus repurchases should favor repurchase as a mechanism for returning cash to shareholders. A second difference between repurchases and dividends stems from the strong reluctance to cut or omit dividend payments once they are initiated. In this regard, dividends represent a commitment to continue to pay out cash in the future, whereas repurchases are more likely associated with a one time disbursement of cash. A final difference between dividends and repurchases is that the timing of repurchases is subject to managerial discretion. The survey evidence in Graham et al. (2005) suggests that managers attempt to repurchase stock when they believe it is currently undervalued.

### *6.1. Empirical evidence on share repurchases*

Similar to dividend increases, announcements of repurchases are generally associated with positive stock price reactions. Vermaelen (1981) and Comment and Jarrell (1991) document abnormal returns of approximately 2 to 3% around announcements of open market repurchases. Average abnormal returns are on the order of 11 to 15% when repurchase tender offers are announced and are approximately 8% around the announcements of Dutch auction repurchases. These studies also report that price increases from buyback announcements are larger when insider wealth is at risk and following negative stock price performance. They also find that stock price increases are increasing in the fraction of shares sought in the repurchase. In addition, a number of studies (e.g., Ikenberry, Lakonishok and Vermaelen, 1995) have found that prices of repurchasing firms continue to drift upward following the repurchase announcements and that this long-run drift is more pronounced in stocks with high book-to-market ratios. The results are generally consistent with the idea that repurchases provide information to market participants that the firm is undervalued.

A number of studies provide evidence on the type of information conveyed by repurchase decisions by examining patterns in earnings following repurchase announcements. Vermaelen (1981) finds that earnings per share increase in the years following fixed price tender offers. Dann, Masulis, and Mayers (1991) confirm Vermaelen's results and also show that the initial market reaction is positively related to the subsequent increase in earnings. They interpret their findings as being consistent with a signaling motive for repurchases. More recently, Nohel and Tarhan (1998) show that the improvement in earnings documented in the prior studies was entirely attributable to firms with high book-to-market ratios (i.e., value firms) and that the improvement in operating performance was positively related to asset sales. They believe their evidence supports the idea that tender offer repurchases are used to control free cash flow problems rather than to signal future earnings performance.

Grullon and Michaely (2000) have reported that earnings performance improves following repurchase announcements using a large sample of repurchases between 1980 and 2000. They document a decline in return on assets in the three-year period following the repurchase announcement, and they also find decreases in capital expenditures and cash reserves. These findings are similar to the documented patterns in earnings following dividend announcements. Grullon and Michaely do, however, find that firm risk declines after repurchase announcements. The cost of capital for repurchasing firms in their sample drops from 16.3% prior to the repurchase to 13.7% in the period following the repurchase. The evidence is not generally consistent with traditional signaling stories. Instead, it appears that firms tend to increase their payouts to shareholders following a decline in their investment opportunities and demand for capital.

Jagannathan, Stephens, and Weisbach (2000) have reported that firms with higher permanent operating cash flows pay dividends, whereas firms with higher temporary nonoperating cash flows tend to use repurchases. Similarly, Guay and Harford (2000) hypothesize that dividend increases will be observed following cash flow shocks with a relatively large permanent component, while repurchases will be used to distribute shocks that are primarily transient. Using a large sample of dividend increases and repurchases, they have found that the post-shock cash flows of dividend-increasing firms do not fully revert to pre-shock levels, while those of repurchasing firms completely revert to pre-shock levels, even settling below them. The stock price reactions to the announcements of both repurchases and dividend increases also show strong evidence that the information in a payout announcement is not only the size of the payout, but also the method used to distribute the cash.

Maxwell and Stephens (2003) find some evidence of negative returns to bondholders around repurchase announcements. The loss to bondholders increases in the size of the repurchase and with the riskiness of the firm's debt. They also find that bonds are more likely to be downgraded following repurchase announcements. Their results are consistent with the view that agency conflicts between shareholders and bondholders are also a determining factor in whether to repurchase shares.

In open market repurchases, issues arise in measuring the amount of repurchasing activity because firms are not obligated to repurchase all of the shares that they initially seek. Stephens and Weisbach (1988) examine several measures of repurchase activity and conclude that over 80% of repurchase programs end within three years and that more than half of the firms in their sample completed their announced repurchase program. However, more than one-tenth of the firms in their sample repurchased less than 5% of the shares they initially sought to repurchase at the announcement. Finally, they showed that the initial market reaction at the announcement of the repurchase is positively related to the intensity of repurchase activity in the following two years.

Another reason that firms may repurchase shares is to avoid the dilution that arises from the exercise of employee stock options. Kahle (2002) documents the fact that repurchase activity is positively correlated with the amount of exercisable stock options held by the firm's employees, but is unrelated to options held by managers. She concludes that managers repurchase both to maximize their own wealth and to fund employee stock option plans. She also finds that the market reacts less positively to announcements of repurchases by firms with high levels of employee options outstanding.

Finally, Barclay and Smith (1988) report that bid-ask spreads widen following repurchase announcements. Their evidence suggests that adverse selection arising from informed trading creates a cost borne by uninformed shareholders. They also show that the additional costs associated with stock repurchases may outweigh their preferential tax treatment relative to dividends and provide a possible explanation of why repurchases are not always substituted for cash dividend payments.

## 7. Alternative theories and new stylized facts

Thaler and Shefrin (1981) and Shefrin and Statman (1984) suggest that various psychological biases related to self-control, prospect theory, and regret aversion can result in preferences for cash dividends relative to homemade dividends. In general, there is relatively little empirical evidence regarding inherent investor preferences for dividends.

One exception is Baker and Wurgler (2004), who relax the assumption of market efficiency in the Miller and Modigliani irrelevance propositions and propose that, if arbitrage behavior is limited, managers have incentives to cater to investor preferences for dividends. Consistent with their argument, they find that firms are more likely to initiate dividend payments when measures of investor demand for dividends are high and to omit dividend payments when investor demand is low. Li and Lie (2005) extend the catering theory to include decreases and increases in existing dividends. Consistent with catering incentives, they find that the decision to change the dividend and the magnitude of the change depend on the premium that the capital market places on dividends.[31]

Fama and French (2002) have documented some new and interesting facts regarding dividends, showing that the percentage of firms paying cash dividends fell from 66.5% in 1978 to 20.8% in 1999. They report that part of the decline can be explained by the dramatic increase in the listing of small, unprofitable firms with strong growth

[31] Baker, Ruback, and Wurgler (2007) provide an overview of behavioral applications to corporate finance.

opportunities. After controlling for firm characteristics, however, they still find that firms are less likely to pay dividends than they were in the past. They argue that the benefits of dividends have likely declined through time in part owing to reduced transactions costs for selling stocks, higher stock option holdings by managers who prefer capital gains to dividends, and improved corporate control mechanisms that limit agency problems.

Hoberg and Prabhala (2005) argue that both the disappearing dividend puzzle and the evidence on catering can be explained by firm risk. They show that risk is a significant determinant of the propensity to pay dividends and that changes in risk can explain about a third of the disappearing dividends documented by Fama and French (2002). They also find little support for the view that dividend policies reflect firms' catering to transient fads for dividends. Absent risk controls, catering matters, but it is insignificant after controlling for risk.

## 8. Conclusion

This chapter surveys the academic research on payout policy—the decision of the firm about whether and how to return cash to its shareholders. More than 40 years have now passed since Miller and Modigliani wrote their seminal paper delineating the conditions under which payout policy can affect firm value. During this time, payout policy has garnered significant attention from both academics and practitioners. Although we now have a better understanding of the factors that should systematically affect firms' payout decisions, many issues remain unsettled and many new questions have been raised.

As an example, the empirical evidence indicates that dividends convey information to the market. The market seems to view dividend increases positively and reacts negatively to decreases in dividends. Nevertheless, our understanding of what information is important and of the mechanism by which the information is conveyed to the market remains incomplete. Are increases in dividends information of superior future performance, or are they conveying information about current profitability or something else? Should we endorse costly signaling models as the description of the equilibrium within which dividends serve as a signaling device? The assumptions required by a costly signaling equilibrium are fairly restrictive. Much of the empirical evidence is inconsistent with costly signaling models. In addition, the dividend decision process described by managers in carefully conducted surveys is also inconsistent with the known signaling models. In summary, we are left with a well-documented empirical regularity—the information content of dividends—without a satisfactory theory to explain it.

An equally important aspect of the dividend decision that has attracted significant attention from financial economists is the effect of taxes on the dividend decision. The generally higher tax rate paid on dividend income should result in investors requiring a higher before-tax risk-adjusted return on dividend-paying stocks. The empirical evidence is, however, mostly inconsistent with this simple prediction. On average, stocks earn higher risk-adjusted returns during the ex-dividend week—there is time-series variation in their expected rate of return. Yet, stocks having higher dividend yields do not earn a higher risk-adjusted rate of return than those with lower dividend yield—there is no cross-sectional variation in expected return. It is difficult to link the higher risk-adjusted return during ex-dividend weeks to taxes. Investors attempting to avoid the dividends would have to sell the stock prior to ex and buy it back after ex, thereby realizing short-term capital gains taxed as dividends. A tax-related explanation of the time-series return variation is even more difficult to establish when one finds no cross-sectional correlation between before-tax risk-adjusted return and taxes. The obvious questions remains—what is the explanation of the higher before-tax rate of return during the ex-dividend week? If taxes are important, why don't stocks exhibit cross-sectional correlation between before-tax risk-adjusted return and dividend yield?

Financial economists provide some explanations for why firms pay dividends even when it is costly to do so. The payment of dividends increases the likelihood that the firm will have to raise external funds. Hence, by paying out dividends, the firm commits to evaluation by external experts in the process of raising funds. By paying out its free cash flows, management reduces its ability to use funds in a suboptimal manner. So, the firm benefits from a better investment decision that offsets the tax-related costs of paying dividends. Finally, the alternative of paying no dividends may be costly too. Without payouts (defined to include share repurchases), the firm would have to reinvest all its net earnings. If corporations face a limited supply of nonnegative NPV projects, they will then be forced to accept bad ones.

In addition, recent empirical evidence indicates that some aspects of limited rationality may bear on the dividend decision. Baker and Wurgler (2004) argue for a "catering theory of dividends" in which firms are more likely to initiate dividend payments in periods when the exogenous demand for dividends is high.

Finally, the survey evidence in Brav et al. (2005) also points to a number of unresolved issues regarding payout policy. Similar to the evidence in the original study by Lintner (1954), Brav et al. find that managers are very reluctant to cut dividends once they are initiated. This reluctance leads to dividends that are sticky, smoothed from year to year, and tied to the long-run profitability of the firm. Beyond maintaining the current dividend level, Brav et al. find that payout policy is a second-order concern for most corporations in the sense that payout policy is considered after investment and liquidity needs are met. In contrast to Lintner's results, Brav et al.'s findings show that present-day managers are more reluctant to increase dividends in tandem with earnings increases and that they do not set dividend policy based on a target percentage of earnings. In addition, repurchases are now used more extensively, and managers view repurchase activity as being more flexible than dividend policy. Executives tend to accelerate repurchase decisions when they view their stock as undervalued and are very conscious of how repurchases affect earnings per share.

Executives believe that dividends are attractive to individual investors but that dividends and repurchases are equally attractive to institutions. However, they find no evidence that payout policy is used as a tool to alter the proportion of institutional investors in the firm. Moreover, despite being aware of the tax advantage of repurchases relative to dividends, managers maintain that this is not an important factor either in their decisions about whether to pay out or increase dividends or in their decision as to the form of the payout—repurchase versus dividends. Surprisingly, in a follow-up survey conducted after the Bush administration announced a proposal to eliminate dividend taxation, more than two-thirds of the executives surveyed said that this would definitely not or probably not affect their dividend decisions. Moreover, among firms not currently paying dividends, 70% say that they never plan to initiate dividends, and more than half say they do not intend to repurchase shares. Among firms that state they will eventually pay out cash to shareholders, the majority maintain that they will use repurchases.

Clearly, the story has not ended. We expect the research on payout policy to continue and to broaden our understanding of the factors that affect this policy. Writers of reviews decades from today will undoubtedly have much more to say.

## References

Ackert, Lucy and Brian Smith, 1993, "Stock Price Volatility, Ordinary Dividends, and Other Cash Flows to Shareholders," *Journal of Finance,* 48(4), 1147–1160.

Aggarwal, Rajesh, 2007, "Executive Compensation," In B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance,* Vol. 1 in Handbook of Finance Series, Elsevier/North-Holland, Amsterdam, Chapter 17 (this volume).

Agrawal, Anup and J. Narayanan, 1994, "The Dividend Policies of All-Equity Firms: A Direct Test of the Free Cash Flow Theory," *Managerial Decision Economics*, 15, 139–148.

Aharony, Joseph and Itzhak Swary, 1980, "Quarterly Dividend and Earnings Announcements and Stockholders' Returns: An Empirical Analysis," *Journal of Finance*, 35(1), 1–12.

Allen, Franklin, Antonio Bernardo and Ivo Welch, 2000, "A Theory of Dividends Based on Tax Clienteles," *Journal of Finance*, 55(6), 2499–2536.

Allen, Franklin and Roni Michaely, 1995, "Dividend Policy," In R. Jarrow et al. (Eds.), *Handbook in Operations Research and Management Science*, Vol. 9, Elsevier/North Holland.

Ambarish, Ramasastry, Kose John and Joseph Williams, 1987, "Efficient Signaling with Dividends and Investments," *Journal of Finance*, 42(2), 321–343.

Amihud, Yakov and Maurizio Murgia, 1997, "Dividends, Taxes, and Signaling: Evidence from Germany," *Journal of Finance*, 52(1), 397–408.

Asquith, Paul and David W. Mullins Jr., 1983, "The Impact of Initiating Dividend Payments on Shareholders' Wealth," *Journal of Business*, 56(1), 77–96.

Auerbach, Alan J., 1979, "Wealth Maximization and the Cost of Capital," *Quarterly Journal of Economics*, 93(3), 433–446.

Bagwell, Laurie, 1992 "Dutch Auction Repurchases: An Analysis of Shareholder Heterogeneity," *Journal of Finance*, 47(1), 71–105.

Bagwell, Laurie Simon, 1991, "Shareholder Heterogeneity: Evidence and Implications," *American Economic Review*, 81(2), 218–221.

Bagwell, Laurie Simon and John Shoven, 1989, "Cash Distributions to Shareholders," *Journal of Economic Perspectives*, 3(3), 129–140.

Baker, Malcolm and Jeffrey Wurgler, 2004, "A Catering Theory of Dividends" *Journal of Finance 59*, 1125–65.

Baker, Malcolm, Richard S. Ruback and Jeffrey Wurgler, 2007, "Behavioral Corporate Finance," In B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, Vol. 1, in Handbook of Finance Series, Elsevier/North-Holland, Amsterdam, Chapter 4 (in press).

Bali, Rakesh and Gailen Hite, 1998, "Ex-dividend Day Stock Price Behavior: Discreteness or Tax-induced Clientele?," *Journal of Financial Economics*, 47, 127–159.

Barclay, M., 1987, "Dividends, Taxes and Common Stock Prices: The Ex-dividend Day Behavior of Common Stock Prices Before the Income Tax," *Journal of Financial Economics*, 14, 31–44.

Barclay, Michael J. and Clifford W. Smith, Jr., 1988, "Corporate Payout Policy: Cash Dividends versus Open-Market Repurchases," *Journal of Financial Economics*, 22(2), 61–82.

Barker C., 1959, "Price Changes of Stock Dividend shares at ex-Dividend Dates" *Journal of Finance*.

Bartov, Eli, 1991, "Open-Market Stock Repurchase as Signals for Earnings and Risk Changes," *Journal of Accounting and Economics*, 14, 275–294.
Bernhardt, Dan, J. Fiona Robertson and Ray Farrow, 1994, "Testing Dividend Signaling Models," Working Paper, Queen's University.
Bernheim, Doug, 1991, "Tax Policy and the Dividend Puzzle," *RAND Journal of Economics*, 22, 455–476.
Bernheim, Doug and Adam Wantz, 1995, "A Tax-based Test of the Dividend Signaling Hypothesis," *American Economic Review*, 85(3), 532–551.
Bernartzi, Shlomo, Gustavo Grullon, Roni Michaely and Richard Thaler, 2002, "Changes in Dividends (Still) Signal the Past," Working Paper, Cornell University, Ithaca, NY.
Bernartzi, Shlomo, Roni Michaely and Richard Thaler, 1997, "Do Changes in Dividends Signal the Future or the Past?," *Journal of Finance*, 52(3), 1007–1043.
Bhattacharya, Sudipto, 1979, "Imperfect Information, Dividend Policy, and 'The Bird in the Hand' Fallacy," *Bell Journal of Economics*, 10(1), 259–270.
Black, Fischer, 1976, "The Dividend Puzzle," *Journal of Portfolio Management*, 2, 5–8.
Black, Fischer and Myron Scholes, 1974 "The Effects of Dividend Yield and Dividend Policy on Common Stock Prices and Returns," *Journal of Financial Economics*, 1, 1–22.
Blume, Marshal E., 1980, "Stock Return and Dividend Yield: Some More Evidence," *Review of Economics and Statistics*, 62, 567–577.
Blume, Marshal E., Jean Crockett and Irwin Friend, 1974, "Stock Ownership in the United States: Characteristics and Trends," *Survey of Current Business*, 16–40.
Boyd, John and Ravi Jagannathan, 1994, "Ex-Dividend Price Behavior of Common Stocks: Fitting Some Pieces of the Puzzle," *Review of Financial Studies*, 7(4), 711–741.
Bradford, David F., 1981, "The Incidence and Allocation Effects of a Tax on Corporate Distributions," *Journal of Public Economics*, 15, 1–22.
Brav, Alon and J. B. Heaton, 1998, "Did ERISA's Prudent Man Rule Change the Pricing of Dividend Omitting Firms?" Working Paper, Duke University.
Brennan, Michael J., 1970, "Taxes, Market Valuation and Financial Policy," *National Tax Journal*, 23, 417–429.
Brennan, Michael J. and Anjan V. Thakor, 1990, "Shareholder Preferences and Dividend Policy," *Journal of Finance*, 45(4), 993–1019.
Brickley, James, 1983, "Shareholders Wealth, Information Signaling, and the Specially Designated Dividend: An Empirical Study," *Journal of Financial Economics*, 12, 187–209.
Campbell, JA, and W. Beranek, 1955, "Stock Price Behavior on Ex-Dividend Dates" *journal of Finance*
Charest, Guy, 1978, "Dividend Information, Stock Returns and Market Efficiency—II," *Journal of Financial Economics*, 6, 297–330.
Chen, Nai Fu, Bruce Grundy and Robert F. Stambaugh, 1990, "Changing Risk, Changing Risk Premiums, and Dividend Yield Effects," *Journal of Business*, 63, S51–S70.
Chowdhry, Bhagwan and Vikram Nanda, 1994, "Repurchase Premia as a Reason for Dividends: A Dynamic Model of Corporate Payout Policies," *Review of Financial Studies*, 7, 321–350.
Christie, William and Vikram Nanda, 1994, "Free Cash Flow, Shareholder Value, and the Undistributed Profits Tax of 136 and 1937," *Journal of Finance*, 49(5), 1727–1754.
Comment, Robert and Gregg Jarrell, 1991, "The Relative Power of Dutch-Action and Fixed-Priced Self-Tender Offers and Open Market share Repurchases," *Journal of Finance*, 46(4), 1243–1271.
Constantinides, George M, 1983, "capital market equilibrium with personal taxes" Economietrica, 51, 611–636.
Constantinides, George M., 1984, "Optimal Stock Trading with Personal Taxes," *Journal of Financial Economics*, 13, 65–89.
Constantinides, George M. and Bruce D. Grundy, 1989, "Optimal Investment with Stock Repurchase and Financing as Signals," *Review of Financial Studies*, 2(4), 445–466.
Dammon, Robert, and Richard Green, 1987, "Tax Arbitrage and the Existence of Equilibrium Prices for Financial assets" *Journal of Finance*

Dann, L., R. Masulis and D. Mayers, 1991, "Repurchase Tender Offers and Earning Information," *Journal of Accounting and Economics*, 14, 217–251.

DeAngelo, Harry and Linda DeAngelo, 1990, "Dividend Policy and Financial Distress: An Empirical Investigation of Troubled NYSE Firms," *Journal of Finance*, 45(5), 1415–1431.

DeAngelo, Harry, Linda DeAngelo, and Douglas Skinner, 1996, "Reversal of Fortune, Dividend Signaling and the Disappearance of Sustained Earnings Growth," *Journal of Financial Economics*, 40, 341–371.

DeAngelo, Harry, Linda DeAngelo and Rene Stulz, 2005, "Dividend Policy and the Earned/Contributed Capital Mix: A Test of the Lifecycle Theory," *Journal of Financial Economics*, forthcoming.

Del Guercio, Diane, 1996, "The Distorting Effect of the Prudent-Man Laws on Institutional Equity Investments," *Journal of Financial Economics*, 40, 31–62.

Dewenter, Kathryn and Vincent Warther, 1998, "Dividends, Asymmetric Information, and Agency Conflicts: Evidence from a Comparison of the Dividend Policies of Japanese and US Firms," *Journal of Finance*, 53(3), 879–904.

Dunbar, Nicholas, 2001, "UK Firms Consider Option Scheme," *Risk* Magazine, 14.

Dunsby, Adam, 1993, "Share Repurchases and Corporate Distributions: An Empirical Study," Working Paper, University of Pennsylvania.

Dybvig, Phil, and Steven Ross, 1986, "Tax Clienteles and asset Pricing" with Stephen Ross, *Journal of Finance*, 41, 751–762.

Eades, Ken, Pat Hess and Han E. Kim, 1984, "On Interpreting Security Returns during the Ex-dividend Period," *Journal of Financial Economics*, 13, 3–34.

Eades, Ken, Pat Hess and Han E. Kim, 1994, "Time-series Variation in Dividend Pricing," *Journal of Finance*, 49, 1617–1635.

Easterbrook, Frank H., 1984, "Two Agency-Cost Explanations of Dividends," *American Economic Review*, 74(4), 650–659.

Eckbo, B. Espen and Savita Verma, 1994, "Managerial Shareownership, Voting Power, and Cash Dividend Policy, *Journal of Corporate Finance*, 1, 33–62.

Elton, Edward and Martin Gruber, 1970, "Marginal Stockholders' Tax Rates and the Clientele Effect," *Review of Economics and Statistics*, 52, 68–74.

Fama, Eugene F. and Harvey Babiak, 1968, "Dividend Policy: An Empirical Analysis," *Journal of the American Statistical Association*, 63(324), 1132–1161.

Fama, Eugene and Kenneth French, 2001, "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay?," *Journal of Financial Economics* 60, 3–43.

Fama, Eugene and Kenneth French, 2000, "Forecasting Profitability and Earnings," *Journal of Business*, 73, 161–175.

Fama, Eugene F. and Kenneth R. French, 1993, "Common Risk Factors in the Returns on Stock and Bonds," *Journal of Financial Economics*, 33, 3–56.

Fama, Eugene F. and James D. Macbeth, 1973, "Risk, Return and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.

Federal Reserve Statistical Release, March 2000, "Flow of Funds Accounts of the United States."

Fenn, G. and N. Liang, 2001, "Corporate Payout Policy and Managerial Stock Incentives," *Journal of Financial Economics*, 60, 45–72.

Fluck, Zsusanna, 1999, "The Dynamics of the Manager-Shareholder Conflict," *Review of Financial Studies*, 12(2), 379–404.

Frank, Murray and Ravi Jagannathan, 1998 "Why Do Stock Prices Drop by Less Than the Value of the Dividend? Evidence from a Country without Taxes," *Journal of Financial Economics*, 47, 161–188.

Gonedes, Nicholas J., 1978, "Corporate Signaling, External Accounting, and Capital Market Equilibrium: Evidence on Dividends, Income, and Extraordinary Items," *Journal of Accounting Research*, 16(1), 26–79.

Gordon, Myron, 1959, "Dividends, Earnings and Stock Prices," *Review of Economics and Statistics*, 41, 99–105.

Green, Richard and Kristian Rydqvist, 1999, "Ex-day Behavior with Dividend Preference and Limitation to Short-term Arbitrage: The Case of Swedish Lottery Bonds," *Journal of Financial Economics*, 53, 145–187.

Grossman, Sanford J. and Oliver D. Hart, 1980, "Takeover Bids, the Free-rider Problem, and the Theory of the Corporation," *Bell Journal of Economics*, 11, 42–54.

Grullon, Gustavo and David Ikenberry, 2000, "What to Know about Stock Repurchase?" *Journal of Applied Corporate Finance*, 13, 31–51.

Grullon, Gustavo and Roni Michaely, 2002, "Dividends, Share Repurchases and the Substitution Hypothesis," *Journal of Finance*, 62(4).

Grullon, Gustavo and Roni Michaely, 2001, "Asymmetric Information, Agency Conflicts and the Impact of Taxation on the Market Reaction to Dividend Changes," Working Paper, Cornell University, Ithaca, NY.

Grullon, Gustavo and Roni Michaely, 2000, "The Information Content of Share Repurchase Programs," Working Paper, Cornell University, Ithaca, NY.

Grullon, Gustavo, Roni Michaely and Bhaskaran Swaminathan, 2002, "Are Dividend Changes a Sign of Firm Maturity?," *Journal of Business*, forthcoming.

Grundy, Bruce, 1985, "Trading Volume and Stock Returns Around Ex-dividend Dates," Working Paper, University of Chicago, Chicago, IL.

Guay and Harford, 2000, "The Cash Flow Permanence and Information Content of Dividend Increases vs. Repurchases," *Journal of Financial Economics*, September 2000.

Handjinicolaou, George and Avner Kalay, 1984, "Wealth Redistributions or Changes in Firm Value: An Analysis of Return to Bondholders and the Stockholders around Dividend Announcements," *Journal of Financial Economics*, 13(1), 35–63.

Hasbrouck, Joel and Irwin Friend, 1984, "Why Do Companies Pay Dividends?: Comment," *American Economic Review*, 74, 1137–1141.

Healy, Paul M. and Krishna G. Palepu, 1988, "Earnings Information Conveyed by Dividend Initiations and Omissions," *Journal of Financial Economics*, 21(2), 149–176.

Hertzel, M. and P. Jain, 1991, "Earnings and Risk Changes Around Stock Repurchase Tender Offers," *Journal of Accounting and Economics*, 14, 252–274.

Hess, Pat J., 1983, "Test of Price Effects in the Pricing of Financial Assets," *Journal of Business*, 56, 537–554.

Hess, Pat J., 1982, "The Ex-dividend Day Behavior of Stock Returns: Further Evidence on Tax Effects," *Journal of Finance*, 37, 445–456.

Hoberg, Gerard and N.R. Prabhala, 2005, "Dividend Policy, Risk, and Catering," Working Paper, University of Maryland.

Hubbard, Jeff and Roni Michaely, 1997, "Do Investors Ignore Dividend Taxation? A Re-examination of the Citizen Utilities Case," *Journal of Financial and Quantitative Analysis*, 32(1).

Ikenberry, David, Josef Lakonishok and Theo Vermaelen, 2000, "Share Repurchases in Canada: Performance and Strategic Trading," *Journal of Finance*, 55, 2373–2397.

Ikenberry, David, Josef Lakonishok and Theo Vermaelen, 1995, "Market Underreaction to Open Market Share Repurchases," *Journal of Financial Economics*, 39, 181–208.

Jagannathan, M. and C. P. Stephens, 2001, "Motives for Open Market Share Repurchases: Under-valuation, Earnings Signaling or Free Cash Flow," Working Paper, University of Missouri-Columbia.

Jagannathan, M., C. P. Stephens and M. S. Weisbach, 1999, "Financial Flexibility and the Choice between Dividends and Stock Repurchases," *Journal of Financial Economics*, 57, 355–384.

Jarrow, Robert, and David Heath, 1988, "Ex-Dividend Stock Price Behavior and Arbitrage Opportunities" *Journal of Business*.

Jensen, Michael C., 1986, "Agency Costs of Free Cash Flow, Corporate Finance and Takeovers," *American Economic Review*, 76(2), 323–329.

Jensen, Michael C. and William H. Meckling, 1976, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3(4), 305–360.

Jensen, Gerald R., Donald P. Solberg and Thomas S. Zorn, 1992, "Simultaneous Determination of Insider Ownership, Debt, and Dividend Policies," *Journal of Financial and Quantitative Analysis*, 27, 247–263.

John, Kose and Avner Kalay, 1982, "Costly Contracting and Optimal Payout Constraints," *Journal of Finance*

John, Kose and Joseph Williams, 1985, "Dividends, Dilution, and Taxes: A Signaling Equilibrium," *Journal of Finance*, 40(4), 1053–1070.

Jolls, Christine, 1998, "The Role of Incentive Compensation in Explaining the Stock Repurchase Puzzle," Working Paper, Harvard Law School, Cambridge, MA.

Kahle, Kathleen, 2002, "When a Buyback Isn't a Buyback: Open-Market Repurchases and Employee Options," *Journal of Financial Economics*, 63, 235–261.

Kalay, Avner, 1982a, "The Ex-dividend Day Behavior of Stock Prices: A Re-examination of the Clientele Effect," *Journal of Finance*, 37, 1059–1070.

Kalay, Avner, 1982b, "Stockholder-Bondholder Conflict and Dividend Constraint," *Journal of Financial Economics*, 14, 423–449.

Kalay, Avner, 1984, "The Ex-Dividend Day Behavior of Stock Prices: A Re-Examination of the Clientele Effect: A Reply," *Journal of Finance*, June 1984.

Kalay, Avner and Roni Michaely, 2000, "Dividends and Taxes: A Reexamination," *Financial Management*, 29(2).

Kang, S. H. and P. Kumar, 1991, "Determinants of Dividend Smoothing: Evidence from Dividend Changes," Working Paper, Carnegie Mellon University.

Karpoff, Jonathan M. and Ralph A. Walkling, 1990, "Dividend Capture in NASDAQ Stocks," *Journal of Financial Economics*, 28(1/2), 39–66.

Karpoff, Jonathan M. and Ralph A. Walkling, 1988, "Short-Term Trading around Ex-Dividend Days: Additional Evidence," *Journal of Financial Economics*, 21(2), 291–298.

Kato, Kiyoshi and Uri Loewenstein, 1995, "The Ex-dividend-day Behavior of Stock Prices: The Case of Japan," *Review of Financial Studies*, 8, 817–847.

Keim, Don, 1985, "Dividend Yields and Stock Returns: Implications of Abnormal January Returns," *Journal of Financial Economics*, 14, 473–489.

King, Mervyn, 1977, *Public Policy and the Corporation*, London, Chapman and Hall.

Koski, Jennifer and Roni Michaely, 2000, "Prices, Liquidity and the Information Content of Trades," *Review of Financial Studies*, 13(3), 659–696.

Kumar, Praveen, 1988, "Shareholder-Manager Conflict and the Information Content of Dividends," *Review of Financial Studies*, 1(2), 111–136.

Lakonishok, Joseph and Theo Vermaelen, 1986, "Tax Induced Trading Around Ex-dividend Dates," *Journal of Financial Economics*, 16, 287–319.

Lakonishok, Joseph and Theo Vermaelen, 1983, "Tax Reform and Ex-dividend Day Behavior," *Journal of Finance*, 38, 1157–1179.

Lambert, Lanen and Larcker, 1989, "Executive Stock Option Plans and Corporate Dividend Policy," *Journal of Financial and Quantitative Analysis*, 24, 409–425.

Lang, Larry H.P. and Robert H. Litzenberger, 1989, "Dividend Announcements: Cash Flow Signaling vs. Free Cash Flow Hypothesis," *Journal of Financial Economics*, 24(1), 181–192.

La Porta, Rafael, Florencio Lopez-De Silanes, Andrei Shleifer and Robert Vishny, 2000, "Agency Problems and Dividend Policy around the World," *Journal of Finance*, 55, 1–33.

Li, Wei and Erik Lie, 2005, "Dividend Changes and Catering Incentives," *Journal of Financial Economics*, forthcoming.

Lie, Erik, 2001, "Financial Flexibility and the Corporate Payout Policy," Working Paper, College of William and Mary, Williamsburg, VA.

Lie, Erik, 2000, "Excess Funds and the Agency Problems: An Empirical Study of Incremental Disbursements," *Review of Financial Studies*, 13(1), 219–248.

Lewellen, Wilbur G., Kenneth L. Stanley, Ronald C. Lease and Gary G. Schlarbaum, 1978, "Some Direct Evidence on the Dividend Clientele Phenomenon," *Journal of Finance*, 33(5), 1385–1399.

Lintner, John, 1956, "Distribution of Income of Corporations among Dividends, Retained Earnings, and Taxes," *American Economic Review*, 46(2), 97–113.

Litzenberger, Robert and Krishna Ramaswamy, 1982, "The Effects of Dividends on Common Stock Prices: Tax Effects or Information Effects?," *Journal of Finance*, 37, 429–443.

Litzenberger, Robert and Krishna Ramaswamy, 1980, "Dividends, Short Selling Restrictions, Tax Induced Investor Clientele and Market Equilibrium," *Journal of Finance*, 35, 469–482.

Litzenberger, Robert and Krishna Ramaswamy, 1979, "The Effects of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence," *Journal of Financial Economics*, 7, 163–195.

Long, John B., Jr., 1978, "The Market Valuation of Cash Dividends: A Case to Consider," *Journal of Financial Economics*, 6(2/3), 235–264.

Long, John B., 1977, "Efficient Portfolio Choice with Differential Taxation of Dividend and Capital Gains," *Journal of Financial Economics*, 5, 25–53.

Loughran, Tim and Jay Ritter, 1995, "The New Issues Puzzle," *Journal of Finance*, 50(1), 23–51.

Lucas, Deborah J. and Robert L. McDonald, 1998, "Shareholder Heterogeneity, Adverse Selection, and Payout Policy," *Journal of Financial and Quantitative Analysis*, 33(2).

Masulis, Ronald W. and Ashok N. Korwar, 1986, "Seasoned Equity Offerings: An Empirical Investigation," *Journal of Financial Economics*, 15(1/2), 91–118.

Michaely, Roni, 1991, "Ex-dividend Day Stock Price Behavior: The Case of the 1986 Tax Reform Act," *Journal of Finance*, 46, 845–860.

Michaely, Roni and Maurizio Murgia, 1995, "The Effect of Tax Heterogeneity on Prices and Volume Around the Ex-dividend Day: Evidence from the Milan Stock Exchange," *Review of Financial Studies.*

Michaely, Roni and Wayne H. Shaw, 1994, "The Pricing of Initial Public Offerings: Tests of the Adverse Selection and Signaling Theories," *Review of Financial Studies*, 7(2), 279–319.

Michaely, Roni, Richard H. Thaler and Kent Womack, 1995, "Price Reactions to Dividend Initiations and Omissions: Overreaction or Drift?," *Journal of Finance*, 50(2), 573–608.

Michaely, Roni and Jean-Luc Vila, 1996, "Trading Volume with Private Valuations: Evidence from the Ex-dividend Day," *Review of Financial Studies*, 9(2), 471–510.

Michaely, Roni and Jean-Luc Vila, 1995, "Investors' Heterogeneity, Prices and Volume Around the Ex-Dividend Day," *Journal of Financial and Quantitative Analysis*, 30, 171–198.

Michaely, Roni, Jean-Luc Vila and Jiang Wang 1997, "A Model of Trading Volume with Tax-induced Heterogeneous Valuation and Transaction Costs," *Journal of Financial Intermediation*.

Miller, James and John McConnell, 1995, "Open-Market Share Repurchase Programs and Bid-ask Spreads on the NYSE: Implications for Corporate Payout Policy," *Journal of Financial and Quantitative Analysis*, 30(3), 365–382.

Miller, Merton, 1987, "The Information Content of Dividends," In J. Bossons, R. Dornbush and S. Fischer (Eds.) *Macroeconomics: Essays in Honor of Franco Modigliani*, MIT Press, Cambridge, MA, 37–61.

Miller, Merton, 1977, "Debt and Taxes," *Journal of Finance*, 32(2), 261–275.

Miller, Merton and Franco Modigliani, 1961, "Dividend Policy, Growth and the Valuation of Shares," *Journal of Business*, 34, 411–433.

Miller, Merton and Kevin Rock, 1985, "Dividend Policy under Asymmetric Information," *Journal of Finance*, 40(4), 1031–1051.

Miller, Merton and Myron Scholes, 1982, "Dividends and Taxes: Empirical Evidence," *Journal of Political Economy*, 90, 1118–1141.

Miller, Merton and Myron Scholes, 1978, "Dividends and Taxes," *Journal of Financial Economics*, 6, 333–264.

Moody's dividend record, various years.

Morgan, Ian G., 1982, "Dividends and Capital Asset Prices," *Journal of Finance*, 37, 1071–1086.

Myers, Stewart C., 2000, "Outside Equity," *Journal of Finance*, 55(3), 1005–1038.

Myers, Stewart C., 1977, "Determinants of Corporate Borrowing," *Journal of Financial Economics*, 5(2), 147–175.

Myers, Stewart C. and Nicholas S. Majluf, 1984, "Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have," *Journal of Financial Economics*, 13(2), 187–221.

Naranjo Andy, M. Nimalendran and Mike Ryngaert, 1998, "Stock Returns, Dividend Yield and Taxes," *Journal of Finance*, 53(6), 2029–2057.

Nissim, Doron and Amir Ziv, 2001, "Dividend Changes and Future Profitability," *Journal of Finance*, 61(6), 2111–2134.

Nohel, T. and V. Tarhan, 1998, "Share Repurchases and Firm Performance: New Evidence on the Agency Costs of Free Cash Flow," *Journal of Financial Economics*, 49, 187–222.

Ofer, Aharon R. and Daniel R. Siegel, 1987, "Corporate Financial Policy, Information, and Market Expectations: An Empirical Investigation of Dividends," *Journal of Finance*, 42(4), 889–911.

Ofer, Aharon R. and Anjan V. Thakor, 1987, "A Theory of Stock Price Responses to Alternative Corporate Cash Disbursement Methods: Stock Repurchases and Dividends," *Journal of Finance*, 42(4), 365–394.

Penman, Stephen H., 1983, "The Predictive Content of Earnings Forecasts and Dividends," *Journal of Finance*, 38(4), 1181–1199.

Perez-Gonzalez, Francisco, 2000, "Large Shareholders and Dividends: Evidence from U.S. Tax Reforms," Working Paper, Harvard University, Cambridge, MA.

Peterson, Pamela, David Peterson and James Ang, 1985, "Direct Evidence on the Marginal Rate of Taxation on Dividend Income," *Journal of Financial Economics*, 14, 267–282.

Pettit, R. Richardson, 1972, "Dividend Announcements, Security Performance, and Capital Market Efficiency," *Journal of Finance*, 27(5), 993–1007.

Poterba, James, 1987, "Tax Policy and Corporate Savings," *Brookings Papers on Economic Activity*, 2, 455–503.

Poterba, James, 1986, "The Market Valuation of Cash Dividends: The Citizens Utilities Case Reconsidered," *Journal of Financial Economics*, 15, 395–406.

Poterba, James and Lawrence H. Summers, 1984, "New Evidence that Taxes Affect the Valuation of Dividends," *Journal of Finance*, 39, 1397–1415.

Poterba, James, Steven Venti and David Wise, 1996, "Do Retirement Saving Programs Increase Saving? Reconciling the Evidence," *Journal of Economic Perspectives*, 10(4).

Richardson, Gordon, Stephan Sefcik and Rex Thompson, 1986, "A Test of Dividend Irrelevance Using Volume Reactions to a Change in Dividend Policy," *Journal of Financial Economics*, 17, 313–333.

Ritter, Jay R., 1991, "The Long Run Performance of Initial Public Offerings," *Journal of Finance*, 46(1), 3–28.

Rosenberg, Bar and Vinay Marathe, 1979, "Tests of Capital Asset Pricing Hypotheses," *Research in Finance*, 1, 115–224.

Sarig, Oded, 1984, "Why Do Companies Pay Dividends?: Comment," *American Economic Review*, 74, 1142.

Shefrin, Hersh M. and Meir Statman, 1984, "Explaining Investor Preference for Cash Dividends," *Journal of Financial Economics*, 13(2), 253–282.

Shleifer Andrei and Robert Vishny, 1986, "Large Shareholders and Corporate Control," *Journal of Political Economy*, 94(3), 461–488.

Sialm, Clemens, 2006, "Investment Taxes and Equity Returns," Working Paper, University of Michigan.

Sialm, Clemens, 2005, "Tax Changes and Asset Pricing: Time-Series Evidence," Working Paper, University of Michigan.

Stephens, Clifford and Michael Weisbach, 1988, "Actual Share Reacquisitions in Open Market Repurchases Programs," *Journal of Finance*, 53(1), 313–333.

Stiglitz, Joseph E., 1983, "Some Aspects of the Taxation of Capital Gains," *Journal of Public Economics*, 21, 257–296.

Stulz, Rene, 1988, "Managerial Control of Voting Rights: Financing Policies and the Market for Corporate Control," *Journal of Financial Economics*, 25–54.

Thaler, Richard H. and Hersh M. Shefrin, 1981, "An Economic Theory of Self-Control," *Journal of Political Economy*, 89(2), 392–406.

Vermaelen, T., 1984, "Repurchase Tender Offers, Signaling and Managerial Incentives," *Journal of Financial and Quantitative Analysis*, 19, 163–181.

Vermaelen, Theo, 1981, "Common Stock Repurchases and Market Signaling: An Empirical Study," *Journal of Financial Economics*, 9(2), 138–183.

Watts, Ross, 1973, "The Information Content of Dividends," *Journal of Business*, 46(2), 191–211.

Weisbenner, Scott, 2000, "Corporate Share Repurchase in the Mid-1990's: What Role Do Stock Options Play," Working Paper, MIT, Cambridge, MA.

Williams, Joseph, 1988, "Efficient Signaling with Dividends, Investment, and Stock Repurchases," *Journal of Finance*, 43(3), 737–747.

Yoon, Pyung S. and Laura Starks, 1995, "Signaling, Investment Opportunities, and Dividend Announcements," *Review of Financial Studies*, 8(4), 995–1018.

Zwiebel, Jeffrey, 1996, "Dynamic Capital Structure under Managerial Entrenchment," *American Economic Review*, 86(5), 1197–1215.

*Chapter 11*

# TAXES AND CORPORATE FINANCE

JOHN R. GRAHAM*

*Fuqua School of Business, Duke University*

## Contents



* I thank Roseanne Altshuler, Alan Auerbach, Alon Brav, Merle Erickson, Ben Esty, Mary Margaret Frank, Michelle Hanlon, Cam Harvey, Steve Huddart, Ravi Jagannathan, Mark Leary, Jennifer Koski, Alan Kraus, Ed Maydew, Bob McDonald, Roni Michaely, Lil Mills, Kaye Newberry, Maureen O'Hara, Jeff Pittman, Michael Roberts, Doug Shackelford, and Terry Shevlin for helpful comments. I also thank Tao Lin, Rujing Meng, and especially Vinny Eng and Krishna Narasimhan for their excellent research assistance. I apologize to those who feel that their research has been ignored or misrepresented. Any errors are mine. A more focused version of some of the material in this chapter appears in Graham (2003). This research is partially funded by the Alfred P. Sloan Research Foundation.

*Handbook of Empirical Corporate Finance, Volume 2*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

## Abstract

This chapter reviews tax research related to domestic and multinational capital structure, debt maturity, payout policy, compensation policy, risk management, earnings management, leasing, pensions, R&D partnerships, tax shelters, transfer pricing, and organizational form. For each topic, the theoretical arguments explaining how taxes can affect corporate decision making and firm value are reviewed, followed by a summary of the related empirical evidence and a discussion of unresolved issues. Tax research generally supports the hypothesis that high-tax rate firms pursue policies that provide tax benefits. Many issues remain unresolved, however, including understanding whether tax effects are of first-order importance, why firms do not pursue tax benefits more aggressively, and whether investor-level taxes affect corporate actions.

## Keywords

capital structure, corporate finance, compensation, dividends, payout policy, taxes

## 1. Introduction

Modigliani and Miller (1958) and Miller and Modigliani (1961) demonstrate that corporate financial decisions are irrelevant in a perfect, frictionless world. Modigliani and Miller (MM) assume that capital markets are perfect, which implies that there are no corporate or personal taxes, among other things. During the past 45 years, research has focused on whether financial decisions become relevant if capital markets are not perfect. The research reviewed in this chapter investigates the consequences of allowing corporate and personal taxation, highlighting the role of corporate and investor taxes in corporate policies and firm value.[1] This role is potentially very important, given the sizable tax rates that many corporations and individuals face (see Fig. 1).

Modigliani and Miller argue that corporate financial policies do not add value in equilibrium, and therefore firm value equals the present value of operating cash flows. Once imperfections are introduced, however, corporate financial policies can affect firm

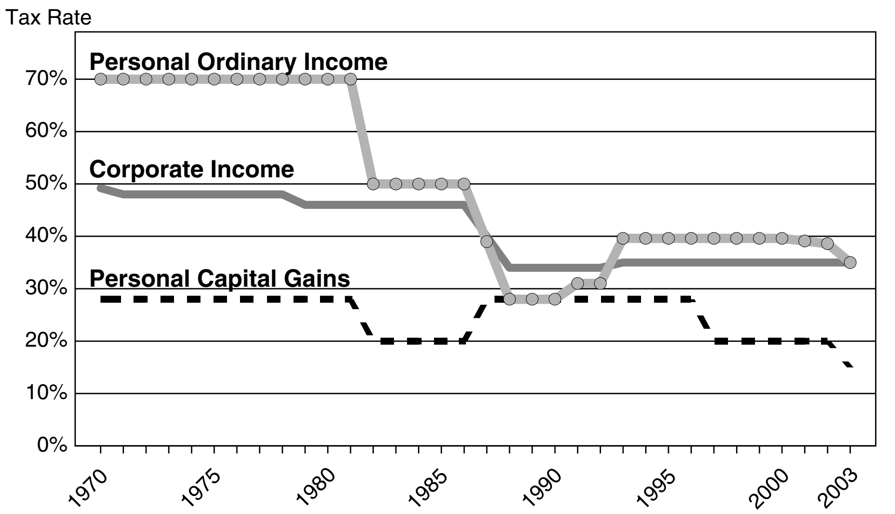



Fig. 1. Corporate and Personal Income Tax Rates. The highest tax bracket statutory rates are shown for individuals and C-corporations. The corporate capital gains tax rate (not shown) was equal to the corporate income tax rate every year after 1987. In May 2003, President Bush signed into law a reduction in the top personal income tax rate to 35% in 2003. This same law reduced top personal tax rates on capital gains and dividends to 15%.
*Source* for pre-2003 numbers: Commerce Clearing House, annual publications.

[1] The interested reader can find excellent reviews of how taxes affect household investment decisions (Poterba, 2001) and the current state of tax research from the perspective of accountants (Shackelford and Shevlin, 2001) and public economists (Auerbach, 2002). Articles reviewing how nontax factors such as agency and informational imperfections affect corporate financial decisions can be found in the other chapters of this handbook.

value, and firms should pursue a given policy until the marginal benefit of doing so equals the marginal cost. A common theme in tax research involves expressing how various tax rules and regulations affect the marginal benefit of corporate actions. For example, when tax rules allow interest deductibility, a \$1 interest deduction provides tax savings of $\$1 \times \tau_C(.)$; $\tau_C(.)$ measures corporate marginal tax benefits and is a function of statutory tax rates, nondebt tax shields, the probability of experiencing a loss, international tax rules about dividend imputation and interest allocation, organizational form, and various other tax rules. A common theme that runs throughout this chapter is the demonstration of how various tax rules affect the $\tau_C(.)$ benefit function, and therefore how they affect corporate incentives and decisions. A second but less common theme in tax research is related to how market imperfections affect costs. Given that this chapter reviews tax research, the emphasis is on research that describes how taxes affect costs and benefits—and the influence of nontax factors is discussed only briefly.

There are multiple avenues for taxes to affect corporate decisions. As outlined above, taxes can affect capital structure decisions, both domestic (Section 2) and multinational (Section 3), organizational form and restructurings (Section 4), payout policy (Section 5), compensation policy (Section 6), risk management (Section 7), and the use of tax shelters (Section 8). For each of these areas, the sections that follow provide a theoretical framework describing how taxes might affect corporate decisions, empirical predictions based on the theory, and summaries of the related empirical evidence. This approach seeks to highlight important questions about how taxes affect corporate decisions, and to summarize and, in some cases, critique the answers that have been thus far provided. Each section concludes with a discussion of unanswered questions and possible avenues for future research. Overall, substantial progress has been made in the investigation of whether and how taxes affect corporate financial decisions, but much work remains to be done. Section 9 concludes and proposes directions for future research.

## 2. Taxes and capital structure—the U.S. tax system

### 2.1. *Theory and empirical predictions*

This section reviews capital structure research that is related to the "classical" tax system found in the United States. (Section 3 reviews multinational and imputation tax systems.) The key features of the classical system are that corporate income is taxed at a rate $\tau_C$, interest is deductible and so is paid out of income before taxes, and equity payout is not deductible but is paid from the residual remaining after corporate taxation. In this tax system, interest, dividends, and capital gains income are taxed upon receipt by investors (at tax rates $\tau_P$, $\tau_{div} = \tau_P$, and $\tau_G$, respectively). Most of the research assumes that equity is the marginal source of funds and that dividends are paid according to a fixed payout policy.[2] To narrow the discussion, it is assumed that regulations or transactions

[2] This assumption implies that retained earnings are not "trapped equity" that is implicitly taxed at the dividend tax rate, even while still retained. See Auerbach (2002) for more on the trapped equity or "new" view.

costs prevent investors from following the tax-avoidance schemes implied by Miller and Scholes (1978), in which investors borrow via insurance or other tax-free vehicles to avoid personal tax on interest or dividend income.

In this framework, the after-personal-tax value to investors of a corporation paying \$1 of interest is \$1$(1 - \tau_P)$. In contrast, if that capital were instead returned as equity income, it would be subject to taxation at both the corporate and personal level, and the investor would receive \$1$(1 - \tau_C)(1[-\tau_E)$. The equity tax rate, $\tau_E$, is often modeled as a blended dividend and capital gains tax rate.[3] The net tax advantage of \$1 of debt payout, relative to \$1 of equity payout, is

$$(1 - \tau_P) - (1 - \tau_C)(1 - \tau_E) \tag{1}$$

If Equation (1) is positive, debt interest is the tax-favored way to return capital to investors, once both corporate and individual taxation are considered. In this case, in order to maximize firm value, a company has a tax incentive to issue debt instead of equity.

Equation (1) captures the benefit of a firm paying out \$1 as debt interest in the current period, relative to paying out \$1 as equity income. If a firm has \$D of debt with coupon rate $r_D$, the net benefit of using debt rather than equity is

$$[(1 - \tau_P) - (1 - \tau_C)(1 - \tau_E)]r_D D \tag{2}$$

Given this expression, the value of a firm with debt can be written as

$$\text{Value}_{\text{with debt}} = \text{Value}_{\text{no debt}} + \text{PV}[(1 - \tau_P) - (1 - \tau_C)(1 - \tau_E)]r_D D \tag{3}$$

where the PV term measures the present value of all current and future interest deductions. Note that Equation (3) implicitly assumes that using debt adds tax benefits but has no other effect on incentives, operations, or value.[4]

Modigliani and Miller (1958) is the seminal capital structure paper. If capital markets are perfect, $\tau_C$, $\tau_P$, and $\tau_E$ all equal zero, and it does not matter whether the firm finances with debt or equity (i.e., $\text{Value}_{\text{with debt}} = \text{Value}_{\text{no debt}}$). That is, the value of the firm equals the value of equity plus the value of debt, but total value is not affected by the proportions of debt and equity. This implication is used as the null throughout the capital structure discussion.

Null hypotheses: Firms do not have optimal tax-driven capital structures. The value of a firm with debt is equal to the value of an identical firm without debt (i.e., there is no net tax advantage to debt).

[3] In mid-2003, Congress passed a law that reduced the tax rate on both dividends and capital gains to 15% for individual investors, thereby simplifying and greatly reducing the level of equity taxation relative to historic levels.

[4] There are other approaches to modeling the tax benefits of debt that do not fit directly into this general framework. For example, Goldstein, Ju, and Leland (2001) have developed a dynamic contingent-claims model in which firms can restructure debt. They estimate that the tax benefits of debt should equal between 8 and 9% of firm value. See Goldstein et al. for references to other contingent-claims models.

In their "correction article," MM (1963) consider corporate income taxation but continue to assume that $\tau_P$ and $\tau_E$ equal zero. In this case, the second term in Equation (3) collapses to PV[$\tau_C r_D D$]: Because interest is deductible, paying \$$r_D D$ of interest saves $\tau_C r_D D$ in taxes each period relative to returning capital as equity. MM (1963) assume that interest deductions are as risky as the debt that generates them and should be discounted by $r_D$.[5] With perpetual debt, MM (1963) argue that the value of a firm with debt financing is

$$V_{\text{with debt}} = V_{\text{no debt}} + \frac{\tau_C r_D D}{r_D} = V_{\text{no debt}} + \tau_C D \tag{4}$$

where the $\tau_C D$ term represents the tax advantage of debt. Note that Equation (4) contains a term that captures the tax benefit of using debt ($\tau_C D$) but no offsetting cost of debt term. Equation (4) has two strong implications. First, corporations should finance with 100% debt because the marginal benefit of debt is $\tau_C$, which is often assumed to be a positive constant. Second, if $\tau_C$ is constant, firm value increases (linearly) with D due to tax benefits

Because the first implication was recognized as extreme, researchers developed models that relax the MM (1958) assumptions and consider costs of debt. In the early models, firms trade off the tax benefits of debt with costs. The first cost proposed in the literature was the cost of bankruptcy, or more generally, costs of financial distress. Kraus and Litzenberger (1973), using a state-preference framework, show that firms should trade off bankruptcy costs with the tax benefits of debt to arrive at an optimal capital structure that involves less than 100% debt. Scott (1976) shows the same thing with continuous variables. The bankruptcy cost solution does not appear empirically to ex ante offset the benefits of debt.[6] Therefore, other papers have proposed non-bankruptcy costs that could be traded off against the tax benefits of debt. For example,

[5] The assumption that debt should be discounted at $r_D$ is controversial because it requires the amount of debt to remain fixed. Miles and Ezzell (1985) demonstrate that if the dollar amount of debt is not fixed but instead is set to maintain a target debt-equity ratio, then interest deductions have equity risk and should be discounted with the return on assets, $r_A$, rather than $r_D$. (Miles and Ezzell, 1985, allow first-period financing to be fixed, which requires adjusting the discount rate by $(1+r_A)/(1+r_D)$). In contrast, Grinblatt and Titman (2002) argue that firms often pay down debt when things are going well and stock returns are high, and do not alter debt when returns are low. Such behavior can produce a low or negative beta for debt and hence a low discount rate for the tax benefits of debt. In either the Miles and Ezzell or Grinblatt and Titman case, however, the value of a levered firm still equals the value of the unlevered firm plus a "coefficient times debt" term—the discounting controversy only affects the coefficient.

[6] Warner (1977) show that direct costs of bankruptcy average no more than 5.3% ex post in railroad bankruptcies. More recently, Andrade and Kaplan (1998) show that the ex-post costs of distress brought about by financing choice amount to 20% of firm value for a group of industrial firms. Miller (1977) note that firms choose optimal debt policy by considering ex-ante costs of distress, indicating that the costs mentioned above need to be multiplied by the conditional probability of distress to measure ex-ante costs. Miller point out that the ex-ante costs of financial distress appear to be very small compared to the apparently large tax benefits of debt.

Jensen and Meckling (1976) introduce agency costs of equity and leverage-related deadweight costs.[7] Myers (1977) introduces underinvestment costs that can result from too much debt.

Regardless of the type of cost, the basic trade-off implications remain similar to those in MM (1963): (1) the incentive to finance with debt increases with the corporate tax rate, and (2) firm value increases with the use of debt (up to the point where the marginal cost equals the marginal benefit of debt). Note also that in these models, different firms can have different optimal debt ratios depending on the relative costs and benefits of debt (i.e., depending on differing firm characteristics).

Prediction 1: All else constant, for taxable firms, value increases with the use of debt because of tax benefits (up to the point where the marginal cost equals the marginal benefit of debt).

Prediction 2: Corporations have a tax incentive to finance with debt that increases with the corporate marginal tax rate. All else equal, this implies that firms have differing optimal debt ratios if their tax rates differ.

Prediction 1 is based directly on Equation (4), whereas Prediction 2 is based on the first derivative of Equation (4) with respect to D.

Miller (1977) argues that personal taxes can eliminate the "100% debt" implication, without the need for bankruptcy or agency costs. (Farrar and Selwyn, 1967, took first steps in this direction.) Miller's argument is that the marginal costs of debt and equity, net of the effects of personal and corporate taxes, should be equal in equilibrium, so firms are indifferent between the two financing sources. In essence, the corporate tax savings from debt is offset by the personal tax disadvantage to investors from holding debt, relative to holding equity. All else equal (including risk), this personal tax disadvantage causes investors to demand higher pretax returns on debt, relative to equity returns. From the firm's perspective, paying this higher pretax return wipes out the tax advantage of using debt financing.

Figure 2 illustrates Miller's point. The horizontal line in Panel A depicts the supply curve for debt; the line is horizontal because Miller assumes that the benefit of debt for all firms equals a fixed constant $\tau_C$. The demand for debt curve is initially horizontal at zero, representing demand by tax-free investors, but eventually slopes upward because the return on debt must increase to attract investors with higher personal income tax rates. By making the simplifying assumption that $\tau_E = 0$, Miller's equilibrium is reached when the marginal investor with $\tau_P^* = \tau_C$ is attracted to purchase debt. In this equilibrium, the entire surplus (the area between the supply and demand curves) accrues to investors subject to personal tax rates less than $\tau_P^*$.

Miller's (1977) analysis has several implications. The first two are new:

[7] Parrino and Weisbach (1999) use simulations to conclude that the agency costs of debt are too small to offset the tax benefits, and Esty (1998) empirically examines the effects of agency costs on capital structure in the banking industry.

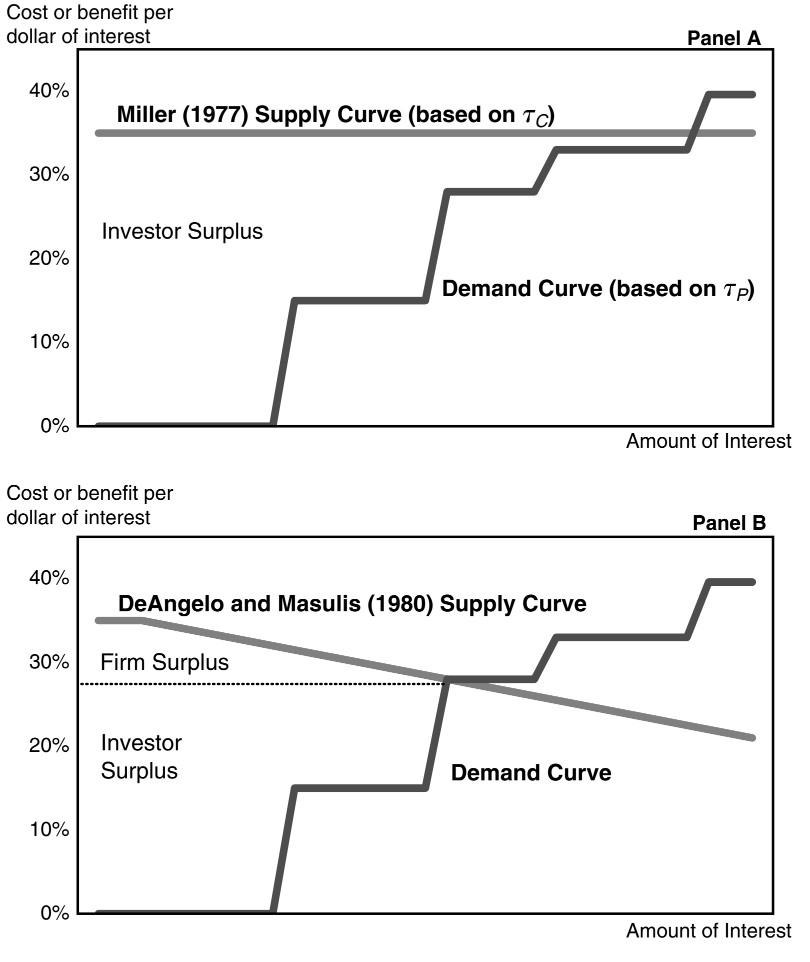



Fig. 2. Equilibrium Supply and Demand Curves for Corporate Debt. The supply curve shows the expected tax rate (and therefore the tax benefit of a dollar of interest) for the firms that issue debt. The demand curve shows the tax rate (and therefore the tax cost of a dollar of interest) for the investors that purchase debt. The tax rates for the marginal supplier of and investor in debt are determined by the intersection of the two curves. In the Miller Equilibrium (panel A), all firms have the same tax rate in every state of nature, so the supply curve is flat. The demand curve slopes upward because tax-free investors are the initial purchasers of corporate bonds, followed by low-tax-rate investors, and eventually followed by high tax-rate-investors. In the Miller Equilibrium, all investors with tax rates less than the marginal investor's (i.e., investors with tax rates of 33% or less in Panel A) are inframarginal and enjoy an "investor surplus" in the form of an after-tax return on debt higher than their reservation return. In Panel B, the supply curve is downward sloping because firms differ in terms of the probability that they can fully utilize interest deductions (or have varying amounts of nondebt tax shields), and therefore have differing benefits of interest deductibility. Firms with tax rates higher than that for the marginal supplier of debt (i.e., firms with tax rates greater than 28% in Panel B) are inframarginal and enjoy "firm surplus" because the benefit of interest deductibility is larger than the personal tax cost implicit in the debt interest rate.

Prediction 3: High personal taxes on interest income (relative to personal taxes on equity income) create a disincentive for firms to use debt.

Prediction 4: The aggregate supply of debt is affected by relative corporate and personal taxes.

The other implications are consistent with the null hypotheses stated above: (1) there is no net tax advantage to debt at the corporate level (once one accounts for the higher debt yields investors demand because of the relatively high personal taxes associated with receiving interest), (2) though taxes affect the aggregate supply of debt in equilibrium, they do not affect the optimal capital structure for any particular firm (i.e., it does not matter which particular firms issue debt, as long as aggregate supply equals aggregate demand), and (3) using debt does not increase firm value.

A general version of Miller's argument (that does not assume $\tau_E = 0$) can be expressed in terms of Equation (3). Once personal taxes are introduced into this framework, the appropriate discount rate is measured after-personal income taxes to capture the (after-personal-tax) opportunity cost of investing in debt. In this case, the value of a firm using perpetual debt is[8]:

$$\begin{aligned} V_{\text{with debt}} &= V_{\text{no debt}} + \frac{[(1-\tau_P) - (1-\tau_C)(1-\tau_E)]r_D D}{(1-\tau_P)r_D} \\ &= V_{\text{no debt}} + \left[1 - \frac{(1-\tau_C)(1-\tau_E)}{(1-\tau_P)}\right] D \end{aligned} \tag{5}$$

If the investor-level tax on interest income ($\tau_P$) is large relative to tax rates on corporate and equity income ($\tau_C$ and $\tau_E$), the net tax advantage of debt can be zero or even negative. Note that Equation (5) is identical to Equation (4) if there are no personal taxes, or if $\tau_P = \tau_E$.

One way that Equation (5) can be an equilibrium expression is for the rightmost term in this equation to equal zero in equilibrium (e.g., $(1-\tau_P) = (1-\tau_C)(1-\tau_E)$), in which case the implications from Miller (1977) are unchanged. Alternatively, the tax benefit term in Equation (5) can be positive, and a separate cost term can be introduced in the spirit of the trade-off models. In this case, the corporate incentive to issue debt and firm value both increase with $[1 - (1-\tau_C)(1-\tau_E)/(1-\tau_P)]$ and firm-specific optimal debt ratios can exist. The bracketed expression specifies the degree to which personal taxes (Prediction 3) offset the corporate incentive to use debt (Prediction 2). Recall that $\tau_P$ and $\tau_E$ are personal tax rates for the marginal investor(s) and therefore are difficult to pin down empirically (more on this in Section 1.4).

DeAngelo and Masulis (1980; hereafter DM) broaden Miller's (1977) model and put the focus on the marginal tax benefit of debt, represented earlier by $\tau_C$. DM argue that $\tau_C(.)$ is not constant and is always equal to the statutory rate. Instead, $\tau_C(.)$ is a function

[8] See Sick (1990), Taggart (1991), or Benninga and Sarig (1997) for derivation of expressions like Equation (5) under various discounting assumptions. These expressions are of the form $V_{\text{with debt}} = V_{\text{no debt}} +$ coefficient*D, with the coefficient an increasing (decreasing) function of corporate (personal income) tax rates.

that decreases in nondebt tax shields (e.g., depreciation and investment tax credits) because nondebt tax shields (NDTS) crowd out the tax benefit of interest. Furthermore, Kim (1989) highlights the fact that firms do not always benefit fully from incremental interest deductions because they are not taxed when taxable income is negative. This implies that $\tau_C(.)$ is a decreasing function of a firm's debt usage because existing interest deductions crowd out the tax benefit of incremental interest.

Modeling $\tau_C(.)$ as a function has important implications because the supply of debt function can become downward sloping (see Panel B in Fig. 2). This implies that using debt has a corporate advantage, as measured by the "firm surplus" of issuing debt (the area above the dotted line but below the supply curve in Panel B). Moreover, high-tax-rate firms supply debt (i.e., are on the portion of the supply curve to the left of its intersection with demand), which implies that tax-driven firm-specific optimal debt ratios can exist (as in Prediction 2), and that the tax benefits of debt add value for high-tax-rate firms (as in Prediction 1). The DeAngelo and Masulis (1980) approach leads to the following prediction, which essentially expands Prediction 2:

Prediction $2'$: All else equal, to the extent that they reduce $\tau_C(.)$, nondebt tax shields and/or interest deductions from already-existing debt reduce the tax incentive to use debt. Similarly, the tax incentive to use debt decreases with the probability that a firm will experience nontaxable states of the world.

### 2.2. *Empirical evidence on whether the tax advantage of debt Increases firm value*

Prediction 1 indicates that the tax benefits of debt add $\tau_C D$ (Equation 4) or $[1-(1-\tau_C)(1-\tau_E)/(1-\tau_P)]D$ (Equation 5) to firm value. If $\tau_C = 40\%$ and the debt ratio is 35%, Equation (4) indicates that the contribution of taxes to firm value equals 14% ($0.14 = \tau_C \times$ debt-to-value). This calculation is an upper bound, however, because it ignores costs and other factors that reduce the corporate tax benefit of interest deductibility, such as personal taxes, nontax costs of debt, and the possibility that interest deductions are not fully valued in every state of the world. This section reviews empirical research that attempts to quantify the net tax benefits of debt. The first group of papers study market reactions to exchange offers, which should net out the various costs and benefits of debt. The remainder of the section reviews recent analyses based on large-sample regressions and concludes by examining explicit benefit functions for interest deductions.

#### 2.2.1. *Exchange offers*

To investigate whether the tax benefits of debt increase firm value (Prediction 1), Masulis (1980) examines exchange offers made during the 1960s and 1970s. Because one security is issued and another is simultaneously retired in an exchange offer, Masulis argues that exchanges hold investment policy relatively constant and are primarily changes in capital structure. Masulis's tax hypothesis is that leverage-increasing (-decreasing) exchange offers increase (decrease) firm value because they increase (decrease) tax deductions.

Note that Masulis implicitly assumes that firms are underlevered. For a company already at its optimum, a movement in either direction (i.e., increasing or decreasing debt) would decrease firm value.

Masulis (1980) finds evidence consistent with his predictions: leverage-increasing exchange offers increase equity value by 7.6%, and leverage-decreasing transactions decrease value by 5.4%. Moreover, the exchange offers with the largest increases in tax deductions (debt-for-common and debt-for-preferred) have the largest positive stock price reactions (9.8% and 4.7%, respectively). Using a similar sample, Masulis (1983) regresses stock returns on the change in debt in exchange offers and finds a debt coefficient of approximately 0.40 (which is statistically indistinguishable from the top statutory corporate tax rate at that era). This is consistent with taxes increasing firm value as in Equation (4) (and is also consistent with some alternative hypotheses discussed below), but it is surprising because such a large coefficient implies near-zero personal tax and nontax costs to debt. That is, the debt coefficient in Masulis (1983) measures the average benefit of debt (averaged across firms and averaged over the incremental net benefit of each dollar of debt for a given firm) *net of the costs*. An average net benefit of 0.40 requires that the costs are much smaller than the benefits for most dollars of debt. For the post-exchange offer capital structure to satisfy the MB = MC equilibrium condition, the benefit or cost curves (or both) must be very steeply sloped near their intersection.

Myers (1984) and Cornett and Travlos (1989) argue that Masulis's (1980) hypothesis is problematic. If firms optimize, they should only adjust capital structure to move toward an optimal debt ratio, whether that involves increasing debt or equity. In other words, increasing debt will not always add to firm value, even if interest reduces tax liabilities. Graham, Hughson, and Zender (1999) point out that if a firm starts at its optimal capital structure, it will only perform an exchange offer if something moves the firm out of equilibrium. They derive conditions under which stock price-maximizing exchanges are unrelated to marginal tax rates because market reactions aggregate tax and nontax informational aspects of capital structure changes. Therefore, nontax reactions might explain Masulis's (1980) results. As described next, several papers have found evidence of nontax factors affecting exchange offer market reactions. It is important to note that these post-Masulis papers do not prove that the tax interpretation is wrong—but they do offer alternative interpretations.

First, some papers find evidence of positive (negative) stock reactions to leverage-increasing (leverage-decreasing) events that are unrelated to tax deductions: Asquith and Mullins (1986), Masulis and Korwar (1986), and Mikkelson and Partch (1986) find negative stock price reactions to straight equity issuance, and Pinegar and Lease (1986) find positive stock price reactions to preferred-for-common exchanges. Second, Mikkelson and Partch (1986) and Eckbo (1986) report that straight debt issuance (without equity retirement) produces a stock price reaction that is indistinguishable from zero. Third, some papers find that exchange offers convey nontax information that affects security prices, perhaps due to asymmetric information problems along the lines suggested by Myers and Majluf (1984) or due to signaling (Ross, 1977) and Leland and Pyle, 1977). For example, Shah (1994) correlates exchange offers with information about reduced future cash flows (for leverage-decreasing offers) and decreased risk (for leverage-increasing offers). Finally, Cornett and Travlos (1989) provide evidence that weakens Masulis's (1983) conclusions. Cornett and Travlos regress event stock returns on the change in debt and two variables that control for information effects (the ex-post change in inside ownership and ex-post abnormal earnings). They find that the coefficient on the change in debt variable is insignificant while the coefficients on the other variables are significant, which implies that the positive stock price reaction is related to positive information conveyed by the exchange.[9] Cornett and Travlos conclude that equity-for-debt exchanges convey information about the future—but find no evidence of increased value due to tax benefits.

Two recent papers examine the exchange of traditional preferred stock for monthly income preferred stock (MIPS). These two securities differ primarily in terms of their tax characteristics, so any market reaction should have minimal nontax explanations. MIPS interest is tax deductible for corporations (like debt interest), and preferred dividends are not. On the investor side, corporate investors enjoy a 70% dividends received deduction (DRD) for preferred dividends, but recipients of MIPS interest receive no parallel deduction.[10] When issuing MIPS to retire preferred, corporations gain the tax benefit of interest deductibility but experience two costs: underwriting costs and possibly an increased coupon due to the personal tax penalty (because investors are fully taxed on MIPS interest in contrast to corporate investors, who receive the DRD on preferred dividends). Engel, Erickson, and Maydew (1999) compare MIPS yields to preferred yields and conclude that the tax benefits of MIPS are approximately \$0.28 per dollar of face value, net of the aforementioned costs. Irvine and Rosenfeld (2000) use abnormal announcement returns to estimate the value at \$0.26. Given that MIPS and preferred are nearly identical in all legal and informational respects, these studies provide straightforward evidence of the positive contribution of taxes to firm value, net of underwriting and personal tax costs.

### 2.2.2. *Cross-sectional regressions*

Fama and French (1998; hereafter FF) attempt to estimate Equation (4) and Prediction 1 directly, by regressing $V_L$ on debt interest, dividends, and a proxy for $V_U$. They argue that a positive coefficient on interest is evidence of positive tax benefits of debt. FF measure $V_L$ as the excess of market value over book assets. They proxy $V_U$ with a collection of control variables, including current earnings, assets, and R&D spending, as well as future changes in these same variables. (All the variables in the regression are deflated by assets.)

[9] Cornett and Travlos do not report whether they get a significant positive tax coefficient (like Masulis, (1983, did) when they exclude the information variables. Therefore, their results could be driven by their use of a sample different from the one Masulis used.

[10] A 70% DRD means that a corporation that owns another firm's stock only pays tax on 30% of the dividends received. Note that evidence in Erickson and Maydew (1998) implies that corporations are the marginal investor in preferred stock (see footnote 27).

If these control variables provide adequate proxy for $V_U$, the regression coefficient on interest will measure the tax benefit of debt (which is hypothesized to be positive). The main difficulty with this approach is that if the control variables measure $V_U$ with error, the regression coefficients can be biased. FF perform a series of regressions on a broad cross section of firms, using both level-form and first-difference specifications. In all cases, the coefficient on interest is either insignificant or negative. Fama and French interpret their results as being inconsistent with debt tax benefits having a first-order effect on firm value. Instead, they argue that interest provides information about earnings that is not otherwise captured by their controls for $V_U$. In other words, $V_U$ is measured with error, which results in the interest coefficient picking up a negative valuation effect related to financial distress or some other cost.

Kemsley and Nissim (2002) attempt to circumvent this measurement problem. They perform a switch of variables, moving the earnings variable (which they assume proxies $V_U$ with error) to the left-hand side of the regression and $V_L$ to the right-side. Therefore, their regression tests the relation $V_U = V_L -$ coeff*D.

When Kemsley and Nissim regress EBIT on $V_L$ and debt, the debt coefficient is negative, which they interpret as evidence that debt contributes to firm value. The coefficient also changes through time in conjunction with changes in statutory tax rates. The Kemsley and Nissim analysis should be interpreted carefully. First, their regression specification can be interpreted as measuring the effect of debt on earnings, just as well as it can be interpreted as a switch-of-variables that fixes a measurement error problem in Fama and French (1998). Second, the debt coefficient has the correct sign for the full sample only in a nonlinear specification in which all the right-hand side variables are interacted with a crude measure of the discount rate. Finally, the coefficient that measures the net benefit of debt has an absolute value of 0.40. While consistent with Masulis (1983), such a large coefficient implies near-zero average debt costs and a near-zero effect of personal taxes.

### 2.2.3. *Marginal benefit functions*

Using a different approach, Graham (2000, 2001) simulates interest deduction benefit functions and uses them to estimate the tax-reducing value of each incremental dollar of interest expense. For a given level of interest deductions, Graham essentially integrates over possible states of the world (i.e., both taxable and nontaxable states) to determine a firm's expected $\tau_C$, which specifies the expected tax benefit of an incremental dollar of interest deduction. Marginal tax benefits of debt decline as more debt is added because the probability increases with each incremental dollar of interest that it will not be fully valued in every state of the world. Using simulation methods (described more fully in Section 1.3.2) and various levels of interest deductions, Graham maps out firm-specific interest benefit functions analogous to the supply of debt curve in Panel B of Figure 2.

By integrating under these benefit functions, Graham (2000) estimates that the tax benefit of debt equals approximately 9 to 10% of firm value during 1980–1994 (ignoring

Table 1

Annual calculations of the mean benefits of debt and degree of debt conservatism

*Before-financing MTR* is the mean Graham (1996) simulated corporate marginal tax rate based on earnings before interest deductions, and *after-financing MTR* is the same based on earnings after interest deductions. *Kink* is the multiple by which interest payments could increase without a firm experiencing reduced marginal benefit on incremental deductions (i.e., the amount of interest at the point at which a firm's marginal benefit function becomes downward sloping, divided by actual interest expense) as in Graham (2000). The *tax benefit of debt* is the reduction in corporate and state tax liabilities occurring because interest expense is tax deductible, expressed as a percentage of firm value. *Money left on the table* is the additional tax benefit that could be obtained, ignoring all costs, if firms with kink greater than one increased their interest deductions in proportion with kink.

| | Before-Financing MTR | After-Financing MTR | Kink | Tax Benefit of Debt | Money Left on Table |
|---|---|---|---|---|---|
| 1980 | 0.415 | 0.324 | 3.10 | 10.1 | 27.7 |
| 1981 | 0.413 | 0.319 | 2.98 | 11.4 | 28.6 |
| 1982 | 0.397 | 0.286 | 2.69 | 11.0 | 23.2 |
| 1983 | 0.388 | 0.282 | 2.68 | 10.7 | 22.5 |
| 1984 | 0.380 | 0.275 | 2.75 | 10.9 | 21.6 |
| 1985 | 0.366 | 0.255 | 2.51 | 11.1 | 21.8 |
| 1986 | 0.356 | 0.241 | 2.39 | 11.6 | 20.5 |
| 1987 | 0.296 | 0.198 | 2.35 | 10.7 | 19.5 |
| 1988 | 0.259 | 0.172 | 2.30 | 9.9 | 16.7 |
| 1989 | 0.258 | 0.169 | 2.24 | 10.6 | 15.8 |
| 1990 | 0.258 | 0.164 | 2.08 | 10.7 | 15.3 |
| 1991 | 0.257 | 0.160 | 1.99 | 9.6 | 11.7 |
| 1992 | 0.258 | 0.165 | 2.07 | 8.7 | 9.7 |
| 1993 | 0.236 | 0.175 | 1.71 | 7.7 | 8.0 |
| 1994 | 0.249 | 0.183 | 1.94 | 7.3 | 8.5 |
| 1995 | 0.260 | 0.207 | 1.99 | 7.8 | 9.8 |
| 1996 | 0.261 | 0.185 | 2.05 | 9.8 | 12.2 |
| 1997 | 0.261 | 0.188 | 2.08 | 9.1 | 10.9 |
| 1998 | 0.252 | 0.165 | 2.00 | 9.5 | 10.7 |
| 1999 | 0.252 | 0.170 | 1.90 | 7.7 | 8.9 |

all costs). Updating Graham's estimates, we find that the tax benefit of debt is 7.8%, 9.8%, 9.1%, 9.5%, and 7.7% of firm value in 1995–1999, respectively (see Table 1). The fact that these figures are less than the 14% estimated (at the beginning of Section 1) with the back of the envelope "$\tau_C D$" calculation reflects the reduced value of interest deductions in some states of the world. When personal taxes are considered, the tax benefit of debt falls to 7–8% of firm value during 1980–1994 (i.e., this is Graham's estimate of the "firm surplus" in Panel B of Fig. 2).

Graham also estimates the "money left on the table" that firms could obtain if they levered up to the point where their last dollar of interest deduction is valued at the full statutory tax rate (i.e., the "kink," which is the point just before incremental tax benefits begin to decline).[11] The money left on the table calculations in Graham (2000, his Fig. 2) is updated. If all firms lever up to operate at the kink in their benefit functions, they could add 10.5% to firm value over the 1995–1999 period (see Table 1). This number can be interpreted either as a measure of the value loss due to conservative corporate debt policy, or as a lower bound for the difficult-to-measure costs of debt that would occur if a company were to lever up to its kink. In the former interpretation, these estimates imply that large tax benefits of debt appear to go unexploited and that large, profitable firms (which would seem to face the lowest costs of debt) are the most conservative in their use of debt.[12] In general, these implications are hard for a trade-off model to explain. Graham (2000), Lemmon and Zender (2001) and Minton and Wruck (2001) try to identify nontax costs that are large enough in a trade-off sense that perhaps these firms are not in fact underlevered.

To sum up, a fair amount of research has found evidence consistent with tax benefits adding to firm value. However, some of this evidence is ambiguous because nontax explanations or econometric issues cloud interpretation. Additional research in three specific areas would be helpful. First, we need more market-based research along the lines of the MIPS exchanges, where tax effects are isolated from information and other factors and therefore the interpretation is fairly unambiguous. Second, additional cross-sectional regression research that investigates the market value of the tax benefits of debt would be helpful in terms of clarifying or confirming the interpretation of existing cross-sectional regression analysis. Finally, if the tax benefits of debt do in fact add to firm value, an important unanswered question is why firms do not use more debt, especially large, profitable firms.[13] We need to better understand whether this implies that some firms are not optimizing, or whether previous research has not adequately modeled costs and other influences.

### *2.3. Empirical evidence on whether corporate taxes affect debt vs. equity policy*

Trade-off models imply that firms should issue debt as long as the marginal benefit of doing so (measured by $\tau_C$) is larger than the marginal cost. $\tau_C(.)$ is a decreasing function of nondebt tax shields, existing debt tax shields, and the probability of

[11] For example, if during 1995–1999 all firms levered up to just before the point of declining benefit, simulations performed for this chapter indicate that the average company would have total tax benefits of debt of around 18% of firm value. That is, by levering up, the typical firm could add interest deductions with tax benefit equal to 10% of firm value, above and beyond their current level of tax benefits.

[12] McDonald (2002) argues that the prevalence of writing puts or purchasing calls on their own shares is also evidence that many firms pass up potential interest deductions. For example, writing a put (which involves implicit borrowing) can be replicated by explicitly borrowing today to purchase a share on the open market and repaying the loan in the future. The cash flows are identical in these two strategies, but the latter results in the firm receiving a tax deduction. The fact that many firms write puts is consistent with them passing up interest tax deductions.

[13] Shyum-Sunder and Myers (1999), Lemmon and Zender (2002), and related papers investigate whether the trade-off model is the correct model of capital structure, which has implications for interpreting these results.

experiencing losses, so the incentive to use debt declines with these three factors (Prediction 2′). In general, high-tax rate firms should use more debt than low-tax rate firms (Prediction 2). The papers reviewed in this section generally use reduced-form cross-sectional or panel regressions to test these predictions, and they ignore personal taxes altogether. For expositional reasons, we start with tests of Prediction 2′.

### *2.3.1. Nondebt tax shields, profitability, and the use of debt*

Bradley, Jarrell, and Kim (1984) perform one of the early regression tests for tax effects along the lines suggested by DeAngelo and Masulis (1980). Bradley et al. regress firm-specific debt-to-value ratios on nondebt tax shields (as measured by depreciation plus investment tax credits), R&D expense, the time-series volatility of EBITDA, and industry dummies.[14] The tax hypothesis is that nondebt tax shields are negatively related to debt usage because they substitute for interest deductions (Prediction 2′). However, Bradley et al. find that debt is positively related to nondebt tax shields, opposite the tax prediction. This surprising finding, and others like it, prompted Stewart Myers (1984) to state in his presidential address to the American Finance Association (p. 588): "I know of no study clearly demonstrating that a firm's tax status has predictable, material effects on its debt policy. I think the wait for such a study will be protracted."

One problem with using nondebt tax shields, in the form of depreciation and investment tax credits, to explain debt policy is that nondebt tax shields are positively correlated with profitability and investment. If profitable (i.e., high-tax rate) firms invest heavily and also borrow to fund this investment, this can induce a positive relation between debt and nondebt tax shields and overwhelm the tax substitution between interest and nondebt tax shields (Dammon and Senbet, 1988; Amihud and Ravid, 1985). Another issue is that nondebt tax shields (as well as existing interest deductions or the probability of experiencing losses) should only affect debt decisions to the extent that they affect a firm's marginal tax rate. Only for modestly profitable firms is it likely that nondebt tax shields have sufficient impact to affect the marginal tax rate and therefore debt policy.[15]

MacKie-Mason (1990) and Dhaliwal, Trezevant, and Wang (1992) address these issues by interacting Non-debt Tax Shields (NDTS) with a variable that identifies firms near "tax exhaustion," at which point the substitution between nondebt tax shields and interest is most important. Both papers find that tax-exhausted firms substitute away from debt when nondebt tax shields are high.[16] Even though these papers find a negative relation

[14] An alternative test would be to match NDTS-intensive firms to companies that are similar in all ways except for their use of nondebt tax shields and to examine whether the NDTS-intensive firms use less debt.

[15] The marginal tax rate for unprofitable firms will be close to zero whether or not the firm has NDTS. The tax rate for highly profitable firms will be near the top statutory rate, unless a firm has a very large amount of NDTS.

[16] Ekman (1995) finds the same for Swedish firms. Trezevant (1992) finds that Compustat PST firms most likely to be tax-exhausted decreased debt usage the most following the 1981 liberalization of tax laws that increased nondebt tax shields.

between the interacted NDTS variable and debt usage, this solution is not ideal. For one thing, the definition of tax exhaustion is ad hoc. Moreover, Graham (1996a) shows that the interacted NDTS variable has low power to detect tax effects and that depreciation and investment tax credits (the usual components of nondebt tax shields) have a very small empirical effect on the marginal tax rate. Ideally, researchers should capture the effects (if any) of nondebt tax shields, existing interest, and the probability of experiencing losses directly in the estimated marginal tax rate, rather than including these factors as stand-alone variables.

A similar issue exists with respect to using profitability as a measure of tax status. Profitable firms usually have high tax rates, and therefore some papers argue that the tax hypothesis implies they should use more debt. Empirically, however, the use of debt declines with profitability, which is often interpreted as evidence against the tax hypothesis (e.g., Myers, 1993). Profitability should only affect the tax incentive to use debt to the extent that it affects the corporate marginal tax rate[17]; therefore, when testing for tax effects, the effects (if any) of profitability should be captured directly in the estimated Marginal Tax Rate (MTR). Researchers would then interpret the stand-alone profitability variable as a control for potential nontax influences.

### *2.3.2. Directly estimating the marginal tax rate*

One of the problems that led to Myers's capital structure puzzle is related to properly quantifying corporate tax rates and incentives. For example, many studies use static MTRs that ignore important dynamic features of the tax code related to net operating losses carryback and carryforwards, investment tax credits and other nondebt tax shields, and the alternative minimum tax. Static MTRs miss the fact that a company might be profitable today but expect to experience losses in the near future. This firm might erroneously be assigned a high current-period tax rate, even though its true economic tax rate is low.[18] Conversely, an unprofitable firm might have a large current economic marginal tax rate if it is expected to soon become and remain profitable (because extra income earned today increases taxes paid in the future: an extra dollar of income today reduces losses that could be carried forward to delay future tax payments, thereby increasing present value tax liabilities).

Shevlin (1987, 1990) uses simulation techniques to capture the dynamic features of the tax code related to net operating loss carrybacks and carryforwards.[19] The first step in simulating an MTR for a given firm-year involves calculating the historic mean and variance of the change in taxable income for each firm. The second step uses this historic

[17] Keep in mind that a marginal tax rate is bound between zero and the top statutory rate, while profitability is not bounded, which can introduce difficulties into interpreting profitability as a proxy for the tax rate.

[18] Scholes and Wolfson (1992) define the *economic marginal tax rate* as the present value of current and future taxes owed on an extra dollar of income earned today, which accounts for the probability that taxes paid today will be refunded in the near future.

[19] Auerbach and Poterba (1987) and Altshuler and Auerbach (1990) simulate tax rates using first-order Markov probabilities that weight the probability of transition between taxable and nontaxable states.

information to forecast future income for each firm. These forecasts can be generated with random draws from a normal distribution, with mean and variance equal to that gathered in the first step; therefore, many different forecasts of the future can be generated for each firm. The third step calculates the present value tax liability along each of the income paths generated in the second step, accounting for the tax-loss carryback and carryforward features of the tax code. The fourth step adds $1 to current-year income and recalculates the present value tax liability along each path. The incremental tax liability calculated in the fourth step, minus that calculated in the third step, is the present value tax liability from earning an extra dollar today, in other words, the economic MTR. A separate marginal tax rate is calculated along each of the forecasted income paths to capture the different tax situations a firm might experience in different future scenarios. The idea is to mimic the different planning scenarios that a manager might consider. The fifth step averages across the MTRs from the different scenarios to calculate the expected economic marginal tax rate for a given firm-year. Note that these five steps produce the expected marginal tax rate for a single firm-year. The steps are replicated for each firm for each year, to produce a panel of firm-year MTRs. The marginal tax rates in this panel vary across firms and can also vary through time for a given firm. The end result is greater cross-sectional variation in corporate tax rates (and hence tax incentives) than implied by statutory rates.

One difficulty with simulated tax rates is that they require a time series of firm-specific data. Moreover, they are usually calculated using financial statement data, even though it would be preferable to use tax return data. With respect to the first problem, Graham (1996b) shows that an easy-to-calculate trichotomous variable (equal to the top statutory rate if a firm has neither negative taxable income nor net operating loss (NOL) carryforwards, equal to one-half the statutory rate if it has one but not the other, and equal to zero if it has both), is a reasonable replacement for the simulated rate. With respect to the tax return issue, Plesko (2003) compares financial-statement-based simulated rates for 586 firms to a static tax variable calculated using actual tax return data. He finds that simulated rates (based on financial statements) are highly correlated with tax variables based on tax return data. Plesko's evidence implies that the simulated tax rates are a robust measure of corporate tax status.

Note that by construction the simulated tax rates capture the influence of profitability on the corporate marginal tax rate. Graham (1996a) extends the simulation approach to directly capture the effects of nondebt tax shields, investment tax credits, and the alternative minimum tax. Graham (1996b) demonstrates that simulated tax rates are the best commonly available proxy for the "true" marginal tax rate (when "true" is defined as the economic tax rate based on realized taxable income, rather than simulations of the future). Using the simulated corporate marginal tax rates, Graham (1996a) documents a positive relation between tax rates and changes in debt ratios (consistent with Prediction 2), as do Graham, Lemmon, and Schallheim (1998) and Graham (1999) for debt levels. Since that time, numerous other studies have also used simulated tax rates to document tax effects in debt decisions. These results help to resolve Myers's (1984) capital structure puzzle; when tax rates are properly measured, it is possible to link tax status with corporate debt policy.

### 2.3.3. Endogeneity of corporate tax status

Even if measured with a very precise technique, tax rates are endogenous to debt policy, which can have important effects on tax research. If a company issues debt, it reduces taxable income, which in turn can reduce its tax rate. The more debt issued, the greater the reduction in the marginal tax rate. Therefore, if one regresses debt ratios on marginal tax rates, the endogeneity of corporate tax status can impose a negative bias on the tax coefficient. This could explain the negative tax coefficient detected in some specifications (e.g., Hovakimian, Opler, and Titman, 2001, and Barclay and Smith, 1995b). Note that endogeneity can affect all sorts of tax variables, including those based on NOLs, or that use an average tax rate (i.e., taxes paid/taxable income).

There are two solutions to the endogeneity problem. MacKie-Mason (1990) proposed the first solution by looking at (0,1) debt versus equity issuance decisions (rather than the debt level) in his influential examination of 1747 issuances from 1977 to 1987. Debt levels (such as debt ratios) are the culmination of many historical decisions, which may obscure whether taxes influence current-period financing choices. Detecting tax effects in the incremental approach only requires that a firm make the appropriate debt-equity choice at the time of security issuance, given its current position, and not necessarily that the firm rebalance to its optimal debt-equity ratio with each issuance (as is implicit in many debt-level studies). To avoid the endogenous effect of debt decisions on the marginal tax rate, MacKie-Mason uses the lagged marginal tax rate to explain current-period financing choice.[20] He finds a positive relation between debt issuance and tax rates. Graham (1996a) follows a similar approach and examines the relation between changes in the debt ratio and lagged simulated MTRs. He finds positive tax effects for a large sample of Compustat firms.[21]

If taxes exert a positive influence on each incremental financing decision, the sum of these incremental decisions should show up in an analysis of current debt levels—if one could fix the endogenous negative effect on tax rates induced by cumulative debt usage.[22] The second approach to fixing the endogeneity problem is to measure tax rates "but for"

[20] Wang (2000) argues that firms do not consider the level of the marginal tax rate when making incremental decisions but rather consider how far the marginal tax rate is from the "optimal MTR." Holding the level of the tax rate constant, Wang shows that companies with tax rates above the optimum are those that use the most debt (an action that should endogenously reduce the marginal tax rate and move it closer to the optimum, essentially reducing MB until it equals MC). The difficulty with this approach is that Wang's "optimal MTR" is ad hoc and is based on the probability of bankruptcy (as measured by Altman's Z-score).

[21] A number of other papers corroborate these results. For example, Shum (1996) finds similar evidence for Canadian firms. Alworth and Arachi (2000) show that lagged after-financing simulated tax rates are positively related to changes in debt for Italian firms. Henderson (2001) finds that changes in total liabilities and changes in long-term debt are both positively related to simulated tax rates in a sample of U.S. banks. Schulman et al. (1996) report that debt levels are positively correlated to tax rates in Canada and New Zealand.

[22] Dittmar (2002) studies corporate spin-offs, which potentially allows her to avoid the endogeneity problem by observing capital structures that are not the end result of a long history of accumulated debt policy decisions. However, it is still the case that past decisions can influence the parent's and/or spun-off unit's new capital structure. Dittmar does not find evidence that corporate tax rates affect spin-off debt ratios.

financing decisions. Graham, Lemmon, and Schallheim (1998) measure tax rates before financing (i.e., based on income before interest is deducted). They find a positive relation between debt-to-value and (endogeneity-corrected) “but-for” tax rates. (They also find a “spurious” negative correlation in an experiment that uses an endogenously affected after-financing tax rate.)

Examining changes in debt answers the question “are incremental decisions affected by tax status?” An alternative approach is to ask: “if tax rates exogenously change, how will a firm alter debt usage?” The Tax Reform Act of 1986 greatly reduced corporate marginal tax rates (see Fig. 1), which in isolation implies a reduction in the corporate use of debt. Givoly, Hahn, Ofer, and Sarig (1992) find that firms with high tax rates prior to tax reform (firms that therefore probably experienced the largest drop in their tax rate) reduce debt the most after tax reform. This finding is somewhat surprising because their corporate marginal tax rate suffers from the negative endogeneity bias described earlier. Moreover, personal taxes are not modeled directly, even though they fell by more than corporate tax rates after the 1986 tax reform.[23] In a paper that examines international evidence during the same time period, Rajan and Zingales (1995) provide weak international evidence that taxes affect debt decisions.

### *2.3.4. Time-series and small-firm evidence of tax effects*

The empirical evidence described thus far confirms cross-sectionally that firms with high tax rates use more debt than those with low tax rates. Presumably, there should also be time-series tax effects. For example, if a firm starts public life with a low tax rate, one would expect increased debt usage if the tax rate increases as the firm matures. There is no known study that documents tax-related time-series effects in debt usage. For example, Graham (1999) uses panel data to document the idea that cross-sectional variation in tax status affects debt usage, but he finds no evidence that time-series variation does.

By studying capital structure decisions among newly formed firms, one might be able to avoid long-lasting effects of past financing decisions. For example, Baker and Wurgler (2001) show that today’s market-to-book ratio and debt-equity issuance decisions continue to affect the firm’s debt ratios for ten or more years. Esty, Qureshi, and Olson (2000) describe various start-up financing issues, including selecting a target debt ratio, as well as how market conditions and collateralization affect the sequence of initial financing choices.

Pittman and Klassen (2001) examine capital structure in the years following an initial public offering (IPO). They perform annual (i.e., years since IPO) cross-sectional regressions and find evidence that taxes have a positive effect on the use of debt in the early years of a firm’s public life—but this relation wanes as the firm ages. Pittman and Klassen attribute this waning to an increase in refinancing transactions costs as firms

[23] Givoly et al. (1992) include lagged dividend yield in their specification to control for personal tax effects, which might allow their tax variable to isolate corporate tax effects. Personal tax effects are examined more fully in Section 1.4.

age. Note that their evidence is not time series in terms of firms altering capital structure as tax rates change through time, though they do link debt policy to firm age. Pittman and Klassen also find that firms use relatively more NDTS as they age.

Almost all capital structure papers study Compustat companies. Ayers, Cloyd, and Robinson (2001) instead examine small companies with less than 500 employees that participated in the 1993 Federal Reserve National Survey of Small Business Finances. A total of 2600 firms meet the Ayers et al. data requirements. The authors regress interest expense divided by pre-interest pre-NDTS income on various variables, including tax expense divided by pre-interest income. They find a positive coefficient on the tax variable in both their outside and inside debt regressions (i.e., interest owed to nonowners and owners, respectively). It is difficult to compare their results to Compusat-based research because Ayers et al. use a different dependent variable than most studies, and they delete firms with a negative value for the dependent variable (which raises statistical issues).

To summarize Section 1.3, once issues related to measuring debt policy and tax rates are addressed, researchers have supplied evidence in response to Myers's (1984) challenge to show that corporate debt usage is positively affected by tax rates. These results are consistent with survey evidence that interest tax deductibility is an important factor affecting debt policy decisions (ranking below only maintaining financial flexibility, credit ratings, and earnings volatility), and is especially important for large industrial firms (Graham and Harvey, 2001). Notwithstanding these empirical results, Myers is still not entirely convinced (Myers et al., 1998); he argues that tax incentives are of "third-order" importance in the hierarchy of corporate decisions. It would be helpful for future research to investigate whether the tax effects on debt versus equity choice are economically important, and if they are not, determine why not.

Several other challenges remain. First, none of the papers cited above provide time-series evidence that firm-specific changes in tax status affect debt policy. It would be quite helpful to examine whether a firm changes its debt policy as it matures and presumably its tax status changes. Second, Fama and French (2001) point out that with few exceptions the panel data examinations do not use statistical techniques that account for cross-correlation in residuals, and therefore, many papers do not allow for proper determination of statistical significance for the tax coefficients. Therefore, it is not clear if all of the tax effects documented above are robustly significant. Finally, most papers ignore the tax cost of receiving interest income from the investor's perspective, an issue that now follows.

### *2.4. Empirical evidence on whether personal taxes affect corporate debt vs. equity policy*

Miller (1977) identifies a puzzle: the benefits of debt seem large relative to expected costs, and yet many firms appear to use debt conservatively. Miller proposes that the personal tax cost of interest income (relative to the personal tax cost of equity) is large enough at the margin to completely offset the corporate tax advantage of debt. The Miller Equilibrium is difficult to test empirically for several reasons, not the least of which is the fact that the identity and tax-status of the marginal investor(s) between debt and equity are unknown. Anecdotally, we can note that the tax rate on interest income ($\tau_P$) was large relative to tax rates on corporate and equity income ($\tau_C$ and $\tau_E$) when Miller wrote his paper, so the Miller Equilibrium was plausible. However, the statutory tax rates shown in Figure 1 imply that Equation (1) has been positive since 1981, so the strict form of the Miller Equilibrium has become less plausible in the last two decades.[24]

From the corporate perspective, the relatively high investor-level taxation of interest leads to a "personal tax penalty" for debt: investors demand a higher risk-adjusted return on debt than on equity. By rearranging Equation (1), the net tax advantage of debt can be represented as

$$\tau_C - [\tau_P - (1 - \tau_C)\tau_E] \tag{6}$$

where $\tau_C$ is the corporate income tax rate, $\tau_E$ is the personal tax rate on equity income, and $\tau_P$ is the personal tax rate on interest income. The bracketed term in Equation (6) accounts for the personal tax penalty: $\tau_P - (1 - \tau_C)\tau_E$.

To quantify the effect of personal taxes in Equation (5), Gordon and MacKie-Mason (1990) and others implicitly assume that investors form clienteles based on firm-specific dividend payout ratios, and therefore that $\tau_E$ is a weighted combination of the tax rates on dividend payout and capital gains income: $\tau_E = (\text{payout})\tau_{\text{div}} + (1 - \text{payout})\tau_{\text{cap gains}}$. This and related papers use historic averages to estimate dividend payout and measure $\tau_{\text{div}}$ as equaling $\tau_P$, where $\tau_P$ is implicitly estimated using the difference between the yield on taxable and tax-free government bonds. $\tau_{\text{cap gains}}$ is often assumed to equal a fraction of the statutory capital gains tax rate (to capture the benefit of reduced effective tax rates due to deferral of equity taxation and omission of equity tax at death).[25]

Given these assumptions, Gordon and MacKie-Mason (1990) estimate that the tax advantage of debt, net of the personal tax penalty, increased following the Tax Reform Act of 1986. Recall that Miller (1977) implies that the aggregate supply of debt is determined by relative corporate and personal tax rates. Gordon and MacKie-Mason document that aggregate corporate debt ratios increased slightly in response to tax reform (consistent with Prediction 4). This is the only known research that investigates this

[24] If the statutory tax rates depicted in Figure 1 are not representative of the tax rates applicable to the marginal investor(s), or if capital gains tax rates are effectively reduced through deferral and/or elimination at death, then the Miller Equilibrium is technically possible even in recent years.

[25] Green and Hollifield (2003) simulate an economy to investigate the degree to which capital gains deferral reduces the effective tax rate on equity income (and therefore, from the company's perspective, increases the personal tax penalty for debt relative to equity). Green and Hollifield find that the ability to defer taxation reduces the implicit tax on capital gains by about 60%. If they were to factor in deferral at death and the lower tax rate on capital gains relative to the rate on dividends and interest, it would reduce the implicit tax rate on capital gains even further. (On the other hand, their calculations ignore the high turnover frequently observed for common stocks and mutual funds, which increases the effective tax rate on equity.) Overall, their evidence suggests that there is a measurable personal tax disadvantage to debt but it does not appear large enough to offset the corporate tax benefits of debt. However, Green and Hollifield find that when coupled with fairly small costs of bankruptcy (e.g., realized bankruptcy costs equal to 3% of pretax firm value), the personal tax penalty is sufficient to offset the corporate tax advantage to debt at the margin and lead to interior optimal debt ratios.

aggregate prediction. Note that Gordon and MacKie-Mason focus on a single point in time, while the Miller Equilibrium has implications for any point in time. Also note that if the marginal investor is taxable at rates like those reflected in Figure 1, then the 2003 reduction in dividend and capital gains tax rates to $\tau_{\text{div}} = \tau_P = 15\%$ should reduce the aggregate amount of debt used in the U.S. economy.

Graham (1999) tests similar predictions using firm-specific data. He finds that between 1989 and 1994 the net tax advantage of the first dollar of interest averaged between 140 and 650 basis points.[26] He finds that the firms for which the net advantage is largest use the most debt in virtually every year. Graham also separately identifies a positive (negative) relation between the corporate tax rate (personal tax penalty) and debt usage. These results are consistent with Predictions 2 and 3.

Campello (2001) assumes that a given firm's debt *and* equity are held by a particular clientele of investors (with the clienteles based on investor tax rates). He investigates the capital structure response to the large reduction in personal taxes (relative to the smaller reduction in corporate tax rates) after the Tax Reform Act of 1986. Campello finds that zero-dividend firms (which presumably have high-tax-rate investors and therefore experienced the largest reduction in the personal tax penalty) increased debt ratios in response to tax reform, while high-dividend payout firms (which presumably have low-tax-rate investors and therefore experienced a small reduction in the personal tax penalty) reduced debt usage relative to peer firms.

### 2.4.1. *Market-based evidence on how personal taxes affect security returns*

The papers we have cited, though consistent with personal taxes affecting corporate financing decisions in the manner suggested by Prediction 3, are not closely tied to market-based evidence about the tax characteristics of the marginal investor between debt and equity. Instead, these papers assume that dividend clienteles exist, and they also make assumptions about the personal tax characteristics of these clienteles based on a firm's payout policy. For example, these papers implicitly assume that there is a certain marginal investor who owns both equity and debt *and* (to estimate $\tau_P$) that this same investor sets prices between taxable and tax-free bonds. The truth is that we know very little about the identity or tax-status of the marginal investor(s) between any two sets of securities, and deducing this information is difficult.

For example, assume that munis yield 7%, Treasuries 10%, and equities 8% (and assume that this equity return has been adjusted to make its risk equivalent to the risk of munis and Treasuries). In a Gordon/MacKie-Mason/Graham type of equilibrium,

[26] I update Graham's (1999) annual tax regressions from his Table 4, Panel B. The tax variable is the tax advantage of debt net of personal taxes, as expressed in Equation (5), with the personal tax penalty based on firm-specific dividend payout ratios. The dependent variable is debt-to-value. The estimated tax coefficients for 1995–1999 are 0.072, 0.046, 0.103, 0.135, and 0.191, respectively, indicating that debt ratios are positively related to net tax incentives. All the tax coefficients are significant at a 1% level, except in 1996 when the *p*-value is 0.026.

$r_{\text{muni}} = r_{\text{Treasury}}(1 - \tau_P) = r_{\text{equity}}(1 - \tau_{\text{equity}}) = 7\%$, which implies that $\tau_P = 30\%$ and $\tau_{\text{equity}} = 12.5\%$. This in turn implies that a large portion of equity returns are expected to come from capital gains (because $\tau_{\text{equity}}$ is so much lower than $\tau_P$). However, things are rarely so simple. First, it is difficult to determine the risk-adjusted equity return.[27] Second, if there are frictions or transactions costs limiting arbitrage between pairs of markets (or if risk adjustments are not perfect), one could observe, say, munis yielding 7%, Treasuries 10%, and equities 12%. In this case, it is not clear which pair of securities should be used to deduce $\tau_P$. If Treasuries and equities are used, the implicit $\tau_P$ could be negative. For example, assume that dividend payout is 15%, that $\tau_{\text{effective cap gains}} = 5\%$, and that $\tau_{\text{equity}}$ is modeled as a weighted average between dividends and retained earnings: $\tau_{\text{equity}} = 0.15(1 - \tau_{\text{div}}) + 0.85(1 - \tau_{\text{effective cap gains}})$, where $\tau_{\text{div}} = \tau_P$. To ensure that $r_{\text{Treasury}}(1 - \tau_P) = r_{\text{equity}}(1 - \tau_{\text{equity}})$, in this example $\tau_P = -30\%$. Clearly, market frictions drive relative returns in this example, so the usual approach cannot be used to deduce the personal tax characteristics of the marginal investor(s).

Williams (2000) points out that when there are more than two assets, different pairs of assets can be arbitraged by different investors, so prices might reflect a mixture of tax characteristics. It is difficult to know which assets are directly benchmarked to each other by the marginal investor(s) and which are "indirectly arbitraged," and it is even difficult to know whether capital gains or income tax rates are priced into security returns.

It would be helpful if future research could quantify the relative importance of personal taxes on security prices, with an eye toward feedback into capital structure decisions. One area in which a fair amount of research has been done along these lines involves determining the investor tax rate implicit between municipals and taxable government bonds. Poterba (1989) finds that the yield difference between high-grade one-year munis and government bonds approximates the top statutory personal tax rate, implying that the marginal investor between these two securities is a highly taxed individual. However, even this experiment is not without difficulty. First, returns on long-term munis and taxables imply a tax rate for the marginal investor that is approxi- mately half that implied by the short-term securities. Chalmers (1998) shows that this holds even when the muni interest payments are prefunded by T-bonds held in "defeasement," and therefore, differences in risk between munis and T-bonds do not explain this conundrum. Green (1993) proposes that taxable bonds might not be "fully taxable" because a portion of their return can come from capital gains (especially for long-term bonds) and also because to some degree the interest income can be offset by investment interest deductions. Mankiw and Poterba (1996) suggest that munis might be benchmarked to equities by one clientele of investors and taxable bonds might be benchmarked to equities by another clientele. In this case, munis and taxables might not be directly benchmarked to each other, which

[27] Gordon and MacKie-Mason (1990) and Graham (1999) avoid the issue of adjusting the equity return. Instead, they assume that $\tau_{\text{div}}$ equals the $\tau_P$ implicit between munis and Treasuries and that $\tau_{\text{effective cap gains}} = 0.25 \times \tau_{\text{statutory capital gains}}$, and they weight these two pieces by the portion of earnings returned as dividends and retained, respectively, to deduce $\tau_{\text{equity}}$. It would be informative if future research could calibrate this approach to market-driven estimates of $\tau_{\text{equity}}$.

could explain the unusual implicit tax rate that is sometimes observed between the two securities.

As an example of trying to link the effects of personal taxes to capital structure issues, consider the implications from Engel et al. (1999) and Irvine and Rosenfeld (2000) about the personal tax penalty.[28] Assume that corporations are the marginal investors in preferred stock but not in debt.[29] Given the similarity of the securities, in equilibrium, we expect their after-investor-tax returns to be equal, within transactions cost bounds: $r_{\text{preferred}}(1-\tau_{\text{DRD}}) = r_{\text{MIPS}}(1-\tau_P)$. Plugging in $r_{\text{preferred}} = 8.14\%$ and $r_{\text{MIPS}} = 8.37\%$ from Engel et al.'s Table 3, and assuming that the marginal corporate investor is taxed at 35% so that $\tau_{\text{DRD}} = 10.5\%$, we can back out the personal tax rate associated with interest income: $0.0814(1-0.105) = 0.0837(1-\tau_P)$ implies that $\tau_P = 13\%$. If we ignore the 30 basis point "yield premium" on MIPS imputed by Engel et al. and use $r_{\text{MIPS}} = 8.67\%$, $\tau_P = 16\%$.

To the extent that results based on MIPS interest carry over to debt interest, finding $\tau_P = 16\%$ for the marginal debt investor is intriguing. First, note that the mean after-financing corporate tax rate in 1993-1999 is approximately 18% (see Table 1), which is a rough estimate of the tax benefit of the *last* dollar of interest deduction (ignoring all costs). If we make Miller's (1977) assumptions that $\tau_E = 0$ and that all firms face the same 18% marginal benefit of debt, then $\tau_P$ should equal 18% (i.e., MC should equal MB), quite close to the $\tau_P = 16\%$ MIPS estimate. As argued by Green and Hollifield (2003), it would only take fairly small costs of bankruptcy to equalize the costs and benefits of debt, creating a environment conducive to an equilibrium with internal optimal debt

[28] Recall that these authors investigate MIPS for preferred exchanges. These two securities are similar in most respects, except that MIPS interest is tax deductible for issuing corporations and preferred dividends are not. On the investor side, corporate investors can take the 70% dividends received deduction (DRD) for preferred dividends, but recipients of MIPS interest receive no parallel deduction.

[29] Erickson and Maydew (1998) provide evidence that corporations are the marginal investors in preferred stock, though they do not precisely identify the numeric value of the marginal investor's tax rate. They study the market reaction to the announced (but never implemented) change in the dividends received deduction (DRD). The DRD allows corporations to deduct a portion of the dividends they receive from other corporations to attenuate "triple taxation" of equity income. Individual investors do not receive the DRD. When the Treasury made a surprise announcement in December 1995 that it was planning to reduce the deduction from 70% to 50%, the typical preferred stock experienced a statistically significant $-1\%$ abnormal return, while there was no reaction among common stocks. This implies that corporations are the marginal investors (i.e., price-setters) in preferred stocks but not in common stocks. One advantage of the Erickson and Maydew study is that they are able to control for risk when examining abnormal returns because they compare a security to itself before and after the exogenous announcement. They are unable to precisely deduce the tax rate of the marginal (corporate) investor, however, because they cannot pinpoint the probability assigned by the market that the Treasury would actually implement the proposal.

While Erickson and Maydew (1998) find no evidence that corporations are the marginal investors in common stocks, Geisler (1999) shows that common-stock holdings by insurance companies vary positively with the allocation of the DRD among insurance companies. (The allocation of DRD can vary across insurance companies for regulatory reasons.) Geisler's evidence is consistent with evidence on clienteles: insurance companies respond to tax incentives to hold common stocks when their tax rate is low (i.e., when their DRD allocation is high).

ratios. However, $\tau_E$ is most likely not zero for the marginal investor in equities. (Green and Hollifield argue that deferral reduces effective $\tau_E$ to about half its statutory level.) Another issue is that the estimated MIPS costs and benefits are average, not marginal. Even if the marginal costs and benefits are equal in an equilibrium like that depicted in Figure 2a, there is a firm surplus/benefit to using debt. Therefore, even if personal tax costs are large enough at the margin to equal marginal benefits, there appear to be tax-driven preferred capital structures for some firms. Presumably, the incremental benefit would be near \$0.35 per dollar for high-tax-rate firms, while the personal tax cost is only half that amount. Only if the nontax costs of debt are large for these high-tax-rate firms could a Miller-type equilibrium hold, in which the benefits of debt are zero for all firms in equilibrium.

In sum, the implicit personal tax costs estimated here suggest that at the margin the tax costs and tax benefits might be of similar magnitude. However, they do not explain cross-sectionally why some inframarginal firms (with large tax benefits of interest) do not use more debt. (More details on this issue are presented in Section 1.4.) One other area in which there has been a fair amount of success—though not unambiguously so—in deducing marginal investor tax characteristics is related to ex-day dividend returns. This discussion is deferred to Section 4, which explores how taxes affect corporate dividend policy.

In the most general sense, any research that shows that personal tax rates affect security returns sheds light on Miller's (1977) claims. Using the CAPM-with-taxes specification, Auerbach (1983) finds evidence that tax-related preferences result in clienteles of investors that purchase stocks based on firm-specific dividend-price ratios. Constantinides (1983) and Dammon, Spatt, and Zhang (2001) investigate how favorable capital gains taxation affects investment and consumption choices. Seida and Wempe (2000) show that individual investors accelerated recognizing capital gains (and delayed losses) in anticipation of the increase in capital gains tax rates associated with the 1986 tax act. For a review of articles related to how personal taxation affects the timing and value of asset sales and purchases, see Poterba (2001).

*2.4.1.1. Tax capitalization* Another group of papers investigates tax capitalization and argue that personal taxes are capitalized into share prices via retained earnings. This in turn affects the relative tax advantage to debt because retained earnings are assumed to be the marginal source of funding. Harris and Kemsley (1999), Collins and Kemsley (2000), and related papers assume that all earnings are eventually paid out as taxable dividends (and none via repurchases or liquidating dividends), which is consistent with the "new view" of the effects of dividend taxation.[30] They argue that (nearly) full dividend taxation is impounded into share prices and that, as a result, no incremental personal tax penalty is imposed when a firm pays a dividend. Therefore, personal taxes are large on interest

[30] See Auerbach (2002) for cites. The "new view" or "trapped equity" assumptions are in contrast to the assumptions made at the beginning of Section 1 that "equity is the marginal source of funds" and that "dividends are paid out according to a fixed payout policy."

income and small on equity income, and the personal tax penalty to debt financing is large.

Harris and Kemsley (1999) regress stock price on variables, including retained earnings, and they infer that retained earnings are penalized at a dividend tax rate of approximately 47%. Collins and Kemsley (2000) argue that reinvesting current earnings leads to investor capital gains taxation when shares are sold, on top of the already impounded dividend taxation. This implies that there is no personal tax penalty for dividend payments (it is already impounded into share prices, and therefore paying a dividend does not lead to further valuation effects). In fact, this leads to the counter-intuitive argument that paying dividends leads to a reduction in future capital gains payments and therefore, *dividend payments are tax advantageous*. This implication only holds if arbitrage by tax-free investors is restricted to the point that personal investors are the marginal price-setters in stocks. Collins and Kemsley find empirical evidence that they interpret as being consistent with their hypotheses. An untested implication of their argument is that there should be a large value gain in deals that result in firms returning capital to investors in any form other than taxable dividends (such as mergers). Research into this area could be informative.

Rather than dividend taxes, an alternative argument is that capital gains taxes on future earnings are impounded into share prices. Consider a shareholder in a nondividend-paying firm and assume that the firm is expected to pay dividends at some point in the distant future. If the market expects that low-tax investors are likely to be the dominant owners of this company when the dividend payments are initiated, the only (future) tax that current investors face is capital gains. In support of this argument, Lang and Shackelford (2000) show that upon announcement that capital gains tax rates were going to decline, stock prices increased most among firms for which capital gains are most important (i.e., firms with the lowest dividend yield). This reaction is opposite that predicted by lock-in models such as Klein (2001), in which, for firms with substantial accrued retained earnings, returns fall when capital gains rates fall because the required return declines along with the tax rate. For further discussion of the tax capitalization literature, see Shackelford and Shevlin (2001).

Overall, the tax status of the marginal investor, and therefore the empirical magnitude of the personal tax penalty, is an open empirical question. This is an important issue. For one thing, failing to control for personal tax considerations can result in an omitted variable bias. For example, personal tax considerations could cause clientele behavior that is correlated with dividend-payout ratios. In a regression that omits personal tax considerations, the dividend-payout coefficient might erroneously be interpreted as supporting a nontax hypothesis. As another example, business students are often taught that the tax advantage of debt is captured by $\tau_C D$ (see Equation 4), which ignores personal tax effects. If it can be demonstrated that personal tax effects are not particularly important, this simplified view of the world might be justified. In contrast, if investor taxes affect security returns in important ways, more care needs to be taken in modeling these effects in corporate finance research. Investigations of personal tax effects face several challenges, not the least of which is that risk differences between securities must be properly controlled to allow one to deduce implicit tax rates from market return data.

### 2.5. Beyond debt vs. equity

#### 2.5.1. Leasing

The discussion thus far has considered the debt versus equity choice; however, it can be extended to leasing arrangements. In certain circumstances, a high-tax-rate firm can have a tax incentive to borrow to purchase an asset, even if it allows another firm to lease and use the asset. With true leases (as defined by the IRS), the lessor purchases an asset and deducts depreciation and (if it borrows to buy) interest from taxable income. The lessee, in turn, obtains use of the asset but cannot deduct interest or depreciation. The depreciation effect therefore encourages low-tax-rate firms to lease assets from high-tax-rate lessors. This occurs because the lessee effectively "sells" the depreciation (and associated tax deduction) to the lessor, who values it more highly (assuming that the lessee has a lower tax rate than the lessor). This incentive for low-tax-rate firms to lease is magnified when depreciation is accelerated, relative to straight-line depreciation. Furthermore, the alternative minimum tax (AMT) system can provide an additional incentive for a lessee to lease, in order to remove some depreciation from its books and stay out of AMT status altogether.

There are other tax effects that can reinforce or offset the incentive for low-tax-rate firms to lease. Lessors with relatively large tax rates receive a relatively large tax benefit of debt, which provides an additional incentive (to borrow to) buy an asset and lease it to the lessee. Moreover, tax incentives provided by investment tax credits (which have existed at various times but are not currently on the books in the United States) associated with asset purchsaes are also relatively beneficial to high-tax-rate lessors. In contrast, the relatively high taxes that the lessor must pay on lease income provide a tax disincentive for firms with high tax rates to be lessors (and similarly the relatively small tax benefit that a low-tax-rate firm obtains from deducting lease expense works against the incentive for low-tax-rate firms to lease rather than buy). The traditional argument is that low-tax-rate firms have a tax incentive to lease from high-tax-rate lessors, though this implication is only true for some combinations of tax rules (e.g., depreciation rules, range of corporate tax rates, existence of investment tax credits or AMT) and leasing arrangements (e.g., structure of lease payments). See Smith and Wakeman (1985) for details on how nontax effects can also influence the leasing decision.

Prediction 5: All else equal, the traditional argument is that low-tax-rate firms should lease assets from high-tax-rate lessors, though this implication is conditional on specifics of the tax code and leasing contract.

There are several complications associated with investigating whether firms lease in response to tax incentives. First, because leasing expense is tax deductible, leasing endogenously reduces a lessee's effective tax rate, which can bias an experiment in favor of detecting tax effects. Similarly, lessor tax rates could be endogenously increased from the effects of lease income. Second, financial statement definitions of leasing are not one-to-one with IRS definitions, making it difficult to use Compustat data to test Prediction 5. Using endogenously affected tax variables, Barclay and Smith (1995b) and Sharpe and Nguyen (1995) find that low-tax-rate firms use relatively many capital leases. However, capital leases do not meet the IRS definition of true leases. (Instead, they are likely a mixture of true leases and conditional sales contracts, the latter of which are treated like debt so that the lessee deducts interest and depreciation.) Therefore, the documented negative relation between capital leases and taxes is hard to interpret because it might be spurious.

Graham, Lemmon, and Schallheim (1998) address the first issue by measuring tax incentives "but-for financing decisions," that is, calculating tax rates using income before debt interest and the implicit interest portion of lease payments are deducted. They address the second issue by focusing on operating leases, which are defined in a manner similar to the IRS definition of true leases. Graham et al. (1998) find that the use of operating leases is negatively related to before-financing tax rates, consistent with Prediction 5, and that capital leases are unrelated to before-financing tax rates. Graham et al. also show that erroneously using an after-financing tax rate would double the magnitude of the negative tax coefficient for operating leases, and spuriously assign a negative tax coefficient to capital lease usage.

Eades and Marston (2001) find that lessors tend to be high-tax-rate firms (consistent with Prediction 5). Finally, O'Malley (1996) finds no evidence that firms systematically lease in response to tax incentives imposed by the AMT. We need research investigating whether the tax benefit of leasing adds to firm value. The jury is still out on whether debt and leasing are substitutes for the lessee (as they might be in the sense considered by DeAngelo and Masulis, 1980, because both lead to tax deductions).

### *2.5.2. Pensions*

Black (1980) assumes that pension plans and the overall company are a single economic entity that should have an integrated financing and investment strategy. Due to interest tax deductions, the cost of corporate borrowing is the after-tax cost of debt. Because they are tax-free entities, defined benefit pension plans (DBs) earn the before-tax rate of interest on bondholdings. Therefore, Black suggests that DBs should increase (decrease) bond (equity) holdings, while the rest of the firm should do the reverse. This action should not increase firm risk because the increase in corporate debt offerings is offset by the increase in bonds held in the pension plan. In an M&M (1963) world, the net effect is that the company earns $\tau_C$ times the amount of bonds held, as in Equation (4). Tepper (1981) argues that there can be a tax advantage to the strategy of corporate borrowing and DBs investing in bonds, even in a Miller (1977) world. In this case, the benefit occurs when the DB is an inframarginal investor in bonds, thereby earning the "extra" return necessary to compensate individual investors for the personal tax penalty associated with interest income (i.e., DBs capture some of the investor surplus depicted in Fig. 2). The Tepper incentive for DBs to hold bonds increases with the difference between personal tax rates on interest and equity income.

Prediction 6: Defined benefit pension plans have an incentive to hold bonds (equity) that increases (decreases) in the corporate tax rate, while the rest of the firm has the reverse incentive.

Myers (2001) finds evidence consistent with the Black (1980) case: she reports that DB bondholdings increase with a simulated corporate marginal tax rate. She does not find evidence consistent with the Tepper argument. In a less direct test of the same incentives, Thomas (1988) finds time-series evidence that firms decrease DB contributions when their tax rate is falling, and cross-sectional evidence that high-tax firms have larger DB funding levels.

Clinch and Shibano (1996) study pension reversions, which occur when a firm terminates an overfunded pension, settles its liabilities, and reverts the excess assets to the firm, all in one year. The reverted assets are taxable in the reversion year. Clinch and Shibano state that firms with the largest tax benefit of reverting do so, and also that firms' time-reversion decisions occur in years with particularly large tax benefits. One nice aspect of the Clinch and Shibano experiment is that their tax variable equals the tax consequence of reverting relative to the tax consequence associated with the next best alternative (e.g., amortizing the excess assets over several years).[31]

### 2.5.3. *Debt maturity*

In the spirit of Modigliani and Miller (1958), Lewis (1990) derives an irrelevance null hypothesis for debt maturity. If corporate taxes are the only market imperfection, Lewis shows that the optimal firm-specific debt policy (i.e., optimal level of promised interest payments) can be achieved by various combinations of short- and long-term debt. This implies that firm value is unaffected by debt maturity structure and that capital market imperfections beyond corporate taxes, like costs to restructuring debt or underinvestment, are needed for debt maturity to matter.

Rather than modeling the simultaneous choice of debt level and maturity structure as in Lewis (1990), Brick and Ravid (1985) assume that firms choose debt level before debt maturity. If the expectations theory of interest rates holds, firms pay the same present value of interest in the long run regardless of debt maturity. However, issuing long-term debt accelerates interest payments, thus maximizing the present value of the interest tax

[31] Chaplinsky and Niehaus (1990) describe the potential tax benefits of Employee Stock Ownership Plans, a form of defined contribution benefit plan. ESOPs offer deferred compensation to employees and a deductible expense to employers. ESOPs are designed to allow firms to borrow to purchase own-company stock on employees' behalf, which provides an interest deduction to the firm. Moreover, half of the interest income received by the lenders is tax-free. Shackelford (1991) finds that lenders keep only 20–30% of the tax benefit associated with this interest, with the remainder being passed along to the ESOP in the form of a lower interest rate on the loan. In late 1989, tax rules changed to restrict the interest exclusion to loans where the ESOP own more than 50% of the stock, which effectively killed the interest exclusion except for a few very unusual cases.

shield. Brick and Ravid (1985) use this logic to argue that debt maturity should increase with the slope in the yield curve.

Prediction 7: Debt maturity increases in the slope in the yield curve.

Most empirical evidence does not support their prediction. Barclay and Smith (1995a) and Stohs and Mauer (1996) include a stand-alone yield curve variable that is either insignificant or has the wrong sign. Guedes and Opler (1996) maintain that the slope of the yield curve should only affect firms with a positive tax rate, and therefore the yield curve variable will interact with the corporate marginal tax rate. Neither Guedes and Opler (using a crude measure of the corporate tax rate) nor Harwood and Manzon (1998, using a simulated corporate tax rate) find a significant coefficient on the yield curve variable. The one exception is Newberry and Novack (1999), who use a dummy variable equal to one during 1992 and 1993 (when the term premium was relatively high) and equal to zero for all other years 1987–1995. Newberry and Novack find a positive coefficient on the yield curve dummy in their public debt regression but not in their private debt analysis.

Kane, Marcus, and McDonald (1985) determine optimal debt maturity in a model that trades off corporate tax benefits with personal tax, bankruptcy, and flotation costs. The implications of their model are that debt maturity decreases with the corporate MTR and increases with the personal tax rate: long maturity implies less frequent recapitalization and relatively low transactions costs, so long-term debt can be desirable even if the net tax benefit is low. Maturity also decreases with the volatility of firm value because volatile firms are more likely to restructure debt.

Prediction 8: Debt maturity decreases with the corporate MTR and the volatility of firm value and increases with the personal tax rate.

Stohs and Mauer (1996) find the following support for Prediction 8: volatile firms generally use shorter term debt. The evidence relating to the tax-rate prediction is weaker. Stohs and Mauer report that debt maturity decreases with corporate tax rates—but their MTR variable is very crude (equal to income tax expense divided by pretax income when this ratio is between 0 and 1, and equal to 0 otherwise). Opler and Guedes (1996) find a negative coefficient on a tax expense divided by assets variable but the wrong sign on an NOL-based tax variable. Finally, Harwood and Manzon (1998) and Newberry and Novack (1999) point to a positive relation between a simulated tax rate variable and debt maturity, opposite the Kane et al. prediction.[32] A positive coefficient makes sense if

[32] Harwood and Manzon's variable equals the Graham (1996a) simulated tax rate divided by the top statutory tax rate. This variable has a large value for firms that do not currently have NOLs and that do not expect to experience a loss in the near future. Harwood and Manzon predict a positive relation between this tax variable and debt maturity. They contend that firms with large values for the tax variable are likely to fully utilize tax deductions in the future, and therefore lock into long-term debt now. In new analysis for this chapter, I perform a more direct test on the hypothesis that uncertainty about future tax-paying status reduces the use of long-term debt. I use the standard deviation of the simulated marginal tax rate to measure uncertainty about tax-paying status, with the standard deviation calculated across the simulated scenarios for any given firm-year. I do not find any relation between debt maturity and uncertainty about tax-paying status.

large simulated MTRs identify firms that use long-term debt because they are relatively likely to be able to deduct interest in current and future periods.

Finally, debt maturity can affect the tax-timing option for firms to opportunely retire debt (e.g., Emery, Lewellen, and Mauer, 1988). If the corporate tax function is convex, the expected present value tax benefit of short-term debt declines with interest rate volatility, while the tax deductions with long-term debt are fixed. Therefore, long-term debt is preferred when interest rates are volatile. Long-term debt also increases the value of the timing option for investors to tax-trade securities (Kim, Mauer, and Stohs, 1995) because option value increases with security maturity and long-term bond prices are more sensitive to changes in interest rates.

Prediction 9: Debt maturity increases with interest rate volatility.

Kim et al. (1995) find that debt maturity increases with interest rate volatility, but Guedes and Opler (1996) do not. Nor do Guedes and Opler find significance for a second variable that interacts interest rate volatility with a corporate MTR variable.

The evidence linking tax incentives to debt maturity is mixed. One factor that makes it difficult to draw general conclusions is that debt maturity is defined differently in various papers. Barclay and Smith (1995a) use a dependent variable measuring the portion of outstanding debt that matures in four or more years; Guedes and Opler (1996) use the log of the term to maturity for new debt issues; Stohs and Mauer (1996) use the book value weighted average of the maturity of a firm's outstanding debt; Newberry and Novack (1999) use the same for new issues; and Harwood and Manzon (1998) use the portion of outstanding debt that is long-term. Another issue that might affect inferences about tax variables is the apparently nonlinear relation between debt maturity and nontax influences (Guedes and Opler, 1996). Unless the nonlinearity of the overall specifi- cation is properly controlled, it might adversely affect the ability to detect tax effects. Finally, the yield curve was never inverted during the periods studied by most of these papers, so the tests of Brick and Ravid (1985) focus on the steepness of the yield curve rather than on the sign.

## 3. Taxes and capital structure—international tax issues

Section 1 reviews capital structure choice in the context of a domestic-only firm operating in a classical tax system (in which interest is tax deductible but equity payments are not). Although much academic research focuses on this paradigm, international tax issues have become more important in recent years. This section reviews how international tax law can affect corporate financing decisions in a multinational firm. The perspective is generally for a firm headquartered in the United States, but many of the implications hold if the firm is headquartered elsewhere.

The general framework is still based on taxes affecting firm value via an expression such as $V_{\text{with debt}} = V_{\text{no debt}} + \tau_C(.)^*\text{D}$. The research in this section demonstrates that multinational tax rules can affect the $\tau_C(.)$ function and therefore the incentive to use both domestic and foreign debt. So as not to let the reader get bogged down in international tax law, this section only sketches the effects of multinational tax incentives. To focus on the central factors that affect multinational firms, several simplifying assumptions (described below) are presented. For a more detailed description of international tax law, see Hines (1996) or Scholes, Wolfson, Erickson, Maydew, and Shevlin (2002) and the references therein.

### *3.1. Tax incentives and financial policy in multinational firms: theory and tax rules*

A multinational corporation can finance its foreign operations with internal equity (i.e., an equity infusion from a parent or subsidiary to an affiliated subsidiary), internal debt (i.e., a loan from the parent to a subsidiary), external funding, or earnings retained by the foreign subsidiary. If internal equity is used, the parent receives its return on equity when the subsidiary repatriates dividends back to the home country. Dividend repatriations based on active operating earnings can usually be deferred indefinitely, until the parent needs an infusion of cash, or to optimize the worldwide tax situation of the firm.[33] In contrast, interest from internal debt is paid according to a fixed schedule. Like a repatriated dividend, interest counts as "worldwide income" on the U.S. tax return of the parent. Unlike a repatriated dividend, the interest is often deductible on the foreign tax return, allowing for a foreign tax deduction analogous to the tax benefit of debt described in Section 1.[34]

Two important items affect the financing choices of U.S.-based multinational firms: foreign tax credits and interest allocation rules. The U.S. government taxes individuals and corporations on the basis of residence or place of incorporation, meaning that they are taxed because they are from the United States, regardless of where they earn income. (Note that the United States only taxes "active foreign source income" at the time of repatriation to the U.S. parent.) At the same time, the government recognizes that income earned abroad is usually taxed by a foreign entity, so the United States offers foreign tax credits to offset taxes paid abroad. If the United States did not offer such credits, the foreign operations of U.S. corporations would face double taxation and therefore have

[33] To illustrate the potential economic importance of repatriations and taxes on such transfers, note that in 2003 the Bush administration proposed reducing the tax on all repatriated income to 5%. The goal was to spur a return of capital to U.S.-domiciled firms in hopes that these firms would productively invest the funds and stimulate the U.S. economy. This provision was eliminated during negotiations with Congress over the tax bill.

[34] There are restrictions to shifting interest deductions abroad by lending from the domestic parent to the foreign subsidiary: thin capitalization rules (i.e., limits on the magnitude of foreign debt ratios), withholding taxes imposed by the foreign government on interest payments and other repatriations, and netting rules that restrict the effect of interest payments on the determination of foreign source income (Newberry and Dhaliwal, 2001, and Scholes et al., 2002). For example, withholding taxes are above and beyond foreign income taxes and are collected by foreign governments on remittances to parent firms.

a tough time competing with foreign corporations. For the purposes of this analysis, the reader should think of the foreign tax rate ($\tau_{\text{For}}$) as a weighted average of tax rates the firm pays in the various countries in which it earns foreign income, with the weights being the relative share of active (i.e., nonpassive) foreign source income repatriated from a particular country.

In simplest terms, if the foreign tax rate is smaller than the U.S. corporate income tax rate ($\tau_{\text{US}}$), a firm receives credit for foreign taxes paid but still must remit to the U.S. government taxes equal to ($\tau_{\text{US}} - \tau_{\text{For}}$)*(foreign source income). Such a firm is called a *deficit credit* firm because it lacks sufficient foreign tax credits (FTCs) to avoid all U.S. taxes. For example, if repatriated foreign earnings are \$200, $\tau_{\text{For}} = 15\%$, and $\tau_{\text{US}} = 35\%$, the firm must pay \$40 in tax to the United States.

In contrast, if $\tau_{\text{For}} > \tau_{\text{US}}$, the firm does not have to pay U.S. taxes because it receives foreign tax credits proportional to $\tau_{\text{For}}$. For example, if $\tau_{\text{For}} = 45\%$ and $\tau_{\text{US}} = 35\%$ and repatriated earnings are \$200, the firm pays \$90 in foreign tax; however, the firm's foreign tax credits are limited to $\text{FTC}_{\text{allow}} = \$70 (= \min[\$200\tau_{\text{US}}, \$200\tau_{\text{For}}])$, which is just enough to shield it from the U.S. tax obligation. The \$20 in unused FTCs can be carried back up to two years or carried forward up to five years to offset taxes on repatriated income (or they can be deducted rather than used as a credit). This firm is an *excess credit* firm because it has more FTCs than it is allowed to use in the current year and accumulates the excess tax credits to potentially shield income in another year. The tax benefit of debt, $\tau_C(.)$, can be modeled as a decreasing function of accumulated FTCs because FTCs can act as nondebt tax shields that are substituted for interest deductions.[35]

Prediction 10: All else equal, the incentive $\tau_C(.)$ to finance with domestic debt decreases with accumulated foreign tax credits for deficit credit firms.[36]

Prediction 10 is a static prediction. Considering the dynamic carryback and carryforward features of the tax code, a dynamic prediction is that the tax incentive to finance

[35] FTCs can affect tax incentives to use debt in a manner that is not reflected in a one-period model. Assume that a multinational firm has accumulated unused FTCs that it has carried forward to the present (or assume that it anticipates receiving excess FTCs sometime in the next two years). If a firm has carried forward FTCs from previous years, it very likely was excess credit, and therefore subject to $\tau_{\text{For}} > \tau_{\text{US}}$ at some point in the past. For the most part, a firm can use these accumulated FTCs only if the foreign tax rate becomes smaller than the U.S. corporate income tax rate. This can occur if there is an exogenous shift in relative tax rates ($\tau_{\text{For}}$ and $\tau_{\text{US}}$) or if a firm repatriates more foreign-source income from low-tax countries, thereby reducing the average $\tau_{\text{For}}$ (i.e., the latter case is an example of a firm endogenously reducing its $\tau_{\text{For}}$). If a firm expects to use accumulated FTCs to reduce taxes, the FTCs compete with interest deductions in a DeAngelo and Masulis sense and reduce the incentive to finance with debt.

[36] Consider a firm with \$1 in pretax foreign earnings that it will repatriate back to the United States to pay investors. Assume that the firm has \$0.15 in accumulated FTCs, $\tau_P = 0.40$, $\tau_E = 0.20$, $\tau_{\text{For}} = 0.20$, and the U.S. corporate tax rate is $\tau_{\text{US}} = 0.35$. Ignoring foreign considerations, $\tau_C = 0.35$ and Equation (1) equals 0.08, so it appears that the firm should finance with domestic debt. However, $\tau_C = 0.20$ once the effect of FTCs is considered (the firm pays \$0.20 in foreign tax and no U.S. tax because the FTCs offset any potential tax owed to the United States); therefore, Equation (1) equals $-0.04$, and the firm should finance with equity. This implication holds for deficit credit firms but not for excess credit firms (because an excess credit firm would not pay U.S. tax at repatriation, regardless of whether they have accumulated FTCs).

with debt decreases with the probability of a firm being deficit credit and the probability of accumulating FTCs.

The second important tax principle affecting multinational corporate financing decisions is the allocation of debt interest between domestic and foreign operations. Via the allocation of domestic interest, the United States limits allowable foreign tax credits, thereby possibly reducing the tax benefit of domestic debt. (The United States does this to limit tax deductions on debt that might possibly be used to finance foreign operations and produce foreign profits.) To implement this policy, the United States allocates domestic interest to foreign operations based on the proportion of total assets that are in foreign subsidiaries. In rough terms, if two-thirds of a company's worldwide assets are held by foreign subsidiaries, then two-thirds of domestic interest deductions are allocated to foreign income when determining the allowable-FTC calculation. Note that this is a U.S. government ruling and does not mean that foreign governments recognize the allocated interest as a deduction against foreign income. Also note that the allocation of a portion of domestic interest abroad technically affects only the allowable-FTC calculation; that is, ignoring FTC, domestic interest deductions are not directly affected.

The interest allocation procedure can reduce the tax incentive for U.S. firms to use domestic debt because $\tau_C(.)$ also declines with the degree of interest allocation. When a firm is an excess credit firm (i.e., $\tau_{\mathrm{US}} < \tau_{\mathrm{For}}$) and taxable on both foreign and domestic operations, the interest allocation procedure reduces the tax benefit of domestic interest deductions by setting $\tau_C(.)$ equal to $\tau_{\mathrm{US}}(.)$*[domestic assets/worldwide assets]. Thus, for excess credit firms the incentive to finance with domestic debt decreases with the proportion of assets held abroad.[37] One implication of the interest allocation rules is that debt policy research cannot assume that financial statement (or Compustat) "domestic interest expense" is fully beneficial to U.S. multinationals.

Table 2 summarizes the tax incentives to use external domestic or foreign debt in a one-period model. The table is self-explanatory, so only the main points need be emphasized here. The model ignores personal taxes, carryforwards, and carrybacks, and assumes that all foreign income is repatriated each year. The worldwide tax liability ($\mathrm{Tax}_{\mathrm{World}}$) is equal to the sum of U.S. tax on worldwide income ($\mathrm{Tax}_{\mathrm{US}}$) and foreign tax on foreign income ($\mathrm{Tax}_{\mathrm{For}}$), less allowable FTCs. The table shows the change in $\mathrm{Tax}_{\mathrm{World}}$ that occurs, for various tax credit and interest allocation situations, when an additional dollar of domestic or foreign interest is deducted.[38]

For the most part, the results in Table 2 are what you would expect without thinking too deeply about the complexities of foreign taxes. If $\mathrm{Tax}_{\mathrm{US}}$ is zero (rows 1 and 3) or

[37] If a U.S. multinational is deficit credit (i.e., $\tau_{\mathrm{US}} > \tau_{\mathrm{For}}$) and taxable both in the United States and overseas, $\tau_C(.) = \tau_{\mathrm{US}}$ and the incentive to use domestic debt is not affected by interest allocation rules. The interest allocation rules limit the amount of deductions a firm is allowed to use to offset repatriated foreign income. When a firm is deficit credit, it pays tax at the rate $\tau_{\mathrm{US}}$ regardless of the amount of FTCs applied to foreign-source income, so reducing allowable FTCs via interest allocation does not affect the current-year tax liability.

[38] This model ignores many techniques by which firms can minimize worldwide taxes. See Scholes et al. (2002) for more information on these alternative mechanisms.

Table 2

Tax incentive to use debt in a U.S. multinational firm with foreign tax credits and allocable domestic interest

Assume that a U.S. multinational firm currently returns \$1 of pre-corporate-tax earnings to its marginal investor as domestic equity. The one-period model in this table shows the tax effect of instead paying the \$1 as foreign interest (rightmost column in each panel) or as \$1 of domestic interest (second-to-rightmost column). The model is adapted from Collins and Shackelford (1992) and assumes that all foreign income ($\mathrm{Inc_{For}}$) is repatriated every year and that tax rules are the same worldwide, except that only the United States allocates interest. The model ignores the AMT, carrybacks and carryforwards, personal taxes, and allocable items other than interest. Because the real-world tax-code is dynamic (i.e., it allows for carrybacks and carryforwards), the one-period nature of this model might overstate (understate) the largest (smallest) tax benefits. Note that foreign losses (i.e., $\mathrm{Inc_{For}} - \mathrm{Int_{For}} < 0$) cannot be repatriated as losses back to the United States. $\mathrm{FTC_{allow}}$ is allowable foreign tax credit (sometimes referred to as $\mathrm{FTC_{limitation}}$), FA is foreign assets net of foreign debt, WA is worldwide assets net of foreign debt, and FSI is foreign source income, which equals $\mathrm{Inc_{For}} - \mathrm{Int_{For}} - \frac{\mathrm{FA}}{\mathrm{WA}}\mathrm{Int_{US}}$.

$$\mathrm{Tax_{World}} = \mathrm{Tax_{US}} + \mathrm{Tax_{For}} - \mathrm{FTC_{allow}} = (\mathrm{Inc_{US}} - \mathrm{Int_{US}} + \mathrm{Inc_{For}} - \mathrm{Int_{For}})\tau_{\mathrm{US}} + (\mathrm{Inc_{For}} - \mathrm{Int_{For}})\tau_{\mathrm{For}} - \mathrm{FTC_{allow}}\text{ , where}$$

$$\mathrm{FTC_{allow}} = \mathrm{Max}\{0, \mathrm{Min}[\mathrm{Tax_{For}}, \mathrm{FSI}\tau_{\mathrm{US}}, \mathrm{Tax_{US}}]\} = \mathrm{Max}\{0, \mathrm{Min}[(\mathrm{Inc_{For}} - \mathrm{Int_{For}})\tau_{\mathrm{For}}, (\mathrm{Inc_{For}} - \mathrm{Int_{For}} - \tfrac{\mathrm{FA}}{\mathrm{WA}}\mathrm{Int_{US}})\tau_{\mathrm{US}}, (\mathrm{Inc_{US}} - \mathrm{Int_{US}} + \mathrm{Inc_{For}} - \mathrm{Int_{For}})\tau_{\mathrm{US}}]\}$$

| | If $\mathrm{Tax_{US}}$ | and $\mathrm{Tax_{For}}$ | then $\mathrm{FTC_{allow}} =$ | and $\mathrm{Tax_{World}} =$ | $\delta(\mathrm{Tax_{World}})\delta(\mathrm{Int_{US}})$ | $\delta(\mathrm{Tax_{World}})\delta(\mathrm{Int_{For}})$ |
|---|---|---|---|---|---|---|
| (1) | $= 0$ | $> 0$ | 0* | $(\mathrm{Inc_{For}} - \mathrm{Int_{For}})\tau_{\mathrm{For}}$ | 0 | $-\tau_{\mathrm{For}}$ |
| (2) | $> 0$ | $= 0$ | 0 | $\mathrm{Inc_{US}} - \mathrm{Int_{US}}\tau\mathrm{US}$ | $-\tau_{\mathrm{US}}$ | 0 |
| (3) | $= 0$ | $= 0$ | 0 | 0 | 0 | 0 |

Otherwise, if $\mathrm{Tax_{US}} > 0$ and $\mathrm{Tax_{For}} > 0$ and

| | if $\mathrm{Inc_{US}}$-$\mathrm{Int_{US}}$ | and $\tau_{\mathrm{US}}$ | then $\mathrm{FTC_{allow}} =$ | referred to as | $\delta$ ($\mathrm{Tax_{World}}$) $\delta(\mathrm{Int_{US}})$ | $\delta(\mathrm{Tax_{World}})$ $\delta(\mathrm{Int_{For}})$ |
|---|---|---|---|---|---|---|
| (4) | $> 0$ | $> \mathrm{Tax_{For}}/\mathrm{FSI}$ | $(\mathrm{Inc_{For}} - \mathrm{Int_{For}})\tau_{\mathrm{For}}$ | deficit credit | $-\tau_{\mathrm{US}}$ | $-\tau_{\mathrm{US}} - \tau_{\mathrm{For}} + \tau_{\mathrm{For}} = -\tau_{\mathrm{US}}$ |
| (5) | $> 0$ | $< \tau_{\mathrm{For}}$ | $(\mathrm{Inc_{For}} - \mathrm{Int_{For}} - \frac{\mathrm{FA}}{\mathrm{WA}}\mathrm{Int_{US}})\tau_{\mathrm{US}}$ | excess credit * | $-\tau_{\mathrm{US}}\ (1 - \frac{\mathrm{FA}}{\mathrm{WA}})$ | $-\tau_{\mathrm{US}} - \tau_{\mathrm{For}} + \tau_{\mathrm{US}} = -\tau_{\mathrm{For}}$ |
| (6) | $< 0$, and $< \mathrm{Inc_{For}} - \mathrm{Int_{For}}$ in absolute value, so some taxes paid | not applicable | $(\mathrm{Inc_{US}} - \mathrm{Int_{US}} + \mathrm{Inc_{For}} - \mathrm{Int_{For}})\tau_{\mathrm{US}}$ | domestic losses but worldwide profits (excess credit)* | 0 | $-\tau_{\mathrm{US}} - \tau_{\mathrm{For}} + \tau_{\mathrm{US}} = -\tau_{\mathrm{For}}$ |

* In a multiperiod model, FTCs above the allowable amount could be carried back or accumulated and carried forward. For example, in the excess credit case with interest allocation (row 5), $\frac{\mathrm{FA}}{\mathrm{WA}}$ of unused FTCs accumulate per incremental dollar of domestic interest.

domestic income is negative (row 6), there is no tax benefit from issuing domestic debt; there is, however, a benefit of $\tau_{\text{For}}$ to deducting \$1 of foreign interest when foreign income is positive (rows 1 and 6).[39] If foreign income is negative but domestic income is positive (row 2), there is no tax incentive to issue foreign debt but an incremental dollar of domestic interest provides a benefit of $\tau_{\text{US}}$.

Two situations are more subtle. If a U.S. multinational is deficit credit (i.e., $\tau_{\text{US}}$ is greater than $\tau_{\text{For}}$) and profitable both in the United States and overseas (row 4), a dollar of domestic or foreign interest produces a tax benefit of $\tau_{\text{US}}$. To see how foreign interest produces a tax benefit proportional to $\tau_{\text{US}}$, consider a case in which a multinational earns \$2 of income in a country with $\tau_{\text{For}} = 45\%$ and \$4 of income in a country with $\tau_{\text{For}} = 25\%$, and assume that $\tau_{\text{US}} = 35\%$. The \$2 of high-tax foreign income produces $\text{Tax}_{\text{For}} = \$0.90$. The firm receives $\text{FTC}_{\text{allow}} = \$0.70$ on this income and has \$0.20 of unused FTCs. The \$4 of low-tax foreign income produces $\text{Tax}_{\text{For}} = \$1$. As a stand-alone item, this income produces \$0.40 of U.S. tax at repatriation [\$4*(35%–25%)]; however, the \$0.20 of extra FTC offsets half of this U.S. tax liability. On net the firm pays the United States \$0.20 in tax on foreign earnings and has a total tax liability of $\text{Tax}_{\text{World}} = \$2.10(\$2.10 = \$0.90$ in high-tax country, \$1.00 in low-tax country, and \$0.20 on income repatriated from low-tax country).[40] If this firm deducts \$1 of interest in the low-tax country, it reduces its tax bill by \$0.35 (\$0.25 reduction in $\text{Tax}_{\text{For}}$ and \$0.10 in U.S. tax owed on that dollar). If the firm uses \$1 of interest in the high-tax country, it reduces its tax bill by \$0.35 (\$0.45 reduction in $\text{Tax}_{\text{For}}$, but \$0.10 less FTC is available to offset taxes owed on the income repatriated from the low-tax country.) Either way, the tax benefit of deducting \$1 of foreign interest is $\tau_{\text{US}}$ when a firm is deficit credit and profitable both in the United States and overseas.

The second subtle situation involves the tax benefit of deducting domestic interest when a firm is excess credit and $\text{Tax}_{\text{US}}$ and $\text{Tax}_{\text{For}}$ are both positive (row 5). In this case, a portion of domestic interest is allocated to foreign-source income, thereby reducing the benefit of a dollar of interest by the ratio of foreign assets to worldwide assets. (Recall that this allocated interest will not reduce $\text{Tax}_{\text{For}}$.) The allocation of domestic interest reduces the incentive of an excess credit firm to issue domestic debt, especially when the firm has substantial foreign assets. Altshuler and Mintz (1995) note that more than 60% of firms were excess credit during the late 1980s, so interest allocation is potentially important.

39 If there is a positive probability that tax losses will be used if carried backward or forward, the tax benefit can be positive even in row (1), (3), or (6). Conversely, if there is a positive probability that losses will occur and be carried back from the future, positive tax benefits might be smaller than those shown in the table. Also, in a more complicated model, one could also net out the personal tax costs associated with interest income. Finally, see Altshuler and Newlon (1993) for the marginal tax costs of repatriations when there are also withholding taxes.

40 In most situations, the income from the high- and low-tax country would be summed and treated as income from one "basket," with $\tau_{\text{For}} = (4 \times 25\% + 2 \times 45\%)/6 = 31.67\%$. The countries are treated separately in this example to highlight how income from one country can lead to FTCs that shield income repatriated from another country.

Prediction 11: Due to interest allocation, the tax benefit of domestic interest deductions declines with the probability that a firm will operate as excess credit and with the proportion of assets held in foreign subsidiaries.

The analysis can be modified to examine the tax incentives associated with the parent supplying the foreign subsidiary with internal debt. The incentive is similar to that for external foreign debt shown in the rightmost column in Table 2, with one difference: with internal debt, the interest is taxable to the parent at rate $\tau_{US}$ when $\text{Tax}_{US} > 0$. Thus, in some cases $\tau_{US}$ should be added in the rightmost column. Specifically, if the debt is internal rather than external, the entries in the rightmost column are $-\tau_{For} + \tau_{US}$, 0, 0, $\tau_{US} - \tau_{For}$, and $\tau_{US} - \tau_{For}$ in rows (1)–(6), respectively. (Recall that a negative term means tax savings.) First consider the deficit credit case (row 4) where the tax incentive to fund a foreign subsidiary with internal debt is nil: there is no tax incentive to use internal debt because the net benefit of deducting in the foreign country is exactly offset by the increased tax in the home country. In the excess credit case (rows 5 and 6), the net tax benefit is $\tau_{For} - \tau_{US}$. For these rows, there is a tax incentive to issue debt increases with $\tau_{For}$, but it is offset by taxes owed by the domestic parent. In row (2), when $\text{Tax}_{For} = 0$ and $\text{Tax}_{US} > 0$, there is a tax disincentive of $\tau_{US}$ per dollar of internal interest: the extra foreign interest does not further reduce $\text{Tax}_{For}$, and yet there is a positive tax liability of $\tau_{US}$ on the remitted interest. In contrast, when $\text{Tax}_{US} = 0$ (rows 1 and 3) using internal rather than external debt does not change the entries in Table 2: there is no tax on the interest received by the parent because the firm otherwise has domestic losses.

Prediction 12: The tax incentive to fund a foreign subsidiary with internal debt generally increases with $\tau_{For}$; however, this incentive is offset in several situations, as shown in Table 2.

Prediction 13: The tax incentive to issue external foreign debt increases with $\tau_{For}$, although this incentive can be affected by the relative taxation of interest and equity income at the investor level.

Note that the incentive to save on foreign taxes might be tempered by investor-level taxes along the lines suggested in Miller (1977).

Other than in this paragraph, the results in this section are derived for the case where the domestic parent operated under a classical tax system in which interest is tax deductible but equity payments are not. If instead there is an imputation or integrated tax system (as in the UK, France, or many other countries), equityholders receive a credit for taxes paid at the corporate level, which partially or fully eliminates the double taxation of equity income. This at least partially reduces the net tax advantage to debt. For example, ignoring personal taxes, Cooper and Nyborg (1999) show that the value of a levered firm in an imputation tax system equals

$$V_{\text{with debt}} = V_{\text{no debt}} + \frac{(\tau_C - \tau_I)}{(1 - \tau_I)} D \tag{7}$$

where $\tau_I$ is the rate of imputation tax. In a full imputation tax system, dividend recipients receive a tax credit for income taxed at the corporate level, which they can use to offset their personal tax liability. If imputation results in a full tax credit at the corporate rate, then $\tau_I = \tau_C$ in Equation (7) and there is no tax advantage to debt. In a partial imputation system, stockholders only receive a partial credit for taxes paid at the corporate level, which is analogous to making equity (at least partially) tax deductible, which in turn reduces the net tax advantage of debt. Whether there has been any research that investigates the following prediction is not known.

Prediction 14: The tax incentive to issue debt decreases with the degree of dividend imputation dictated by the tax law under which a company operates

### *3.2. Empirical evidence related to multinational tax incentives to use debt*

Testing multinational tax hypotheses is difficult because the data are hard to obtain and noisy. Most of the international capital structure tests are based on implications found in row (4) and especially row (5) of Table 2. Table 3 summarizes some empirical evidence related to multinational debt policy.

With respect to Prediction 11 (due to interest allocation, the tax benefit of domestic interest deductions declines with the probability that a firm will operate as excess credit and with the proportion of assets held in foreign subsidiaries), Froot and Hines (1995) observe that debt usage is reduced for excess credit firms, with the reduction proportional to the fraction of assets that are foreign. Altshuler and Mintz (1995) also show that the use of foreign debt increases with the proportion of assets held overseas (presumably because domestic interest would be allocated abroad). Newberry (1998) and Newberry and Dhaliwal (2001) find that the likelihood of issuing domestic debt is highest when a firm is not excess credit and when less interest is allocated abroad. A related prediction is that firms shift away from debt financing when interest is allocated abroad. Collins and Shackelford (1992) show that firms increase their use of preferred stock when domestic interest allocation is unfavorable. Froot and Hines (1995) point out that, unlike interest, lease payments are not allocable, and they show that excess credit firms rely more heavily on leasing.

Several papers provide evidence with respect to Prediction 12 (the tax incentive to fund a foreign subsidiary with internal debt increases with $\tau_{\mathrm{For}}$) and Prediction 13 (the tax incentive to issue external foreign debt increases with $\tau_{\mathrm{For}}$, although this incentive can be affected by the relative taxation of interest and equity income at the investor level). Examining a cross section of countries with differing foreign tax rates, Desai (1997) indicates that the net internal debt infusion into foreign subsidiaries increases with $\tau_{\mathrm{For}}$ (Prediction 12). Newberry and Dhaliwahl (2001) find that the propensity to issue bonds in foreign markets increases in $\tau_{\mathrm{For}}$ (Prediction 13). Hines (1995) demonstrates that royalty payments increase when they are a cheaper form of repatriation than are dividends. Finally, Grubert (1998) finds that an increase in the price of one form of

Table 3
Summary of predictions and empirical evidence for multinational capital structure

| Prediction | Empirical Evidence |
|---|---|
| Firm uses less debt when it has accumulated FTCs | None |
| Excess credit firms should have less incentive than deficit credit firms to use domestic debt.<br><br>The incentive for excess credit firms to use domestic debt declines with the proportion of assets that are foreign.<br><br>The incentive to use foreign debt increases in the foreign tax rate. | Debt usage declines when firm is excess credit. The reduction is increasing in the fraction of assets that are foreign (Froot and Hines, 1995).<br><br>Likelihood of issuing domestic debt is highest when deficit credit and decreases as FTC limitations increase (Newberry, 1998, and Newberry and Dhaliwal, 2001).<br><br>Excess credit firms' use of foreign debt increases in $\tau_{\mathrm{For}}$ and in the share of foreign assets (Altshuler and Mintz, 1995).<br><br>Debt ratios of foreign affiliates increase in $\tau_{\mathrm{For}}$ (Desai, 1997, and Altshuler and Grubert, 2000). |
| If domestic losses, use foreign debt. | U.S. multinationals borrow in foreign subsidiary when they have domestic NOL carryforwards (Newberry and Dhaliwal, 2001). |
| Use a different financing source than domestic debt, especially when foreign assets are substantial. For example, use leases instead of debt because lease payments are not allocated to foreign operations. | Weak evidence that excess credit firms lease more than other firms (Froot and Hines, 1995).<br><br>U.S. firms' incentive to finance with preferred stock rather than debt increases with proportion foreign assets (Collins and Shackelford, 1992, and Newberry, 1998). |
| Use internal debt infusion rather than internal equity to finance foreign subsidiary, especially when $\tau_{\mathrm{For}}$ is high.<br>Similarly, finance via royalty agreement rather than with equity. | Net internal borrowing by subsidiary from parent increases in $\tau_{\mathrm{For}}$ (Desai, 1997).<br><br>Increase royalty payments when cheaper than repatriating dividends (Hines, 1995). |
| Use transfer pricing to increase (decrease) cash flow to low (high) tax affiliate. | Multinationals overinvoice low-tax affiliates (Lall, 1973).<br><br>Foreign-controlled U.S. firms' U.S. tax expense is inversely related to difference between the U.S. and global tax rate (Mills and Newberry, 2000). |
| Repatriate dividends when excess credit.<br><br>Repatriation for deficit credit firms negatively related to $\tau_{\mathrm{US}} - \tau_{\mathrm{For}}$. | Excess credit firms repatriate more than deficit credit firms, and repatriation by deficit credit firms is inversely related to the cost of doing so (Hines and Hubbard, 1990) |

*(Continued)*

Table 3 (*Continued*)

| Prediction | Empirical Evidence |
|---|---|
| Firm uses less debt when it has accumulated FTCs | None |
| Remit dividends from high- and low-foreign-tax firms simultaneously, to reduce potential domestic taxes. | Most repatriated dividends are "cross-credited" (Altshuler and Newlon, 1993). |
| Borrow via U.S. subsidiary that is less than 80% owned by multinational parent. | Example: Ford Motor Co. set up domestic financing subsidiary of which it owned 75% (Scholes et al., 2002). |
| Use triangle arrangements between subsidiaries in foreign jurisdictions with different tax burdens to reduce domestic taxes owed on remittances. | Low-foreign-tax subsidiaries invest in high-tax affiliate subsidiaries, which in turn remit funds to U.S. parent at low or zero domestic tax liability; or low-foreign-tax subsidiaries are capitalized by high-tax affiliate subsidiary, so repatriations from high-tax subsidiary are assigned a foreign tax rate that is a mixture of the low- and high-tax rates (Altshuler and Grubert, 2000). |

remittance does not reduce total payments. Firms hold the total constant and substitute between different forms of remittance, such as dividends, interest, or royalties.

There is no known research that explicitly investigates Prediction 10 (the incentive $\tau_C(.)$ to finance with domestic debt decreases with accumulated foreign tax credits for deficit credit firms) and Prediction 14 (the tax incentive to issue debt decreases with the degree of dividend imputation).

Other than Altshuler and Mintz (1995), most papers use very general specifications to test for foreign tax effects or the influence of interest allocation. For example, when they are considered at all, separate terms indicating excess credit status, $\tau_{\mathrm{For}}$, or the ratio of foreign to worldwide assets are used, rather than interacting the variables in the manner suggested by the theory. Also, the sharper predictions are often ignored. Finally, the existence of any multinational tax research that directly links the tax benefits of debt to firm value is not known. To the extent that data are available, variation across countries in tax rules and incentives provides a rich and under-researched environment within which to investigate how variation in tax rules affects $\tau_C(.)$ and, therefore, the financing decisions of multinational firms.

### *3.3. Other predictions and evidence about multinational tax incentives*

Interest allocation can be avoided altogether if the domestic borrowing is performed by a domestic subsidiary that is less than 80% owned by the parent (although this subsidiary must allocate interest on its own books). I am unaware of any systematic research investigating this issue. Scholes et al. (2002) present an example describing how Ford Motor Co. implemented this strategy.

Besides directly altering where and whether it issues debt, there are many related mechanisms by which a firm might respond to multinational tax law. A company might alter its transfer prices (the prices at which goods and services are transferred between related entities) to shift income from the high-tax to the low-tax affiliate. Although transfer prices are supposed to be "arms-length prices," the rules are vague enough to allow wiggle room. Properly designed, transfer pricing allows for tax-free dividend repatriation. Consistent with this means of reducing overall taxes, Lall (1973) reports that multinational firms overinvoice their low-tax Colombian subsidiaries. Mills and Newberry (2000) find that shifting income to foreign operations increases the difference between the U.S. tax rate and the global tax rate. Alternatively, multinational firms can use "triangle schemes" in which one subsidiary is capitalized by or invested in by another affiliate subsidiary (Altshuler and Grubert, 2000). These schemes allow firms to optimally mix remittances from high- and low-tax subsidiaries in ways that reduce domestic taxes on foreign-source income.

More generally, firms can time dividend repatriation to coincide with low overall tax cost to the parent and subsidiary. In particular, deficit credit firms owe U.S. tax when they repatriate dividends, so they have the incentive to delay repatriation. In contrast, excess credit firms often do not owe additional tax upon repatriation. Taking debt versus equity choices as given, Hines and Hubbard (1990) find that excess credit firms repatriate more than do deficit credit firms and that repatriation by deficit credit firms is inversely related to the tax cost of doing so. Altshuler and Newlon (1993) show that most repatriated dividends are "cross-credited"; that is, the parent firm simultaneously receives payments from both high- and low-foreign-tax subsidiaries, and can use the extra credits from one source to offset potential domestic taxes from another.

## 4. Taxes, LBOs, corporate restructuring, and organizational form

### *4.1. Theory and predictions*

Under perfect capital markets, an MM analysis implies a null hypothesis that organizational form and restructurings are irrelevant to firm value. However, imperfections in the tax, legal, and information environments can create situations in which the form of the organization or restructuring can matter.

#### *4.1.1. Leveraged buyouts*

There is a tax incentive for corporations to use substantial leverage in the management buyout process. This flows directly from the predictions in Section 1 that high-tax-rate firms have incentive to use debt and that the associated tax benefits add to firm value.

Leveraged buyouts (LBOs) are particularly interesting because they lead to a much larger increase in leverage than do most debt issuances. LBOs also can provide an opportunity to mark assets to market, thereby increasing depreciation and the associated tax savings.

Prediction 15: All else equal, the tax incentive to perform a highly levered buyout increases with the firm's expected post-deal tax rate, $\tau_C(.)$.

### *4.1.2. Distressed reorganizations and chapter 11*

Tax incentives can affect distressed reorganizations. Distressed firms with substantial accumulated net operating losses (NOLs) have incentive to file Chapter 11 because it facilitates reducing debt ratios (Gilson, 1997). Chapter 11 allows the firm that emerges from bankruptcy to have unlimited use of the pre-filing NOLs to shield future income, as long as there is no change in ownership (i.e., a large change in the ensuing two years in ownership of the firm's equity). Reducing the debt ratio during reorganization preserves debt capacity and decreases the likelihood of precipitating an ownership change by future equity issuances.

Prediction 16: The tax incentive for a firm to file Chapter 11 (versus a workout), to better facilitate reducing its debt ratio in reorganization, increases with the firm's accumulated NOL carryforwards and its expected post-deal tax rate.

### *4.1.3. C-corporations vs. S-corporations*

Taxes affect organizational form in general, not just reorganizations. When an entity operates as a common "C-corporation," revenues returned to investors as equity are taxed at both the firm and investor levels. The firm-level taxation is at the corporate income tax rate, and the investor taxation is at the personal equity tax rate. The equity rate is often relatively low because equity income can be deferred or taxed at the relatively low capital gains rate. In contrast, partnership income is passed-through and taxed only at the investor level, at ordinary income tax rates. The tax burden is often disadvantageous to corporate form. For example, at current maximum statutory federal tax rates (Fig. 1), in 2002 an investor would have received \$0.604 in partnership income; in contrast, corporate equity payments would have returned only approximately \$0.52 (assuming equity is taxed at a 20% capital gains tax rate). There are, however, nontax benefits to corporate form that outweigh the tax costs for many firms. Gordon and MacKie-Mason (1994) argue that these nontax benefits are large, annually equaling about 4% of equity value. See Scholes et al. (2002) and Gordon and MacKie-Mason (1997) for details about nontax costs and benefits of corporate form. See Shelley, Omer, and Atwood (1998) for a discussion of the costs.

Prediction 17: All else equal, the tax incentive to operate as a C corporation (versus a partnership or S-corp) increases in $[(1 - \tau_P) - (1 - \tau_C)(1 - \tau_E)]$.

#### 4.1.4. Divestitures and asset sales

Tax incentives can also affect the valuation, purchase, and sale of assets. Alford and Berger (1998) argue that high-tax-rate firms prefer spin-offs when they shed assets that lead to taxable gains because spin-offs can be structured to avoid taxes to both the seller and buyer. In contrast, all else equal, sales are preferred when the transaction results in a loss because this loss can be deducted against corporate income. Moreover, when a firm sells an asset, the deal can be structured to benefit the seller or purchaser, possibly by financing the deal with debt (Erickson, 1998).

Prediction 18: There is a tax incentive for high-tax firms to shed assets in spin-offs when the deal is profitable and via sales when the deal is not profitable. When a firm acquires assets, high-tax firms have the incentive to use "taxable deals" financed with debt.

#### 4.1.5. R&D partnerships

Leasing allows a low-tax-rate firm to "sell" tax deductions to high-tax-rate lessors. Analogously, research and development limited partnerships (RDLPs) allow low-tax firms to sell start-up costs and losses to high-tax-rate investing partners.

Prediction 19: All else equal, low-tax-rate R&D firms should form research partnerships with high-tax-rate investors.

### 4.2. Empirical evidence

Kaplan (1989) and others investigate tax benefits in leveraged buyouts. LBOs provide large interest tax deductions and also can provide an opportunity for asset value to be stepped up to market value. Note that the tax benefit of $1 of interest does not necessarily equal the top statutory tax rate. The net benefit is less than the top rate if all of the LBO interest expense cannot be deducted in the current year, if there is a personal tax penalty on interest income, or if there are nontax costs to debt. Assuming that the net tax benefit of $1.00 of interest is $0.15 and that LBO debt is retired in eight years, Kaplan estimates that the tax benefit of interest deductions equals 21% of the premium paid to LBO target shareholders.[41] Kaplan also estimates that among firms electing to step up asset value, the incremental depreciation tax benefit equals 28% of the premium. It is not known if there is any research that explicitly investigates whether the probability of choosing a highly levered form of reorganization increases with the expected post-deal MTR (Prediction 15).

Gilson (1997) shows that firms in Chapter 11 reduce their debt ratios more when pre-filing NOLs are large (Prediction 16). He concludes that firms file Chapter 11 (versus a

[41] Graham (2000) accounts for the declining marginal benefit of incremental interest deductions and estimates that the gross tax benefit of debt equaled approximately one-fourth of firm value in the mid-1980s RJR Nabisco and Safeway LBOs.

workout) in part because of tax incentives: Chapter 11 status offers smaller transactions costs to reducing the debt ratio, thereby minimizing the chance of an ownership change that would result in the loss of pre-filing NOLs.

Research centered on tax reforms has linked taxes with organizational form. The Tax Reform Act of 1986 (TRA86) set corporate tax rates above personal income tax rates, and also equalized capital gains and ordinary tax rates, providing a natural environment to test Prediction 17. These tax-rate changes made partnerships attractive by greatly increasing the tax disadvantage of operating as a corporation. Scholes et al. (2002) point out that there was a huge increase in the formation of S-corporations (which are taxed as partnerships) following TRA86. Gordon and MacKie-Mason (1997) show that the increased corporate tax disadvantage due to TRA86 resulted in a reduction in the portion of aggregate profits paid via (and assets held in) corporate form. However, the economic importance of this reduction was modest. Finally, Guenther (1992) investigates how corporations responded to the 1981 Economic Recovery Tax Act reduction in personal income tax rates, which increased the tax disadvantage for corporations. He finds that firms altered policies that contribute to the double taxation of equity payout: firms reduced dividends and instead returned capital by increasing the use of debt, share repurchases, and payments in mergers (which are often taxed as capital gains).

Ayers, Cloyd, and Robinson (1996) study small firms and find that entities choose to operate as S-corps, rather than C-corps, when they experience losses in their early years of operation. These losses can immediately be passed through to S-corp investors, while C-corps must carry losses forward to offset future corporate income. The experiment of studying small firms is especially telling because small firms can generally choose between the S- or C-corp form with little difference in cost or nontax considerations; therefore, the choice highlights tax incentives. Interpreting this result as strong tax evidence is somewhat clouded, however, because Ayers et al. do not find that the choice between C-corp and proprietorship/partnership form is affected by tax losses (though nontax considerations can affect this choice). Erickson and Wang (2002) contend that S-corps can be sold for more than C-corps because of favorable tax treatment. Finally, Hodder, McAnally, and Weaver (2001) conclude that banks convert to S-corp status to eliminate double taxation of dividends and to reduce the onerous burden of the AMT. Research investigating organizational form choices using micro firm- and owner-specific tax information would be helpful. Such papers would most likely require accessing confidential tax returns.

Scholes and Wolfson (1990) describe tax incentives that encouraged merger and acquisition activity in the early 1980s (following the 1981 tax act) and discouraged these activities after TRA86. They provide aggregate evidence that M&A activity surged in the early 1980s and declined in 1987, consistent with tax incentives. See Scholes et al. (2002) for details of how acquisitions vary along the tax dimension depending on whether the deal involves C- or S-corporations, subsidiaries, spin-offs, carve-outs, and so on.

Alford and Berger (1998) show that firms trade off tax and nontax considerations when choosing between spin-offs and asset sales (Prediction 18). They estimate tax benefits as a means of determining the size and nature of nontax costs and argue that adverse selection, moral hazard, and agency costs are all traded-off against tax benefits to influence how firms structure their deals. Erickson (1998) also demonstrates that the structure of deals is affected by tax concerns. He shows that the probability that a sale is structured as a "taxable deal," financed with tax-deductible debt, increases with the acquirer's tax rate; however, he finds no evidence that seller tax characteristics affect deal structure. Erickson and Wang (2000) find that the price of subsidiary sales can be affected by tax considerations. These authors show that premiums (and seller abnormal stock returns) increase when the sale is structured to allow a step-up in subsidiary basis, so that the acquiring firm receives additional depreciation tax benefits. Thus, contrary to a Modigliani and Miller perfect markets null hypothesis, tax considerations affect both the pricing and structure of asset sales.

While taxes appear to affect the structure and price of some deals, the tax-minimizing form is not always selected. Hand and Skantz (1998) maintain that issuing new shares in equity carve-outs can avoid tax liabilities that occur when a firm issues secondary shares (at a price above the firm's tax basis in the shares). The authors determine that, relative to issuing new shares, secondary carve-outs increase tax liabilities by an amount equal to 11% of the carve-out IPO proceeds. Hand and Skantz are not able to identify benefits associated with secondary carve-outs that are large enough to offset the increased tax payment. Maydew, Schipper, and Vincent (1999) find that incremental taxes incurred when firms perform taxable sales (rather than tax-free spin-offs) amount to 8% of the value of divested assets. The authors argue that firms incur these tax costs (1) because they are smaller than the financial reporting benefits (e.g., larger financial statement earnings) and (2) when selling firms are cash-constrained (sales provide a cash inflow; swaps do not).

Shevlin (1987) investigates whether firms that perform R&D via partnerships have lower tax rates than firms that do R&D in-house (Prediction 19). Two notable features of Shevlin's careful experimental design are his use of simulated tax rates, and his specification of many explanatory variables in "as-if" form (i.e., defining right-hand-side variables for all firms as if they funded R&D in-house, to avoid the endogenous choice of in-house versus RDLP possibly affecting the variables' values). Shevlin shows that tax rates exert a significant, negative influence on the probability of choosing an RDLP in two out of three as-if regressions. Using an NOL dummy to measure tax incentives, Beatty, Berger, and Magliolo (1995) find that low-tax firms are more likely to finance R&D via a financing organization both before and after TRA86.

The Research and Experimentation Tax Credit has also influenced corporate R&D spending. In his economically weighted regressions, Berger (1993) finds a positive market reaction to announcements affirming the tax credit. His regression coefficients indicate that three-fourths of the benefit of the credit accrues to shareholders, with the remaining one-fourth increasing product price and therefore flowing to employees or suppliers. This latter finding implies that the tax credit creates an implicit tax in the form of higher prices for tax-favored R&D activity and that this implicit tax offsets some of the intended benefit from the credit (in other words, some of the R&D tax credit is passed along in the form of higher prices to suppliers of R&D inputs). Berger also detects a negative market reaction among firms that do not use the credit themselves but compete with firms that do. Swenson (1992) finds evidence consistent with low-tax-rate firms pursuing firm-specific R&D tax credits less aggressively than they are pursued by high-tax-rate firms.

Overall, this research indicates that tax considerations affect the structure and pricing of research and development activity in the United States. The cited papers investigate R&D spending associated with pre-TRA86 tax rules. It is not know, if there is any similar research that investigates the influence of the tax credit on R&D activity based on post-TRA86 rules (under which the credit is based on the R&D-to-sales ratio, rather than on nominal R&D spending). Moreover, the R&D tax credit has temporarily expired several times since 1986. It would be interesting to know whether these expirations have affected real R&D activity.

## 5. Taxes and payout policy

Modern dividend research began with Lintner's (1956) field interviews with 28 firms. Lintner found that dividends are stable, appear to adjust toward an earnings-payout target, and are rarely reduced. Miller and Modigliani (1961) provide the theoretical foundation of payout policy and conclude that dividend policy is irrelevant in a frictionless world with perfect capital markets. Research since that time has explored how market imperfections create an environment in which payout policy affects firm value. This section highlights the tax incentives related to corporate payout policy. For brevity, I narrow the discussion to payout issues that parallel those in Section 1 or that shed light on unresolved capital structure issues (e.g., whether personal taxes affect security prices). For broad reviews of the various tax and nontax imperfections that can lead to payout policy affecting firm value and corporate decisions, see Allen and Michaely (1995, 2001) and Poterba (2001).

### *5.1. Theory and empirical predictions*

Miller and Modigliani (1961) argue that in a perfect economic environment, firm value is determined by operating cash flows, not by whether a company retains or pays out profits, or by the form of payout. This line of reasoning produces the null hypotheses for this section.

Null hypotheses: Firm value is not affected by payout policy. Taxes do not affect corporate payout decisions.

Allen and Michaely (2001) show that the null can also hold if different classes of investors are taxed differently and firms have differing payout policies, as long as the marginal price-setter is tax-free.

Alternatively, firms can have a tax incentive to return equity capital via share repurchases rather than dividends if dividends are taxed more heavily than are capital gains for the marginal investor(s). Financial executives' statements that repurchases are a "tax efficient means of returning capital to investors" support this point of view (though Brav, Graham, Harvey, and Michaely, 2003, conclude that taxes only play a second-order role in the choice between returning capital as dividends or repurchases).

If dividends are taxed more heavily than repurchases, there can be a negative valuation of dividends (relative to repurchased shares) (e.g., the CAPM with corporate and investor taxation in Brennan, 1970, or Auerbach and King, 1983). All else equal, if a firm were to increase dividends, the pretax return on its stock would need to increase so that after-tax returns did not change. This effect increases as dividend taxation increases relative to capital gains taxation.

Prediction 20: All else equal, tax effects imply that firm value is negatively related to (1) the portion of payout dedicated to dividends, and (2) dividend taxation relative to capital gains taxation. Analogously, required pretax stock returns increase with dividend payout and relative dividend taxation.

Nontax factors also can lead to negative (e.g., reduced funds to pursue positive NPV projects) or positive (e.g., signaling or agency alleviation) dividend valuation (see Allen and Michaely, 2001).

Note that dividend clienteles, in which high-tax-rate investors own stocks with low-dividend payouts, can occur under the null or Prediction 20. Under the null, firms can have different payout policies that do not affect value, even if some investors are taxed more heavily on dividends (capital gains) and have a tax preference for capital gain (dividend) income. Similar clienteles can form under Prediction 20, based on the relative taxation of dividends and capital gains for different groups of investors.

To the extent that transactions are not costless, clientele tax characteristics can affect security prices. For example, the price of a stock changes from $P_{\mathrm{cum}}$ to $P_{\mathrm{ex}}$ as the stock goes ex-dividend. If the firm issues a dividend Div, its investors receive $\mathrm{Div}(1-\tau_{\mathrm{div}})$ but simultaneously avoid capital gains taxes of the amount $(P_{\mathrm{cum}} - P_{\mathrm{ex}})\tau_{\mathrm{cap\,gain}}$. With risk neutrality, continuous prices, and no transactions costs, and clienteles that do not vary before and after ex days, Elton and Gruber (1970) show that $(P_{\mathrm{cum}} - P_{\mathrm{ex}})(1-\tau_{\mathrm{cap\,gain}}) = \mathrm{Div}(1-\tau_{\mathrm{div}})$ in equilibrium, and therefore

$$\frac{P_{\mathrm{cum}} - P_{\mathrm{ex}}}{\mathrm{Div}} = \frac{(1-\tau_{\mathrm{div}})}{(1-\tau_{\mathrm{cap\,gain}})} \tag{8}$$

where $(P_{\mathrm{cum}} - P_{\mathrm{ex}})/\mathrm{Div}$ is referred to as the ex-day premium.

Prediction 21: The ex-day premium reflects the relative taxation of dividends and capital gains for a given stock's clientele of investors.

Allen and Michaely (2001) call dividend clienteles "static" if they do not vary through time. Alternatively, if there are advantages to trade among differentially-taxed investors, dividend clienteles might be dynamic, which can lead to changes in the composition of the clientele around certain dates. Dynamic clienteles might lead to abnormally high volume around ex days. For example, low-dividend-tax investors might buy stocks just before ex day, capture the dividend, then sell the stock after it goes ex dividend. Through this route, taxes might lead to ex-day behavior that produces trading volume but where the ex-day premium is close to one. Thus, Prediction 21 is a joint prediction about clienteles being static as well as tax effects.

Payout effects should vary with the tax rules of the country under consideration. For example, assuming static clienteles, the ex-day premium should increase with the degree of dividend imputation in a given country (because a tax refund for corporate taxes paid is attached to dividends in imputation countries, which reduces the effect of dividend taxation). The premium can be greater than one if imputation makes dividends tax-favored relative to capital gains (Bell and Jenkinson, 2002).

### 5.2. *Empirical evidence on whether firm value is negatively affected by dividend payments*

Black and Scholes (1974) test Prediction 20 by adding dividend yield as a right-hand side variable in the market model. They conclude that firm value is not related to dividends. In contrast, Litzenberger and Ramaswamy (1979) find a significant, positive dividend-yield coefficient. Kalay and Michaely (2000) emphasize that the positive dividend effect should show up in cross-sectional (because of cross-firm variation in dividend-payout) long-run returns (i.e., returns for stocks held long enough to qualify for capital gains treatment). They point out that Litzenberger and Ramaswamy (1979) use monthly returns and allow high-dividend yield firms to be considered zero-dividend in nondividend months. Kalay and Michaely (2000) do not find cross-sectional or long-run return evidence that high-dividend stocks earn a tax premium. Kalay and Michaely imply that the effect identified by Litzenberger and Ramaswamy occurs for short-run returns, perhaps only during the ex-dividend week.

Fama and French (1998) test Prediction 20 by regressing (changes in) firm value on (changes in) dividends and "firm value if no dividends."[42] If personal taxes reduce the value of dividends, and one could design a clean statistical experiment that isolates tax effects, there should be a negative coefficient on the dividend variable in this specification. In contrast, Fama and French find a positive coefficient, which probably occurs because either their proxy for "firm value if no dividends" is measured with error and/or nontax effects overwhelm the tax influence of dividends. For example, if firms use dividends to signal quality, dividend payments might be positively correlated with firm value. Or if dividends are priced by tax-free investors, one would not expect a negative

[42] As discussed in Section 1.2.2, FF regress the excess of market value over book assets on dividends, interest, and a collection of variables that are proxy for $V_{\mathrm{U}}$, with all variables deflated by assets. The variables that are proxy for $V_{\mathrm{U}}$ include current earnings, assets, R&D spending, and interest, as well as future changes in earnings, assets, R&D, interest, and firm value. $V_{\mathrm{U}}$ is probably measured with error, which clouds interpretation of FF's results.

influence of dividends on firm value. Fama and French conduct the only known study that directly regresses firm value on dividend variables in an attempt to determine the tax effect of dividends.[43]

### 5.3. *Evidence on whether ex-day stock returns and payout policy are affected by investor taxes*

#### 5.3.1. *Dividend clienteles*

Because Prediction 21 is based on the existence of static dividend clienteles, we start by reviewing dividend clientele research. Blume, Crocket, and Friend (1974), Pettit (1977), and Chaplinsky and Seyhun (1987) provide weak evidence that investors hold stocks such that dividend yield is inversely related to personal tax rates; Lewellen, Stanley, Lease, and Schlarbaum (1978) find no such evidence. However, these studies have poor measures for tax, risk, and wealth effects and therefore are hard to interpret. Auerbach (1983) concludes that tax-related preferences result in clienteles of investors that purchase stocks based on firm-specific dividend-price ratios. Scholz (1992) uses self-reported data from the 1983 Survey of Consumer Finances. This survey contains information on retail investor stock holdings, a sophisticated measure of the investor's relative dividend and capital gains tax rates, household wealth, and self-declared risk preferences. Scholz finds a negative relation between the dividend yield for an investor's stockholdings and the relative taxation of dividends, which is consistent with a general preference for dividends by low-tax investors. Graham and Kumar (2004) investigate stockholdings and trades from brokerage house investors during 1991–1996. They find that retail investors as a group prefer nondividend-paying stocks and that institutions prefer dividend-paying stocks. Within the class of retail investors, Graham and Kumar also report evidence of dividend clienteles. Low-income (i.e., low-tax-rate) and older investors prefer dividend-paying stocks, and within the class of dividend-paying stocks, older low-tax-rate investors prefer high-yield stocks. They also show that high-income retail investors decreased their dividend holdings when dividend tax rates increased in 1993.

According to Strickland (1996), mutual funds and money managers hold low-dividend-yield portfolios, while untaxed institutions such as pension funds show no preference. Dhaliwal, Erickson, and Trezevant (1999) find that the percentage of shares owned by institutional investors increases by about 600 basis points in the year after a firm initiates paying a dividend.[44] Overall, there is weak evidence that the preference for dividends decreases with income tax rates—but no direct evidence that this preference leads to static tax-based clienteles.

[43] Another approach to study whether personal taxes affect asset prices investigates tax capitalization. See Section 1.4.1.

[44] See Del Guercio (1996) and Brav and Heaton (1997) for evidence that institutional investors favor high-dividend stocks for nontax reasons like prudent-man regulations.

Several papers link corporate actions to the (assumed) tax characteristics of their investors. Pérez-González (2000) classifies firms by whether their largest shareholder is an individual or an institution and finds that the individual pays 30% fewer dividends than the institution. He also shows that when tax reform increases (decreases) the taxation of dividends relative to capital gains, firms with large retail shareholders decrease (increase) dividend payout. Poterba and Summers (1985) find a similar result for aggregate dividend behavior in the UK from 1950 to 1983. Lie and Lie (1999) also conclude that investor-level taxes affect payout policy. They find that firms with low-dividend payout (and presumably high-tax-rate investors) use self-tender-offer share repurchases more often than they use special dividends; these firms also use open-market repurchases more often than they increase regular dividends.

Allen and Michaely (1995) point out that the trading volume around ex days provides evidence about whether clienteles are static (which would imply that trading only occurs between investors in the same tax bracket, who always hold stocks with the same dividend characteristics) or dynamic (in which case there might be advantages to trade among differentially taxed investors, potentially involving dividend-capture or arbitrage by low-dividend-tax investors). In the static case, there should be no abnormal volume because there are no abnormal advantages to trade around the ex day. Grundy (1985), Lakonoshok and Vermaelen (1986), and Michaely and Vila (1996) find evidence of abnormal trading volume on the ex day, which is consistent with dynamic tax-related trading on the ex day.[45]

### 5.3.2. *Ex-day premia and returns*

Elton and Gruber (1970) note that the ex-day premium was 0.78 on average in the 1960s, which they interpret to imply that dividends are priced at a 22% disadvantage relative to capital gains (Prediction 21). Moreover, the premium ranged from 0.70 (for the lowest dividend-yield decile of stocks) to 1.18 (for the highest decile), which is consistent with the highest (lowest) tax-rate investors purchasing the lowest (highest) dividend-yield stocks. The Elton and Gruber evidence is consistent with personal taxes affecting stock prices via dividend payout and dividend clienteles. Their findings are strengthened by Barclay's (1987) evidence that the premium was 1.0 in the early 1900s, before the advent of personal income taxes.

Interpreting the ex-day phenomenon presents several complications. Kalay (1982) points out that absent transactions costs and risk, arbitrage by tax-free investors should push the premium to 1.0. Kalay argues that transactions costs are too large for individual investors to be the marginal price-setters, but instead zero-tax-rate institutions

[45] Koski and Michaely (2000) find that abnormal volume can be quite large on ex days due to nontax activity. In their case, Japanese insurance companies captured dividends for regulatory reasons, using nonstandard settlement procedures that allowed them to buy just before and sell just after the ex day. Note, however, that this form of nonstandard settlement ended in 1989, so it cannot explain abnormal ex-day volume in recent years.

might fulfill that role at ex-day. Kalay's findings suggest that inferring tax rates from ex-day returns is complicated by transactions costs and the effect of institutional traders. Consistent with this view, Michaely (1991) finds that the mean premium equaled approximately 1.0 in both 1986 (when capital gains tax rates were much lower than dividend tax rates for wealthy individuals) and in 1987–1988 (when statutory dividend and capital gains tax rates were nearly equal), and was relatively invariant across dividend-yield deciles during these years. Michaely's evidence is not consistent with retail investor taxation affecting stock prices, suggesting that prices might have been set by institutional investors in the mid-1980s.[46]

Bali and Hite (1998) contend that discrete stock prices lead to patterns consistent with those observed by Elton and Gruber (1970). Suppose a $0.20 dividend is paid, and, during the era when stock prices were divisible by one-eighth, the stock price drops by the largest increment less than the dividend: $0.125. This implies an ex-day premium of 0.625, which occurs in the absence of personal tax effects. Moreover, this effect is strongest for low-dividend stocks. Bali and Hite's (1998) argument might explain some of the observed ex-day phenomenon; however, it does not explain abnormal volume on the ex day, which Michaely and Vila (1996) consider evidence of tax-motivated trading.

Frank and Jagannathan (1998) argue that dividends are a nuisance and that market-makers are well situated to handle their collection and reinvestment. Therefore, investors unload the stock cum-dividend to market-makers, who are compensated for handling the dividend by the dividend itself. This is especially true for low-dividend stocks, for which the nuisance remains relatively the same but for which the reward for handling the dividend is smallest. The implication is that prices should fall by less than the dividend in part because transactions are at the bid when the market-maker buys the stock on the cum date and are at the ask when the market-maker sells the stock ex dividend, and in part due to reduced demand on the cum date. They present evidence consistent with their arguments on the Hong Kong exchange, where the average premium is approximately one-half during 1980–1993, even though dividends and capital gains are not taxed at the personal level. Kadapakkam (2000) strengthens this argument by showing that when the nuisance of handling dividends (i.e., cumbersome physical settlement procedures) was greatly reduced with the onset of electronic settlement, the premium in Hong Kong became indistinguishable from 1.0.

Graham, Michaely, and Roberts (2003) cast doubt on price discreteness (Bali and Hite) or bid-ask bounce (Frank and Jagannathan), explaining ex-day pricing in the United States. Graham et al. note that price discreteness and bid-ask bounce were greatly reduced as pricing increments changed from 1/8ths to 1/16ths (in 1997) to decimals (in 2001) on the New York Stock Exchange. According to the price discreteness and bid-ask bounce hypotheses, the ex-day premium should have moved closer to one as the pricing grid

[46] This discussion ignores the effect of risk (see Michaely and Vila, 1995) and transactions costs (see Boyd and Jagannathan, 1994; Michaely and Vila, 1996; and Michaely, Vila, and Wang, 1996) on ex-day behavior. For example, Boyd and Jagannathan (1994) regress capital return on dividend yield and find a slope coefficient of one and a negative intercept. They interpret the negative intercept as a measure of transactions costs.

became finer. In contrast, the ex-day premium got smaller (further from one), which is inconsistent with the price discreteness and bid-ask bounce hypotheses. Graham et al. do find evidence consistent with the original Elton and Gruber tax hypothesis, however. They find that the ex-day premium fell in conjunction with the 1997 reduction in capital gains tax rates.

Graham and Kumar (2004) observe that low-tax rate (and older) investors purchase stocks just before they go ex dividend, especially high-dividend-yield stocks, while high-tax-rate (and younger) investors wait until after the ex day. This is consistent with clienteles of investors who have a preference for dividends actively acquiring dividends around the ex day. Graham and Kumar also find that among low-market-capitalization stocks, for which retail investors are plausibly the marginal price-setters, the ex-day premium is lower in absolute magnitude in stocks for which high-tax-rate investors dominate. This is consistent with the implication from Elton and Gruber (1970) that investor tax rates are impounded into ex-day stock returns, and in particular, low-tax-rate investors' tax rates are impounded into the ex-day returns of high-dividend-yield stocks.

Overall, it is not possible to unambiguously interpret the ex-dividend day evidence in terms of personal taxes, though some recent evidence is compelling. Green and Rydqvist (1999) provide convincing evidence of personal taxes being impounded into asset prices. Swedish lottery bonds are taxed like common stock with tax-free dividends (i.e., the coupon is tax-free and capital gains are taxed). Therefore, one would expect prices to be bid up cum-coupon (as high-tax-rate investors purchase the bonds) and drop after the coupon is paid (with the drop leading to a capital loss deduction, which reduces taxes in proportion to the capital gains rate). Because the coupon is tax-free, the ratio of price drop to coupon should be greater than 1.0, reflecting the personal tax advantage of the coupon. Moreover, regulations prohibit coupon capture or arbitrage of the type that might be expected to force the ratio to 1.0, and unlike the case of stock dividends, frictions and price discreteness work in the opposite direction of the proposed tax effect. Green and Rydqvist (1999) state that the ratio of price drop to coupon averages 1.30 for Swedish lottery bonds, implying that the relative tax advantage of coupons relative to capital gains is impounded into bond prices. They also find that this implicit tax rate declined as tax reform reduced the top statutory personal tax rate during the 1980s and 1990s. Florentsen and Rydqvist (2002) report that the ratio averages about 1.46 for similar lottery bonds in Denmark.

McDonald (2001) investigates ex-day behavior in Germany, where the dividend imputation tax system attached to most dividends a tax credit for corporate taxes (until this feature was repealed in late 2000). This tax credit means that dividends are more valuable to German investors than are capital gains, all else equal. McDonald shows that tax considerations imply that the ex-day premium should be 1.43 under these conditions. In his empirical work, McDonald shows that the average premium is 1.26, indicating that about 60% of the dividend tax credit is impounded into the ex-day price. He also finds that 55% (35%) of the tax credit is reflected in futures (options) prices. Finally, McDonald demonstrates that there is abnormal volume for the six days leading up to and including the ex day, and that abnormal volume increases in the dividend yield. This is consistent with foreigners, who do not enjoy the German tax credit, selling the stock just before ex day.

Bell and Jenkinson (2002) investigate the effects of a July 1997 tax reform in the UK. Prior to 1997, the imputation tax credit attached to dividends was such that tax-free investors received a full tax refund, even though they did not pay taxes on the dividend. In other words, a \$1 dividend was worth more than \$1 to tax-free investors. The tax reform eliminated imputation credits for tax-free investors, implying that a \$1 dividend is now worth only \$1 to these investors. Bell and Jenkinson show that tax-free institutions like pension funds own the majority of UK equities and that they therefore are the marginal price-setters. Bell and Jenkinson find that dividend valuation decreased after the tax reform effectively reduced the imputation tax benefit of dividends. It would be helpful if there were more such research that exploits the rich variation in tax codes around the world.

Overall, some ex-day papers provide clear evidence that personal taxes affect asset prices. This conclusion is not unambiguous across all papers, however, because of potential nontax explanations for abnormal ex-day returns. If these alternative hypotheses completely explain ex-day returns, then in these circumstances personal taxes are not impounded into stock prices. However, even if tax rates do not appear to affect stock returns directly, tax considerations might still affect financial markets if they increase trading volume.

The payout results have implications for capital structure research. If the marginal investor in equities is tax-free but the debt price-setter is not, then the personal tax penalty for using debt might be quite large. If the marginal investor in equities *and* debt is tax-free, there is no personal tax penalty associated with debt financing. Finally, if the marginal price-setter for equities is taxable and his tax rate is impounded into stock returns, this reduces the personal tax penalty on debt relative to the Miller (1977) scenario. Understanding the tax characteristics of the marginal price-setter(s) in various securities is an important issue for future research.

## 6. Taxes and compensation policy

### *6.1. Theory and empirical predictions*

An MM-perfect capital markets analysis would lead to a null hypothesis that compensation policy does not affect firm value absent market imperfections. There has been a great deal of research investigating how agency costs and informational asymmetry can drive a wedge between employee objectives and shareholder wealth, as well as how compensation policy can improve the situation.[47] Another group of papers

[47] See Murphy (1999) for a broad review of compensation research, including pay-performance sensitivity and linking salary, bonuses, and stock compensation to firm performance. See Core, Guay, and Larcker (2002) for a review that focuses on using equity compensation to align executive and shareholder incentives.

investigates how the tax code can affect the choice of when and how to pay employees. Analogous to Miller's (1977) arguments about capital structure, the study by Scholes et al. (2002) argues that to understand compensation policy, one must consider the tax implications for both the employer and employee. Scholes et al. show how different tax rates for the firm and its employees, or changing tax rates for either party, produce trade-offs between salary and bonuses, deferred compensation, compensatory loans, pension contributions, fringe benefits, and stock option compensation. This section reviews compensation research that is most closely linked to taxes and corporate finance: the choice of salary versus equity compensation, the choice between incentive stock options (ISOs) and nonqualified stock options (NQOs), and, linking back to Section 1 of this chapter, the trade-off between compensation deductions and debt tax shields.

The first tax issue is straightforward. Salary payments lead to an immediate deduction that reduces tax liabilities, while employee stock options lead to a corporate deduction only when the options are eventually exercised (if then—see below). Ignoring incentives and other nontax issues, the first compensation prediction is

Prediction 22: All else equal, the tax preference of paying salary compensation instead of option compensation increases with the corporation's tax rate because salary expense is deducted immediately and option expense is delayed.

The second tax issue involves the choice between paying employees with incentive versus nonqualified stock options. ISOs and NQOs are similar in most respects other than tax treatment, allowing researchers to isolate how tax imperfections affect corporate compensation decisions.[48] With ISOs, the firm never gets a tax deduction, and the employee pays capital gains tax on the amount the share price exceeds the grant price when the stock is eventually sold (assuming that the option is exercised at least 12 months after grant and the share of stock is sold at least 12 months after exercise). With NQOs, on the exercise date the firm gets a deduction equal to the amount by which the price upon exercise exceeds the grant price, and the employee pays ordinary income taxes on this same amount. The tax trade-off between incentive and nonqualified stock options amounts to comparing the relatively light burden of the employee paying capital gains taxes for ISOs to the net NQO benefit (i.e., the corporate deduction less the cost to the employee of paying taxes sooner and at a higher rate with NQOs).

Prediction 23: All else equal, when the corporation is taxed at a higher rate than the employee on ordinary income, nonqualified options are preferred to incentive stock options because they lead to lower "all parties" taxation of option compensation. Incentive stock options are generally preferred if the corporation has a low tax rate relative to the employees.

[48] Stock appreciation rights are similar except that the net benefit is paid in cash, not shares of stock. With stock appreciation rights, the employee pays tax at ordinary personal tax rates on the cash benefit when it is paid, and the firm contemporaneously deducts the cash benefit.

This section also investigates whether deductions from employee stock options serve as nondebt tax shields that substitute for the use of interest tax deductions by corporations. DeAngelo and Masulis (1980) argue that firms with substantial nondebt tax shields will use less debt. Among papers investigating this hypothesis, most find weak or no evidence that the traditional measure of nondebt tax shields (depreciation) crowds out debt tax shields (see Section 1). Section 7.2 reviews recent research that examines whether option deductions might serve the role of nondebt tax shields as laid out in DeAngelo and Masulis. It is not formally stated as a prediction because it is already stated in Prediction $2'$.

Finally, restricted stock is a form of compensation that appears to be gaining popularity as a substitute for stock options (e.g., Microsoft's recent public declaration that it will begin using restricted stock extensively). With restricted stock, the employee is granted the shares of stock but is restricted from selling the shares for a prespecified period. Unless the employee elects (via Section 83(b) of the Internal Revenue Code) to pay ordinary taxes on the shares at the time of the grant, the employee pays ordinary income tax when the restrictions are lifted (typically after a vesting period of several years expires). The company receives a deduction of the same dollar amount upon which the employee pays tax, at the time the employee pays tax. As of this writing, one key difference relative to stock options is that with restricted stock the company *must* take a charge to earnings that is spread over the restriction period (the charge is fixed at the time restricted stock is granted and is based on APB Ruling #25 or the fair value as determined by FASB Ruling #123), while an earnings charge is not required with stock options. The other key difference is that stock options have little value unless the stock price increases, whereas restricted stock is worth the value of a share of stock, and so can have substantial value even if the price falls somewhat after the stock is granted. It is not clear whether there is any empirical research that comprehensively investigates restricted stock.

### *6.2. Empirical evidence*

Several papers that investigate whether corporate and employee tax status affect compensation choice have drawn mixed conclusions. Hall and Liebman (2000) assume that all firms pay the top statutory tax rate, and they report that the use of executive options increased as the corporate tax rate declined from the 1970s to the 1980s. This finding is consistent with Prediction 22 (the tax benefit of options increases as corporate tax rates fall because the forgone opportunity to deduct salary expense immediately is less important). However, when allowing for cross-sectional differences in tax rates and annual fixed effects, the Hall and Liebman tax coefficient becomes insignificant.

In contrast, Core and Guay (2001) examine stock option plans for employees other than the top five executives. Nonexecutives hold two-thirds of outstanding compensation options. Core and Guay find that high-tax-rate firms grant fewer options, consistent with Prediction 22, but that low-tax-rate firms grant more options. Finally, Klassen and Mawani (2000) find that among Canadian firms option use decreases with the corporate marginal tax rate, as in Prediction 22. (Note that option compensation is not deductible for Canadian corporations, which only strengthens the incentive to deduct salary expense immediately.) Overall, then, the evidence is weakly consistent with Prediction 22.

Several papers investigate whether corporate and employee tax status affect the choice between incentive and nonqualified options. Austin, Gaver, and Gaver (1998) assume that executives are taxed at the highest statutory rate and investigate whether high-tax-rate firms use NQOs. Austin et al., using five different variables to measure the corporate tax rate, show that none of the variables are statistically related to the form of option plan. This conclusion is generally consistent with the finding by Madeo and Omer (1994) that low-rather than high-tax-rate firms switched from ISOs to NQOs following the 1969 tax act, opposite the tax prediction. Apparently, at present no research provides evidence unambiguously consistent with Prediction 23.

Consistent with personal tax incentives, Huddart (1998) finds that some employees accelerated the NQO option exercise in 1992, prior to the anticipated 1993 increase in upper income personal tax rates (from 31 to 39.6%). However, he concludes that only one in five employees took this action, indicating that nontax factors more than offset personal tax incentives in many situations. Goolsbee (1999) finds that in aggregate an abnormally large number of options were exercised in 1992, prior to the tax increase. Hall and Liebman (2000) note that Goolsbee defines abnormal based on a linear trend in exercise activity. When they instead consider the number of vested options and recent changes in stock prices, Hall and Liebman do not find that employees accelerated the options exercise in anticipation of the personal tax-rate increase; nor do they find a delay in exercise in anticipation of personal tax-rate reductions in the 1981 and 1986 tax acts.

Matsunga, Shevlin, and Shores (1992) conclude that tax factors affect the disqualification of ISOs. An ISO plan is disqualified (i.e., treated as an NQO plan for tax purposes) if an employee sells her stock less than 12 months after exercising incentive stock options. A company might want to disqualify an ISO plan to receive the corporate deduction associated with NQOs if the corporate tax-rate increases relative to the personal tax rate and/or if the ordinary personal rate falls relative to the capital gains tax rate, both of which happened after the 1986 tax reform. Matsunga et al. perform a careful “all parties” tax analysis and conclude that firms with the largest net benefit of disqualification were the firms most likely to disqualify.

Overall, there is only modest evidence that taxes are a driving factor affecting corporate or employee compensation decisions. This is perhaps surprising because popular press articles indicate that the size of the corporate deduction provided by NQOs is huge, completely eliminating corporate taxes for many large, profitable firms in the late 1990s (e.g., *New York Times*, June 13, 2000).[49]

Footnote information from financial statements can help us understand whether the magnitude of option compensation deductions is sufficient to affect overall corporate tax planning and to determine whether these deductions are inversely correlated with interest

[49] Hanlon and Shevlin (2002) present evidence about options deductions for NASDAQ 100 firms. Hanlon and Shevlin provide an excellent summary of the accounting issues related to options deductions.

deductions (and therefore might explain why some firms use little debt). Information on the exercise and grant prices for all options exercised from 1996 to 1998 by employees of Fortune 500 firms has been gathered (see Table 4). Assuming that all of the options are nonqualified implies that the corporate options deduction equals the difference between the exercise and grant prices of the exercised options. Note that these deductions appear on tax returns and reduce taxes owed to the government; they do not appear as a deduction on financial statements,[50] nor are they collected by Compustat. One could multiply these deductions by $\tau_C$ to estimate their tax value.

The average (median) Fortune 500 firm had \$85 (\$16) million of annual deductions resulting from employees exercising stock options during 1996–1998 (Panel A of Table 4). These numbers are skewed: the firm at the 90th (95th) percentile had \$185 (\$379) million in deductions. As a percentage of financial statement tax expense, the deductions average 50%. As a percentage of the amount of interest it takes to lever a firm up until there are declining benefits associated with incremental deductions (i.e., levering up to the kink in the Graham [2000] benefit functions discussed in Section 1.2.3), the option deductions average 49%. Panel B of Table 4 shows the numbers for some specific firms. In the years shown in the table, option deductions are larger than interest deductions for Dell Computer, Intel, Dollar General Corporation, General Motors, and Circuit City. Moreover, options deductions are larger than tax expense for Intel, GM, and Circuit City.

Overall, the magnitude of the compensation deductions are large for some firms; however, they are moderate for many companies and therefore do not appear to provide the final answer to the puzzle of why some firms appear to be underlevered. Nonetheless, Panel C of Table 4 reveals that compensation deductions appear to substitute for interest deductions and so at least partially address the puzzle. The Pearson (Spearman) correlation coefficient between the magnitude of option deductions and the degree to which a firm appears to be conservatively levered (as measured by the amount of interest it would take to lever up to the kink in the benefit function) is 0.33 (0.46).[51]

[50] Option deductions do not reduce financial statement tax expense because the deductions are not treated as a permanent expense. Instead, the deductions are added to stockholders' equity.

Note that the discussion in this section applies to the vast majority of firms because they elect to present their stock option information using the intrinsic value method and therefore do not expense options and reduce net income or earnings per share, but instead present the information in the financial statement footnotes (and never expense the option compensation to reduce net income). In 2002 and 2003, several dozen firms announced that they would begin expensing option costs on financial statements (which has the effect of reducing net income on financial statements). At press time, most other firms appear to be in a holding pattern, waiting to see whether accounting standards change with respect to reporting option compensation, so it is not possible to say whether more firms will elect to expense options in their financial statements. Note that the tax rules have not changed regarding options—this footnote simply discusses whether firms report options expense as a net-income-reducing item on financial statements.

[51] One shortcoming of this analysis is that I measure the tax benefit of *realized* compensation deductions, which are not necessarily the same as the deductions that managers expect ex ante, when they plan their capital structure. Nor do I distinguish between ISOs and NQOs, although Hall and Liebman (2000) note that NQOs account for 95% of option grants. Future research should address these issues.

Table 4
Corporate tax deductions resulting from option compensation for Fortune 500 firms, 1996–1998

| Panel A Fortune 500 1996–1998 | Annual Option Deduction ($ million) | Deduction/ Interest Expense | Deduction/ Tax Expense |
|---|---|---|---|
| Mean | 85.2 | 9.371 | 0.495 |
| 25th percentile | 3.8 | 0.030 | 0.029 |
| Median | 16.1 | 0.153 | 0.097 |
| 75th percentile | 58.1 | 0.585 | 0.283 |
| 90th percentile | 184.7 | 1.800 | 0.632 |
| 95th percentile | 378.6 | 4.088 | 1.136 |

| Panel B Specific Firms | Option Deduction ($million) | Interest Expense ($million) | Tax Expense ($million) | Deduction/ Interest Expense | Deduction/ Tax Expense |
|---|---|---|---|---|---|
| Dell Comp. (1997) | 468.6 | 3.0 | 424.0 | 156.19 | 1.11 |
| Intel Corp. (1998) | 1185.7 | 40.0 | 3069.0 | 29.64 | 0.39 |
| Dollar Gen. (1997) | 57.8 | 3.7 | 87.2 | 15.36 | 0.66 |
| GM (1998) | 157.1 | 72.8 | −44.7 | 2.16 | −3.51 |
| Circuit City (1998) | 27.2 | 9.1 | −15.0 | 3.00 | −1.81 |

*Option Deduction* is the dollar amount of option compensation expense that a firm can deduct from its taxable income in a given year, which is calculated as the number of options exercised in a given year times the difference between the weighted-average exercise price and the weighted-average grant price. This calculation treats all exercised options as if they are nonqualified options. *Deduction/Interest* is the option compensation deduction divided by interest expense, where *interest expense* is from financial statements. *Deduction/Tax Expense* is the compensation deduction divided by tax expense, where *tax expense* is from financial statements.

Two recent papers investigate whether option deductions displace the use of debt along the lines suggested in DeAngelo and Masulis (1980)—that is, to explore whether option deductions serve as a form of nondebt tax shield that might substitute for interest deductions (Section 1.1). Graham, Lang, and Shackelford (2002) find that the magnitude of option deductions is large enough to reduce the median MTR for NASDAQ 100 and S&P 100 firms from 34% (when option deductions are ignored) to 26% (when option deductions are considered) in 2000.[52] Documenting a reduction in MTRs is important because, as argued in Section 1.3.1, nondebt tax shields should reduce the use of debt to the extent that the NDTS alter the marginal tax rate. Graham et al. find that debt ratios are positively related to tax rates and negatively related to the amount by which option deductions reduce marginal tax rates (consistent with Prediction 2′). Similarly, Kahle and

[52] In light of the large reduction in tax rates for some firms, it is surprising that (for tax reasons) some of these firms do not use more incentive and fewer nonqualified stock options. One reason might be restrictions on the total amount of incentive stock options that can be granted in a given year.

Shastri (2002) find that long- and short-term debt ratios are negatively related to the size of tax benefits from option exercise. Finally, Graham et al. show that firms that appear to use debt conservatively when option deductions are ignored appear significantly less underlevered when options are considered.

Overall, the evidence is consistent with managers substituting away from debt when their firm has substantial option deductions. It would be interesting for future research to investigate whether other nondebt tax shields play this role (e.g., R&D tax credits or foreign tax credits), especially in eras during which option deductions were less prevalent. One "secretive" source of such deductions is tax shelters, which are investigated in Section 8.

## 7. Taxes, corporate risk management, and earnings management

If capital markets were perfect, there would be no benefit to corporate hedging because investors would be able to achieve the same outcome by hedging on personal account. The null hypothesis is therefore that corporate hedging does not increase firm value. And yet, the corporate use of derivatives (presumably) to hedge has increased enormously in the past decade. For example, OTC swaps increased from $11 trillion in 1994 to $69 trillion by 2001 (http://www.isda.org/index.html). A large number of corporate finance research papers investigate which market imperfections create situations that can make corporate hedging advantageous.

Theory suggests that hedging to reduce volatility can reduce the expected costs of bankruptcy (Smith and Stulz, 1985), reduce underinvestment costs by shifting funds into states where they would otherwise be scarce (Froot, Scharfstein and Stein, 1993), help offset conservative decision making that results from employee risk-aversion (Tufano,1996), and reduce the effects of information asymmetry between managers, investors, and the labor market (DeMarzo and Duffie, 1991, and Breeden and Viswanathan, 1998). Though narrower in scope, taxes can also provide an incentive to hedge. This section reviews imperfections in the tax code that can lead to corporate hedging being beneficial and also explores how similar imperfections can provide an incentive to manage earnings.

### *7.1. Theory and empirical predictions*

Smith and Stulz (1985) show that if the function that maps taxable income into tax liabilities is convex, a firm can reduce its expected tax liability by hedging to reduce income volatility. The tax function is generally convex because corporate income tax rates are progressive, though the degree of progressivity for positive income is small. The main form of progressivity occurs because profits are immediately taxed at a positive rate, while the tax-reducing effect of losses is effectively spread through time via tax-loss carrybacks and carryforwards and is only valuable in states in which the firm is profitable. Due to the time value of money, therefore, the tax function is convex because the present value tax benefit of \$1 in losses is less than the tax cost of \$1 in profits.[53] With a convex tax function, firms have incentive to use derivatives to shift taxable income from good to bad states in order to reduce volatility and expected tax liabilities.

Prediction 24: All else equal, the corporate incentive to hedge increases with the degree of tax schedule convexity.

The second tax incentive to hedge involves increasing debt capacity by reducing income volatility. To the extent that increased debt capacity leads to greater debt usage, it also results in greater tax benefits and firm value. Alternatively, increased debt capacity might go unexploited, thereby reducing expected bankruptcy costs (Smith and Stulz, 1985). Ross (1997) and Leland (1998) argue that the former effect dominates and therefore that hedging increases firm value via the tax benefits of debt.

Prediction 25: There is a tax incentive to hedge because it increases debt capacity. When firms use this extra debt capacity, the tax benefits of debt increase.

Hedging with derivatives transfers income across states within a given time-period. In contrast, earnings management is usually regarded as smoothing income through time. Like the hedging case, tax function convexity can provide an incentive to smooth income.[54] However, tax incentives to smooth are more unidirectional: All else equal, companies prefer to delay paying taxes due to the time value of money. Moreover, if tax rates are expected to fall, tax incentives to delay income are strengthened. The following prediction summarizes three conditions that can lead to a convexity-like incentive to smooth that works against the incentive to delay income recognition:

Prediction 26: Unless one or more of the following conditions are met, there exists a tax incentive to delay recognition of taxable income: (1) the tax function is progressive, (2) net operating loss carryforwards and other deductions are less than fully valued due to limitations on use and the time value of money, and/or (3) tax rates are expected to increase.

### *7.2. Empirical evidence*

Many empirical papers measure tax function convexity using variables based on the existence of NOL or tax credit carryforwards (Prediction 24). These papers regress

[53] The logic is that the government effectively holds a call option on corporate tax liabilities and writes a put on corporate tax refunds, the net value of which can be reduced by reducing volatility. That is, present value considerations from delayed tax refunds reduce the value of the government's written put, so on net the governmental call option on tax collections is more valuable. Reducing volatility reduces the value of the call and the government's claim on corporate earnings.

[54] There are numerous nontax explanations for earnings management. Schipper's (1989) review notes that firms might manage earnings to reduce required returns by lowering earnings variance, to impress outside investors who value stock via earnings multiples, because executive compensation is tied to accounting numbers and because insiders cannot credibly convey private information via other means.

corporate derivative usage on a proxy for convexity and several nontax right-hand-side variables, and generally do not find evidence that convexity affects the corporate use of derivatives (e.g., Nance et al., 1993, for Fortune 500 types of firms or Tufano, 1996, for gold-mining firms). Rather than proxying for convexity, Graham and Smith (1999) explicitly map out tax functions and find that they are convex for about half of Compustat firms. They also report that the average among these firms could save approximately $125,000 in expected tax liabilities by reducing income volatility by 5%. Graham and Rogers (2001) compare this explicit measure of tax function convexity to derivatives usage for a broad cross section of firms and find no evidence that firms hedge in response to tax function convexity. In contrast, Dionne and Garand (2000), using regression coefficients from Graham and Smith (1999) to estimate convexity, show that hedging among gold-mining firms is positively related to estimated convexity.

Graham and Rogers (2001) use simultaneous equations to investigate the joint hedging/capital structure decision and to determine whether firms hedge to increase debt capacity (Prediction 25). In one equation, they regress derivatives usage on variables, including debt ratios, that explain corporate hedging and in the other equation they regress debt ratios on variables, including derivatives usage, that explain debt policy. Graham and Rogers find that hedging leads to greater debt usage. For the average firm, hedging with derivatives increases the debt ratio by 3% and adds tax shields equal to 1.1% of firm value.

Overall, the empirical evidence suggests that the tax incentive to hedge because of tax function convexity is weak at best. The statistical evidence is stronger that the tax incentive to increase debt capacity leads to greater hedging—though the economic importance of this effect appears to be only moderate.

In terms of earnings management, very little research directly investigates the conditions that can lead to a tax incentive to smooth earnings, particularly with respect to the three conditions in Prediction 26. Scholes, Wilson, and Wolfson (1992) state that firms delayed recognizing income in 1986 in anticipation of lower future tax rates. Barton (2001) regresses a measure of earnings management (i.e., discretionary accruals) on a crude convexity variable. Barton's measure of convexity is the excess of a firm's marginal tax rate over its average tax rate (i.e., tax expense divided by taxable income); a positive number indicates a progressive tax function. Barton finds that the absolute value of discretionary accruals is positively related to this measure of convexity, which he interprets as evidence of income smoothing in response to tax incentives. Similarly, using a NOL-based convexity variable, Pincus and Rajgopal (2002) find that profitable oil and gas firms use derivatives to smooth income in response to tax incentives. As with the corporate hedging evidence, tax incentives appear to be a second-order consideration rather than a dominant influence on earnings management.

## 8. Tax shelters

Tax shelters offer a means of reducing taxes that may displace traditional sources of corporate tax deductions. Three common characteristics of shelters are that they reduce tax liability without greatly altering financial statement information, they are shrouded in secrecy, and they are often shut down once detected by the Treasury. Tax shelters can take many different forms, and the current "hot product" is always evolving. They usually exploit glitches in the tax system such as asymmetric domestic and foreign tax treatment or a situation in which income is allocated beyond economic income. In the short-run, before detection, shelters can create a money pump for some firms, with benefits far exceeding transactions costs and the probability-weighted cost of audit/detection. One could imagine a long-run equilibrium in which the benefits of shelters are competed away or greatly reduced, but, as a class, their secretive nature and the proliferation of new products appears to make "short-run" benefits continue unabated for those who participate.

One type of shelter, the high-basis low-value variety, involves an untaxed foreign investor and a taxable domestic corporation both participating in a deal. The untaxed investor is allocated a large portion of the income from the deal and then exits the transaction in a manner that leaves a large economic loss. The corporation can deduct the loss against taxable income. To get a feel for the magnitude of the benefit, Bankman (1999) presents an example in which the corporation contributes \$11 million to a deal and receives \$10 million in property and a \$40 million deductible loss. Therefore, the company effectively pays \$1 million (plus maybe \$3 million in transactions costs and a small expected cost of being caught) for a tax benefit of $\$40\tau_C$ million.

Some recent research investigates tax shelters. Desai (2002) compares taxable income reported on financial statements to actual tax collections and detects a growing wedge between these two series. He argues that traditional explanations such as accelerated depreciation, stock options, and earnings management explain only a portion of the wedge. Desai concludes (p. 1) that new "enhanced opportunities for avoiding and evading taxes through cheaper, more sophisticated, and less transparent mechanisms" (i.e., tax shelters) explain at least one-third of the book-tax income gap as of 1998, and that the portion of the wedge explained by shelters is growing. Graham and Tucker (2004) examine a sample of nearly 50 tax shelter firms. The tax shelters in their sample are huge, producing deductions that average nearly 9% of asset value. These authors find that companies that use tax shelters use less debt than do their nonshelter peers, which is consistent with shelter deductions serving as a nondebt tax shield that is substituted for debt, in the spirit of DeAngelo and Masulis (1980). See Graham and Tucker for references to other recent tax shelter research.

Some forms of shelters, such as the tax-deductible preferred stock (MIPS) discussed in Section 1, receive positive rulings from the Treasury and go on to become accepted financial transactions. Further discussion of tax shelters is beyond the scope of this chapter. The interested reader is directed to Bankman (1999), the source for much of the tax shelter discussion in this section.

## 9. Summary and suggestions for future research

This chapter reviews research related to how taxes affect corporate activities. The research often finds that taxes affect corporate financial decisions—but the magnitude of the effect is not always large.

With respect to capital structure, there is cross-sectional regression evidence that high-tax-rate firms use debt more intensively than do low-tax-rate firms. There is also evidence that debt tax benefits add to firm value. However, much additional research is needed to improve our understanding of capital structure tax effects.

One gap in our knowledge is the lack of time-series evidence about whether firm-specific changes in tax status affect debt policy. Another important area for future research is to isolate the market value of the tax benefits of debt for the broad cross section of firms. Additional research is also needed to explain the apparently conservative debt policy of many firms. Such analysis might investigate whether nondebt tax shields substitute for interest deductions—and help solve the "conservative leverage puzzle." Two such nondebt tax shields are employee stock option deductions and accumulated foreign tax credits. Recent research indicates that employee stock option deductions help to (partially) explain apparent underleverage in some firms. Keep in mind, however, that nondebt tax shields should only affect tax incentives to the extent that they affect the corporate marginal tax rate.

We have also only scratched the surface regarding tax-related leasing research. There is currently not much analysis about whether taxes affect the pricing and structure of lease (or other financial) contracts, about whether leases and debt are substitutes for the lessee, or about how lessor tax rates affect leasing. There is also little research into the effect of relative corporate and personal taxes on the aggregate demand and supply of debt. Unambiguous evidence about whether taxes affect debt maturity choices is also lacking. Finally, all of this research should emphasize robust statistical treatment of standard errors and the economic importance of tax effects, in light of the statement by Myers et al. (1998) that taxes are of third-order importance in the hierarchy of corporate decisions.

Though intriguing in theory, the profession has made only modest progress in documenting whether investor taxes affect asset prices and in turn affect the costs and benefits of corporate policies. There is strong evidence that personal taxes drive a wedge between corporate and municipal bond yields. There is also plausible evidence that the personal tax penalty on MIPS interest income is only modest, which might imply that the personal tax penalty on debt is only modest (relative to using equity)—but this implication needs to be verified. Several papers assume that companies have clienteles of investors that have similar tax characteristics, and then link these companies' policies to the assumed investor tax rates. It would be helpful to make these linkages more direct. In general, we need more market evidence about the importance of personal taxes affecting asset prices, the effective equity tax rate for the marginal investor(s), and information related to the identity of the marginal investor(s) between different securities. One level deeper, we also need evidence that corporate policies are altered in response to these investor tax influences on security prices. Some of this evidence will be difficult to come by and might require access to confidential information or data from countries with unique data or institutional settings.

Progress has been made relating multinational tax considerations to corporate financing decisions, especially in terms of the use of debt by affiliated foreign entities when foreign tax rates are high. However, there is a need for research that highlights capital structure comparisons between classical and other tax systems and direct tests of multinational tax incentives, including the interaction of explanatory variables when appropriate (e.g., excess credit status interacted with interest allocation considerations). It would be helpful if excess (or deficit) credit tax position were measured more precisely than simply using current-period average tax rates.

Several studies link corporate payout policy to tax considerations. In particular, the ex-day stock return and volume evidence is consistent with investor tax considerations influencing asset markets. The Green and Rydqvist (1999) study of Swedish lottery bonds stands out in terms of presenting clean ex-day evidence documenting personal tax effects and serves as a model for future research that isolates tax effects. Unique insights into some payout issues might be provided by comparing payout policy in classical versus other tax systems. In addition, there currently is no convincing evidence that the interaction of investor tax characteristics and payout policy affects firm value and stock returns. Finally, there is a need for direct evidence that tax-based investor clienteles exist (i.e., that investors hold certain securities because of the investor's tax status and the form of payout)—because many of the payout hypotheses implicitly assume that such clienteles exist.

Some recent evidence documents tax-motivated compensation payments (i.e., the choice between salary and options paid to nonexecutive employees), risk management (i.e., hedging to increase debt capacity and the tax benefits of debt), and earnings management. However, we need more "all parties, all deductions" research in these areas, as well as analysis of whether these forms of nondebt tax shields are substitutes for each other or for debt interest. We also need compensation studies based on firm- and employee-specific tax rates and the choice between ISO and NQO plans. Finally, to date there have been few direct tests of whether earnings management is related to progressive tax schedules, less than full valuation of accumulated NOLs and other deductions, and/or expectations of changes in future tax rates.

Some studies have provided documentation that firms choose organizational form based on relative corporate and personal tax rates, that asset sales are structured in response to tax considerations, and that corporate bankruptcy and highly levered restructurings have tax implications. However, we need more evidence about the choice of corporate form using firm-specific data, evidence that firms choose ex ante to perform highly leveraged buyouts in response to tax incentives, and, in general, more evidence about tax incentives affecting corporate reorganizations, spin-offs, and other forms of restructuring.

Finally, although it is convenient for academic research to investigate these tax issues one by one, there is potential for large gains from investigating how these various policies and tax incentives interact from the perspective of a corporate financial manager or tax planner. Along these lines, some recent progress has been made investigating tax shelters. Additional studies that integrate the murky world of tax shelters into the overall tax planning environment would be helpful, though much of this research might end up being case studies. Overall, there are numerous important areas in which careful research can contribute to our understanding of how the imperfections created by taxes affect corporate decisions and firm value.

## References

Alford, A. and P. Berger, 1998, "The Role of Taxes, Financial Reporting, and Other Market Imperfections in Structuring Divisive Reorganizations," Working Paper, Wharton School, University of Pennsylvania.

Allen, F. and Michaely, R., 2001, "Payout Policy," Working Paper, Cornell University.

Allen, F. and Michaely, R., 1995, "Dividend Policy," In R. Jarrow, V. Maksimovic, and W.T. Ziemba (Eds.), *Handbooks in Operations Research and Management Science*, Vol. 9, Elsevier/North-Holland, 793–837.

Altshuler, R. and A. J. Auerbach, 1990, "The Significance of Tax Law Asymmetries: An Empirical Investigation," *Quarterly Journal of Economics*, 105,1, 61–86.

Altshuler, R. and H. Grubert, 2000, "Balance Sheets, Multinational Financial Policy, and the Cost of Capital at Home and Abroad," Working Paper, Rutgers University, J Pub Econ, 2003.

Altshuler, R. and J. Mintz, 1995, "U.S. Interest-Allocation Rules: Effects and Policy," *International Tax and Public Finance*, 2, 7–35.

Altshuler, R. and T. S. Newlon, 1993, "The Effects of U.S. Tax Policy on the Income Repatriation Patterns of U.S. Multinational Corporations," In Alberto Giovannini, R. Glenn Hubbard, and Joel Slemrod (Eds.), *Studies in International Taxation*, University of Chicago Press, Chicago.

Alworth, J. and G. Arachi, 2000, "The Effect of Taxes on Corporate Financing Decisions: Evidence from a Panel of Italian Firms," *International Tax and Public Finance*, 8, 353–376.

Andrade, G. and S. N. Kaplan, 1998, "How Costly Is Financial (Not Economic) Distress? Evidence from Highly Levered Transactions That Became Distressed," *Journal of Finance*, 53, 1443–1493.

Asquith, P. and D. Mullins, 1986, "Equity Issues and Offering Dilution," *Journal of Financial Economics*, 15, 61–89.

Auerbach, A. J., 2002, "Taxation and Corporate Financial Policy," In A. Auerbach and M. Feldstein (Eds.), *Handbook of Public Economics*, Vol. 3, Elsevier/North-Holland, Amsterdam, 1251–1292.

Auerbach, A. J., 1983, "Stockholder Tax Rates and Firm Attributes," *Journal of Public Economics*, 21, 107–127.

Auerbach, A. J. and M. A. King, 1983, "Taxation, Portfolio Choice, and Debt-Equity Ratios: A General Equilibrium Model," *Quarterly Journal of Economics*, 98, 587–609.

Auerbach, A. J. and J. M. Poterba, 1987, "Tax-Loss Carry forwards and Corporate Tax Incentives," In Martin Feldstein (Ed.), *The Effects of Taxation on Capital Accumulation*, University of Chicago Press, Chicago, 305–338.

Austin, J., J. Gaver and K. Gaver, 1998, "The Choice of Incentive Stock Options vs. Nonqualified Options: A Marginal Tax Rate Perspective," *Journal of the American Taxation Association*, 20 (Fall), 1–21.

Ayers, B., C. Cloyd and J. Robinson, 2001, "The Influence of Income Taxes on Use of Inside and Outside Debt by Small Business," *National Tax Journal*, 54, 27–56.

Ayers, B., C. Cloyd and J. Robinson, 1996, "Organizational Form and Taxes: An Empirical Analysis of Small Businesses," *Journal of the American Taxation Society*, 18 (Supplement), 49–67.

Baker, M. and J. Wurgler, 2001, "Market Timing and Capital Structure," *Journal of Finance* forthcoming.

Bali, R. and G. L. Hite, 1998, "Ex-Dividend Day Stock Price Behavior: Discreteness or Tax-Induced Clienteles?" *Journal of Financial Economics*, 47, 127–159.

Bankman, J., 1999, "The New Market for Corporate Tax Shelters," *Tax Notes*, 83, 1775

Barclay, M. J, 1987, "Dividends, Taxes, and Common Stock Prices: The Ex-Dividend Day Behavior of Common Stock Prices Before the Income Tax," *Journal of Financial Economics*, 19, 31–43.

Barclay, M. J. and C. W. Smith Jr., 1995a, "The Maturity Structure of Corporate Debt," *Journal of Finance*, 50, 609–631.

Barclay, M. J. and C. W. Smith, Jr., 1995b, "The Priority Structure of Corporate Liabilities," *Journal of Finance*, 50, 899–917.

Barton, J., 2001, "Does the Use of Financial Derivatives Affect Earnings Management Decisions?" forthcoming in *The Accounting Review*, 76, 1–26.

Beatty, Anne, Philip G. Berger and Joseph Maglio, 1995, "Motives for Forming Research and Development Financing Organizations," *Journal of Accounting and Economics*, 19, 411–442.

Bell, L. and T. Jenkinson, 2002, "New Evidence on the Impact of Dividend Taxation and on the Identity of the Marginal Investor," *Journal of Finance*, 57, 1321–1346.

Benninga, S. and O. H. Sarig, 1997, "Corporate Finance: A Valuation Approach," McGraw-Hill, New York.

Berger, Philip, 1993, "Explicit and Implicit Tax Effects of the R & D Tax Credit," *Journal of Accounting Research*, 31 (2), 131–171.

Black, F., 1980, "The Tax Consequences of Long-run Pension Policy," *Financial Analysts Journal*, 14(7), 21–28.

Black, F. and M. Scholes, 1974, "The Effects of Dividend Yield and Dividend Policy on Common Stock Prices and Returns," *Journal of Financial Economics*, 1, 1–22.

Blume, M., J. Crockett and I. Friend, 1974, "Stock Ownership in the United States: Characteristics and Trends," *Survey of Current Business*, 54 (11), 16–40.

Boyd, J. and R. Jagannathan, 1994, "Ex-Dividend Price Behavior of Common Stocks," *Review of Financial Studies*, 7, 711–741.

Bradley, M., G. Jarrell and E. H. Kim, 1984, "On the Existence of an Optimal Capital Structure: Theory and Evidence," *Journal of Finance*, 39, 857–878.

Brav, A., J. R. Graham, C. Harvey and R. Michaely, 2005, "Payout Policy in the 21st Century," *Journal of Financial Economics* 77, 483–527.

Brav, A. and J. B. Heaton, 1997, "The Economic Effects of Prudent Man Laws: Empirical Evidence from Stock Ownership Dynamics," Working Paper, Duke University.

Breeden, D. and S. Viswanathan, 1998, "Why Do Firms Hedge? An Asymmetric Information Model," Working Paper, Duke University.

Brennan, M. J., 1970, "Taxes, Market Valuation and Corporate Financial Policy," *National Tax Journal*, 23, 417–427.

Brick, I. and S. A. Ravid, 1985, "On the Relevance of Debt Maturity Structure," *Journal of Finance*, 40, 1423–1437.

Campello, M., 2001, "Taxes and Capital Structure: Do Investors' Taxes Matter? Evidence from the Tax Reform Act of 1986," Working Paper, Arizona State University.

Chalmers, J., 1998, "Default Risk Cannot Explain the Muni Puzzle: Evidence from Municipal Bonds that Are Secured by U.S. Treasury Obligations," *Review of Financial Studies*, 11, 281–308.

Chaplinsky, S. and Gregory Niehaus, 1990, "The Tax and Distributional Effects of Leveraged ESOPs," *Financial Management*, 19, 29–38.

Chaplinsky, S. and H. N. Seyhun, 1987, "Tax Rationality and the Demand for Dividends," Working Paper, University of Michigan.

Clinch, G. and T. Shibano, 1996, "Differential Tax Benefits and the Pension Reversion Decision," *Journal of Accounting and Economics*, 21 (1), 69–106.

Collins, J. and D. Kemsley, 2000, "Capital Gains and Dividend Capitalization in Firm Valuation: Evidence of Triple Taxation," *Accounting Review*, 75, 405–427.

Collins, J. and D. Shackelford, 1992, "Foreign Tax Credit Limitations and Preferred Stock Issuances," *Journal of Accounting Research* (Supplement 1992), 103–124.

Constantinides, G., 1983, "Capital Market Equilibrium with Personal Tax," *Econometrica*, 51, 611–636.

Cooper, I. A. and K. G. Nyborg, 1999, "Discount Rates and Tax," Working Paper, London Business School.

Core, J. and W. Guay, 2001, "Stock Options Plans for Non-Executive Employees," *Journal of Financial Economics*, 61, 253–287.

Core, J., W. Guay and D. Larcker, 2002, "Executive Compensation and Incentives: A Survey," forthcoming in *Economic Policy Review*, Federal Reserve Bank of New York.

Cornett, M. and N. Travlos, 1989, "Information Effects Associated with Debt for Equity and Equity for Debt Exchange Offers," *Journal of Finance*, 44, 451–468.

Dammon, R. and L. Senbet, 1988, "The Effect of Taxes and Depreciation on Corporate Investment and Financial Leverage," *Journal of Finance*, 43, 357–373.

Dammon, R., C. Spatt and H. Zhang, 2001, "Optimal Consumption and Investment with Capital Gain Taxes," *Review of Financial Studies*, 14, 583–616.

DeAngelo, H. and R. W. Masulis, 1980, "Optimal Capital Structure under Corporate and Personal Taxation," *Journal of Financial Economics*, 8, 3–29.

Del Guercio, D., 1996, "The Distorting Effect of the Prudent-Man Laws on Institutional Equity Investments," *Journal of Financial Economics*, 40, 31–62.

DeMarzo, P. and D. Duffie, 1991, "Corporate Financial Hedging with Proprietary Information," *Journal of Economic Theory*, 53, 261–286.

Desai, Mihir A., 2002, "The Divergence Between Book and Tax Income," 2002, Working Paper, Harvard Business School.

Desai, M. A., 1997, "A Multinational Perspective on Capital Structure Choice and Internal Capital Markets," Working Paper, Harvard University.

Dhaliwal, D., M. Erickson and R. Trezevant, 1999, "A Test of the Theory of Tax Clienteles for Dividend Policies," *National Tax Journal*, 52, 179–194.

Dhaliwal, D., R. Trezevant and S. Wang, 1992, "Taxes, Investment-Related Tax Shields And Capital Structure," *Journal of the American Taxation Association*, 14, 1–21.

Dionne, G. and M. Garand, 2000, "Risk Management Determinants Affecting Firms' Values in the Gold Mining Industry: New Empirical Results," HEC Working Paper.

Dittmar, A. K., 2002, "An Analysis of Capital Structure and Investment Policies: Evidence from Corporate Divestitures," forthcoming in *Journal of Business*.

Dotan, A. and S. A. Ravid, 1985, "On the Interaction of Real and Financial Decisions of the Firm under Uncertainty," *Journal of Finance*, 40, 501–517.

Eades, K. and F. Marston, 2001, "Incentives for Leasing: Evidence from the Largest U.S. Lessees and Lessors," Working Paper, University of Virginia.

Eckbo, E., 1986, "Valuation Effects of Corporate Debt Offerings," *Journal of Financial Economics*, 15, 119–151.

Ekman, E., 1995, "Taxation and Corporate Financial Policy," Working Paper, Uppsala University, Sweden.

Elton, E. J. and M. J. Gruber, 1970, "Marginal Stockholder Tax Rates and the Clientele Effect," *Review of Economics and Statistics*, 52, 68–74.

Emery, D., W. Lewellen and D. Mauer, 1988, "Tax-Timing Options, Leverage, and the Choice of Corporate Form," *Journal of Financial Research*, 11, 99–110.

Engel, E., M. Erickson and E. Maydew, 1999, "Debt-Equity Hybrid Securities," *Journal of Accounting Research*, 37, 249–274.

Erickson, M., 1998, "The Effect of Taxes on the Structure of Corporate Acquisitions," *Journal of Accounting Research*, 36, 279–298.

Erickson, M. and E. Maydew, 1998, "Implicit Taxes in High Dividend Yield Stocks," *Accounting Review*, 73, 435–458.

Erickson, M. and S. Wang, 2002, "The Effect of Organization Form on Acquisition Price," Working Paper, University of Chicago.

Erickson, M. and S. Wang, 2000, "The Effect of Transaction Structure on Price: Evidence from Subsidiary Sales," *Journal of Accounting and Economics*, 30, 59–97.

Esty, B., 1998, "The Impact of Contingent Liability on Commercial Bank Risk Taking," *Journal of Financial Economics*, 47, 189–218.

Esty, B., F. Qureshi and W. Olson, 2000, "Iridium LLC," Harvard Business School Case Study No. 200–039, Rev. 3/27/01.

Fama, E. F. and K. R. French, 2001, "Testing Trade-off and Pecking-Order Predictions about Dividends and Debt," *Review of Financial Studies*, 15, 1–33.

Fama, E. F. and K. R. French, 1998, "Taxes, Financing Decisions, and Firm Value," *Journal of Finance*, 53 (3), 819–843.

Farrar, D. and L. Selwyn, 1967, "Taxes, Corporate Policy, and Return to Investors," *National Tax Journal*, 20, 444–454.

Florentsen, B. and K. Rydqvist, 2002, "Ex-Day Behavior When Investors and Professional Traders Assume Reverse Roles—The Case of Danish Lottery Bonds," *Journal of Financial Intermediation*, 11, 152–175.

Frank, M. and R. Jagannathan, 1998, "Why Do Stock Prices Drop by Less Than the Value of the Dividend? Evidence from a Country Without Taxes," *Journal of Financial Economics*, 47, 161–188.

Froot, K. and J. Hines, Jr., 1995, "Interest Allocation Rules, Financing Patterns, and the Operations of U.S. Multinationals," In Martin Feldstein, James Hines Jr., and R. Glenn Hubbard (Eds.), *The Effects of Taxation on Multinational Corporations*, University of Chicago Press.

Froot, K., D. Scharfstein and J. Stein, 1993, "Risk Management: Coordinating Corporate Investment and Financing Policies," *Journal of Finance*, 48, 1629–1658.

Geisler, F., 1999, "Equity Security Investments: Evidence on Tax-Induced Dividend Clienteles," *Journal of American Taxation Association*, 22, 1, 1–17.

Gilson, Stuart C., 1997, "Transactions Costs and Capital Structure Choice: Evidence from Financially Distressed Firm," *Journal of Finance*, 52, 111–133.

Givoly, D., C. Hahn, A. Ofer and O. H. Sarig, 1992, "Taxes and Capital Structure: Evidence from Firms' Response to the Tax Reform Act of 1986," *Review of Financial Studies*, 5, 331–355.

Goldstein, R., N. Ju and H. Leland, 2001, "An EBIT-Based Model of Dynamic Capital Structure," *Journal of Business*, 74, 483–512.

Goolsbee, A., 1999, "What Happens When You Tax The Rich? Evidence From Executive Compensation".

Gordon, R. H. and J. K. MacKie-Mason, 1997, "How Much Do Taxes Discourage Incorporation?" *Journal of Finance*, 52(2), 477–505.

Gordon, R. H. and J. K. MacKie-Mason, 1994, "Tax Distortions to the Choice of Organizational Form," *Journal of Public Economics*, 55, 279–306.

Gordon, R. H. and J. K. MacKie-Mason, 1990, "Effects of the Tax Reform Act of 1986 on Corporate Financial Policy and Organizational Form," In: J. Slemrod (Ed.), *Do Taxes Matter*? MIT Press, Cambridge, MA, 91–131.

Graham, J. R., 2003, "Taxes and Corporate Finance: A Review," *Review of Financial Studies*, 16, 1074–1128.

Graham, John R., 2001, "Estimating the Tax Benefits of Debt," *Journal of Applied Corporate Finance*, 14, 42–54.

Graham, J. R., 2000, "How Big Are the Tax Benefits Of Debt?" *Journal of Finance*, 55, 1901–1941.

Graham, J. R., 1999, "Do Personal Taxes Affect Corporate Financing Decisions?" *Journal of Public Economics*, 73, 147–185.

Graham, J. R., 1996a, "Debt and the Marginal Tax Rate," *Journal of Financial Economics*, 41, 41–74.

Graham, J. R., 1996b, "Proxies for the Marginal Tax Rate," *Journal of Financial Economics*, 42, 187–221.

Graham, J. R. and C. Harvey, 2001, "The Theory and Practice of Corporate Finance: Evidence from the Field," *Journal of Financial Economics*, 60, 187–243.

Graham, J. R. and A. Kumar, 2004, "Do Dividend Clienteles Exist? Evidence on Dividend Preferences of Retail Investors," Working Paper, Duke University.

Graham, J. R., E. Hughson and J. Zender, 1999, "Stock Market Reactions to Capital Structure Changes: Theory and Evidence," Working Paper, Duke University.

Graham, J. R., M. Lang and D. Shackelford, 2002, "Employee Stock Options, Corporate Taxes and Debt Policy," *Journal of Finance*, 59, 1585–1618.

Graham, J. R., M. Lemmon and J. Schallheim, 1998, "Debt, Leases, Taxes, and the Endogeneity of Corporate Tax Status," *Journal of Finance*, 53, 131–162.

Graham, J. R., R. Michaely and M. Roberts, 2003, "Do Price Discreteness and Transactions Costs Affect Stock Returns? Comparing Ex-Dividend Pricing Before and After Decimalization," *Journal of Finance*, 58, 2613–2637.

Graham, J. R. and D. Rogers, 2001, "Do Firms Hedge in Response to Tax Incentives?" *Journal of Finance*, 57, 815–839.

Graham, J. R. and C. W. Smith Jr., 1999, "Tax Incentives to Hedge," *Journal of Finance*, 54, 2241–2262.

Graham, J. R. and A. Tucker, 2004, "Tax Shelters and Corporate Debt Policy," Working Paper, Duke University.

Green, R. C., 1993, "A Simple Model of the Taxable and Tax-Exempt Yield Curves," *Review of Financial Studies*, 6, 233–264.

Green, R. C. and B. Hollifield, 2003, "The Personal-Tax Advantages of Equity," *Journal of Financial Economics*, 67, 175–216.

Green, R. C. and K. Rydqvist, 1999, "Ex-Day Behavior with Dividend Preference and Limitations to Short-Term Arbitrage: The Case of Swedish Lottery Bonds," *Journal of Financial Economics*, 53, 145–187.

Grinblatt, M. and S. Titman, 2002, *Financial Markets and Corporate Strategy*, Irwin McGraw-Hill, Boston, MA.

Grubert, H., 1998, "Taxes and the Division of Foreign Operating Income Among Royalties, Interest, Dividends and Retained Earnings," *Journal of Public Economics*, 68, 2, 269–290.

Grundy, B. D., 1985, "Trading Volume and Stock Returns Around Ex-Dividend Dates," Working Paper, University of Chicago.

Guedes, J. and T. Opler, 1996, "The Determinants of the Maturity Structure of Corporate Debt Issues," *Journal of Finance*, 51, 1809–1833.

Guenther, J. A., 1992, "Taxes and Organizational Form: A Comparison of Corporations and Master Limited Partnerships," *The Accounting Review*, 67, 17.

Hall, B. and J. Liebman, 2000, "The Taxation of Executive Compensation," In James Poterba (Ed.), *Tax Policy and the Economy*, MIT Press, Cambridge, MA, 14.

Hand, J. R. M. and T. R. Skantz, 1998, "Market Timing Through Equity Carve-Outs," Working Paper, University of North Carolina at Chapel Hill.

*Handbook of Corporate Finance*, 2004, B. Espen Eckbo (Ed.), Elsevier/North-Holland.

Hanlon, M. and T. Shevlin, 2002, "Accounting for Tax Benefits of Employee Stock Options and Implications for Research," *Accounting Horizons*, 16, 1–16.

Harris, T. and D. Kemsley, 1999, "Dividend Taxation in Firm Valuation: New Evidence," *Journal of Accounting Research,* 37, 275–291.

Harwood E. and G. B. Manzon Jr., 1998, "Tax Clienteles and Debt Maturity," Working Paper, Boston College.

Henderson, C. B., 2001, "The Financing Choices of Large Commercial Bank," Working Paper, University of Texas, Austin.

Hines, J. R., 1996, "Altered States: Taxes and the Location of FDI in America," *American Economic Review*, 86, 1076–1094.

Hines, J. R., 1995, "Taxes, Technology, Transfer, and the R & D Activities of Multinational Firms," In Martin Feldstein, James R. Hines Jr., and R. Glenn Hubbard (Eds.), *The Effect of Taxation on Multinational Corporations*, University of Chicago Press, Chicago, 225–248.

Hines, J. R. and R. G. Hubbard, 1990, "Coming Home to America: Dividend Repatriations by U.S. Multinationals," In Assaf Razin and Joel Slemrod (Eds.), *Taxation in the Global Economy*, University of Chicago Press, Chicago.

Hodder, L., M. L. McAnally and C. D. Weaver, 2001, "The Influence of Tax and Non-Tax Factors on the Organizational Form of Banks," Working Paper, University of Texas Austin.

Hovakimian, H., T. Opler and S. Titman, 2001, "The Debt-Equity Choice," *Journal of Financial and Quantitative Analysis*, 36, 1–24.

Huddart, S., 1998, "Patterns of Stock Option Exercise in the United States," In J. Carpenter and D. Yermack (Eds.), *Executive Compensation and Shareholder Value*, New York University and Kluwer Academic Publishers, Norwell, MA.

Huddart, S., 1998, "Tax Planning and the Exercise of Employee Stock Options," *Contemporary Accounting Research*, 15 (2), 203–216.

Irvine, P. and J. Rosenfeld, 2000, "Raising Capital Using Monthly-Income Preferred Stock: Market Reaction and Implications for Capital-Structure Theory," *Financial Management*, 29 (2), 5–20.

Jensen, M. and W. Meckling, 1976, "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure," *Journal of Financial Economics*, 7, 305–360.

Kadapakkam, P., 2000, "Reduction of Constraints on Arbitrage Trading and Market Efficiency: An Examination of Ex-Day Returns in Hong Kong after Introduction of Electronic Settlement," *Journal of Finance*, 55, 2841–2861.

Kahle, K. and K. Shastri, 2002, "Firm Performance, Capital Structure, and the Tax Benefits of Employee Stock Options," Working Paper, University of Pittsburgh.

Kalay, A., 1982, "The Ex-Dividend Behavior of Stock Prices: A Reexamination of the Clientele Effect," *Journal of Finance*, 37, 1059–1070.

Kalay, A. and Michaely, R., 2000, "Dividends and Taxes: A Re-Examination," *Financial Management*, 29, 55–75.

Kane, A., A. J. Marcus and R. L. McDonald, 1985, "Debt Policy and the Rate of Return Premium to Leverage," *Journal of Financial and Quantitative Analysis*, 20, 479–499.

Kaplan, S., 1989, "Management Buyouts: Evidence on Taxes as a Source of Value," *Journal of Finance*, 44 (3), 611–632.

Kemsley, D. and D. Nissim, 2002, "Valuation of the Debt-Tax Shield," *Journal of Finance*, 57, 2045–2073.

Kim, E. H., 1989, "Optimal Capital Structure in Miller's Equilibrium," In S. Bhattacharya and G. Constantinides (Eds.), *Financial Markets and Incomplete Information*, Rowman and Littlefield, Totowa, NJ, 36–48.

Kim C. S., D. C. Maurer and M. H. Stohs, 1995, "Corporate Debt Maturity Policy and Investor Tax-Timing Option: Theory and Evidence," *Financial Management*, 24, 33–45.

Klassen, K. and A. Mawani, 2000, "The Impact of Financial and Tax Reporting Incentives on Option Grants to Canadian CEOs," *Contemporary Accounting Research*, 17, 227–262.

Klein, P., 2001, "The Capital Gain Lock-In Effect and Long-Horizon Return Reversal," *Journal of Financial Economics*, 59, 1, 33–62.

Koski, J. L. and R. Michaely, 2000, "Prices, Liquidity, and the Information Content of Trades," *Review of Financial Studies*, 13, 659–696.

Kraus, A. and R. H. Litzenberger, 1973, "A State-Preference Model of Optimal Financial Leverage," *Journal of Finance*, 28, 911–922.

Lakonishok, J. and T. Vermaelen, 1986, "Tax-Induced Trading Around the Ex-Day," *Journal of Financial Economics*, 16, 287–319.

Lall, S., 1973, "Transfer Pricing by Multinational Manufacturing Firms," *Oxford Bulletin of Economics and Statistics*, 35, 173–195.

Lang, M. and D. Shackelford, 2000, "Capitalization of Capital Gains Taxes: Evidence from Stock Price Reactions to the 1997 Rate Reductions," *Journal of Public Economics*, 76, 69–85.

Leland, H. E., 1998, "Agency Cost, Risk Management and Capital Structure," *Journal of Finance*, 53, 1213–1243.

Leland, H. E. and D. H. Pyle, 1977, "Informational Asymmetric, Financial Structure, and Financial Intermediation," *Journal of Finance*, 32, 2371–2387.

Lemmon, M. and J. Zender, 2002, "Debt Capacity and Tests of Capital Structure Theories," Working Paper, University of Utah.

Lemmon, M. and J. Zender, 2001, "Looking under the Lamppost: An Empirical Examination of the Determinants of Capital Structure," Working Paper, University of Utah.

Lewellen, W., K. Stanley, R. Lease and G. Schlarbaum, 1978, "Some Direct Evidence on the Dividend Clientele Hypothesis," *Journal of Finance*, 33, 1385–1399.

Lewis, C., 1990, "A Multiperiod Theory of Corporate Financial Policy under Taxation," *Journal of Financial and Quantitative Analysis*, 25, 25–43.

Lie, E. and H. Lie, 1999, "The Role of Personal Taxes in Corporate Decisions: An Empirical Analysis of Share Repurchases and Dividends," *Journal of Financial and Quantitative Analysis*, 34, 534–552.

Lintner, John, 1956, "Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes,"*American Economic Review*, 46 (2), 97–113.

Litzenberger, R. H. and K. Ramaswamy, 1979, "The Effect of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence," *Journal of Financial Economics*, 7, 163–700.

MacKie-Mason, J. K., 1990, "Do Taxes Affect Corporate Financing Decisions?" *Journal of Finance*, 45 (5), 1990, 1471–1493.

McDonald, R., 2001, "Cross-Border Investing with Tax Arbitrage: The Case of German Dividend Tax Credits," *Review of Financial Studies* 14, 617–657.

McDonald, R. L., 2002, "The Tax (Dis)Advantage of a Firm Issuing Options on Its Own Stock," forthcoming in *Journal of Public Economics*.

Madeo, S. and T. Omer, 1994, "The Effect of Taxes on Switching Stock Option Plans: Evidence from the Tax Reform Act of 1969," *Journal of the American Taxation Association*, 16 (Fall), 24–42.

Mankiw, N. Gregory and James M. Poterba, 1996, "Stock Market Yields and the Pricing of Municipal Bounds," NBER Working Paper 5607.

Masulis, R. W., 1983, "The Impact of Capital Structure Change on Firm Value: Some Estimates," *Journal of Finance*, 38, 107–126.

Masulis, R. W., 1980, "Stock Repurchase by Tender Offer: An Analysis of the Causes of Common Stock Price Changes," *Journal of Finance*, 35, 305–319.

Masulis, R. W. and A. Korwar, 1986, "Seasoned Equity Offerings: An Empirical Investigation," *Journal of Financial Economics*, 15, 91–118.

Matsunaga, S., T. Shevlin and D. Shores, 1992, "Disqualifying Dispositions of Incentive Stock Options: Tax Benefits Versus Financial Reporting Costs," *Journal of Accounting Research*, 30 (Supplement), 37–76.

Maydew, E., K. S. Schipper and L. Vincent, 1999, "The Impact of Taxes on the Choice of Divestiture Method," *Journal of Accounting and Economics,* 28, 117–150.

Michaely, R., 1991, "Ex-Dividend Day Stock Price Behavior: The Case of the 1986 Tax Reform Act," *Journal of Finance*, 46 (3), 845–860.

Michaely, R. and J. Vila, 1995, "Investors' Heterogeneity, and Volume Around the Ex-Dividend Day," *Journal of Finance and Quantitative Analysis*, 30 (2), 171–198.

Michaely, Roni and Jean-Luc Vila, 1996, "Trading Volume with Private Valuations: Evidence from the Ex-dividend Day," *Review of Financial Studies*, 9, 471–510.

Michaely, R., J. Vila and J. Wang, 1996, "A Model of Trading Volume with Tax-Induced Heterogeneous Valuation and Transaction Costs," *Journal of Financial Intermediation*, 5, 340–371.

Mikkelson, W. and M. Partch, 1986, "Valuation Effects on Security Offerings and the Issuance Process," *Journal of Financial Economics*, 15, 31–60.

Miles, James and John Ezzell, 1985, "Reformulating Tax Shield Valuation: A Note," *Journal of Finance*, 40, 1485–1492.

Miller, M. H., 1977, "Debt and Taxes," *Journal of Finance*, 32, 261–275.

Miller, M. and F. Modigliani, 1961, "Dividend Policy, Growth and the Value of the Firm," *Journal of Business*, 34, 411–433.

Miller, M. H. and M. Scholes, 1978, "Dividends and Taxes," *Journal of Financial Economics*, 6, 333–364.

Mills, L. and K. Newberry, 2000, "The Influence of Tax and Non-Tax Costs on Book-Tax Reporting Differences: Public and Private firms," Working Paper, University of Arizona at Tucson.

Minton, B. and K. H. Wruck, 2001, "Financial Policy, Investment Opportunities and the Demand for External Financing: Evidence from Low Debt Firms," Working Paper, Ohio State University.

Modigliani, F. and M. H. Miller, 1958, "The Cost of Capital, Corporation Finance and the Theory of Investment," *American Economic Review*, 48, 261–297.

Modigliani, F. and M. H. Miller, 1963, "Corporate Income Taxes and the Cost of Capital: A Correction (in Communications)," *American Economic Review*, 53, 433–443.

Murphy, Kevin, 1999, "Executive Compensation," In O. Ashefelter and D. Card (Eds.), *Handbook of Labor Economics*, Vol. 3, Elsevier/North-Holland, Amsterdam.

Myers, M. M., 2001, "The Impact of Taxes on Corporate Defined Benefit Plan Asset Allocation," *Journal of Accounting Research*, 40, 1163–1190.

Myers, S. C., 1993, "Still Searching for the Optimal Capital Structure," *Journal of Applied Corporate Finance*, 6, 4–14.

Myers, S. C., 1984, "The Capital Structure Puzzle," *Journal of Finance*, 39, 575–592.

Myers, S. C., 1977, "Determinants of Corporate Borrowing," *Journal of Financial Economics*, 3, 799–819.

Myers, S. C. and Nicholas S. Majluf, 1984, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics*, 13, 187–221.

Myers, Stewart, John McConnell, Alice Peterson, Dennis Soter, and Joel Stern, 1998, "Vanderbilt University Roundtable on the Capital Structure Puzzle," *Journal of Applied Corporate Finance*, 11, 8–24.

Nance, D. R., C. W. Smith Jr. and C. W. Smithson, 1993, "On the Determinants of Corporate Hedging," *Journal of Finance*, 48, 267–284.

Newberry, K., 1998, "Foreign Tax Credit Limitations and Capital Structure Decisions," *Journal of Accounting Research*, 36, 157–166.

Newberry, K. and D. S. Dhaliwal, 2001, "Cross-Jurisdictional Income Shifting by U.S. Multinationals: Evidence from International Bond Offerings," *Journal of Accounting Research*, 39, 643–362.

Newberry K. J. and G. F. Novack, 1999, "The Effect of Taxes on Corporate Debt Maturity Decisions: An Analysis of Public and Private Bond Offerings," *Journal of the American Taxation Association*, 21, 2.

O'Malley, Michael P., 1996, "The Effects of Taxes on Leasing Decisions: Evidence from Panel Data," Working Paper, Board of Governers of the Federal Reserve System.

Opler, T. and J. Guedes, 1996, "The Determinants of the Maturity of Corporate Debt Issues," *The Journal of Finance*, Vol. 51, No. 5 (December), pp. 1809–1833.

Parrino, Robert and Michael Weisbach, 1999, "Measuring Investment Distortions arising from Stockholder-Bondholder Conflicts," *Journal of Financial Economics*, 53, 3–42.

Pérez-González, F., 2000, "Large Shareholders and Dividends: Evidence from U.S. Tax Reforms," Working Paper, Harvard University, Cambridge, MA.

Pettit, R. R., 1977, "Taxes, Transaction Costs and the Clientele Effect of Taxes," *Journal of Financial Economics*, 8, 419–436.

Pincus, M. and R. Rajgopal, 2002, "The Interaction between Accrual Management and Hedging: Evidence from Oil and Gas Firms," *The Accounting Review*, 77, 127–160.

Pinegar M. J. and R. C. Lease, 1986, "The Impact of Preferred-for-Common Exchange Offers on Firm Value," *Journal of Finance*, 41, 795–814.

Pittman, J. and K. Klassen, 2001, "The Influence of Firm Maturation on Firms' Rate of Adjustment to Their Optimal Capital Structures," *Journal of the American Taxation Association*, 23, 70–94.

Plesko, G. A., 2003, "An Evaluation of Alternative Measures of Corporate Tax Rates," *Journal of Accounting and Economics*, 35, 201–226.

Poterba, J. M., 2001, "Taxation and Portfolio Structure: Issues and Implications," in Luigi Guiso, Michael Haliassos, and Tullio Jappelli (Eds.), *Household Portfolios*, MIT Press, Cambridge, MA.

Poterba, J. M., 1989, "Lifetime Incidence and the Distributional Burden of Excise Taxes," *American Economic Review*, 79, 325–330.

Poterba, J. M. and L. Summers, 1985, "The Economic Effects of Dividend Taxation," In E. Altman and M. Subramanyam (Eds.), *Recent Advances in Corporate Finance*, Richard Irwin, Homewood, IL.

Rajan, R. G. and L. Zingales, 1995, "What Do We Know about Capital Structure Choice? Some Evidence from International Data," *Journal of Finance*, 50, 1421–1460.

Ross, M. P., 1997, "Corporate Hedging: What, Why and How?" Working Paper, University of California, Berkeley.

Ross, S., 1977, "The Determination of Financial Structure: The Incentive Signaling Approach," *Bell Journal of Economics*, 8, 23–40.

Schipper, K., 1989, "Commentary on Earnings Management," *Accounting Horizons*, 3, 91–102.

Scholes, M. and M. Wolfson, 1992, "Taxes and Business Strategy: A Planning Approach," Prentice-Hall, Englewood Cliffs, NJ.

Scholes, M. and M. Wolfson, 1990, "The Effects of Changes in Tax Laws on Corporate Reorganization Activity," *Journal of Business*, 63, 141–164.

Scholes, M., P. Wilson and M. Wolfson, 1992, "Firms Responses to Anticipated Reductions in Tax Rates: The Tax Reform Act of 1986," *Journal of Accounting Research*, 30 (Supplement), 161–191.

Scholes, M., M. Wolfson, M. Erickson, E. Maydew and T. Shevlin, 2002, "Taxes and Business Strategy," 2nd edition, Prentice-Hall: Englewood Cliffs, NJ.

Scholz, J., 1992, "A Direct Examination of the Dividend Clientele Hypothesis," *Journal of Public Economics*, 49, 261–285.

Schulman, Craig, Deborah Thomas, Keith Sellers, and Duane Kennedy, 1996, "Effects of Tax Integration and Capital Gains Tax on Corporate Leverage," *National Tax Journal*, 49, 31–54.

Scott, J. H., Jr., 1976, "A Theory of Optimal Capital Structure," *Bell Journal of Economics*, 7, 33–54.

Seida J. and W. Wempe, 2000, "Do Capital Gain Tax Rate Increases Affect Individual Investors' Trading Decisions?" *Journal of Accounting and Economics*, 30, 1, 33–57.

Shackelford, D., 1991, "The Market for Tax Benefits: Evidence from Leveraged ESOPs," *Journal of Accounting and Economics*, 14, 117–146.

Shackelford, D. and T. Shevlin, 2001, "Empirical Tax Research in Accounting," *Journal of Accounting and Economics*, 31, 321–387.

Shah, K., 1994, "The Nature of Information Conveyed by Pure Capital Structure Changes," *Journal of Financial Economics*, 36, 89–126.

Sharpe, S. and H. Nguyen, 1995, "Capital Market Imperfection and the Incentive to Lease," *Journal of Financial Economics*, 39, 271–294.

Shelley, M., T. Omer and T. J. Atwood, 1998, "Capital Restructuring and Accounting Compliance Costs: The Case of Publicly Traded Partnerships," *Journal of Accounting Research*, 36, 365–378.

Shevlin, T., 1990, "Estimating Corporate Marginal Tax Rates with Asymmetric Tax Treatment of Gains and Losses," *Journal of the American Taxation Association*, 12, 51–67.

Shevlin, T., 1987, "Taxes and Off-Balance Sheet Financing: Research and Development Limited Partnerships," *The Accounting Review*, 62, 480–509.

Shum, P. M., 1996, "Taxes and Corporate Debt Policy in Canada: An Empirical Investigation," *Canadian Journal of Economics*, 29, 556–572.

Shyum-Sunder L. and S. Myers, 1999, "Testing Static Tradeoff Against Pecking Order Models of Capital Structure," *Journal of Financial Economics*, 51, 219–224.

Sick, G. A., 1990, "Tax-Adjusted Discount Rates," *Management Science*, 36, 1432–1450.

Smith, C.W., Jr. and R. M. Stulz, 1985, "The Determinants of Firms' Hedging Policies," *Journal of Financial and Quantitative Analysis*, 20, 391–405.

Smith, Clifford and L. MacDonald Wakeman, 1985, "Determinants of Corporate Leasing Policy," *Journal of Finance*, 40, 895–908.

Stohs, M. and D. Mauer, 1996, "The Determinants of Corporate Debt Maturity Structure," *Journal of Business*, 69, 279–312.

Strickland, D., 1996, "Determinants of Institutional Ownership: Implications for Dividend Clienteles," Working Paper, Ohio State University.

Swenson, C. W., 1992, "Some Tests of the Incentive Effects of the Research and Experimentation Tax Credit," *Journal of Public Economics*, 49, 203–218.

Taggart, R., 1991, "Consistent Valuation and Cost of Capital Expressions with Corporate and Personal Taxes," *Financial Management*, 20, 8–20.

Tepper, I., 1981, "Taxation and Corporate Pension Funding Policy," *Journal of Financial Economics*, 3–51.

Thomas, J. K., 1988, "Corporate Taxes and Defined Benefit Pension Plans," *Journal of Accounting and Economics*, 10, 199–237.

Trezevant, R., 1992, "Debt Financing and Tax Status: Tests of the Substitution Effect and the Tax Exhaustion Hypothesis Using Firms' Responses to the Economic Recovery Tax Act of 1981," *Journal of Finance*, 47 (4), 1557–1568.

Tufano, P., 1996, "Who Manages Risk? An Empirical Examination of Risk Management Practices in the Gold Mining Industry," *Journal of Finance*, 51, 1097–1137.

Wang, S., 2000, "Optimal Expected Corporate Marginal Tax Rate and Leverage Adjustments," Working Paper, Cornell University.

Warner, Jerold B., 1977, "Bankruptcy Costs: Some Evidence," *Journal of Finance*, 32, 337–347.

Williams, M., 2000, "Tax Clienteles, Implicit Taxes, and Tax Capitalization: Interaction of Multiple Clienteles in Financial Markets," Working Paper, University of California, Los Angeles.

*Chapter 12*

# TRADE-OFF AND PECKING ORDER THEORIES OF DEBT*

MURRAY Z. FRANK

*Carlson School of Management, University of Minnesota*

VIDHAN K. GOYAL

*Department of Finance, Hong Kong University of Science and Technology*

## Contents



* We thank Long Chen, Sudipto Dasgupta, Espen Eckbo, Chris Hennessy, Jay Ritter, Michael Roberts, Tan Wang, and Jaime Zender for comments and helpful discussions. Murray Frank thanks Piper Jaffray for financial support. Vidhan Goyal thanks the Research Grants Council of the Hong Kong Special Administrative Region for financial support (Project No. 6489/06H). The literature on capital structure is vast, containing hundreds of papers, and so we cannot hope to cover it all. Inevitably, we have left out a great deal, much of it very important. As such, we apologize in advance for our inevitable oversights. Even more than usual, it must be stressed that this survey reflects only our views.

*Handbook of Empirical Corporate Finance, Volume 2*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

## Abstract

Taxes, bankruptcy costs, transactions costs, adverse selection, and agency conflicts have all been advocated as major explanations for the corporate use of debt financing. These ideas have often been synthesized into the trade-off theory and the pecking order theory of leverage. This chapter reviews these theories and the related evidence and identifies a number of important empirical stylized facts. To understand the evidence, it is important to recognize the differences among private firms, small public firms, and large public firms. Private firms use retained earnings and bank debt heavily; small public firms make active use of equity financing; and large public firms primarily use retained earnings and corporate bonds. The available evidence can be interpreted in several ways. Direct transaction costs and indirect bankruptcy costs appear to play important roles in a firm's choice of debt. The relative importance of the other factors remains open to debate. No currently available model appears capable of simultaneously accounting for all of the stylized facts.

## Keywords

capital structure, leverage, corporate financing, trade-off theory, pecking order theory, agency costs, partial adjustment, taxes, bankruptcy costs

## 1. Introduction

How do firms finance their operations? How should firms finance their operations? What factors influence these choices? How do these choices affect the rest of the economy? These are important questions that have guided researchers for a long time. At one time, the complexity of the problem of financing was thought to be so great as to defy development of reasonable theories. Half a century ago, Weston (1955) even felt the need to argue whether it was possible to develop reasonable theories about these matters.[1] Since then, a remarkably large number of ideas and theories have been proposed to answer these questions.

In a particularly influential treatment of the problem, Myers (1984) considers a contest between two perspectives on corporate debt. He calls one of these hypotheses the trade-off theory, which states that firms balance tax savings from debt against deadweight bankruptcy costs. He calls the other hypothesis the pecking order theory, which states that, due to adverse selection, firms first look to retained earnings, then to debt, and only in extreme circumstances to equity for financing. In this review, we consider the literature and evidence that have developed out of Myers's contest.

According to Myers (1984), these theories have at least two key implications. The key implication of the trade-off theory is that leverage exhibits target adjustment so that deviations from the target are gradually eliminated. The key prediction of the pecking order theory is the strict ordering of financing. Myers presents these two theories as broad organizing frameworks that can potentially help account for many facts. But it is also possible to view both theories as part of a much broader set of factors that determine the capital structure of a firm. Many scholars seem inclined to view both theories in this more limited way.

The claim that leverage exhibits target adjustment is in fact neither necessary nor sufficient for a firm to be balancing tax savings against bankruptcy costs. Accordingly, target adjustment is better viewed as being a separate hypothesis. We use the term *static trade-off theory* for the hypothesis that bankruptcy and taxes are the key factors that determine leverage within a static model.

Drawing a distinction between the static trade-off theory and the target adjustment hypothesis is important for both theoretical and empirical reasons. Target adjustment can be implied by a variety of dynamic theories. These theories may reflect taxes and bankruptcy costs, but they may also have other causes. Static models do not make

[1] Many teachers of business finance are skeptical about the existence or possibility of theories of financial policy. To support their position, they emphasize both the wide range of influences on financial decisions and their varying importance from one situation to another. The subjectivity of many factors makes them dependent on the psychological makeup of the decision-maker. Furthermore, it is well known that in the analysis of business cases, two or more solutions usually appear equally defensible. Those who hold to the skeptical position therefore assert that there can be no science of business finance since experts cannot arrive at unique answers (Weston, 1955, p. 130).

predictions about dynamics without some auxiliary assumptions. The target adjustment hypothesis receives much clearer empirical support than does either the static trade-off theory or the pecking order theory.

The empirical literature supports a number of generalizations that appear to be robust and particularly important for an understanding of actual leverage. To draw attention to these key facts, we have highlighted these generalizations using the label "stylized fact." For ease of reference, these stylized facts are also collected in an appendix.

A model that incorporated all of the stylized facts would be ideal. Not only is such a unifying model not currently available but, even worse, the standard versions of the available models actually contradict some of the known facts. Different models have problems with different facts.[2]

A particularly important problem for the standard static trade-off theory is provided by the historical record. In the static trade-off theory, the desire to limit tax payments motivates a firm to use debt financing (see Modigliani and Miller, 1963, for an extreme version). As discussed in Section 3.1, it is quite difficult to match the observed leverage ratios in particular decades with the corporate tax rates in those decades. Even more remarkably, corporate income taxes are only about a century old. Debt financing was common long before the introduction of the corporate income taxes.[3] Thus, we know that taxes do not provide a complete justification for the use of debt financing. This does not, of course, imply that taxes can be safely ignored when analyzing modern corporate use of debt.

A particularly important problem for the standard version of the pecking order theory concerns the use of equity financing. Firms issue too much equity (Frank and Goyal, 2003) and at the wrong times (Fama and French, 2005; Leary and Roberts, 2007). In the pecking order, it is the financing deficit that drives debt issues. Empirically, however, other factors appear to be more important (Frank and Goyal, 2003).

Thus the standard versions of both the trade-off theory and the pecking order theory appear to be inadequate. Both approaches need to be improved to account for the known facts.

Recently, proponents of the trade-off approach have focused on developing dynamic structural trade-off models. An attractive feature of these models is that they try to provide a unified framework that can simultaneously account for many facts. Examples include Leary and Roberts (2007), Hennessy and Whited (2005), Ju et al. (2005), and Strebulaev (2007).

Proponents of the pecking order theory have focused recently on the development of a satisfactory notion of debt capacity (see, e.g., Lemmon and Zender, 2004) and on more complex adverse selection models (see, e.g., Halov and Heider, 2005).

[2] Myers (2003) argues that a satisfactory unifying model is unlikely to become available in the foreseeable future. We are somewhat more hopeful.

[3] "For example, [in 1731] the French consul in Genoa wrote: 'Lack of confidence keeps money in short supply; so those who usually do business on credit, which means most of the merchants in the city, are doing very little. The best purses are shut'" (Braudel, 1982, pp. 397–398).

Bankruptcy costs and direct transaction costs clearly play at least some role, and leverage is stationary over the long run. There is room for reasonable differences of opinion regarding the relative importance of many factors, including taxation, adverse selection, and various agency conflicts.

## 2. Theory

### *2.1. Kinds of theories*

Disagreements over the merits of financial theories stem, in part, from different views of the role that theory plays. It is, therefore, helpful to recognize the different kinds of theory.

One kind of theory represents a point of view. A point-of-view theory is not an explicit model, but rather a set of principles that guide the development of specific models and tests. Both the pecking order and the trade-off theories can be understood as point-of-view theories. Each provides a guide for the development of models and tests. But neither is tied to a specific model formulation.

A second kind of theory is an illustrative theory. An illustrative theory shows how a certain idea can be expressed in a coherent manner. The point of this theory is to show an idea in as clear and as simple a manner as possible. Accordingly, strong assumptions are often made to solve specific models in a closed form.

A third kind of theory, a unifying model, is presented as a means of tying together a variety of observations in a coherent manner. A unifying model is supposed to integrate many facts to show that these facts stem from a common underlying structure. Often, though not always, these models have to be numerically calibrated because they do not have closed-form solutions.

Fourth is normative theory, which is intended to offer advice to someone. At this stage of development, very little of the theory in finance is intended as advice to chief financial officers. However, some studies, such as that by Graham (2000), in which he argues that many firms could increase value by levering up, seem very interesting from a normative perspective. In the next few years, this kind of analysis will likely become much more common in corporate finance.

Both the pecking order and the trade-off theories provide points of view. Both have been illustrated in specific models with particular simplifying assumptions, and both are often presented as unifying theories.

Moving from a point of view to a specific model requires making assumptions. When the pecking order and the trade-off theories are formulated as specific models, they are easy to reject on a variety of dimensions. But not all rejections of a model are equally serious. The model may still provide a very useful way to think about the data. Even if a model is rejected, it may still fit the evidence better than any other available model. How to balance formal rejection versus insight is not easy. What should count as evidence against a particular point of view and what should not? Reasonable people can answer this question in differing ways.

Advocates of the trade-off point of view tend to take rational optimizing behavior particularly seriously. Advocates of the pecking order point of view tend to take the dominance of retained earnings and debt over equity with particular seriousness. These are not inherently conflicting considerations, however.

### 2.2. *The Modigliani-Miller theorem*

The theory of business finance in a modern sense starts with the Modigliani and Miller (1958) capital structure irrelevance proposition.[4] Until then, there was no generally accepted theory of capital structure. Modigliani and Miller's initial assumption is that the firm has a particular set of expected cash flows. When the firm chooses a certain proportion of debt and equity to finance its assets, all that it does is to divide up the cash flows among investors. Investors and firms are assumed to have equal access to financial markets, which allows for homemade leverage. The investor can create any leverage that was wanted but not offered, or the investor can get rid of any leverage that the firm took on but was not wanted. As a result, the leverage of the firm has no effect on the market value of the firm.

The Modigliani and Miller paper led subsequently to both clarity and controversy. As a matter of theory, capital structure irrelevance can be proved under a range of circumstances. There are two fundamentally different types of capital structure irrelevance propositions. The first, the classic arbitrage-based irrelevance proposition, provides settings in which arbitrage by investors keeps the value of the firm independent of its leverage. In addition to the original Modigliani and Miller paper, important contributions include papers by Hirshleifer (1966) and Stiglitz (1969).

A second kind of capital structure irrelevance is associated with multiple equilibria. In models of this kind, equilibrium conditions pin down the aggregate amount of debt and equity in the market. But the model does not specify how these aggregate quantities get divided up among the firms. The classic paper is by Miller (1977) in which consideration of both personal and corporate tax determines an economywide leverage ratio, but there are multiple equilibria in which debt is issued by different firms. A similar kind of firm-level capital structure irrelevance is found in Auerbach and King (1983).

The 1958 paper stimulated serious research devoted to disproving irrelevance as a matter of theory or as an empirical matter. This research has shown that the Modigliani–Miller theorem fails under a variety of circumstances. The most commonly used elements include consideration of taxes, transaction costs, bankruptcy costs, agency conflicts, adverse selection, lack of separability between financing and operations, time-varying financial market opportunities, and investor clientele effects. Alternative models use

[4] As is common with important contributions to knowledge, there is some dispute on the origin of the idea. Williams (1938) makes a relatively clear statement of the idea but does not present an explicit arbitrage-based proof. Rubinstein (2003) presents an interesting discussion of the history of ideas.

differing elements from this list. Given that so many different ingredients are available, it is not surprising that many different theories have been proposed. Covering all of these theories would go well beyond the scope of this chapter. Harris and Raviv (1991) provided a survey of the development of this theory as of 1991.

As an empirical proposition, the Modigliani–Miller irrelevance proposition is not easy to test. With debt and firm value both plausibly endogenous and driven by other factors such as profits, collateral, and growth opportunities, we cannot establish a structural test of the theory by regressing value on debt.[5] But the fact that fairly reliable empirical relations between a number of factors and corporate leverage exist, while not disproving the theory, does make it seem an unlikely characterization of how real businesses are financed.

What then is one to make of the theorem? A popular defense has been as follows. "While the Modigliani–Miller theorem does not provide a realistic description of how firms finance their operations, it provides a means of finding reasons why financing may matter." This description provides a reasonable interpretation of much of the theory of corporate finance up to perhaps the 1980s. Accordingly, it influenced the early development of both the trade-off theory and the pecking order theory. However, as the next two sections show, current progress in capital structure theory is not based on reexamining the list of assumptions that generate the Modigliani–Miller theorem to find a previously unrelaxed assumption.

### 2.3. The trade-off theory

The term *trade-off theory* describes a family of related theories. In all of these theories, a decision maker running a firm evaluates the various costs and benefits of alternative leverage plans. Often it is assumed that an interior solution is obtained so that marginal costs and marginal benefits are balanced.

The original version of the trade-off theory grew out of the debate over the Modigliani–Miller theorem. When corporate income tax was added to the original irrelevance proposition Modigliani and Miller (1963), this created a benefit for debt in that it served to shield earnings from taxes. Since the firm's objective function is linear and there is no offsetting cost of debt, this implied 100% debt financing.

To avoid this extreme prediction, an offsetting cost of debt is needed. The obvious candidate is bankruptcy. Kraus and Litzenberger (1973) provide a classic statement of the theory that optimal leverage reflects a trade-off between the tax benefits of debt and the deadweight costs of bankruptcy. According to Myers (1984), a firm that follows the trade-off theory sets a target debt-to-value ratio and then gradually moves toward the target. The target is determined by balancing debt tax shields against costs of bankruptcy.

Several aspects of Myers's definition merit discussion. First, the target is not directly observable. It may be imputed from evidence, but that depends on adding a structure. Different papers add that structure in different ways.

[5] Fama and French (1998) and Kemsley and Nissim (2002) provide related discussions.

Second, the tax code is much more complex than that assumed by the theory. Depending on which features of the tax code are included, different conclusions regarding the target can be reached. Graham (2003) provides a useful review of the literature on the tax effects.

Third, bankruptcy costs must be deadweight costs rather than transfers from one claimant to another. The nature of these costs is also important. Are these fixed costs? Do they increase with the size of the bankruptcy? Are the costs one-time costs such as a lawyer's fees, or are they permanent costs such as the cost of a damaged reputation? Haugen and Senbet (1978) provide a useful discussion of bankruptcy costs.

Fourth, for the analysis to work transaction costs must take a specific form. For the adjustment to be gradual rather than abrupt, the marginal cost of adjusting must *increase* when the adjustment is larger. This assumed form of adjustment cost is rather surprising since one expects to see large fixed costs and perhaps roughly constant marginal costs. This implies a very different adjustment path. Leary and Roberts (2005) describe the implications of alternative adjustment cost assumptions.

For these reasons, we break Myers's definition into two parts. We call the first part the static trade-off theory, and the second part, target adjustment behavior.

**Definition 2. 1.** A firm is said to follow the static trade-off theory if the firm's leverage is determined by a single period trade-off between the tax benefits of debt and the deadweight costs of bankruptcy.

**Definition 2. 2.** A firm is said to exhibit target adjustment behavior if the firm has a target level of leverage and if deviations from that target are gradually removed over time.

### *2.3.1. The static trade-off theory*

Bradley et al. (1984) provide the standard presentation of the static trade-off theory. The assumed tax structure is not intended to be strictly realistic. For instance, the tax code contains important dynamic aspects that cannot be properly represented in a single-period model. However, the model does contain some important elements of the actual U.S. tax code.

Investors are risk-neutral and face a *progressive* tax rate on end-of-period wealth from bonds. Dividends and capital gains are taxed at a single constant rate. Risk-neutrality induces the investor to invest in whichever security offers the better expected after-tax deal.

The firm faces a constant marginal tax rate on end-of-period wealth. It can deduct both interest and principle payments, but the investor must pay taxes as these payments are received. Nondebt tax shields exist, but they cannot be arbitraged across firms or across states of nature. If the firm is unable to make the promised debt payment, then it incurs deadweight financial distress costs, so that "the pie shrinks."

Let $\tau_c$ = the constant marginal tax rate on corporate income, $\tau_{pb}$ = the progressive tax rate on investor bond income, $\tau_{ps}$= the tax rate on investor equity income, $X$ = the end-of-period value of the firm before taxes and debt payments, $k$ = the fraction of end-of-period value that is lost if the firm defaults on debt, $B$ = the end-of-period payment promised to bondholders, $\phi$ = the total after-tax value of nondebt tax shields if fully used, $r_f$ = the risk-free, tax-free rate of return, $f(X)$ = the probability density of $X$, and $F(\cdot)$ = the cumulative probability density function.

The following table describes the returns to stockholders and bondholders in various states defined by the level of corporate earnings. The column "Total" indicates the firm's total earnings, denoted as $X$. If the earnings are negative, then both debt and equity give up their claims; no debt is paid. If the earnings are positive but not enough to cover the promised debt payment, $B$, then equity defaults, and debt takes over. A deadweight loss of $kX$ is used up in the process.

| Total | State | Debt | Equity | Tax | Loss |
|---|---|---|---|---|---|
| $X$ | $X < 0$ | 0 | 0 | 0 | 0 |
| $X$ | $0 < X < B$ | $X(1-k)$ | 0 | 0 | $kX$ |
| $X$ | $B < X < B+\phi/\tau_c$ | $B$ | $X - B$ | 0 | 0 |
| $X$ | $X > B+\phi/\tau_c$ | $B$ | $X - B - \tau_c(X - B) + \phi$ | $\tau_c(X - B) - \phi$ | 0 |

If earnings are large enough for equity not to default, there remains the question of whether the earnings are low enough that the nondebt tax shield is sufficient to cover the tax liability. Thus, the last states in the table differ with respect to taxation. In the last state (high income), the firm is able to utilize fully the nondebt tax shield ($\phi$), and equityholders therefore receive $X - B - \tau_c(X - B) + \phi$. In the penultimate range of states, income is not sufficiently high and nondebt tax shields are not fully utilized. As a result, no tax is paid and equityholders receive $X - B$. The dividing line occurs as the point where income is just sufficient to have $X - B = X - B - \tau_c(X - B) + \phi$. Obviously, this can be rearranged to $X = B+\phi/\tau_c$, which defines the boundary as shown above.

The market value of debt is found by integrating the bondholder after-tax returns across different states:

$$V_B = \left(\frac{1-\tau_{pb}}{1+r_f}\right)\left[\int_B^{\infty} Bf(X)dX + \int_0^B X(1-k)f(X)dX\right] \tag{1}$$

The market value of equity can be obtained by integrating the stockholder after-tax returns across different states:

$$V_S = \left(\frac{1-\tau_{ps}}{1+r_f}\right)\left[\int_{B+\phi/\tau_c}^{\infty}[(X-B)(1-\tau_c)+\phi]f(X)dX + \int_B^{B+\phi/\tau_c}(X-B)f(X)dX\right] \tag{2}$$

Adding together $V_S$ and $V_B$ gives an expression for the value of the firm, that is, $V = V_S + V_B$. It is assumed that the firm's choice of leverage, $B$, is determined by maximizing $V$.

The assumption that $B$ is chosen to maximize $V$ is conventional, but it is not innocuous. For instance, the firm might be maximizing managerial welfare, or the welfare of a particular set of large shareholders who have control. Such agency conflicts are assumed away. As usual, in an optimization problem, the optimal value might be found at either an interior point or on a boundary. If the optimal solution is interior, then it is provided by a first-order condition determined by differentiating $V$ with respect to $B$ and setting it equal to zero, that is, $\partial V/\partial B = 0$. For this model we have

$$\frac{\partial V}{\partial B} = \left(\frac{1-\tau_{pb}}{1+r_f}\right)\left\{[1-F(B)]\left[1-\frac{(1-\tau_c)(1-\tau_{ps})}{(1-\tau_{pb})}\right] - \frac{(1-\tau_{ps})\tau_c}{1-\tau_{pb}}[F(B+\phi/\tau_c)-F(B)] - kBF(B)\right\} \tag{3}$$

The first term in this expression represents the marginal net tax benefit of debt; the second term represents the increase in the probability of wasting interest tax shields when earnings are less than tax shields; and the third, the marginal increase in expected costs of distress. The firm's decision involves a trade-off between the marginal tax advantage of debt and the marginal leverage-related costs.[6] The main predictions from the model are found by redifferentiating the first-order condition with respect to each of the parameters of interest. They show that:

(1) An increase in the costs of financial distress ($k$) reduces the optimal debt level.
(2) An increase in nondebt tax shields ($\phi$) reduces the optimal debt level.
(3) An increase in the personal tax rate on equity ($\tau_{ps}$) increases the optimal debt level.
(4) At the optimal capital structure, an increase in the marginal bondholder tax rate ($\tau_{pb}$) decreases the optimal level of debt.
(5) The effect of risk ($\sigma$) is ambiguous, even if uncertainty is assumed to be normally distributed. For "reasonable" parameter values, Bradley et al. (1984) show that the relation between the debt ratio and volatility is negative.

Although most of the predictions are intuitively reasonable, it is surprising that the effect of risk on leverage is ambiguous. This ambiguity between leverage and risk is also found in a variety of other models.

[6] This model nests a number of prior theories as special cases. To get the model in Miller (1977), set $\tau_{ps} = k = \phi = 0$, so that $\partial V/\partial B = ([1-F(B)]\tau_c - [1-F(B)]\tau_{pb})/(1+r_f)$. The firm term is the marginal tax advantage of debt obtained by multiplying the corporate tax rate ($\tau_c$) with the probability that the firm will not default ($[1-F(B)]$). The second term is the tax premium the firm expects to pay to bondholders. To get the model in Deangelo and Masulis (1980), set $\tau_{ps} = 0$, and find $\partial V/\partial B = [[1-F(B)](\tau_c - \tau_{pb}) - \tau_c[F(B+\phi/\tau_c) - F(B)] - (1-\tau_{pb})kBf(B)]/(1+r_f)$. If either $\phi$ or $k$ is positive, $\tau_c$ will be greater than $\tau_{pb}$, and the first term, which represents the net tax advantage of debt, will be positive.

Tests of this model face the problem that the main elements of the model are not directly observable. Instead, proxies are used. When, for instance, Bradley et al. (1984) find an unexpected sign on nondebt tax shields, it is unclear whether the problem is a defect in the theory or in the proxy.

This trade-off model is static, although firms in the real world operate over many periods. Thus, testing the theory with data requires making auxiliary assumptions. Two aspects of static modeling are particularly important in tests of the theory—the role of retained earnings and the interpretation of mean reversion.

By construction, there are no retained earnings in the model. How should one interpret retained earnings? At one level, it can be argued that retained earnings are direct evidence that a one-period model is inappropriate. Although there is some truth to such a claim, it is fairly harsh. Theories are always simplifications. Retained earnings represent inside equity, and profitable firms automatically create this kind of equity. Unless the firm takes some offsetting action, the more profitable a firm is, the lower its leverage will be. This kind of equity creation is conceptually rather different from a secondary equity issue.

Again, by construction, and as already discussed, this theory says nothing about mean reversion. The model has a solution for leverage, but there is no room in the model for the firm ever to be anywhere but at the solution. Thus, the model contains no notion of target adjustment. This is why we separate the static trade-off theory from the target adjustment hypothesis. Evidence for or against mean reversion is not evidence of the applicability of the static trade-off theory. These are separate questions. Predictions about dynamics arise from dynamic models.

These two unmodeled aspects of the theory have been very influential in forming the profession's view of the trade-off theory and have resulted in considerable dissatisfaction with it. Some scholars have reacted by turning away from taxation and bankruptcy costs as key features altogether (e.g., see Jensen and Meckling, 1976, and Myers, 1984); for many years, this alternative line of research dominated corporate finance scholarship. In the last few years, some scholars have been returning to consideration of taxation and bankruptcy costs, but with an explicit treatment of the fact that firms last longer than a single period, which leads to the dynamic trade-off theory.

### *2.3.2. The dynamic trade-off theory*

Constructing models that recognize the role of time requires specifying a number of aspects that are typically ignored in a single-period model. Of particular importance are the roles of expectations and adjustment costs. In a dynamic model, the correct financing decision typically depends on the financing margin that the firm anticipates in the next period. Some firms expect to pay out funds in the next period, whereas others expect to raise funds. If funds are to be raised, they may take the form of debt or equity. More generally, a firm undertakes a combination of these actions. Before discussing individual papers, we present two illustrative examples of the potential importance of dynamics in financing decisions.

*2.3.2.1. First example* Consider a highly profitable firm. Instead of raising funds, it plans to distribute money to its shareholders. It can distribute funds today, or it can retain these funds and distribute them one period later. Which should the firm do? The answer depends on the tax rates and on rates of return that the firm can earn relative to the returns that the shareholders can obtain directly. (Evidently, we are far from the world of Modigliani and Miller, 1958, at this point.) Given that the firm is profitable, its investment opportunities are likely to be better than those of its shareholders. This may lead to situations when it is better for a firm to retain funds, even though it faces a higher tax rate than do its shareholders. The more profitable the firm, presumably the more likely this is true. The example suggests that more profitable firms should retain more earnings than should less profitable firms. Since retained earnings are equity, in this example we might expect to see more profitable firms have lower leverage.

*2.3.2.2. Second example* Consider a firm that has more funds than it wishes to invest in this period. The firm anticipates investing in a year or two, at which time it will need funds. In a tax-free world, the firm could pay out the excess cash to its shareholders today, and later, when funds are needed, it could issue new equity. But taxes create a wedge. Paying out money causes shareholders to pay taxes. With taxes, financing round trips can be expensive. Thus, distributing funds and then raising new equity subsequently imposes a tax liability on shareholders that could have been avoided had the firm retained the funds. Hence, taxes can directly motivate firms to retain earnings.

These two examples are not complete theories. They are merely illustrations of the fact that dynamic trade-off models depart from static trade-off models in interesting ways.

The early attempts to model the dynamic trade-off appeared to be technically difficult, and not all that promising at a time when adverse selection and agency considerations were center stage in the literature. Currently, scholars are starting to work through the technical problems that are present in dynamic models with uncertainty and bankruptcy.[7] The dynamic models contain features that allow the trade-off theory to provide a better account of how firms finance their operations than had been previously thought.

An important precursor to modern dynamic trade-off theories was Stiglitz (1973), who examines the effects of taxation from a public finance perspective. Stiglitz's model is not a trade-off theory since he took the drastic step of assuming away uncertainty. This, of course, simplifies things immensely. His analysis, which allows for both personal and corporate taxes, highlights an interesting asymmetry in the tax code. Money paid in to the

[7] Due to technical difficulties, these current papers make important simplifying assumptions that probably experience loss of generality. Unfortunately, at this time, in contrast to the literature on asset pricing, we do not yet have a common workhorse model from which individual papers can naturally be developed.

firm is not taxed, but money paid out is taxed.[8] For reasonable parameter values, Stiglitz's basic result is that it pays to finance as much investment as possible through retained earnings and to finance the excess of investment over retained earnings with debt. The observed leverage ratio is thus a "fortuitous outcome of the profit and investment history of the firm" (Stiglitz, 1973, p. 32). In other words, the solution is essentially what we might now call the pecking order.

The first dynamic models to consider the tax-savings versus bankruptcy cost trade-off are by Kane et al. (1984) and Brennan and Schwartz (1984). Both analyzed continuous time models with uncertainty, taxes, and bankruptcy costs, but no transaction costs. Since firms react to adverse shocks immediately by rebalancing costlessly, firms maintain high levels of debt to take advantage of the tax savings. These models reinforced Miller (1977) in that the trade-off theory predicts much higher debt levels than those typically observed in most firms (see Section 3.2.3).

To avoid the unrealistically rapid rebalancing problem, Fischer et al. (1989) introduced transaction costs into the analysis of dynamic capital structures. Because of transaction costs, the firm allows its capital structure to drift much of the time. When its leverage gets too far out of line, the firm undertakes a discrete rebalancing. They assumed that the rebalancing takes place at an upper and at a lower limit so that recapitalization takes the form of an "($s$, $S$)" policy. When the firm earns profits, it pays down debt. If the lower leverage limit is reached, the firm recapitalizes. If the firm loses money so that debt increases, it will again permit the drift until the boundary is reached. Accordingly, when we look at a large panel of data, most of the data reflects drift rather than active rebalancing. This can account for the empirical observation that profits and leverage are negatively related.

Fischer et al. (1989) solve the model numerically. Their simulations suggest that even small transaction costs can lead to delay in rebalancing and wide variations in the debt ratio. The numerical solutions have a number of reasonable features. The tax advantage of debt is increasing in the corporate tax rate and decreasing in the personal tax rate. Greater volatility is associated with an increased range over which the firm permits leverage to fluctuate and with a reduction in the target to which the firm recapitalizes when boundaries are reached. Thus, volatility is negatively associated with average leverage. Leary and Roberts (2005) show that the Fischer et al. model accounts for a number of aspects of firm leverage dynamics. More controversially, in the Fischer et al. model, good operating performance will eventually cause the firm to hit the refinancing barrier, at

[8] There is "a basic asymmetry (which arises even in our idealized tax structure) between payments to shareholders and receipts from them. Payments to shareholders are taxed, so reductions in dividends or in shares purchased back from shareholders reduce the taxes paid, but receipts from shareholders are not taxed. Accordingly, if the firm is not paying out any dividends, using all of its retained earnings for investment, and financing the excess of investment over retained earnings by debt, an attempt to increase the equity by reducing the new debt issue and increasing the new equity issue will have disadvantageous tax effects; there will be no reduction in taxation on 'equity account' this period but an increase in corporate profit taxes paid in future periods because of the reduction in interest payments" (Stiglitz, 1973, p. 7).

which point it loads up on debt.[9] Thus, good performance is eventually followed by debt issues.

In order to understand the recent dynamic trade-off literature, it is helpful to classify the papers according to their assumptions. One important dividing line is the treatment of investment. Classical analysis such as Modigliani and Miller (1958) and Kraus and Litzenberger (1973) takes the firm's cash flows as exogenous. Many trade-off models such as Kane et al. (1984), Fischer et al. (1989), Goldstein et al. (2001) and Strebulaev (2007) follow this tradition of keeping the firm's cash flows exogenous. However, it is quite likely that investment and thus the firm's cash flows depend on how the firm finances its operations. Thus, some papers consider investment together with financing. Notably, this is done by Brennan and Schwartz (1984), Mello and Parsons (1992), Mauer and Triantis (1994), Titman and Tsyplakov (2007), and Hennessy and Whited (2005).

If the firm's earnings are stochastic but unrelated to leverage, then one must decide how to model the excess cash in good times. Generally, it is assumed that excess cash will be paid out to the shareholders. Many papers do not give the firm a choice of how much to pay out versus how much to retain. For example, Brennan and Schwartz (1984) and Titman and Tsyplakov (2007) assume that the firm pays out all funds. This assumption obviously limits the ability of the theory to speak to the empirically important issue of retained earnings. Both Stiglitz (1973) and Hennessy and Whited (2005) are more satisfactory in this respect.

The tax system assumptions differ across papers. Kane et al. (1984), Fischer et al. (1989), and Titman and Tsyplakov (2007) omit consideration of taxes on corporate payouts. To include linear tax on distributions, Goldstein et al. (2001) and Strebulaev (2007) appear to be effectively assuming that the shareholders get a tax rebate when contributing equity. By contrast, Stiglitz (1973) and Hennessy and Whited (2005) both capture the fundamental distinction that generally taxes are due on distributions from firms to investors, but not on funds that investors provide to firms.

Dynamic trade-off models can also be used to consider the option values embedded in deferring leverage decisions to the next period. Goldstein et al. (2001) observe that a firm with low leverage today has the subsequent option to increase leverage. Under their assumptions, the option to increase leverage in the future serves to reduce the otherwise optimal level of leverage today. Strebulaev (2007) analyzed a model quite similar to that of Fischer et al. (1989) and Goldstein et al. (2001). Again, if firms optimally finance only periodically because of transaction costs, then the debt ratios of most firms will deviate from the optimum most of the time. In the model, the firm's leverage responds less to short-run equity fluctuations and more to long-run value changes.

[9] Some have objected that persistently profitable firms do not often go out to load up on debt. However, such firms do undertake significant new debt when engaging in mergers and acquisitions. Profitable firms do seem more likely to undertake such actions. The best way to think about the relation between leverage and M&A activity probably deserves more attention.

Hennessy and Whited (2005) consider the interaction of financing and investment in a model with corporate and personal taxes, financial distress costs, and equity flotation costs. In contrast to many of the earlier papers, the firm is not obliged to pay out funds, and so it allows for an explicit analysis of the kinds of dynamic considerations discussed in the two examples presented earlier in this section. They find that optimal leverage is path dependent and that profitable firms tend to be less highly levered. Another closely related paper that considers the interaction between investment and financing decisions is by Tserlukevich (2006), who develops a dynamic structural trade-off model with real frictions. In this model, while equity values increase when firms experience positive demand shocks, investment is delayed because of real frictions, and consequently leverage goes down. For tax reasons, firms issue debt only when investments are made. Incorporating these real frictions in a dynamic model of financing decisions, Tserlukevich is able to replicate several stylized facts such as the mean reversion in leverage and the inverse relation between leverage and profitability without relying on transaction costs.

Hennessy and Whited (2005), Fischer et al. (1989) as further developed by Leary and Roberts (2005), and Strebulaev (2007) all seem capable of accounting for the results of Baker and Wurgler (2002) and Welch (2004). The models of Goldstein et al. (2001) and Hennessy and Whited (2005) also help to resolve the high-debt problem identified by Kane et al. (1984) and Brennan and Schwartz (1984).[10]

Lewellen and Lewellen (2006) argue that if a firm repurchases shares, the tax that must be paid by the shareholders depends on the capital gains that have been incurred since they bought in originally. Accordingly, the optimal financing of the firm may depend on how frequently their shares are turned over in the stock market. Firms with many long-term shareholders may be more reluctant to trigger a tax bill for their shareholders than would firms with many short-term shareholders.[11]

Certain ideas are fairly general in dynamic models. The optimal financial choice today depends on what is expected to be optimal in the next period. In the next period, it may be optimal to raise funds or to pay them out. When raising new funds, it might be optimal to raise them in the form of debt or in the form of equity. In each case, what is expected to be optimal in the next period will help to pin down the relevant comparison for the firm in the current period. By stressing different costs, different dynamic models lead to somewhat different conclusions.

The idea that the rate of return in the hands of the firm needs to be compared to the rate of return in the hands of the investor is fairly general and seems to transcend specific

[10] A variety of other dynamic agency models has been proposed such as Morellec (2004) and Mello and Parsons (1992). The optimal capital structure can also be considered from an optimal dynamic contracting perspective as in Atkeson and Cole (2005). Titman and Tsyplakov (2007) study the difference between firm value maximization and equity value maximization.

[11] Green and Hollifield (2003) provide a careful analysis of many of the complexities that arise due to personal taxes. However, they simplify the analysis by assuming that corporate debt policy keeps the interest payments fixed from period to period.

models. The fact that transaction costs and taxes can create wedges that lead to times when money should be left in whichever hands it currently is in is also quite general and recurs across many models.

As stated earlier, there is an asymmetry created by the fact that payments by firms to investors trigger taxation, but payments by investors to firms do not do so. This asymmetry seems to be fairly basic, and it is likely to arise in many dynamic models. Both taxes and transactions costs can create wedges such that shocks are not undone. This leads to path-dependent solutions in a variety of models.

Much of the work on dynamic trade-off models is fairly recent, and so any judgments on their results must be somewhat tentative. This work has already fundamentally altered our understanding of mean reversion, the role of profits, the role of retained earnings, and path dependence. As a result, the trade-off class of models now appears to be much more promising than it did even just a few years ago.

### *2.4. The pecking order theory*

The pecking order theory comes from Myers (1984), who in turn was influenced by the earlier institutional literature, including the book by Donaldson (1961). Myers (1984) argues that adverse selection implies that retained earnings are better than debt and debt is better than equity. This ranking was motivated with reference to the adverse selection model in Myers and Majluf (1984). The ordering, however, stems from a variety of sources, including agency conflicts and taxes.[12]

**Definition 2. 3.** Myers (1984): A firm is said to follow a pecking order if it prefers internal to external financing and debt to equity if external financing is used.

This definition can be interpreted in different ways. What does it mean to "prefer" internal financing? Does this mean that the firm uses all available sources of internal finance before using any debt or equity issues? Or does it mean that, "other things equal," the firm will mostly use internal financing before using external financing? If the verb "prefer" is interpreted strictly, the theory is more testable. If "prefer" is interpreted in the "other things equal" way, then any test of the theory rests on the specification of "other things equal."

Most firms hold some internal funds (cash and short-term investments) even when they are raising outside funds. This is so obvious that it is rarely considered in tests of the pecking order. It is implicitly assumed that these funds are held for reasons that are outside the theory, such as for transactions. Accordingly, almost all discussions maintain some version of an "other things equal" interpretation of the relative use of internal and external funds.

[12] Baker et al. (2007) point out that excessive managerial optimism might also be used to generate a version of the pecking order if it were the only distortion.

A second problem for the definition concerns the preference of debt over equity. As we will see, initial claims for the theory tended to rest on a strict interpretation in which equity is never issued if debt is feasible. As it has become increasingly clear that this strict interpretation is not only more refutable, but actually refuted, proponents of the pecking order theory have moved increasingly to the "other things equal" interpretation. Different papers invoke different empirical versions of "other things equal." Of course, the more a test depends on the other things, the less the data are explained by the pecking order itself.

At what point is equity introduced? The strict interpretation suggests that after the initial public offering (IPO), equity should never be issued unless debt has for some reason become infeasible. This leads to the notion of a "debt capacity." The debt capacity serves to limit the amount of debt within the pecking order and to allow for the use of equity. Obviously, this raises the problem of defining the debt capacity. The literature provides no agreed-upon definition. Several recent papers have used factors commonly employed in tests of the trade-off theory to define the debt capacity. Of course, this leads to difficulties in interpreting the results.

Pecking order models can be derived based on adverse selection considerations, agency considerations, or other factors. Two common features underlie pecking order theories. The first feature is the linearity of the firm's objective function; this helps because it means that costs tend to drive the results to corner solutions. The second feature is the model's relative simplicity. The pecking order hierarchy is a relatively simple structure. A model that is complex is unlikely to have such a simple solution. When many things are factored in, a more complex range of things tends to happen. Thus, it seems that the pecking order is generally more likely to emerge from an illustrative model than it is from a unifying model.

While this section describes pecking order models based on adverse selection and agency costs, Section 2.3.2 shows that tax considerations alone can also generate pecking order behavior. It is also possible to have other features that lead to a financing hierarchy. To the best of our knowledge, no one has tried to distinguish among the alternative possible sources of pecking order behavior.

### *2.4.1. Adverse selection*

The most common motivation for the pecking order is adverse selection, developed by Myers and Majluf (1984) and Myers (1984). The key idea is that the owner–manager of the firm knows the true value of the firm's assets and growth opportunities. Outside investors can only guess these values. If the manager offers to sell equity, then the outside investor must ask why the manager is willing to do so. In many cases, the manager of an overvalued firm will be happy to sell equity, while the manager of an undervalued firm will not. Our presentation follows Cadsby et al. (1990).

Consider an original owner/operator of a firm and potential investors. Everyone is risk-neutral, and there are no transaction costs and no discounting. All financing is through equity. The firm has some existing assets, and it decides whether or not to invest in a project. If the project is to be undertaken, then the potential investors compete in an auction for the right to finance the project. The auction is for a share of equity in the firm that the investor demands in exchange for the necessary funding of the project. Accordingly, financing is breakeven given the beliefs of the investors.

The firm has assets in place, denoted by $A_i$, and access to a positive net present value (NPV) project that offers a net payoff denoted by $B_i$. The subscript $i$ refers to the firm's type, which can be either type $H$ (high) or type $L$ (low). The sum of the assets in place plus the net value of the project is greater for a type $H$ firm than it is for a type $L$ firm. The two types are equally likely. The firm knows the true worth of both its assets and the project. The investors can only guess about the firm's type. In order to undertake the project, the firm would need to raise $I > 0$ from the investor.

If the project is not undertaken, then the firm's value (denoted $V_i$) is just $V_i = A_i$. If the project is undertaken, $V_i$ must be shared with the outside investor. The investor's share of the firm is denoted $s$, so the original owner gets $(1 - s)V_i$. An auction is held among the risk-neutral investors for the right to provide $I$ in exchange for $sV_i$. The winner of the auction expects to break even.

There is a unique pooling equilibrium in which both type $H$ and type $L$ firms undertake the new projects if and only if $(I/V_L) < (B_H + I)/V_H$. The investor gets a share denoted $s^*$, where $s^* = I/(0.5V_H + 0.5V_L)$.

The pooling equilibrium conditions allow the investor only to expect to break even on average since both types of firm will undertake the project. Under the parameter value restriction, the new project is sufficiently lucrative that the high-type firm wishes to go ahead, despite the fact that the investor is only financing the project on average terms. Thus, all players are willing to follow the suggested strategies.

There is a unique separating equilibrium in which a type $L$ firm undertakes the project and a type $H$ firm does not, if and only if $(B_H + I)/V_H < I/(0.5V_H + 0.5V_L)$. The investors get a share, $s^* = I/V_L$.

In this case, only the low-type firm goes ahead with the project. The investor knows that a low-type firm is being financed and therefore demands terms that reflect this fact. If a high-type firm were to go ahead, the investor would demand the same unattractive terms required from low-type firms. As such, the high-type firm finds it better to simply forgo the project altogether. The parameter values are such that the suggested strategies reflect each player's self-interest given how all the other players are acting.

Both a pooling and a separating equilibrium exist simultaneously when $I/(0.5V_H + 0.5V_L) < (B_H + I)/V_H < I/V_L$. The investor's shares depend on whether the equilibrium is pooling or separating. The investor always expects only to break even. Cadsby et al. (1990) point out that in the overlapping region, there is also a semi-separating equilibrium.[13]

[13] Cadsby et al. (1990) conducted experimental tests of the model. The model predicted well. Cadsby et al. (1998) also considered an extended version of the model in which the firm could advertise its type. In this case, the predictions of the model were not as good.

In the pooling equilibrium, the asymmetric information does not cause the valuable project to be lost. But if the value of the assets in place is quite high relative to the value of the positive NPV of the project, then the firm chooses not to raise any outside funds.

In this model, internal financing when feasible would always work. That is to say, such financing would avoid all asymmetric information problems. External equity is sometimes too expensive, and the firm will even give up positive NPV projects to avoid it. This is part of the pecking order hierarchy.

As in Myers and Majluf (1984), debt is not formally included in the analysis. If debt were available and risk free, it would work as well as internal financing. If debt is available and risky, then Myers (1984) argues *intuitively* that it ought to fall somewhere between retained earnings and equity, thus creating the pecking order.

The formal analysis of a model with risky debt is not as simple as it seems when reading Myers (1984). When both debt and equity financing are feasible, there are often multiple equilibria, and it is not clear how to select among them. Noe (1988) provides an important analysis of the problem. Cadsby et al. (1998) provide experimental tests of some of the equilibrium selection arguments that have been invoked in financial theory. Path dependence and learning seem to play more important roles than do formal equilibrium selection criteria.

The subsequent theoretical literature has considered many versions of adverse selection problems. Generally, the results are not as elegant as the standard pecking order suggests. For example, the adverse selection model of Myers and Majluf (1984) assumes one-sided asymmetric information in which a firm selects securities for cash. However, if information asymmetry is two sided (as in Eckbo et al., 1990), there are several possible equilibria leading to the firm's preference for stock, or a combination of stock and cash over pure cash. Thus, in mergers with two-sided information asymmetry, firms sometimes actually prefer stock transactions over cash transactions.

Dybvig and Zender (1991) show that properly designed managerial compensation contracts (with compensation tied to the value of the firm) could solve adverse selection problems. However, in practice, one rarely observes managerial compensation contracts that are linked to firm value; they are mostly tied to equity value. Viswanath (1993) considers a world with more than one period, and he finds that the results depend on how the first- and the second-period uncertainties are related. Ravid and Spiegel (1997) consider adverse selection with no assets in place to start with. This results in entrepreneur and investor splitting the proceeds. In their setting, as in the examples discussed above, firms will use riskless debt before turning to equity financing.

Eckbo and Masulis (1992) and Eckbo and Norli (2004) extend the basic adverse selection model to allow for current shareholder participation in equity issues and underwriter quality certification. Adverse selection would be less severe if current shareholders were allowed to participate in the equity issue. In their model, firms that expect a high proportion of their current shareholder to take-up new issues face low adverse selection and prefer to issue uninsured rights. Firms with expectations of low current shareholder take-up prefer to issue equity using "firm-commitment" underwritten offerings. Firms with expectations of intermediate current shareholder take-up prefer to issue equity using standby rights. This implies what might be termed a pecking order of equity flotation method choices.

Halov and Heider (2005) argue that the standard pecking order is a special case of adverse selection. When there is adverse selection about firm value, firms prefer to issue debt over outside equity and standard pecking order models apply. However, when there is asymmetric information about risk, adverse selection arguments for debt apply and firms prefer to issue external equity over debt. Thus, adverse selection can lead to a preference for external debt or external equity, depending on whether asymmetric information problems concern value or risk.

The main conclusion is that adverse selection models can be a bit delicate. It is possible to construct equilibria with a pecking order flavor. But adverse selection does not imply pecking order as a general situation.

### 2.4.2. *Agency theory*

The idea that managers prefer internal financing to external financing is, of course, old (e.g., Butters, 1949). Traditionally, the argument was that outside financing required managers to explain the project details to outside investors and therefore expose themselves to investor monitoring. Managers dislike this process and prefer retained earnings to external financing. But there is no direct prediction about the relative use of debt versus equity when seeking external financing. These ideas were subsequently developed into agency theories, with Jensen and Mecking (1976) being a prominent contribution.

Myers (2003) points out that some versions of agency theory imply a financing hierarchy. Agency costs of equity, for example, could result in a pecking order. Consider a simple and conventional example of the agency cost of equity that follows Jensen and Meckling (1976). The firm is owned and run by an entrepreneur. The entrepreneur has $R$ dollars; if she invests all of $R$, then her return is $V(R)$ with $V' > 0 > V''$. Her consumption of desirable perks is the difference between $R$ and the amount she chooses to invest. Let that amount of investment be $I$. With no outside financing, her problem is:

$$\max_{I} \quad V(I) + (R - I) \tag{4}$$

$$s.t. \quad I \leqslant R \tag{5}$$

This gives the obvious first-order condition, $V' = I$, if the constraint is not binding. Let $I^*$ denote the solution to this first-order condition. This gives her a payoff of $V(I^*) + R - I^*$.

What happens if the constraint is binding so that $I^* > R$? Then, outside financing is interesting. Assume that financing is with riskless debt. Then, the entrepreneur asks for $I^* - R$ and promises to repay $D$. The entrepreneur invests optimally and repays properly. There is no distortion. If internal financing is inadequate, then risk-free external debt does not cause any distortions.

Introduction of equity into the model requires a notion of exogenous debt capacity that becomes binding at some point. For simplicity, we directly assume that outside financing takes the form of equity, $E$, and that the entrepreneur cannot commit to not consuming the perks. The outsiders will get a fraction, $1 - s$, of the firm. The amount raised will be $E = (1 - s)V(I)$. Thus, the problem for the entrepreneur is now:

$$\max_{I} \quad sV(I) + R + E - I \tag{6}$$

$$s.t. \quad I \leqslant R + E \tag{7}$$

The associated first-order condition is $sV'(I) = 1$. The solution is denoted $I^{**}$. As long as $s < 1$, then $I^{**} < I^{*}$ and the entrepreneur is underinvesting. She bears the full cost of any perks not consumed, and she must share the benefits.

Obviously, this underinvestment is inefficient. Use of internal financing would result in higher welfare. Thus, retained earnings are preferred. Debt is just as good in this simple model. Equity is inefficient. We, therefore, have a version of the pecking order.

Jensen and Meckling (1976) also identified an agency problem of debt called risk shifting. The idea is that if the firm is operated on behalf of equity, only cash flows in nonbankrupt states matter. The firm will therefore tend to accept projects that are too risky but have large payoffs in good states. It is clear that this kind of behavior is sometimes observed when a firm is in desperate circumstances, but the general importance of this kind of risk-taking behavior is under debate (see Parrino and Weisbach, 1999).

If both kinds of agency conflicts are at work, then their relative importance is unclear. One might imagine that they balance at an interior optimum as in the trade-off theory. However, the details of conflicting investment incentives can lead to complex problems, as suggested by Berkovitch and Kim (1990). Eventually, dynamic agency models such as Morellec (2004) and Atkeson and Cole (2005) and dynamic trade-off models such as those discussed in Section 2.3.2 are likely to go a long distance toward closing theoretical gaps between the various approaches to leverage. It seems likely to happen over the next few years.

## 3. Evidence

The available evidence on capital structure is organized in six parts, as follows: evidence on financing decisions at the aggregate level; a review of cross-sectional evidence on capital structure; an examination of evidence on leverage changes and a discussion of the tests of the pecking order theory and tests of mean reversion; evidence on capital structure changes from event studies; evidence from natural experiments; and the evidence from surveys of corporate managers.

### *3.1. Financing decisions at the aggregate level*

How has leverage changed at the aggregate level? How do firms finance imbalances between investments and internal cash flow? Do they issue debt or equity? Who holds debt and equity claims in the economy? Who are the major issuers and purchasers of debt and equity claims? The aggregate data help answer these questions. In addition, they provide an understanding of the differences between private and public firms since the U.S. Flow of Funds data include both. By contrast, much capital structure research examines publicly traded firms included in the Compustat files. Comparing the Compustat data to the U.S. Flow of Funds data reveals many similarities in capital structure decisions of private and public firms. But there are also important differences. However, the greatest differences appear when we examine the financing behavior of small and large public firms.[14]

### *3.1.1. Balance sheet*

The aggregate balance sheet data in Table 1 show the remarkable stability of leverage ratios over the last half century.[15] Debt neither vanishes from corporate accounts nor explodes to overwhelm equity. Aggregate leverage seems to be quite stationary. The evidence from before the 1950s is much sketchier, but what is known reinforces this basic sense of stability. Wright (2004) provides a useful compilation of data about the corporate sector from 1900 onward. Aggregate debt and aggregate equity both grow decade by decade. While leverage fluctuates during 1900–2002, it stays within rather narrow bounds. It is remarkable how similar leverage ratios are to each other in year 1900 and in year 2002. This is despite phenomenal changes in many features of the business environment during this period.

**Stylized Fact 1** *Over long periods of time, aggregate leverage is stationary.*

Similar persistence at the firm level is reported in Lemmon et al. (2007). The persistence in leverage ratios places important limits on theory. It means that a satisfactory theory must account for why firms keep leverage stationary. Or the theory must explain

[14] The aggregate series for the U.S. nonfarm nonfinancial sectors are taken from the Federal Reserve's Flow of Funds Statistics (Federal Reserve, 2003). The data cover the period from 1945 to 2002. The level data are taken from Table L.102, the balance sheet data are taken from Table B.102, and the flows from Table F.102. To examine public and private firms separately, an aggregated annual series for publicly traded firms is constructed first. The difference between the series for the entire economy and the series for the publicly traded firms provides the series for the private firms. The publicly traded firm series is constructed by aggregating Compustat-listed firms. Excluded are firms identified on the database as private and foreign-incorporated firms and firms with SIC codes less than 1000 and between 6000 and 7000.

[15] Interestingly, the asset structure shows significant changes, with large increases in the proportion of financial assets and decline in tangible assets. We observe that public firms hold considerably more tangible assets than do private firms. The large decline in tangible assets for the aggregate economy is primarily driven by private firms. We also find that, in recent years, public firms have noticeably higher book leverage than do private firms.

Table 1

Common-size balance sheets

This table presents average balance sheets for the aggregate U.S. Nonfarm Nonfinancial Corporate Business. The data is constructed using the Federal Flow of Funds (March 2003 release). The value of each balance-sheet item is calculated as a percentage of the replacement value of total assets and then averaged over available years in each decade.

| Balance-sheet item | 1945–1949 | 1950s | 1960s | 1970s | 1980 | 1990s | 2000+ |
|---|---|---|---|---|---|---|---|
| Tangible assets | 0.78 | 0.77 | 0.74 | 0.73 | 0.68 | 0.58 | 0.49 |
| Financial assets | 0.22 | 0.23 | 0.26 | 0.27 | 0.32 | 0.42 | 0.51 |
| Total assets | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Commercial paper | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 |
| Municipal securities | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 |
| Corporate bonds | 0.08 | 0.09 | 0.10 | 0.10 | 0.08 | 0.12 | 0.13 |
| Bank loans n.e.c. | 0.04 | 0.04 | 0.06 | 0.06 | 0.05 | 0.05 | 0.04 |
| Other loans/advances | 0.01 | 0.01 | 0.01 | 0.02 | 0.03 | 0.04 | 0.04 |
| Mortgages | 0.03 | 0.03 | 0.03 | 0.04 | 0.03 | 0.02 | 0.02 |
| Total debt | 0.16 | 0.17 | 0.21 | 0.22 | 0.21 | 0.25 | 0.25 |
| Trade payables | 0.06 | 0.07 | 0.09 | 0.08 | 0.06 | 0.07 | 0.08 |
| Taxes payable | 0.03 | 0.03 | 0.02 | 0.01 | 0.01 | 0.00 | 0.00 |
| Misc. liabilities | 0.01 | 0.01 | 0.02 | 0.06 | 0.15 | 0.18 | 0.20 |
| Liabilities | 0.26 | 0.28 | 0.34 | 0.37 | 0.43 | 0.51 | 0.53 |
| Net worth | 0.74 | 0.72 | 0.66 | 0.63 | 0.57 | 0.49 | 0.47 |
| Liabilities + net worth | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $\frac{Debt}{Debt + MV\ of\ equity}$ | 0.36 | 0.32 | 0.27 | 0.40 | 0.45 | 0.32 | 0.32 |

why the environment serves to maintain the leverage despite managerial indifference.[16] The market leverage ratios increased during the 1970s and the 1980s, perhaps caused by increases in the mergers and acquisitions and in leveraged buyout activity,[17] but they subsequently fell to the long-term average of about 0.32.

[16] For example, this poses problems for Welch (2004), who argues that changes in leverage are not undone. Frank and Goyal (2004) show that shocks in the equity market are cushioned by adjustments in the debt market in a manner that maintains the leverage ratio.

[17] In a trade-off theory, if taxes affect the mix between debt and equity, then it is possible that the Tax Reform Act of 1986 increased the attractiveness of debt during the 1980s; see Givoly et al. (1992).

**Stylized Fact 2** *Over the past half century, the aggregate market-based leverage ratio has been about 0.32. There have been surprisingly small fluctuations in this ratio from decade to decade.*

Following Myers (1984), it may seem that the stability of aggregate leverage is consistent with the trade-off theory. In fact, there is too much stability for the simple version of tax versus bankruptcy theory. For most of the 1950s and the 1960s, the top corporate tax rate was roughly 50% (see Taylor, 2003); in the 1990s, it was around 35%. Despite this large difference in tax rates, the market leverage ratio averaged 0.32 in both the 1950s and the 1990s, while, in the 1960s, it averaged 0.27. Have bankruptcy costs really fluctuated in just the right manner to account for this evidence? It seems difficult to imagine. This evidence, though not a proof, is certainly a serious warning sign for the trade-off theory.[18] The remarkable stability of leverage ratios also poses a problem for the pecking order theory. Leverage should fluctuate as the financing deficit ebbs and flows according to the standard pecking order theory. In order to account for this evidence, something must be added to the basic pecking order theory.

### *3.1.2. Cash flow statements*

Aggregate cash flow data (reported in Table 2 and plotted in Figure 1) show that dividends, capital expenditures, and net debt issues all fluctuate, but were rather stable over the last half of the twentieth century. The fact that aggregate dividends have not changed much contradicts some of the recent literature that finds declining dividends from U.S. firms. It is possible that the aggregate data mask a lot of heterogeneity in the dividend decisions of firms and that large increases in dividends by a certain sector of the economy offset increasing numbers of nondividend-paying firms.

Net debt issues finance a large part of the financing deficit. Equity issues are negative, and debt issues exceed the financing deficit during the last two decades, suggesting that firms issued debt to finance debt for equity swaps. While such swaps do take place, it is likely that debt-financed takeovers contribute more significantly to explaining these patterns.

Considerable heterogeneity exists between small public firms and large public firms and between private and public firms. Figure 2 plots the flow variables for *large public firms* (defined as firms whose book assets are in the top one-third of all publicly traded firms each year). For these firms, capital expenditures and internal funds are highly correlated. Their debt issues track financing deficits.[19]

[18] The recent dynamic models that are supportive of the trade-off theory tend to use relatively recent data. For example, Hennessy and Whited (2005) study data from 1993 to 2001.

[19] Faulkender and Petersen (2006) and Lemmon and Zender (2004) draw a distinction between firms with and without a credit rating. Those with a credit rating have easier access to public debt markets and thus use more debt financing. Empirically, it is likely that the large firms considered in Figure 2 are generally the firms with good credit ratings.

Table 2

Common-size statement of sources and use of funds

This table presents funds flow data for U.S. Nonfarm Nonfinancial Corporate Business. The value of each flow item is calculated as a percentage of the replacement value of total assets and then averaged over available years in each decade.

| Fund flow statement item | 1945–1949 | 1950–1959 | 1960–1969 | 1970–1979 | 1980–1989 | 1990–1999 | 2000–2002 |
|---|---|---|---|---|---|---|---|
| Profit before taxes | 0.075 | 0.066 | 0.059 | 0.044 | 0.026 | 0.030 | 0.019 |
| Taxes | 0.028 | 0.032 | 0.026 | 0.017 | 0.010 | 0.011 | 0.007 |
| Profit after taxes | 0.047 | 0.034 | 0.033 | 0.026 | 0.016 | 0.020 | 0.012 |
| Depreciation | 0.020 | 0.027 | 0.037 | 0.033 | 0.039 | 0.039 | 0.038 |
| Internal Funds-US op. | 0.067 | 0.061 | 0.069 | 0.059 | 0.055 | 0.059 | 0.050 |
| Foreign earnings ret. abroad | 0.001 | 0.001 | 0.002 | 0.003 | 0.003 | 0.005 | 0.006 |
| Inventory valuation adjustment | −0.009 | −0.002 | −0.001 | −0.006 | −0.002 | 0.000 | 0.000 |
| Net capital transfers | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Total internal funds | 0.059 | 0.060 | 0.070 | 0.056 | 0.056 | 0.064 | 0.055 |
| Financing deficit | | | | | | | |
| +Dividends | 0.015 | 0.014 | 0.015 | 0.010 | 0.009 | 0.014 | 0.014 |
| +Capital expenditures | 0.050 | 0.050 | 0.058 | 0.055 | 0.05 | 0.051 | 0.045 |
| +Change in working capital | 0.002 | 0.008 | 0.008 | 0.009 | 0.003 | 0.010 | 0.007 |
| +Discrepancy | 0.009 | 0.004 | 0.008 | 0.006 | 0.008 | −0.003 | −0.003 |
| −Internal Funds | 0.059 | 0.060 | 0.070 | 0.056 | 0.056 | 0.064 | 0.055 |
| Financing deficit | 0.017 | 0.016 | 0.019 | 0.023 | 0.014 | 0.009 | 0.009 |

*(Continued)*

Table 2 *(Continued)*

| Fund flow statement item | 1945–1949 | 1950–1959 | 1960–1969 | 1970–1979 | 1980–1989 | 1990–1999 | 2000–2002 |
|---|---|---|---|---|---|---|---|
| Net equity issues | 0.003 | 0.003 | 0.001 | 0.003 | −0.007 | −0.004 | −0.004 |
| Net debt issues | 0.014 | 0.012 | 0.017 | 0.021 | 0.021 | 0.013 | 0.012 |
| Net funds issued | 0.017 | 0.016 | 0.019 | 0.023 | 0.014 | 0.009 | 0.009 |
| Sources of debt financing | | | | | | | |
| Commercial paper issues | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 | 0.001 | −0.002 |
| Municipal securities issues | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 | 0.000 | 0.000 |
| Corporate bonds issues | 0.007 | 0.006 | 0.007 | 0.008 | 0.008 | 0.009 | 0.011 |
| Bank loans increase | 0.004 | 0.004 | 0.006 | 0.005 | 0.005 | 0.002 | −0.001 |
| Other loans increase | 0.000 | 0.001 | 0.002 | 0.003 | 0.004 | 0.001 | 0.002 |
| Mortgages issued | 0.002 | 0.001 | 0.002 | 0.004 | 0.001 | 0.000 | 0.002 |

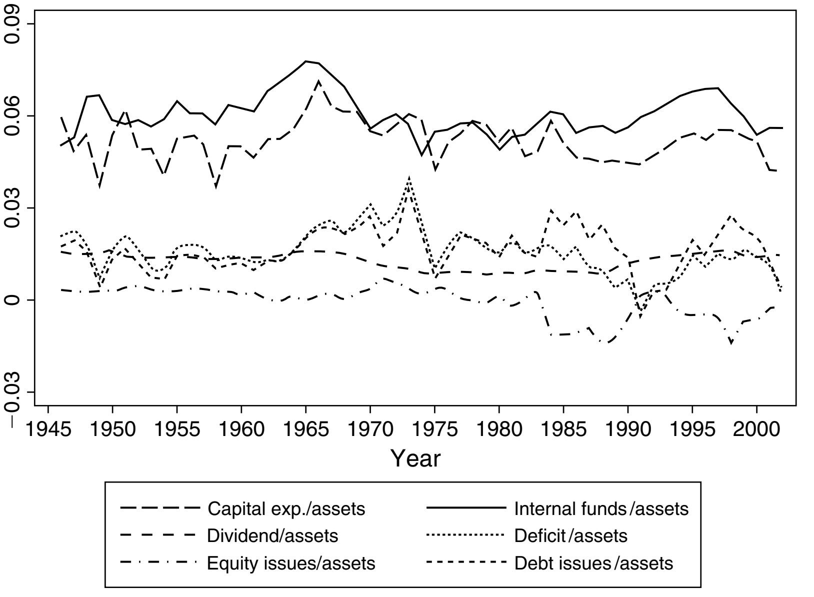



Fig. 1. **Aggregate Federal Flow of Funds Data:** Aggregate data from Funds Flow Statements (March 2003) release is used to construct capital expenditure to assets, internal funds to assets, dividend to assets, deficit to assets, net equity issued to assets, and net debt issued to assets, 1945 to 2002. The data are for the nonfarm nonfinancial corporate sector of the U.S. economy. The deficit is calculated as cash dividends plus investments plus change in working capital plus discrepancy minus internal funds.

The flow variables for *small public firms* are shown in Figure 3. Small public firms are defined as firms in the bottom one-third in terms of book assets among publicly traded firms each year. For these firms, capital expenditures exceed internal funds, and the net equity issuances and the financing deficit appear to be strongly correlated. Debt issues are fairly minor. Thus, important differences exist between large and small firms in how they finance their deficits. Large firms issue debt to finance deficits, while small firms issue equity.

Figure 4 plots the data for *private firms*. Capital expenditures tracked internal funds up until the mid-1980s, but since then internal funds have exceeded capital expenditures. Debt issues and deficits show a closer relation than do equity issues and deficits.

Several other differences between public and private firms are worth noting. Public firms are more profitable, invest more, and use more external financing (particularly equity). Private firms seem to have been increasing their dividends over time.[20]

[20] This is somewhat puzzling. The aggregate Flow of Funds data say that dividend ratios have not changed. Evidence from large public firms suggests some decline, as observed by Fama and French (2001). Since small public firms pay almost no dividends, adding up the evidence seems to imply that private firms must be making up the difference through increased dividends. It would be nice to have direct evidence of such a change.

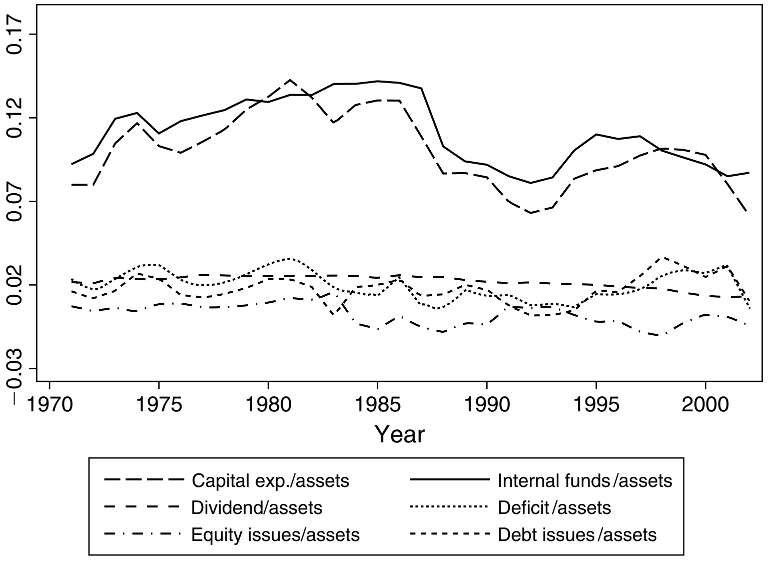



Fig. 2. **Large Public Firms from Compustat:** Average capital expenditures to assets, internal funds to assets, dividend to assets, deficits to assets, net equity issued to assets, and net debt issued to assets, 1971–2002. The sample comprises large publicly traded U.S. firms on the Compustat files (in the top one-third by book assets each year). Financial firms and regulated utilities are excluded. The deficit is calculated as cash dividends plus investments plus change in working capital minus internal cash flow. Net debt issued is long-term debt issuance minus long-term debt redemption. Net equity issued is the issue of stock minus the repurchase of stock. The variables are constructed using data from Compustat funds flow statements.

Figures 1 to 4 clearly illustrate that equity financing is more important for small public firms than it is for either private firms or large public firms. Presumably for many small public firms, the desire to issue equity easily induced them to go public in the first place.

**Stylized Fact 3** *At the aggregate level, capital expenditures are very close to internal funds. This is true for large public firms and private firms; this is not true for small public firms.*

**Stylized Fact 4** *At the aggregate level, the financing deficit is very close to debt issues. This holds for large public firms and for private firms; this does not hold for small public firms. For small public firms, financing deficits very closely match equity issues.*

**Stylized Fact 5** *Aggregate dividends are very smooth and almost flat as a fraction of total assets for all classes of firms. There has been remarkable stability in the aggregate*

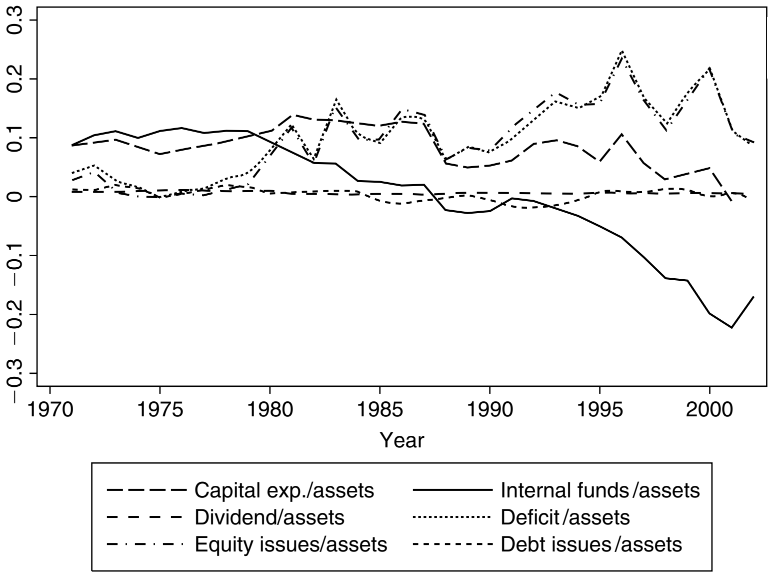



Fig. 3. **Small Public Firms from Compustat:** Average capital expenditure to assets, internal funds to assets, dividend to assets, deficit to assets, net equity issued to assets, and net debt issued to assets, 1971–2002. The sample comprises small publicly traded U.S. firms on the Compustat files (in the bottom one-third by book assets each year). Financial firms and regulated utilities are excluded. The deficit is calculated as cash dividends plus investments plus change in working capital minus internal cash flow. Net debt issued is long-term debt issuance minus long-term debt redemption. Net equity issued is the issue of stock minus the repurchase of stock. The variables are constructed using data from Compustat funds flow statements.

*dividend rate over time. Large public firms pay higher dividends than do small public firms. Many small firms pay no dividends.*

The evidence from before the 1950s is, as usual, quite a bit sketchier than is the subsequent evidence. Taggart (1985) suggests that stock issues were more important before World War II relative to those during the 1960s and 1970s. He also reports that, during the 1930s, aggregate total corporate debt declined, as internal funds exceeded the uses of funds. Presumably, this reflects the struggle of firms to stay afloat during the Depression. Graham and Narasimhan (2004) provide an interesting study of the Depression period.

### *3.1.3. Holdings of corporate financial claims*

According to the Fisher Separation Theorem (Hirshleifer, 1958), with complete markets, investors are unanimous about how the firm should be run. It does not matter who provides the firms with funding. However, when markets are incomplete, it is well known that

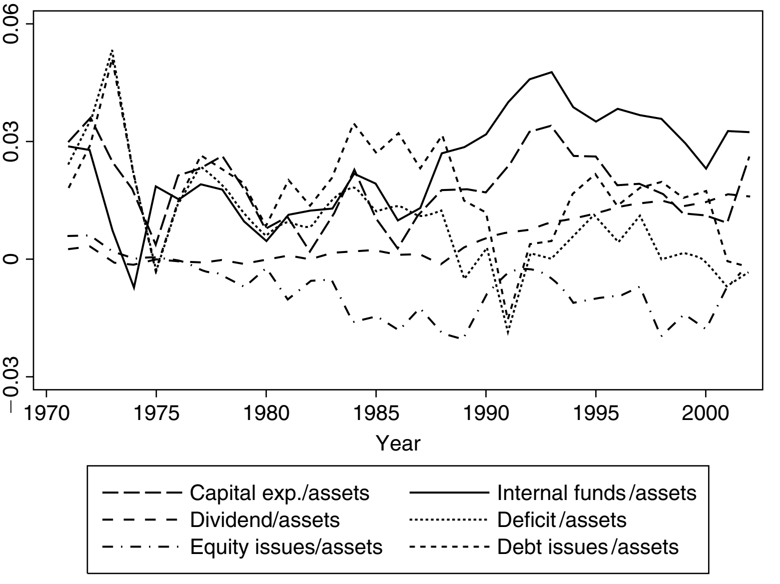



Fig. 4. **Private sector (Difference Series):** Private-sector series is the difference series between aggregate values for the nonfarm nonfinancial corporate sector for the U.S. economy from the Funds Flow statements and nonfarm nonfinancial publicly traded sample from the Compustat. The difference series is used to compute capital expenditure to assets, internal funds to assets, dividend to assets, deficit to assets, net equity issued to assets, and net debt issued to assets, 1971–2002.

differences of opinion can matter and investor clientele effects may be important.[21] Accordingly, it is of interest to examine what we know about the providers of funds to different sectors of the economy.

Corporations raise funds from the rest of the economy. These funds come in the form of equity and debt. By definition, all debt and equity are owned either directly or indirectly. Indirect ownership happens through a variety of institutional forms, including banks, insurance companies, pension funds, and mutual funds. The markets for corporate debt and equity must reconcile investor demands with the willingness of firms to supply debt or equity. In the Flow of Funds data, debt and equity claims can be viewed as being issued by three major sectors of the economy: nonfinancial U.S. corporations, U.S. financial firms, and the rest of the world. The claims are purchased and held by six major sectors: households, governments, the rest of the world, banks, insurance companies, and funds.

Table 3 provides aggregate data on the issuers of and investors in bonds and equity. First consider the bond market. Five major sectors held most of the debt issued by the

[21] The idea that investor clienteles might play an important role in capital structure goes back at least to Schwartz (1959). He suggested thinking about optimal capital structure in terms of monopsonistic discrimination against outside investors. The focus of the profession at this time was on making sense of Modigliani and Miller (1958), and Schwartz's paper did not receive much attention.

Table 3

Levels of Securities Outstanding by Sector

This table presents aggregate debt and equity issued and held by different sectors of the economy. Bonds issued and held are reported as a fraction of total bonds outstanding. Equity issued and held is reported as a fraction of total equity outstanding.

| | 1945–1949 | 1950–1959 | 1960–1969 | 1970–1979 | 1980–1989 | 1990–1999 | 2000–2002 |
|---|---|---|---|---|---|---|---|
| Bond Issued by | | | | | | | |
| Corporate | 0.896 | 0.884 | 0.809 | 0.762 | 0.665 | 0.501 | 0.443 |
| Rest of world | 0.091 | 0.061 | 0.071 | 0.080 | 0.083 | 0.095 | 0.087 |
| Financial firms | 0.014 | 0.055 | 0.120 | 0.158 | 0.253 | 0.403 | 0.470 |
| Bond Holdings by | | | | | | | |
| Households | 0.229 | 0.107 | 0.102 | 0.153 | 0.070 | 0.145 | 0.137 |
| Government | 0.000 | 0.000 | 0.000 | 0.000 | 0.005 | 0.013 | 0.015 |
| Rest of world | 0.010 | 0.007 | 0.008 | 0.026 | 0.123 | 0.132 | 0.196 |
| Banks | 0.126 | 0.085 | 0.048 | 0.109 | 0.121 | 0.087 | 0.085 |
| Insurance firms | 0.610 | 0.774 | 0.807 | 0.683 | 0.632 | 0.502 | 0.396 |
| Funds | 0.026 | 0.027 | 0.034 | 0.030 | 0.048 | 0.122 | 0.172 |
| Equity Issued by | | | | | | | |
| Corporate | 0.885 | 0.876 | 0.845 | 0.876 | 0.860 | 0.775 | 0.698 |
| Rest of world | 0.008 | 0.008 | 0.008 | 0.010 | 0.024 | 0.083 | 0.106 |
| Financial firms | 0.107 | 0.116 | 0.147 | 0.114 | 0.116 | 0.141 | 0.196 |

*(Continued)*

Table 3 *(Continued)*

| | 1945–1949 | 1950–1959 | 1960–1969 | 1970–1979 | 1980–1989 | 1990–1999 | 2000–2002 |
|---|---|---|---|---|---|---|---|
| Equity Holdings by | | | | | | | |
| Households | 0.917 | 0.885 | 0.824 | 0.615 | 0.514 | 0.490 | 0.392 |
| Government | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 | 0.008 |
| Rest of world | 0.023 | 0.022 | 0.024 | 0.037 | 0.062 | 0.068 | 0.103 |
| Banks | 0.002 | 0.003 | 0.014 | 0.112 | 0.076 | 0.034 | 0.021 |
| Insurance firms | 0.032 | 0.052 | 0.087 | 0.188 | 0.293 | 0.280 | 0.277 |
| Funds | 0.027 | 0.039 | 0.051 | 0.048 | 0.055 | 0.124 | 0.199 |
| Loans to Corporate/ Total bank loans | 0.747 | 0.674 | 0.632 | 0.563 | 0.614 | 0.621 | 0.575 |

business sector in the 1990s: households (14% of aggregate debt), insurance companies (40%), the rest of the world (20%), pension and mutual funds (17%) and banks (9%). Over time, insurance companies have dramatically reduced their holdings of corporate debt, while the pension and mutual funds and the rest of the world have increased their holdings of corporate debt.

Bond issuers have also changed. The U.S. nonfinancial corporate sector issued almost 90% of outstanding debt in the 1950s. The financial sector played a minor role, and its outstanding debt as a fraction of total debt was only about 1%. But, by the 1990s, debt issued by the U.S. nonfinancial corporate sector declined to only about 44%; the financial sector debt now exceeds the debt from the nonfinancial corporate sector.

The equity issued by financial firms has also grown relative to that issued by nonfinancial corporate firms. Growth in equity issued in the United States by the rest of the world is even more impressive. In the 1950s, equity issued by the rest of the world as a fraction of total outstanding equity was less than 1%; by the 1990s, this number had increased to about 10%. About 70% of equity in the 1990s was issued by the nonfinancial corporate sector and the financial sector issued another 20%.

The direct holdings of equity by households have declined sharply. Equity ownership by insurance companies and by funds has increased, as has foreign ownership of U.S. equity. In the 1990s, corporate equity was held heavily by households (39% of the aggregate equity outstanding), pension and mutual funds (20%), insurance firms (28%), and the rest of the world (10%). Banks and governments do not hold much corporate equity.

**Stylized Fact 6** *Over the past half century, there has been a large decrease in direct holding of corporate securities by households and a corresponding huge increase in financial intermediation of such claims.*

This growth in intermediation may be important. It is possible that in incomplete markets, mutual funds and insurance firms have different views about the appropriate rate at which they discount the future. Mutual funds may be more interested in the short term, while insurance firms may be more interested in the long term. Since 1980, the importance of insurance firms as financial intermediaries has been relatively constant, but the importance of mutual funds increased explosively over that period.[22]

The change in intermediation is also important with respect to taxation. Many intermediaries pay no tax on dividends or capital gains. Depending on the type of account held by the ultimate owner, there may even be no applicable personal tax to pay. This reduces the effect of taxes on capital structures (see McDonald, 2004). If taxes were the

[22] There is also some evidence that the traditional role of the bank loan has been changing, at least for large firms. In many cases, banks seem to have resold the loans in a secondary market. Sufi (2007) reports that in recent years more than 15% of nonfinancial U.S. corporate debt has come by means of a loan syndicate rather than from a single source.

full story, then all equity ought to flow into tax-advantaged accounts. Although there has been a significant flow in that direction, the flow has not resulted in all equity being held in tax-advantaged accounts.

#### *3.1.4. Flows of corporate financial claims*

Another perspective on capital structure decisions comes from examining the aggregate flows of corporate financial claims. Table 4 reports debt and equity issuances and purchases by various sectors of the economy. Consistent with the increasing share of financial sector debt, financial firms have become significant issuers of corporate debt. By the 1990s, financial sector debt issuances exceeded those by the nonfinancial corporate sector. The debt issuances by the rest of world have declined over time. Corporate debt is bought heavily by insurance firms, funds, and the rest of the world.

The net equity issuances by the U.S. nonfinancial corporate sector have been negative since the 1980s. The net issuances are defined as new equity issues less repurchases less cash-financed takeovers. Both stock repurchases and cash-financed takeovers became more important in the 1980s and in the 1990s. The negative net equity issuances imply that new equity issues by U.S. firms, together with positive net equity issuances by the rest of the world, have not been enough to offset aggregate repurchases and takeover distributions. At times, cash-financed takeover distributions have been more important than share repurchases (Wright, 2004; Holmstrom and Kalplan, 2001). The connection between mergers and acquisitions (M&A) activity and leverage deserves more attention since M&A activity is an important method by which firms exit. According to Maksimovic and Phillips (1998), it avoids at least the direct costs of bankruptcy.

Corporate equity is bought mostly by the rest of the world and mutual funds. Households have been the net sellers of corporate equity since the 1960s. Households get equity in several ways: They get it through entrepreneurship when they create a firm that did not previously exist. They get it when retained earnings increase the value of their existing equity holdings. They get it as compensation for labor, for example, through employee stock ownership plans, stock options, and stock grants.

**Stylized Fact 7** *Households have been net suppliers of corporate equity since the 1960s. Corporations have been net buyers of equity since the 1980s. Most equity is no longer held directly. Insurance companies, mutual funds, and pension funds now hold more direct equity and debt than do households.*

### *3.2. Leverage differences between firms*

Cross-sectional tests of capital structure theories examine whether debt ratios vary across firms as predicted by the theory. Two strands can be distinguished. The first strand, the bulk of this literature, is concerned with determining which factors are correlated with leverage. This literature is fairly extensive and includes contributions by Bradley et al. (1984), Long and Malitz (1985), Titman and Wessels (1988), Crutchley and Hansen

Table 4
Flows of Securities by Sector
Bond issues and purchases are divided by lagged bonds outstanding. Equity issues and purchases are divided by lagged equity outstanding.

| | 1945–1949 | 1950–1959 | 1960–1969 | 1970–1979 | 1980–1989 | 1990–1999 | 2000–2002 |
|---|---|---|---|---|---|---|---|
| Total Bond Issuance/ Total bonds outstanding | 0.096 | 0.081 | 0.079 | 0.101 | 0.127 | 0.114 | 0.102 |
| Bonds Issued by | | | | | | | |
| Corporates | 0.090 | 0.068 | 0.059 | 0.068 | 0.074 | 0.040 | 0.042 |
| Rest of world | 0.000 | 0.004 | 0.007 | 0.011 | 0.006 | 0.016 | 0.000 |
| Financial firms | 0.007 | 0.010 | 0.013 | 0.022 | 0.046 | 0.059 | 0.06 |
| Bonds Purchased by | | | | | | | |
| Households | −0.018 | 0.001 | 0.013 | 0.017 | 0.005 | 0.024 | 0.006 |
| Government | 0.000 | 0.000 | 0.000 | 0.000 | 0.002 | 0.002 | 0.001 |
| Rest of world | −0.001 | 0.000 | 0.001 | 0.003 | 0.024 | 0.018 | 0.036 |
| Banks | 0.009 | 0.001 | 0.003 | 0.013 | 0.018 | 0.005 | 0.009 |
| Insurance firms | 0.106 | 0.077 | 0.059 | 0.065 | 0.067 | 0.042 | 0.028 |
| Funds | 0.000 | 0.002 | 0.003 | 0.002 | 0.011 | 0.024 | 0.023 |
| Total Equity Issuance/ Total equity outstanding | 0.012 | 0.012 | 0.004 | 0.009 | −0.018 | 0.004 | 0.003 |

*(Continued)*

Table 4 *(Continued)*

| | 1945–1949 | 1950–1959 | 1960–1969 | 1970–1979 | 1980–1989 | 1990–1999 | 2000–2002 |
|---|---|---|---|---|---|---|---|
| Equity Issues by | | | | | | | |
| Corporates | 0.010 | 0.010 | 0.002 | 0.006 | −0.023 | −0.005 | −0.004 |
| Rest of world | 0.000 | 0.001 | 0.002 | 0.000 | 0.001 | 0.008 | 0.004 |
| Financial firms | 0.002 | 0.002 | 0.000 | 0.002 | 0.004 | 0.002 | 0.003 |
| Equity Purchased by | | | | | | | |
| Households | 0.007 | 0.003 | −0.007 | −0.007 | −0.034 | −0.017 | −0.015 |
| Government | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 |
| Rest of world | −0.001 | 0.001 | 0.001 | 0.002 | 0.003 | 0.002 | 0.007 |
| Banks | 0.000 | 0.000 | 0.001 | 0.000 | −0.004 | −0.003 | −0.001 |
| Insurance firms | 0.003 | 0.005 | 0.007 | 0.016 | 0.014 | 0.006 | 0.002 |
| Funds | 0.003 | 0.002 | 0.002 | −0.002 | 0.003 | 0.014 | 0.009 |

(1989), Smith and Watts (1992), Rajan and Zingales (1995) and Frank and Goyal (2007a). The survey by Harris and Raviv (1991) has also been influential. A second strand is concerned with the debt conservatism puzzle. This puzzle lies in the allegation that many (or all) firms have lower leverage than would maximize firm value from a static trade-off perspective. Contributions include Miller (1977), Graham (2000) and Ju et al. (2005).

### *3.2.1. Leverage definition and other econometric issues*

In testing which factors are correlated with leverage, it is necessary to define leverage. Many different empirical definitions have been used. A key issue has been whether to examine book leverage (debt divided by total assets) or market leverage (debt divided by the sum of book debt plus the market value of equity). Early empirical work tended to focus on book leverage. Myers (1977) argued that managers focus on book leverage because debt is better supported by assets in place than it is by growth opportunities. Book leverage is also preferred because financial markets fluctuate a great deal and managers are said to believe that market leverage numbers are unreliable as a guide to corporate financial policy.

The subsequent literature has given more attention to a market-based measure of leverage. Welch (2004), for example, argues that the book value of equity is primarily a "plug number" that is used to balance the left-hand side and the right-hand side of the balance sheet rather than a managerially relevant number. He further objects that it can even be negative.[23]

Conceptually, book and market leverage ratios are different. The book measure looks backward, measuring what has taken place. In contrast, markets are generally assumed to be forward looking. Thus, there is no reason these two concepts should match (see Barclay et al., 2006). As Table 1 shows, we do see a difference. Market-based leverage has been rather more stable over the decades than has book leverage.

Cross-sectional leverage studies must also confront other empirical issues. One issue is how to deal with the panel structure of the data. Typically, studies examine a large number of firms over a limited number of years. Thus, one has a panel in which the errors are unlikely to be independent. How should one adjust for this lack of independence? Different papers handle the problem in different ways. The panel data textbook by Baltagi (2001) provides particularly helpful coverage of this problem. Petersen (2007) presents a useful discussion that is more directly tied to corporate finance applications. In practice, it is relatively common for studies to try more than one method of correction and then only focus on results that are robust across methods.

[23] Welch (2004) seems to prefer the use of interest coverage ratios. Frank and Goyal (2007a) show that, as an empirical issue, interest coverage ratios are not attractive. They lead to empirically fragile results. A focus on such fragile results might then serve to obscure the robust evidence that is obtained with the more popular measures.

A second issue is how to deal with incomplete data in the panel. Many firms have only incomplete records. Typical studies drop firms that lack the necessary data items. Current versions of the standard econometric packages encourage this practice since they do it more or less automatically. The user might be almost unaware that firms are dropped. However, this practice has its drawbacks. It creates a bias if the missing data are related to the process being studied. To get around this problem, statisticians such as Little and Rubin (2002) often recommend a process of "multiple imputation."[24] Essentially, the observed data are used to make a best guess about the value of the unobserved data. This is done several times to reflect the fact that there is uncertainty about the imputed values. Such a procedure can help mitigate the bias. Frank and Goyal (2007a) found that the main leverage factors discussed in this section are robust to whether or not one carries out multiple imputation. However, many of the minor factors are not robust in this respect.

A third common problem for such studies is how to deal with outliers. The standard data sources such as Compustat have a nontrivial number of observations that seem quite anomalous. For instance, data items that by definition cannot be negative are sometimes coded as negative. Sometimes data items are coded in ways that result in the balance sheet not balancing or the cash flow identities not matching up. In some cases, a firm will have a value of some variable that is several orders of magnitude too large to be plausibly correct.

These "outliers" happen in too many cases for it to be practical to chase down and correct them from the original sources such as company annual reports. So how should we deal with this problem? Several approaches have been used. Some studies simply ignore the issue. This is potentially serious since we know that outliers can generate seriously misleading conclusions. Most studies take some steps to deal with the problem. Three kinds of corrections are particularly common: rule of thumb truncations, winsorization, and robust regressions.[25]

The most common method of dealing with the problem is to use some rule of thumb to remove data that is so extreme that "it cannot possibly be correct." Different studies employ different rules of thumb, so that one study might remove firms with a market-to-book ratio that is reported to exceed 5, while another study might remove only those that exceed 10. Many studies include some kind of minimum firm-size criterion. When a study uses a variety of such plausible, but more or less arbitrary rules of thumb, it is difficult to be sure exactly how sensitive the results are to these truncations. In particular, multiple truncation rules might interact in surprising ways.

[24] Currently, SAS, S-plus, and R all provide precoded routines to carry out multiple imputation. Some users have coded Stata routines to carry out multiple imputation as well, and these are easy to obtain. Multiple imputation is not yet common in corporate finance, but it seems quite likely that in the next few years this will become standard practice.

[25] Data problems are much more common for small firms than they are for large firms. Some studies restrict attention to large firms, and thus they largely avoid the main data problems. However, the results are then conditional on the firm size filter. It is disappointingly common for papers to use such a filter but then ignore the fact that the data have been filtered when discussing the results. As shown earlier, there are important differences in how small and large firms finance themselves.

Recently, it has become more common to use winsorization, in which the most extreme tails of the distribution are replaced by the most extreme value that has not been removed. It is particularly common to winsorize each tail at 0.5% or 1%. In essence, this procedure amounts to saying: "I do not believe the data are correct, but I know that the data exist. So instead of completely ignoring the data item, I will replace it with something a bit more reasonable." This procedure has the advantage that it is more systematic than pure rules of thumb, and it is then easier to have consistency across papers. This kind of approach can be viewed as some type of a "poor man's" Bayesian method. A prior is being imposed, but the full Bayesian machinery is not being used.

Another fairly common method is to run robust regressions. Most statistical packages currently include one or more types of robust procedures. The statistics literature contains quite a few alternative robust procedures.[26] It is fairly common for empirical papers to use some type of robust regression procedure. Typically, the results are not reported as the main results. Instead, they tend to be relegated to footnotes.

A final issue concerns the assumptions to be made about the debt market. It is now common for papers to study the ratio of book debt to the sum of book debt plus market equity. This is often accompanied by an apology to the reader. Book debt is being studied due to the inconvenient fact that a large number of firms do not have market-traded debt. Thus, only book debt numbers are available. Of course, this fact itself is informative and deserves attention. Why do so many firms have traded equity, while only relatively large firms have traded debt? Why is it that so much corporate equity is traded on organized exchanges, while so little corporate debt is exchange trade? These are not merely nuisance issues for empiricists. They are first-order facts about corporate debt that deserve attention.[27]

### *3.2.2. Leverage factors*

The capital structure literature identifies a large number of cross-sectional variables that appear related to debt ratios. However, Frank an Goyal (2007a) show that only a small number of factors are actually empirically robust and financially significant. This section summarizes the predictions and evidence for factors that exhibit the most robust correlation with leverage.

*3.2.2.1. Leverage and growth* The static trade-off theory predicts a negative relation between leverage and growth. Growth firms lose more of their value when they go into distress. Several agency theories also predict a negative relation between leverage and growth. For example, the underinvestment problem is more severe for growth firms, leading these firms to prefer less debt. The underinvestment problem arises because firms

[26] The popular package Stata includes the command "rreg," which provides basic robust regression functionality. S-plus and R both provide much more complete sets of robust procedures.

[27] It does not seem too hard to imagine that the high fixed costs of entering public debt markets may play a role.

with risky debt have an incentive to underinvest in positive net present value projects since shareholders bear the entire cost of the project but receive only a fraction of the increase in firm value; part of it goes to debtholders (see Myers, 1977).

As growth options increase, asset substitution problems also become more severe. In high-growth firms, it is easier for stockholders to increase project risk, and it is harder for debtholders to detect such changes. Thus, debt is more costly for firms with high-growth opportunities. Agency costs of free cash flow are less severe for growth firms (see Jensen, 1986), and this also leads to the prediction that high-growth firms should have less debt. Debt mitigates agency costs of free cash flow when firms have fewer growth opportunities. The discipline that debt provides is less valuable for firms with good growth opportunities.[28] In summary, both the tax-bankruptcy cost trade-off and the agency theories are generally interpreted as predicting an inverse relation between the leverage ratio and growth opportunities.

By contrast, the pecking order theory predicts that firms with more investments—holding profitability fixed—should accumulate more debt over time. Thus, according to the pecking order theory, growth opportunities and leverage are expected to be positively related.

The relation between leverage and growth features in many different cross-sectional studies, including those by Bradley et al. (1984), Long and Malitz (1985), Smith and Watts (1992), Barclay et al. (1995), Barclay et al. (2006), and Frank and Goyal (2007a). The ratio of market value of assets to book value of assets is a commonly used proxy for growth opportunities. The studies generally conclude that leverage is negatively related to market-to-book ratios, which is consistent with trade-off theories. Rajan and Zingales (1995) show that the negative relation between leverage and market-to-book ratios exists in all G7 countries.

*3.2.2.2. Leverage and firm size* Static trade-off theory is generally interpreted as predicting that large firms will have more debt since larger firms are more diversified and have lower default risk. Larger firms are also typically more mature firms. These firms have a reputation in debt markets and consequently face lower agency costs of debt. Hence, the trade-off theory predicts that leverage and firm size should be positively related.

The pecking order theory is usually interpreted as predicting an inverse relation between leverage and firm size. The argument is that large firms have been around longer and are better known. Thus, large firms face lower adverse selection and can more easily issue equity compared with small firms where adverse selection problems are severe. There is an important caveat here. Large firms also have more assets and thus the adverse selection may be more important if it impinges on a larger base. Thus, the pecking order prediction for firm size is ambiguous. Cross-sectional tests of the relation between leverage and firm size find the relation to be robustly positive.

[28] The board of directors may provide a more direct method of dealing with this problem as in Yen (2005).

*3.2.2.3. Leverage and tangibility of assets* Tangibility of assets is typically measured by the ratio of fixed assets to total assets. Some studies construct a measure of collateralizable assets measured as the ratio of inventory plus net property plant and equipment to total assets and find a positive relation between leverage and the extent to which a firm's assets are collateralizable.

Tangible assets are easier to collateralize, and they suffer a smaller loss of value when firms go into distress. Thus, from a trade-off perspective, tangibility has an important effect on the costs of financial distress. In addition, tangibility makes it difficult for shareholders to substitute high-risk assets for low-risk ones. Agency costs of debt are therefore lower for firms with more tangible assets. Both the static trade-off and agency theories predict a positive relation between leverage and tangibility of assets. Under the pecking order, Harris and Raviv (1991) argue that the low information asymmetry associated with tangible assets makes equity less costly, resulting in a negative relation between leverage and tangibility.

The relation between debt and tangibility of assets is reliably positive. Inventory is sometimes included and sometimes excluded in measures of tangibility. Empirically, inventory seems to help to explain the use of short-term debt much more than it helps to explain the use of long-term debt.[29]

*3.2.2.4. Leverage and profitability* Static trade-off theory predicts that profitable firms should have more debt. Expected bankruptcy costs are lower, and interest tax shields are more valuable for profitable firms. Furthermore, firms that generate higher profits relative to investments benefit from the discipline that debt provides in mitigating the free cash flow problem (Jensen, 1986). The pecking order theory argues that firms prefer internal finance over external funds. Thus, according to the pecking order theory, with investments and dividends fixed, more profitable firms should become less levered over time.

The empirical studies typically find a negative relation between profitability and leverage. This negative relation is traditionally interpreted as being consistent with the pecking order theory and inconsistent with the trade-off theory. However, the theory is not quite so simple. Profitability also proxies for growth opportunities. If profitability is a less noisy proxy for growth than the market-to-book ratio, the negative sign on profitability is consistent with the predictions of the static trade-off theory. More importantly, as discussed in Section 2.3.2, in a dynamic model, the traditional interpretation might not be valid. Indeed, both of the examples in that section illustrate that there are good reasons in a trade-off model for leverage to be negatively related to leverage. Thus, the trade-off theory predictions on profit are ambiguous.[30]

[29] This has sometimes led to ideas of maturity matching. We do not pursue this idea here because to do so would require a more complete treatment of the various clauses and contingencies on debt contracts. Such studies are worthwhile, but they do not speak directly to Myers's contest.

[30] Chen and Zhao (2005) argue that neither transaction costs nor tax reasons can properly explain the negative relation between leverage and profitability. Thus, the best way to think about the relation between leverage and profits is not yet entirely settled.

*3.2.2.5. Leverage and industry median debt ratios* Industry leverage is a powerful predictor of firm leverage. Presumably, at least from a trade-off perspective, much of the power comes from the fact that industry leverage reflects a number of otherwise omitted common factors (Frank and Goyal, 2007a). Product market interactions are also important. As a result, the industry median leverage is likely to be a proxy for the target capital structure, albeit a noisy one. Hovakimian et al. (2001) find that firms adjust their debt ratios toward industry median debt ratios. Mackay and Phillips (2005) provide a recent analysis of industry effects on leverage and show that there is significantly more variation in leverage within industries than across industries.

*3.2.2.6. Leverage and expected inflation* Taggart (1985) argues that features of the tax code suggest a positive relation between debt and expected inflation. The real value of tax deductions on debt is higher when inflation is expected to be high. Thus, the trade-off theory suggests a positive relationship between leverage and expected inflation. A positive relation can also arise if managers time their debt. If managers are timing, then they will issue debt when expected inflation is high relative to current interest rates. Compared with the 1970s and the early 1980s, in recent years, inflation has not figured prominently in the academic literature on capital structure. However, the effects continue to be present in the data. Frank and Goyal (2007a) show that there is a robust positive relation between leverage and expected inflation.

Our interpretation of the evidence from cross-sectional tests of capital structure is summarized as follows.

**Stylized Fact 8** *There is a core set of six reliable factors that are correlated with cross-sectional differences in leverage. Leverage is positively related to median industry leverage, tangibility, log of assets, and expected inflation. Leverage is negatively related to market-to-book and profits.*

Over time, we have acquired a better understanding of the factors that are empirically related to leverage. These reliable factors together explain about 30% of the cross-sectional variation in leverage ratios. Lemmon et al. (2007) question the relative importance of traditional leverage factors when compared to fixed effects. They show that a firm's fixed effects explain almost 60% of the variation in the leverage ratio. However, it is not clear what lies behind the unidentified components of leverage. Interpreting the evidence has also remained difficult. Many variables could reasonably be interpreted as representing different theories of capital structure. Moreover, empirical specifications are linear even when some models contain nonlinearities. Many tests are static, even though the data are generated by the dynamics of the firm's financing decisions. Thus, cross-sectional variation in debt ratios may arise because either optimal ratios differ or the actual ratios diverge from optimal ones.

### *3.2.3. Debt conservatism*

Since at least Miller (1977), there has been some concern about the seemingly low leverage of firms given the substantial tax benefits of debt. Miller argued that bankruptcy costs appear to be too small to offset the large tax subsidies of debt: "the great emphasis on bankruptcy costs in recent discussions of optimal capital structure policy seems to me to have been misplaced. . . . the supposed trade-off between tax gains and bankruptcy costs looks suspiciously like the recipe for the fabled horse-and-rabbit stew—one horse and one rabbit" (Miller, 1977, p. 264). Dynamic trade-off models of Kane et al. (1984) and Brennan and Schwartz (1984) considerably strengthened the idea that firms are underlevered relative to the predictions of the trade-off theory.

A number of studies have attempted to quantify bankruptcy costs. Direct bankruptcy costs are indeed small (see Warner, 1977). Maksimovic and Phillips (1998) find that assets are often reshuffled between firms, and so direct bankruptcy costs may not be very high. Indirect bankruptcy costs (Titman, 1984) are likely to be much larger, but they have been difficult to quantify. A recent attempt to estimate bankruptcy costs by Andrade and Kaplan (1998) finds that, for a sample of 31 highly leveraged transactions, bankruptcy costs are between 10 and 23% of firm value.

From a different point of view, Molina (2005) observes that many estimates of default costs, such as that by Warner (1977), are ex-post estimates. Default is endogenous to the leverage decision. Molina uses the firms' past market valuations and marginal tax rates as instruments to estimate the effect of increasing leverage on the default probability. Ex-ante costs of financial distress can be obtained by multiplying this estimated effect of leverage on the firms' default probability with the previous estimates of ex-post costs of financial distress. Molina finds that the ex-ante costs of financial distress are comparable to the current estimates of the tax benefits of debt.

Graham (2000) estimates tax-rate functions to determine how aggressively firms use debt. He finds that a significant number of Compustat firms are surprisingly conservative in their use of debt. These are generally large, profitable, and liquid firms, the very firms that are expected to face lower costs of distress. They could have levered more. Graham concludes that the capital structures of a significant number of U.S. publicly traded firms are leaving significant sums of money on the table.[31]

As Almeida and Philippon (2007) point out, however, most debt conservatism calculations focus on expected costs of financial distress rather than the risk-adjusted costs of financial distress. This may matter. Bankruptcy occurs more commonly in bad times than in good. This is the opposite of what an insurance motive would call for. Accordingly, the utility cost must be magnified by risk considerations. Almeida and Philippon argue that this effect can be large enough to fully account for the results in Graham (2000).

[31] Note, however, that this sounds suspiciously like the first example in Section 2.3.2. Also, these firms may have high-growth options and assets that are largely intangible. If so, then the agency costs of debt may be particularly large for this sample.

Another problem for tests of taxes is that important element of corporate taxes known as tax shelters. Information about these is very hard to find since the Internal Revenue Service (IRS) treats tax investigations confidentially. Graham and Tucker (2006) studied the results of an exhaustive search of tax court records and financial news stories and identified 44 tax-sheltering cases involving in 43 firms between 1975 and 2000. They found that firms with tax shelters use less debt as predicted by the static trade-off theory. Many scholars suspect that Graham and Tucker (2006) are only observing the "tip of the iceberg." Unfortunately, we have no direct way of knowing the actual significance of such tax shelters.

Several recent papers attempt to reconcile the observed capital structures to those predicted by models. Minton and Wruck (2001) examine low-leverage firms and find that the low-leverage is largely transitory. These firms appear to be stockpiling financial slack or debt capacity, which is used later to make acquisitions and capital expenditures. Minton and Wruck's evidence seems quite similar to Example 2 in Section 2.3.2.

Morellec (2004) presents a contingent claims model with manager–stockholder conflicts. The model can generate the low-debt ratios observed in practice. In another recent paper, Ju et al. (2005) present a dynamic framework that provides estimates of optimal capital structures based on a calibrated contingent-claims model. They show that firms are not underlevered relative to the predictions of their model. Maximizing share value for a firm that is calibrated to be similar to the median Compustat firm results in an optimal debt-to-capital ratio of about 15%, which is below the median Compustat debt-to-capital value of about 23%. Their results contradict the view that firms are conservative in debt financing. Their results also show that moderate deviations of capital structure from optimal values have a very small impact on firm value. Thus, in the presence of transaction costs, it may be optimal for firms to let their capital structure deviate from the target by substantial amounts. Hennessy and Whited (2005) and Strebulaev (2007) also dispute the claim that firms are underlevered relative to the predictions of dynamic trade-off models. Their models also appear to be capable of accounting for the observed corporate debt levels.

Debt conservatism has also been examined from a behavioral perspective.[32] Behaviorists frequently report that overconfidence is the single most important deviation from rationality. Hackbarth (2007) presents a model in which an overconfident manager chooses higher debt levels than does a rational manager. Malmendier et al. (2005) report that, as an empirical matter, overconfident CEOs are more likely than other CEOs to raise debt (rather than equity) to cover financing deficits. They do not report on the magnitude of the effect. Instead of resolving the puzzle of why firms are underlevered, these behavioral studies deepen the debt conservatism puzzle just as the rational models are coming to grips with the problem.

While the problem of debt conservatism has attracted a certain amount of attention, it is not a first-order problem for the trade-off theory. There are a variety of ways to generate 'low' leverage in simulations of quite conventional trade-off models.

[32] See Baker et al. (2007) for a review of behavioral approaches to corporate finance problems.

### 3.3. Studies of leverage changes

Leverage can change due to the firm's active decision to issue or repurchase securities. Leverage can also change when the firm's circumstances or its stock price changes. Many studies therefore examine changes in leverage (e.g., Shyam-Sunder and Myers, 1999, and Frank and Goyal, 2003). Some studies examine changes in equity (e.g., Fama and French, 2005, and Leary and Roberts, 2007). Frank and Goyal (2004) examine both changes in debt and changes in equity in a two-equation Vector Autoregression (VAR) system.

Before turning to the individual studies, it is useful to examine the raw data on leverage changes. Table 5 shows the leverage transitions from one year to the next and reports market leverage adjustments. Book leverage adjustment transitions are essentially identical and thus omitted.

The bottom row of Table 5 indicates how common each leverage category is in the overall data. Many firms have leverage ratios between 0 and 10%. As leverage increases, the number of firms declines. Only 1.2% of firms have leverage greater than 90%. Large changes in leverage are quite rare in the sense that it is common for a firm to remain within the same category from one year to the next. When the firm leaves a particular category, it typically moves to an adjacent leverage category. It is rather rare for a high-leverage firm to cut leverage dramatically.

#### 3.3.1. Tests of the pecking order

Changes in debt have played an important role in assessing the pecking order theory. This is because the financing deficit is supposed to drive debt according to this theory. Shyam-Sunder and Myers (1999) examine how debt responds to short-term variation in investment and earnings. The theory predicts that when investments exceed earnings, debt grows, and when earnings exceed investments, debt falls. Dividends are assumed to be sticky in the short term.

Tests of the pecking order theory define financing deficit as investments plus change in working capital plus dividends less internal cash flow. The theory predicts that in a regression of net debt issues on the financing deficit, the estimated slope coefficient should be one. The slope coefficient indicates the extent to which new debt issues are explained by financing deficits.[33] Shyam-Sunder and Myers find strong support for this prediction in a sample of 157 large firms. The coefficient is 0.75 with an $R^2$ of 0.68 (see column 2 of their Table 2). They interpret this evidence to imply that the "pecking order

[33] Chirinko and Singha (2000) use several examples to illustrate that Shyam-Sunder and Myers's tests have low power. They show that the slope coefficient could be less than one for a firm that strictly follows the pecking order. This may happen because equity issues, while at the bottom of the financing hierarchy, are still a substantial percentage of external financing. Chirinko and Singha also show that the coefficient on deficit could be close to one even when a firm violates the pecking order model; that is, it issues equity before issuing debt or it issues debt and equity in fixed proportions.

Table 5

Leverage Transition Rates

Leverage transition rates for the untrimmed market leverage ratios (D) for the period 1950–2000. The data are from the Compustat files. Market leverage is defined as the ratio of book value of debt divided by book debt plus market value of equity. The row number is the group that the firm leverage belongs to in year $t$. The column number is the group that the firm's leverage belongs to in year $t+1$. The cell entries measure percentages. Exit is defined as not a missing value in year $t$, but a missing value in year $t+1$. Due to the lagging involved, the number of exits excludes the last two years.

| | $D < 0$ | $0 < D \leq 0.1$ | $0.1 < D \leq 0.2$ | $0.2 < D \leq 0.3$ | $0.3 < D \leq 0.4$ | $0.4 < D \leq 0.5$ | $0.5 < D \leq 0.6$ | $0.6 < D \leq 0.7$ | $0.7 < D \leq 0.8$ | $0.8 < D \leq 0.9$ | $D > 0.9$ | Exit | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $D \leq 0$ | 71.1 | 14.7 | 4.2 | 2.0 | 0.9 | 0.6 | 0.4 | 0.3 | 0.1 | 0.0 | 0.0 | 5.7 | 100.0 |
| $0 < D \leq 0.1$ | 6.8 | 64.7 | 14.3 | 5.0 | 2.1 | 1.0 | 0.5 | 0.3 | 0.1 | 0.1 | 0.0 | 5.2 | 100.0 |
| $0.1 < D \leq 0.2$ | 1.6 | 20.1 | 39.9 | 19.4 | 7.8 | 3.4 | 1.5 | 0.7 | 0.3 | 0.1 | 0.0 | 5.3 | 100.0 |
| $0.2 < D \leq 0.3$ | 0.7 | 5.4 | 21.1 | 33.2 | 18.8 | 8.5 | 3.8 | 1.7 | 0.8 | 0.2 | 0.1 | 5.9 | 100.0 |
| $0.3 < D \leq 0.4$ | 0.5 | 2.3 | 7.8 | 20.7 | 30.4 | 18.3 | 8.4 | 3.5 | 1.4 | 0.4 | 0.1 | 6.3 | 100.0 |
| $0.4 < D \leq 0.5$ | 0.4 | 1.1 | 3.0 | 8.5 | 20.4 | 29.6 | 18.4 | 7.8 | 3.4 | 1.1 | 0.3 | 6.1 | 100.0 |
| $0.5 < D \leq 0.6$ | 0.3 | 0.7 | 1.6 | 3.6 | 9.1 | 20.2 | 28.8 | 18.0 | 7.2 | 2.7 | 0.5 | 7.4 | 100.0 |
| $0.6 < D \leq 0.7$ | 0.3 | 0.6 | 0.9 | 1.9 | 4.3 | 9.5 | 21.2 | 28.9 | 17.2 | 6.1 | 1.4 | 7.9 | 100.0 |
| $0.7 < D \leq 0.8$ | 0.4 | 0.5 | 0.6 | 1.0 | 2.2 | 4.5 | 10.3 | 21.4 | 29.5 | 16.1 | 4.0 | 9.6 | 100.0 |
| $0.8 < D \leq 0.9$ | 0.8 | 0.6 | 0.5 | 0.8 | 1.4 | 2.3 | 4.3 | 9.8 | 21.0 | 31.4 | 15.0 | 12.3 | 100.0 |
| $D > 0.9$ | 1.2 | 0.9 | 0.3 | 0.8 | 0.9 | 1.2 | 1.9 | 3.7 | 7.5 | 23.1 | 42.6 | 16.0 | 100.0 |
| Total | 9.3 | 20.2 | 13.2 | 11.6 | 10.1 | 8.8 | 7.3 | 5.6 | 3.9 | 2.4 | 1.2 | 6.4 | 100.0 |

is an excellent first-order descriptor of corporate financing behavior" (Shyam-Sunder and Myers, 1999, p. 242).

The evidence in Shyam-Sunder and Myers is based on a small sample of 157 firms. These are large firms that traded continuously during the 1971–1989 period. The question is, what are the broad patterns of financing activity for a large cross section of firms?

Frank and Goyal (2003) examine the broad applicability of the pecking order theory. Their evidence, based on a large cross section of U.S. publicly traded firms over long time periods, shows that some firms make heavy use of external financing. On average, net equity issues track the financing deficit more closely than do net debt issues. These facts do not match the claims of the pecking order theory. The greatest support for pecking order is found among large firms, which might be expected to face the least severe adverse selection problem since they receive much better coverage by equity analysts. Even here, the support for pecking order is declining over time. They conclude that the pecking order theory does not explain broad patterns in the data.

**Stylized Fact 9** *Firms frequently adjust their debt. The financing deficit plays a role in these decisions. The traditional cross-sectional factors, however, are more important than the financing deficit.*

Lemmon and Zender (2004) attempt to reconcile the findings presented by Fama and French (2002) and Frank and Goyal (2003) with those presented by Shyam-Sunder and Myers (1999). According to Lemmon and Zender, the idea of debt capacity is important in understanding the rejections of the pecking order theory. Consideration of debt capacity suggests that, when unconstrained by debt capacity, firms issue debt, but that, when constrained, they issue equity. These tests require a workable definition of debt capacity. If debt capacity is defined as the point when adding more leverage reduces firm value, then debt capacity is similar to the concept of target leverage as defined by the trade-off theory of capital structure. Thus, finding that firms use debt to fill the financing deficit when they are below their debt capacity may not sharply distinguish two theories. Lemmon and Zender operationalize the concept of debt capacity by focusing on firms with rated debt. They argue that firms with debt ratings are unconstrained by debt capacity, while firms without debt ratings are constrained. Lemmon and Zender find, as expected, that the coefficient on financing deficit in net debt regressions is significantly larger for firms with rated debt and smaller for firms with no debt rating. They also show that firms with no debt rating are small, high-growth firms and that they use equity to finance their deficits. These results are consistent with those in Fama and French (2002) and Frank and Goyal (2003). The interpretation, however, is different. While Frank and Goyal suggest that these firms face more asymmetric information problems, and thus the pecking order predicts that they should issue debt, Lemmon and Zender state that these firms are debt capacity constrained and therefore issue equity.

Another attempt to reconcile the evidence in Frank and Goyal (2003) with the predictions of adverse selection arguments is described in Halov and Heider (2005). The 2005 paper argues that when there is greater asymmetric information about risk, debt has a more severe adverse selection problem and firms would only issue equity. To test these arguments, Halov and Heider use asset volatility as a proxy for asymmetric information about risk and divide firms into deciles based on asset volatility. They show that as asset volatility increases, firms use more equity to finance their deficits. The interpretation of these results rests on the assumption that differences in asset volatility deciles capture differences in asymmetric information about cash flow variance. The mean of the distribution is common knowledge. Thus, small, young, high-growth firms will issue equity to finance their deficits if these firm have more asymmetric information about risk and less asymmetric information about value.

Helwege and Liang (1996) found that the use of external financing by firms that undertook IPOs in 1983 did not match the pecking order prediction that financing deficit is the critical factor. Leary and Roberts (2007) find that when firms use external finance, less than 40% match the pecking order's predictions. The pecking order accurately identifies less than 20% of the observed equity issuances. They study whether these rejections are due to debt capacity or to time-varying adverse selection; they conclude that these suggestions do not account for the evidence.

Fama and French (2005) consider equity issuances. Most firms actually issue and/or retire equity in most years. Equity is issued, through many mechanisms and not only through seasoned equity offerings (SEOs). Many issues by large firms are fairly small. Violations of the pecking order are routine. More than half of the firms violate the pecking order when issuing equity. Gomes and Phillips (2005) find that half the equity issues are in the public market, whereas half are private issues. The pecking order provides a better account of the public issues and has difficulty accounting for the private issues.

Korajczyk et al. (1990) find that debt ratios do not rise prior to equity issues. There is also evidence that stock issues are typically followed by debt issues, and therefore leverage changes induced by equity issuances are only temporary (Eckbo and Masulis, 1995; Alti, 2006).

**Stylized Fact 10** *After an IPO, equity issues are more important for small firms than for large firms. Many large firms infrequently issue significant amounts of equity. When larger firms do issue, the issues can be large. Many small firms issue equity fairly often.*

### *3.3.2. Tests of mean reversion*

Static trade-off theory predicts a target debt ratio that depends on the tax benefits of debt and the costs of financial distress. By relying on adjustment costs, this theory may suggest a target adjustment process. As shown in Table 1, the aggregate data show that, in the U.S. economy as a whole, leverage is quite stable. Something must be causing such stability. It could be caused by the mean-reverting actions of individual firms, or it could be caused by the process of firm entry and exit.

Empirical tests of target adjustments focus on two related questions. First, does firm-level leverage revert to a target? Second, what do firms do when actual debt ratios deviate from the target?

Since the target is not observable, it must be estimated or its effects must be imputed. Early studies take a long-term average as the target. These early papers estimate the target debt ratio as the average debt ratio across a sample period. Examples include Taggart (1977), Marsh (1982), Jalilvand and Harris (1984), and Shyam-Sunder and Myers (1999). The approach assumes that firm characteristics that affect leverage remain unchanged over time. However, it is quite likely that the target changes over time as firm characteristics change. For example, if firms issue equity after increases in stock prices, one interpretation is that this action is inconsistent with firms targeting debt ratios because it moves them further from their target. However, an alternative interpretation is that stock price change reflects improvements in a firm's investment opportunity set. The improvement in growth prospects lowers the target debt ratio, and the equity issuance decision is a rational response of the firm to move toward its new target ratio.

Recent studies therefore adopt a two-step procedure in which an equation for the target is estimated first and the fitted value is then substituted into the adjustment equation. Examples of this approach include papers by Hovakimian et al. (2001), Fama and French (2002), Korajczyk and Levy (2003), and Kayhan and Titman (2007). For example, in the first stage in Hovakimian et al. (2001), leverage is regressed on a vector of variables presumed to affect leverage targets (similar to those described in Section 3.2). In the second stage, a logit regression predicting a firm's financing choice is estimated as a function of the difference between the actual leverage and the estimated target leverage and other variables affecting the deviation of the actual debt ratio from the target. Their results suggest that firms adjust toward target debt ratios. They issue debt when actual debt ratios are below the target, and they reduce debt when actual debt ratios are above the target. However, adjustments are stronger and more significant for debt reductions than they are for debt issuances. It is not clear why firms adjust more quickly when they are overlevered but not when they are underlevered.

The literature commonly agrees that leverage exhibits mean reversion. Mean reversion in leverage is in fact not surprising in light of the evidence that leverage has been quite stationary over long periods of time. But there is significant disagreement over how rapidly the adjustment takes place. Fama and French (2002) report estimates of adjustment speeds based on a two-step procedure of between 7 and 10% for dividend payers and between 15 and 18% for dividend nonpayers. They conclude on the one hand that reversion occurs at "a snail's pace." On the other hand, Alti (2006) and Leary and Roberts (2005) report evidence that reversion is quite fast and is mostly accomplished in two to three years.

It is also possible to substitute the target equation into the adjustment equation and then estimate the resulting structure, as in more recent papers (see Flannery and Rangan, 2006; Lemmon et al., 2007; Frank and Goyal, 2007b; and Huang and Ritter, 2007). Let $Lev_{i,t}$ denote the leverage ratio for firm $i$ at time $t$ and $Lev^*_{i,t}$ denote its target leverage. The standard partial adjustment equation is

$$\Delta Lev_{i,t} = \lambda(Lev^*_{i,t} - Lev_{i,t-1}) + \varepsilon_{i,t} \tag{8}$$

where $\lambda$ is the speed of adjustment and $\varepsilon_{i,t} \sim N(0, \sigma_i^2)$. In other words, deviations from target leverage are corrected at a rate of $\lambda$ each year. The target leverage can be specified as a function of firms' fixed effects, $\alpha_i$, and time-varying firm characteristics, $X_{i,t-1}$. The parameters to be estimated are represented by $\beta$. Thus,

$$Lev^*_{i,t} = \alpha_i + \beta X_{i,t-1} \tag{9}$$

By substituting Equation (9) in (8), we get the reduced-form specification:

$$Lev_{i,t} = (\lambda\beta)X_{i,t-1} + (1-\lambda)Lev_{i,t-1} + \lambda\alpha_i + \varepsilon_{i,t} \tag{10}$$

The estimated speed of adjustment in a dynamic panel data model with fixed effects is quite sensitive to the econometric procedures. Different papers employ different procedures and reach different conclusions. Estimations using pooled OLS ignoring the fixed effects biases the adjustment speeds downward. To mitigate these biases, Flannery and Rangan (2006) employ the firm fixed effects in Equation (10) and show rapid speeds of adjustment (about 34% per year). However, if the time dimension is short, adjustment speeds estimated using firm fixed effects are biased upward. An alternative is to employ the system generalized method of moments (GMM) as in Lemmon et al. (2007). Huang and Ritter (2007) point out that system GMM estimates are also biased upward if the dependent variable is highly persistent. They instead use a long-differencing estimator and report that leverage adjustment speeds are modest. In this case, the adjustment speeds are larger than those reported by Fama and French (2002) but lower than those reported by Flannery and Rangan (2006).

Another question is whether the target to which firms are adjusting is time varying or time invariant. Lemmon et al. (2007) show that while firms are actively managing leverage and exhibiting mean reversion, the rebalancing is largely toward a relatively time-invariant target. They argue that this target even predates the firm's IPO. Leverage ratios are relatively stable, and much of the variation in leverage is cross-sectional rather than time-series. This is in contrast to the conclusion by Hovakimian et al. (2001) and Flannery and Rangan (2006), who maintain that firms are largely adjusting to time-varying leverage targets. Frank and Goyal (2007b) provide some evidence that this firm persistence may actually result from its correlation with the managerial team. When managerial teams change, the change affects leverage. But such changes only happen every few years, and so they are highly correlated with firm fixed effects.

Shyam-Sunder and Myers (1999) have argued that mean reversion is not necessarily incompatible with the pecking order. They argue that in a pecking order world in which firms do not have leverage targets, leverage may appear to be mean reverting. This happens when capital investments are lumpy and positively serially correlated, free cash flows vary over the business cycle, and the average debt ratio is taken as the target. Chang and Dasgupta (2007) show that even with random financing and with no apparent target, leverage may appear to be mean reverting. They show that this mechanical rebalancing can lead to mean reversion in simulated data where the financing is purely random.

Transaction costs are potentially quite important. As pointed out by Leary and Roberts (2005), different forms of adjustment costs are likely to induce different patterns of leverage changes. Fama and French (2005) argue that there are many different ways to issue equity and that these are associated with differing levels of transaction costs. It is likely that the transaction costs that a firm faces when issuing a security are generally not the same as the transaction costs associated with repurchasing that same security. Thus, it is likely that asymmetries should be found. (Of course, as pointed out by Stiglitz, 1973, there is also an important asymmetry in the tax code.) Chen and Zhao (2005) find some evidence consistent with asymmetry.

The problem of estimating adjustments is not unique to corporate finance. It arises in several fields. Caballero and Engel (2003) consider the effect of using standard econometric techniques to estimate adjustment speeds when the adjustment is lumpy and infrequent. They report that the standard practice of estimating the speed of adjustment with partial-adjustment autoregressive moving average (ARMA) procedures substantially overestimates this speed. Since the adjustment speed methods used in the capital structure literature are closely related methods, the extent to which their concerns affect current estimates is unclear. We summarize the evidence this way.[34]

**Stylized Fact 11** *Corporate leverage is mean reverting at the firm level. The speed at which this happens is not a settled issue.*

Mayor and Sussman (2004) examine the financing of unusually large projects. Large firms finance investment spikes with debt and small firms with equity. New equity issues are associated with small, loss-making firms. Mayer and Sussman also observe a tendency toward readjustment to previous levels of leverage after the spike. Kayhan and Titman (2007) also show that firms behave as if they have a target debt ratio. Their evidence suggests that investment needs, cash flows, and stock returns lead to transitory deviations from leverage targets, but firms gradually undo these deviations.

Frank and Goyal (2004) study a simple framework in which possible interactions between shocks to debt and equity are allowed to have both their own effects and cross effects on subsequent issuing decisions. To do this, a two-equation VAR system in which cointegration is permitted is estimated. They find that shocks to equity value are followed by offsetting actions in the debt market. The fact that the offsetting actions take place in the debt market is directly relevant to studies of equity market timing and seems to contradict Welch (2004).

[34] Caballero and Engel explain the bias as follows: "In linear models, the estimated speed of adjustment is inversely related to the degree of persistence in the data. That is, a larger first order correlation is associated with lower adjustment speed. Yet this correlation is always zero for an individual series that is adjusted discretely (and has i.i.d. shocks), so that the researcher will conclude, incorrectly, that adjustment is infinitely fast. To see that this crucial correlation is zero, first note that the product of current and lagged changes in the variable of concern is zero when there is no adjustment in either the current or the preceding period. This means that any non-zero serial correlation must come from realizations in which the unit adjusts in two consecutive periods. But when the unit adjusts in two consecutive periods, and whenever it acts it catches up with all accumulated shocks since it last adjusted, it must be that the later adjustment only involves the latest shock, which is independent from the shocks included in the previous adjustment" (p. 1).

**Stylized Fact 12** *At the aggregate level, mean reversion of leverage mainly happens through debt market actions.*

### *3.3.3. Exit*

Bankruptcy has been heavily studied in its own right. Here we only make a few observations about the connection between exit and leverage. Bankruptcy and financial distress play a crucial role in the trade-off theory. Generally, it is thought that firms in trouble are highly levered. Table 5 provides some descriptive evidence. In every leverage category, more than 5% of firms exit from one year to the next. As expected, exit is much more common for high-leverage firms than for low-leverage firms.

Bankruptcy is only one of many ways for a firm to exit. Table 6 decomposes the reasons for firm exits as listed by Compustat. By far, the most common reason for exit is either an acquisition or a merger. This alone accounts for more than half of the identified cases. Actual identified bankruptcies and liquidations are surprisingly infrequent. However, it seems likely that the "other" category includes a fair number of otherwise unidentified bankruptcies.

**Stylized Fact 13** *Mergers and acquisitions are more common reasons for exit than are bankruptcies and liquidations.*

Table 6

Distribution of Compustat Deletion Codes

The table provides a distribution of Compustat deletion codes for the period 1950–2000. Companies can exit in various ways. Compustat footnote 35 provides some evidence on the relative importance of alternative exit mechanisms. This footnote appears to be somewhat incomplete and so should be regarded as suggestive rather than definitive. Frequency gives the raw count of the number of times the particular reason for deletion was listed. Fraction gives each reason divided by the total.

| Reason for deletion | Frequency | Fraction |
|---|---|---|
| Acquisition or merger | 3176 | 0.589 |
| Bankruptcy—Chapter 11 | 368 | 0.068 |
| Liquidation—Chapter 7 | 186 | 0.035 |
| Reverse acquisition (from 1983 onward) | 50 | 0.009 |
| No longer fits original file format (from 1978 onward) | 28 | 0.005 |
| Leveraged buyout | 91 | 0.017 |
| Now a private company | 320 | 0.059 |
| Other (no longer files with SEC etc.) | 1170 | 0.217 |
| Total | 5389 | 1.000 |

Firms that are financially distressed often take steps to try to mitigate their problems. Asquith et al. (1994) studied a group of firms that issued junk bonds and then got into financial distress. In addition to restructuring their finances, it was quite common for these firms to sell assets. Such asset sales appear to be limited by the state of the industry. Maksimovic and Phillips (1998) argue that, in fact, industry conditions are the key factor in asset redeployment. More controversially, they suggest that firms in Chapter 11 face only minor bankruptcy costs.[35]

Exit by merger is quite different from exit by liquidation. Liquidated assets are likely to fetch low prices when sold piecemeal in second-hand asset markets. However, firms that are targets of takeovers receive large premiums. When a firm is getting into trouble, it might look for someone to acquire the firm as a going concern. To the extent that this takes place, exit through takeovers mitigates the importance of direct bankruptcy costs.

### *3.3.4. The effect of current market conditions*

How important are current market conditions for leverage adjustment choices? As a matter of theory, they might or might not be important. Intuitively, it seems that market conditions ought to matter. For instance, when taking out a home mortgage, it would seem natural to look at the current term structure relative to historical norms, before deciding on the term to maturity and whether to take a fixed rate or a floating rate mortgage. Perhaps CFOs do the same thing. Such behavior may not be so crazy: "naïve investors, who judge bonds by their yields to maturity and buy long bonds when their yields are relatively high, have tended to earn superior returns in the postwar period in the United States" (Campbell et al., 1997, pp. 423–424).

Thus, it is not too surprising that the evidence points to a role for current market conditions. As shown in Section 3.2, both the market-to-book ratio and the expected inflation rate are significant factors in standard panel regressions. Both of these can be interpreted as an effect of market conditions. Frank and Goyal (2004) find that aggregate corporate debt adjusts in reaction to the current market-to-book ratio. But neither the market-to-book ratio nor the current interest rate seems to affect the aggregate long-term leverage ratio. There is good evidence that IPOs come in waves (Ritter and Welch, 2002), as do mergers (Andrade et al., 2001).

The issue of exactly how market conditions matter and how long lived the effects are is much more controversial. At one extreme, Baker and Wurgler (2002) suggest that the effects are long lived. They report that firms that issued equity when market conditions were good have lower leverage for a decade or more. They interpret this finding as support for the claim that leverage is determined by the attempt of firms to time the equity market. Chang et al. (2006) argue that information asymmetry affects a firm's incentives to time the market. They show that firms with low information asymmetries (the ones with greater analyst coverage) have lower incentives to time the market. Firms

[35] Debate about the magnitude of bankruptcy costs goes back to Warner (1977). See also Andrade and Kaplan (1998).

followed by fewer analysts make infrequent but larger issues of equity. From a different point of view, Welch (2004) contends that shocks to a firm's equity are not undone, and so the effects of equity shocks on leverage are effectively permanent.

These ideas have been sharply challenged in several recent papers. First, it has been shown that such long-lived effects are consistent with dynamic trade-off theories such as Hennessy and Whited (2005) and Strebulaev (2007). Similarly, these models are able to replicate the regression results reported by Welch (2004). Thus, as a matter of theory, the evidence taken at face value does not contradict the dynamic trade-off theory.

But not everyone is willing to take the evidence at face value. Alti (2006), Leary and Roberts (2007), and Kayhan and Titman (2007) have reported evidence that "market timing" effects are present, but that they have largely dissipated after a couple of years. Huang and Ritter (2007) dispute this claim. This debate is intimately connected to the tests of target adjustment. Recall that in Section 3.3.2 we found that the rate of mean reversion is not a settled issue. Thus, it is not surprising that the durability of the market impact is also best regarded as an open issue.

The idea that managers do not react to equity shocks appears to be simply incorrect. A number of papers, including Hovakimian et al. (2001), show that good equity returns are commonly followed by further equity issues. Frank and Goyal (2004) state that equity shocks induce offsetting debt market reactions. Strebulaev (2007) points out that, in an optimizing model, managers should react to long-term changes, but not to every little blip in the market. Thus, the evidence for market condition effects appears to be compatible with fairly conventional trade-off models with varying leverage targets. The need for a completely new market timing theory as a competitor to the conventional theories is, as yet, not established.[36]

**Stylized Fact 14** *Market conditions have some effect on leverage decisions. The magnitude and durability of these effects are not settled issues.*

### *3.4. Market valuation of leverage changes*

#### *3.4.1. Predictions*

When a firm issues, repurchases, or exchanges one security for another, it changes its capital structure. What are the valuation effects of these changes?

Under the trade-off theory, firms will only take action if they expect benefits. An implication of the theory is that the market reaction to both equity and debt securities will be positive. But the interpretation is not that easy. The market response to a leverage change confounds two pieces of information: the revelation of the fact that the firm's

[36] The proponents of market timing theory have not directly developed an explicit model that might then be tested on other dimensions. A new theory is normally expected to account for the facts that existing theories can already explain. So far, the proponents of market timing theory have not attempted to do so.

conditions have changed, necessitating financing, and the effect of the financing on security valuations. The information contained in security issuance decisions could be either good news or bad news. It would be good news if the firm is issuing securities to take advantage of a promising new opportunity that was not previously anticipated. It might be bad news if the firm is issuing securities because the firm actually needs more resources than anticipated to conduct operations. A firm may also issue securities now in anticipation of a change in future needs. This implies that the trade-off theory by itself places no obvious restrictions on the market valuation effects of issuing decisions. Everything depends on the setting.

Jung et al. (1996) suggest an agency perspective and argue that equity issues by firms with poor growth prospects reflect agency problems between managers and shareholders. If this is the case, then stock prices would react negatively to news of equity issues.

The pecking order theory is usually interpreted as predicting that securities with more adverse selection (equity) will result in more negative market reactions. Securities with less adverse selection (debt) will result in less negative or no market reaction. This does, of course, still rest on some assumptions about market anticipation.

### *3.4.2. Evidence*

Announcements of ordinary debt issues generate no market reaction on average (see Eckbo, 1986; Antweiler and Frank, 2006). The lack of a market reaction to corporate debt issues is robust to various attempts to control for partial anticipation. Announcements of convertible debt issues result in mildly negative stock price reactions (see Dann and Mikkelson, 1984; Mikkelson and Partch, 1986). Announcements of equity issues result in significant negative stock price reactions (see Asquith and Mullins Jr., 1986; Masulis and Korwar, 1986; Antweiler and Frank, 2006). The announcement effects are positive when common stock is repurchased (see Masulis, 1980b; Dann, 1981; Antweiler and Frank, 2006). Equity issues by utilities generate less negative reactions than those by industrial issuers. Exchange of common stock for debt/preferred stock generates positive stock price reactions, while exchange of debt/preferred stock for common stock generates negative reactions (Masulis, 1980a).

Summarizing the event study evidence, Eckbo and Masulis (1995) conclude that announcements of security issues typically generate a nonpositive stock price reaction. The valuation effects are the most negative for common stock issues, slightly less negative for convertible debt issues, and least negative (zero) for straight debt issues. The larger the issue, the more negative the effects.

**Stylized Fact 15** *Announcements of corporate debt issues and debt repurchases have little, if any, effect on the market value of the firm.*

**Stylized Fact 16** *Announcements of equity issues are generally associated with a drop in the market value of the firm. Announcements of equity repurchases are generally associated with an increase in the market value of the firm.*

The negative market reaction to equity issues and the lack of a market reaction to debt issues are consistent with adverse selection arguments. Indeed, there are other interpretations. Jung et al. (1996) show that firms without valuable investment opportunities experience a more negative stock price reaction to equity issues than do firms with better investment opportunities. Thus, agency cost arguments could also explain the existing evidence on security issues. Further support for the agency view comes from the finding that firms without valuable investment opportunities issuing equity invest more than similar firms issuing debt and that firms with low managerial ownership have worse stock price reactions to new equity issue announcements than do firms with high managerial ownership.

The impact of equity issues appears to differ between countries. Several studies find positive market reaction to equity issues around the world (see Eckbo et al., 2007, for a summary). To understand this evidence, Eckbo and Masulis (1992) and, more recently, Eckbo and Norli (2004) have examined stock price reactions to equity issues conditional on a firm's choice of flotation method. Firms can issue equity using uninsured rights, standby rights, firm-commitment underwriting, and private placements. The stock price reactions to equity issues depend on the flotation method. For U.S. firms, Eckbo and Masulis (1992) find that the average announcement-period abnormal returns are insignificant for uninsured rights offerings and significantly negative for firm-commitment underwritten offerings. Eckbo and Norli (2004) in a study of equity issuances on the Oslo Stock Exchange, find that uninsured rights offerings and private placements result in positive stock price reactions, while standby rights offerings generate negative market reactions. These papers interpret the effect of the flotation method as reflecting different degrees of adverse selection problems.

### *3.5. Natural experiments*

A problem with cross-sectional tests is that financial policy decisions are made jointly with investment and payout policy decisions. Thus, it is difficult to make causal inferences about debt ratios. A natural idea is thus to look for plausibly exogenous changes in a firm's environment and then see how leverage responds. This method differs from studies of leverage changes because the defining criterion is a change in the firm's environment. The literature contains a number of such studies.

Blanchard et al., (1994) examined a sample of 11 firms that received a large cash windfall without any change in marginal $q$. They found that firms increased their long-term debt following the cash windfall. The pecking order predicts an increase in debt if firms have attractive investment opportunities and borrow more money to undertake these projects. Since these firms did not have such opportunities, the increase in debt following cash windfalls is inconsistent with the pecking order theory.

The agency theories predict that managers expand firms when possible. Firms are able to increase debt because cash windfalls increase a firm's debt capacity. The increase in long-term debt is therefore potentially consistent with the predictions of the agency theories.

It is often suggested that cash can be viewed as negative debt. Suppose that this is correct. Then, the cash windfall is a reduction in leverage. If the original leverage was optimal, then the firm needs to increase its debt (or repurchase equity) in response to the windfall. Thus, the behavior observed by Blanchard et al., (1994) seems quite compatible with the trade-off theory perspective. The fact that the adjustment takes place in the debt market rather than the equity market is consistent with the aggregate evidence of Frank and Goyal (2004).

To examine how exogenous shocks affect firms' financing decisions, the Undistributed Profits Tax in 1936–1937 provides an interesting historical case. Christie and Nanda (1994) and Calomiris and Hubbard (1995) focus on the behavior of firms around the introduction of the tax on undistributed profits, which was introduced by the Roosevelt administration in 1936 but was abolished in 1938 following strong protests by businesses. Calomiris and Hubbard (1995) show that firms increased their debt after of the undistributed profits tax was introduced. This finding is consistent with firms increasing the amount of debt to reduce taxes on retained profits. They also show that the firms that paid the highest taxes (and lowest dividends) had high debt ratios both before and after the introduction of the tax. These are small firms with arguably high costs of external financing.

In 1986, a tax reform legislation reduced both corporate and personal marginal tax rates. Givoly et al., (1992) report that firms with high tax rates prior to the tax reform reduced their debt the most after the reform. On its own, this observation seems compatible with the trade-off theory because these are presumably the firms that receive the largest tax reductions. Graham (2003) suggests that this result is a bit surprising given the endogeneity bias of the tax rates and given the fact that the personal tax rate drops were not modeled in the analysis. This historical episode might be worth further research.

In 2003, there was a large cut in individual dividend income taxes. This event provides an alternative angle to consider the question of whether taxes affect the nature of corporate financing. In a model like that proposed by Stiglitz (1973) or Hennessy and Whited (2005), when such taxes are cut, more firms should find it attractive to pay dividends. Chetty and Saez (2005) show that there was a significant increase in dividend payments following the tax cut, along several dimensions as predicted in the tax-based theories.

Goyal et al. (2002), in their examination of the U.S. defense industry during the 1980–1995 period, found that growth opportunities increased substantially for U.S. weapons manufacturers during the Reagan defense buildup of the early 1980s and declined significantly with the end of the cold war and the associated defense budget cuts in the late 1980s and the early 1990s. It seems quite unlikely that changes in corporate debt policies altered the U.S. defense buildup. Thus, it is reasonable to consider firm leverage as reacting to the defense budgets rather than causing them.

Goyal et al. (2002) examine how the level and structure of corporate debt changed in a sample of defense firms relative to a benchmark sample over this period. As growth opportunities declined, weapons manufacturers, which were most affected by the decline in defense budgets, increased the level of debt in their capital structures. New debt issued increased significantly for weapons manufacturers during the low-growth period.

In addition, weapons manufacturers lengthened the maturity structure of their debt, decreased the ratio of private debt to total debt, and decreased the use of senior debt. Their evidence suggests that growth opportunities play a prominent role in corporate debt policies.

Baggs and Brander (2006), studying the effect of the North American Free Trade Agreement (NAFTA) on the leverage of Canadian firms, find that when domestic tariff protection is reduced, corporate profits decline at the affected firms and their leverage increases. When foreign tariffs decline, profits tend to rise and leverage declines. The results are interesting, but the interpretation is not simple. It is not clear whether the main force is the realized effect on profits or the anticipated effect that operates through growth opportunities.

Dittmar (2004) and Mehrotra et al. (2003) examine the capital structure choices that firms make when engaging in spin-offs. Spin-offs are interesting since in essence it is at this point that a capital structure must be selected. In most respects, firms allocate leverage based on attributes shown to be important in cross-sectional studies. Thus, firms with higher tangibility of assets are allocated more leverage. Assets with lower liquidation costs have more leverage. Differences in leverage between the parent and the subsidiary are negatively related to variability in the industry's operating income. In one respect, however, the results differ from cross-sectional evidence. Differences in leverage are positively related to differences in profitability.

Gilson (1997) examined the capital structures of firms emerging from financial distress and argues that Chapter 11 bankruptcy helps firms overcome transaction costs and thus permits financially distressed firms to reduce leverage. This provides direct evidence that it may not be easy for a firm to restructure its finances outside of bankruptcy, even when the actual underlying business remains valuable. This is interesting evidence of the possible nature of the transaction costs associated with high-debt levels.

**Stylized Fact 17** *The natural experiments papers are generally easy to understand from the perspective of trade-off theory.*

This stylized fact characterizes the available studies, but, from the perspective of the trade-off theory, perhaps the most important natural experiment was the introduction of the corporate income tax in 1909. It would be nice to know more than we currently do about how firms reacted. It would not be too surprising if this natural experiment proves difficult to interpret under either the pecking order or the trade-off theory.

### *3.6. Surveys*

While large sample studies offer cross-sectional variation and statistical power, they have the disadvantage that researchers cannot ask qualitative questions. Natural experiments and clinical studies provide excellent details but typically use small samples. The survey approach provides a balance: it typically uses moderately large samples and has the ability to ask qualitative questions. Despite these benefits, the survey approach remains rare in corporate finance.

The most common criticism of the survey approach is that it measures beliefs rather than actions; the approach implicitly assumes that the manager's beliefs reflect reality. Different executives within a given firm might answer the same question in different ways. Perhaps more importantly, surveys rely on language. Ideas can be expressed in different words. In some instances, the words sound attractive, while in others they sound unattractive. For instance, it is difficult to imagine any manager agreeing that he or she is employing a "cash-burning signal." Yet, that same manager might agree that the reason a particular expensive action was worthwhile was that it proved to the market "how serious" the firm really is in some respect. As a result, considerable care is needed if the survey is to truly measure what it intends to measure. Theorists may use language somewhat differently than do practitioners.

An important contribution to the recent survey approach is by Graham and Harvey (2001).[37] Their study, which presents responses from U.S. CFOs, reveals that firms value financial flexibility in making debt decisions. This desire for financial flexibility seems inconsistent with the pecking order theory since dividend-paying firms (firms with relatively less information asymmetry) value flexibility the most. Graham and Harvey also find that firms that perceive their stock to be undervalued are reluctant to issue equity. But they find that large and dividend-paying firms are more likely to delay equity issuance because of undervaluation. Again, these results are surprising from the perspective of the pecking order theory.

The survey evidence suggests that CFOs consider the tax advantages of debt to be moderately important. The tax advantages are more important for firms likely to be paying more taxes (large, regulated, and dividend-paying firms). Managers show a great deal of concern about credit ratings and earnings volatility in making debt decisions. In terms of whether firms have a target debt ratio, almost 44% of the CFOs responded that they have a tight or somewhat tight target capital structure. About 34% responded that they have a flexible target and only 19% said that they have no target ratio.

Overall, the CFOs' ranking of the top three factors affecting capital structure choice—financial flexibility, credit ratings, and earnings volatility—is consistent with the view that debt decisions are influenced by a desire to avoid getting the firm into distress. This does suggest a concern to avoid bankruptcy costs or financial distress costs. The terminology is sufficiently vague that this evidence is consistent with a wide range of possible sources of such costs.

Two other recent surveys of European managers confirm the findings of Graham and Harvey (2001). Both Bancel and Mittoo (2004) and Brounen et al. (2004) report that European managers also rank financial flexibility as the most important factor in determining their firm's debt policy. In Europe, as in the United States, some firms report

[37] References to earlier surveys can be found in Graham and Harvey (2001).

having a target capital structure. But the target is flexible in most cases. Brounen et al. find that the tax advantage of interest expense ranked as the fourth most important factor after financial flexibility, credit rating, and earnings volatility.

Thus, different surveys now seem to provide relatively similar rankings of factors, despite the fact that the firms are operating in countries with very different institutions. This finding is somewhat reassuring. The evidence is relatively difficult to interpret in terms of the standard theories. Some of what is reported seems consistent with each theory, and other reports seem inconsistent with each theory. The stress on financial flexibility is interesting but open to a variety of interpretations. In our view, the survey evidence is of interest, but it is best regarded as being suggestive rather than providing definitive tests.

## 4. Conclusion

According to the standard trade-off theory, taxes and bankruptcy account for the corporate use of debt. According to the standard pecking order theory, adverse selection accounts for the corporate use of debt. There are good reasons to question the standard versions of both theories.

The trade-off theory focuses on taxes and bankruptcy costs. Until quite recently, this theory has been the dominant one in corporate finance textbooks, while, at the same time, the theory was in serious disrepute among most finance scholars. Recently, that has changed somewhat. Some of the most prominent objections to the trade-off theory have become less compelling in light of more recent evidence and an improved understanding of some aspects of the dynamic environment.

The suggestions that firms use too little debt relative to the trade-off theory were asserted with particular force by Miller (1977). However, a number of direct attempts to quantify Miller's idea, such as that by Ju et al. (2005), find that the observed debt levels are not surprising when somewhat realistic structures and parameters are considered.

Similarly, many scholars, such as Myers (1984) and Fama and French (2002), regard the lack of a positive correlation between profits and debt as a problem for the trade-off theory. However, this objection is also compatible with fairly standard dynamic trade-off models, as discussed in Section 2.3.2.

Despite the improved fortunes of the trade-off theory, it cannot be the full story. The U.S. corporate income tax did not begin until 1909 when it was introduced at a 1% rate. The use of debt contracts by businesses has a much longer history than does the corporate income tax. Thus, while taxes probably play an important role, there must be more to it.

The pecking order theory also has serious problems on a number of dimensions. Firms that have cash on hand actually issue debt. Frank and Goyal (2003) show that financing deficits does not wipe out the effect of the conventional factors. Firms routinely issue equity when they should not do so (Fama and French, 2005; Leary and Roberts, 2007). As shown in Section 3.1, leverage has been quite stationary over the recent decades.

With the standard versions of both approaches having clear flaws, it is perhaps not surprising that active research is underway.[38] We have a clear sense of the failings of the theories, but naturally there are differences of opinion on how best to make progress.

A lot of evidence seems consistent with bankruptcy affecting financing. The importance of collateral is quite strong in the data. A variety of other facts are also easy to interpret in this light. Direct transaction costs also seem to play a role. For instance, the differences in the use of debt and equity by small firms when compared to the use by large firms seems easy to understand in terms of direct transaction costs. The importance of retained earnings is quite consistent, with both transaction costs or taxes playing a role.

How important are the various agency conflicts relative to each other? relative to adverse selection? relative to taxes? We do not really know. A further problem is that currently there is little research that examines capital structures within a general equilibrium context. It would be nice to have calibrated versions of models along the lines of Auerbach and King (1983) or McDonald (2004). Ideally, the model might help provide an account of the kind of evidence presented in Tables 3 and 4. The remarkable growth in the role of financial intermediaries deserves much more attention than it has received in the literature on capital structure.

Many recent papers have focused on the dynamic aspects of leverage. This literature has already seriously altered our understanding of corporate capital structure. There is good reason to believe that this area of research will continue to be a productive one over the next few years. Inclusion of agency conflicts and adverse selection in these dynamic models will undoubtedly prove interesting and will help close the gap between the approaches.

Where does this leave Myers's contest? As one might have hoped, in the two decades since his address, there have been significant improvements in our understanding of the theory and marked improvements in our knowledge of the facts. Perhaps the most serious problem at this time is the lack of a satisfactory unifying model. We are not aware of any current model that is capable of simultaneously accounting for the main stylized facts, but it would be nice to have one.

## 5. Appendix: the stylized facts

1. Over long periods of time, aggregate leverage is stationary.
2. Over the past half century, the aggregate market-based leverage ratio has been about 0.32. There have been surprisingly small fluctuations in this ratio from decade to decade.
3. At the aggregate level, capital expenditures are very close to internal funds. This is true for large public firms and private firms; this is *not* true for small public firms.

[38] Since the standard versions of both theories have serious weaknesses, Fama and French (2002) have suggested that perhaps we should revert to the Modigliani-Miller theorem. However, that is not really credible either. Far too many systematic patterns hold up across countries and across time periods.

4. At the aggregate level, the financing deficit is very close to debt issues. This holds for large public firms and for private firms; this does *not* hold for small public firms. For small public firms, financing deficits very closely match equity issues.
5. Aggregate dividends are very smooth and almost flat as a fraction of total assets for all classes of firms. There has been remarkable stability in the aggregate dividend rate over time. Large public firms pay higher dividends than do small public firms. Many small firms pay no dividends.
6. Over the past half century, there has been a large decrease in direct holding of corporate securities by households, and a corresponding huge increase in financial intermediation of such claims.
7. Households have been net suppliers of corporate equity since the 1960s. Corporations have been net buyers of equity since the 1980s. Most equity is no longer held directly. Insurance companies, mutual funds, and pension funds now hold more equity and debt than households hold directly.
8. There is a core set of six reliable factors that are correlated with cross-sectional differences in leverage. Leverage is positively related to median industry leverage, collateral, log of assets, and expected inflation. Leverage is negatively related to market-to-book and profits.
9. Firms frequently adjust their debt. The financing deficit plays a role in these decisions. The traditional cross-sectional factors are, however, more important than the financing deficit.
10. After an IPO, equity issues are more important for small firms than for large firms. Many large firms infrequently issue significant amounts of equity. When larger firms do issue, the issues can be large. Many small firms issue equity fairly often.
11. Corporate leverage is mean reverting at the firm level. The speed at which this happens is not a settled issue.
12. At the aggregate level, mean reversion of leverage mainly happens through debt market actions.
13. Mergers and acquisitions are more common reasons for exit than are bankruptcies and liquidations.
14. Market conditions have some effect on leverage decisions. The magnitude and durability of these effects are not settled issues.
15. Announcements of corporate debt issues and debt repurchases have little, if any, effect on the market value of the firm.
16. Announcements of equity issues are generally associated with a drop in the market value of the firm. Announcements of equity repurchases are generally associated with an increase in the market value of the firm.
17. The natural experiments papers are generally easy to understand from the perspective of trade-off theory.

## References


Almeida, Heitor and Thomas Philippon, 2007, "The Risk-adjusted Cost of Financial Distress," *Journal of Finance*, Forthcoming.

Alti, A., 2006, "How Persistent Is the Impact of Market Timing on Capital Structure?" *Journal of Finance*, 61, 1681–1710.

Andrade, G. and S. N. Kaplan, 1998, "How Costly Is Financial (not economic) Distress? Evidence from Highly Leveraged Transactions that Became Distressed," *Journal of Finance*, 53, 1443–1493.

Andrade, G., M. Mitchell and E. Stafford, 2001, "New Evidence and Perspectives on Mergers," *Journal of Economic Perspectives*, 15, 103–120.

Antweiler, W. and M. Z. Frank, 2006, "Do U.S. Stock Markets Typically Overreact to Corporate News Stories?," Working Paper, UBC and University of Minnesota.

Asquith, P., R. Gertner and D. Scharfstein, 1994, "Anatomy of Financial Distress: An Examination of Junk-bond Issuers," *Quarterly Journal of Economics*, 109, 625–658.

Asquith, P. and D. W. Mullins Jr., 1986, "Equity Issues and Offering Dilution," *Journal of Financial Economics*, 15, 61–89.

Atkeson, A. and H. Cole, 2005, "A Dynamic Theory of Optimal Capital Structure and Executive Compensation," Working Paper, NBER 11083.

Auerbach, A. J. and M. A. King, 1983, "Taxation, Portfolio Choice, and Debt-Equity Ratios: A General Equilibrium Model," *Quarterly Journal of Economics*, 98, 587–610.

Baggs, J. and J. Brander, 2006, "Trade Liberalization, Profitability, and Financial Leverage," *Journal of International Business Studies*, 37, 196–211.

Baker, M., R. S. Ruback, and J. Wurgler, 2007, "Behavioral Corporate Finance," In B. E. Eckbo (ed.) *Handbook of Corporate Finance–Empirical Corporate Finance, Vol. 1*. In: *Handbook of Finance Series, Chapter 4*, Elsevier/North-Holland, Amsterdam.

Baker, M. and J. Wurgler, 2002, "Market Timing and Capital Structure," *Journal of Finance*, 57, 1–32.

Baltagi, B. H., 2001. *Econometric Analysis of Panel Data*, John Wiley & Sons, New York.

Bancel, F. and U. R. Mittoo, 2004, Cross-country determinants of capital structure choice: A survey of european firms. *Financial Management*, 103–132.

Barclay, M. J., E. Morellec and C. W. Smith, 2006, "On the Debt Capacity of Growth Options," *Journal of Business*, 79, 37–59.

Barclay, M. J. Jr., C. W. Smith and R. Watts, 1995, "The Determinants of Corporate Leverage and Dividend Policies," *Journal of Applied Corporate Finance*, 7, 4–19.

Berkovitch, E. and E. H. Kim, 1990, "Financial Contracting and Leverage Induced Over- and Under-investment Incentives," *Journal of Finance*, 45, 765–794.

Blanchard, O., F. Lopez de Silanes and A. Shleifer, 1994, "What Do Firms Do with Cash Windfalls?" *Journal of Financial Economics*, 36, 337–360.

Bradley, M., G. Jarrell and E. H. Kim, 1984, "On the Existence of an Optimal Capital Structure: Theory and Evidence," *Journal of Finance*, 39, 857–877.

Braudel, F., 1982, *Civilization and Capitalism, 15th–18th century. Volume 2. The Wheels of Commerce*, Harper & Row, New York.

Brennan, M. J. and E. S. Schwartz, 1984, "Optimal Financial Policy and Firm Valuation," *Journal of Finance*, 39, 593–607.

Brounen, D., A. de Jong and K. Koedijk, 2004, "Corporate Finance in Europe: Confronting Theory with Practice," *Financial Management*, 71–101.

Butters, J. K., 1949, "Federal Income Taxation and External vs. Internal Financing," *Journal of Finance*, 4, 197–205.

Caballero, R. J. and E. Engel, 2003, "Adjustment Is Much Slower than You Think," Working paper, MIT.

Cadsby, C. B., M. Z. Frank and V. Maksimovic, 1998, "Equilibrium Dominance in Experimental Financial Markets," *Review of Financial Studies*, 11, 189–232.

Cadsby, C. B., M. Z. Frank and V. Maksimovic, 1990, "Pooling, Separating, and Semiseparating Equilibria in Financial Markets: Some Experimental Evidence," *Review of Financial Studies*, 3, 315–342.

Calomiris, C. W. and R. G. Hubbard, 1995, "Internal Finance and Investment: Evidence from the Undistributed Profits Tax of 1936–1937," *Journal of Business*, 68, 443–482.

Campbell, J. Y., A. W. Lo and A. C. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton N. J.

Chang, X. and S. Dasgupta, 2007, "Target Behavior and Financing: how Conclusive is the Evidence?," Working Paper, University of Melbourne and Hong Kong University of Science and Technology.

Chang, X., S. Dasgupta and G. Hillary, 2006, "Analyst Coverage and Financing Decisions," *Journal of Finance*, 61, 3009–3048.

Chen, L. and X. Zhao, 2005, "Firm Financing Decisions," Working Paper, Michigan State University and Kent State University.

Chetty, R. and E. Saez. "Dividend Taxes and Corporate Behavior: Evidence from the 2003 Dividend Tax Cut," *Quarterly Journal of Economics*, 120, 791–833.

Chirinko, R. and A. Singha, 2000, "Testing Static Trade-off Against Pecking Order Models of Capital Structure: A Critical Comment," *Journal of Financial Economics*, 58, 417–425.

Christie, W. G. and V. Nanda, 1994, "Free Cash Flow, Shareholder Value, and the Undistributed Profits Tax of 1936 and 1937," *Journal of Finance*, 49, 1727–1754.

Crutchley, C. E. and R. S. Hansen, 1989, "A Test of the Agency Theory of Managerial Ownership, Corporate Leverage, and Corporate Dividends," *Financial Management*, 18, 36–46.

Dann, L. and W. Mikkelson, 1984, "Convertible Debt Issuance, Capital Structure Change and Financing-Related Information: Some New Evidence," *Journal of Financial Economics*, 13, 157–186.

DeAngelo, H. and R. Masulis, 1980, "Optimal Capital Structure under Corporate and Personal Taxation," *Journal of Financial Economics*, 8, 3–29.

Dittmar, A., 2004, "Capital Structure in Corporate Spinoffs," *Journal of Business*, 77, 9–43.

Donaldson, C., 1961, *Corporate Debt Capacity*, Harvard University Press, Cambridge, MA.

Dybvig, P. and J. Zender, 1991, "Capital Structure and Dividend Irrelevance with Asymmetric Information," *Review of Financial Studies*, 4, 201–220.

Eckbo, B. E., 1986, "Valuation Effects of Corporate Debt Offerings," *Journal of Financial Economics*, 15, 119–151.

Eckbo, B. E., R. Giammarino and R. Heinkel, 1990, "Assymetric Information and the Medium of Exchange in Takeovers: Theory and Tests," *Review of Financial Studies*, 3, 651–675.

Eckbo, B. E. and R. W. Masulis, 1995, "Seasoned Equity Offerings: A Survey," In Robert A. Jarrow, Vojislav Maksimovic, and William Ziemba (eds.) *Finance*, Vol. 9 of Handbooks in Operation Research and Management Science, North-Holland, Chapter 31, 1017–1072.

Eckbo, B. E. and R. W. Masulis, 1992. "Adverse Selection and the Rights Offer Paradox," *Journal of Financial Economics*, 32, 293–332.

Eckbo, B. E., R. W. Masulis and O. Norli, 2007, "Security Offerings," In B. E. Eckbo (ed.) *Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1*. Chapter 6. in Handbook of Finance Series, Elsevier/North-Holland, Amsterdam.

Eckbo, B. E. and O. Norli, 2004, "The Choice of Seasoned-Equity Selling Mechanisms: Theory and Evidence," Working Paper, Dartmouth.

Fama, E. and K. R. French, 2005, "Financing Decisions: Who Issues Stock?" *Journal of Financial Economics*, 76, 549–582.

Fama, E. and K. R. French, 2002, "Testing Trade-Off and Pecking Order Predictions About Dividends and Debt," *Review of Financial Studies*, 15, 1–33.

Fama, E. and K. R. French, 2001, "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay?" *Journal of Financial Economics*, 60, 3–43.

Fama, E. and K. R. French, 1998. "Taxes, Financing Decisions and Firm Value," *Journal of Finance*, 53, 819–843.

Faulkender, M. and M. A. Petersen, 2006, "Does the Source of Capital Affect Capital Structure?" *Review of Financial Studies*, 19, 45–79.

Federal Reserve, 2003, Flow of Funds Accounts, release March 6, 2003.

Fischer, E., R. Heinkel and J. Zechner, 1989, "Dynamic Capital Structure Choice: Theory and Tests," *Journal of Finance*, 44, 19–40.

Flannery, M. and K. Rangan, 2006, "Partial Adjustment Towards Target Capital Structures," *Journal of Financial Economics*, 79, 469–506.

Frank, M. Z. and V. K. Goyal, 2007a, "Capital Structure Decisions: Which Factors are Reliably Important?" Working Paper, University of Minnesota and Hong Kong University of Science and Technology.

Frank, M. Z. and V. K. Goyal, 2007b, "Corporate leverage: How much do managers really matter?" Working Paper, University of Minnesota and Hong Kong University of Science and Technology.

Frank, M. Z. and V. K. Goyal, 2004, "The Effect of Market Conditions on Capital Structure Adjustment," *Finance Research Letters*, 1, 47–55.

Frank, M. Z. and V. K. Goyal, 2003, "Testing the Pecking Order Theory of Capital Structure," *Journal of Financial Economics*, 67, 217–248.

Gilson, S. C., 1997, "Transactions Costs and Capital Structure Choice: Evidence from Financially Distressed Firms," *Journal of Finance*, 52, 161–196.

Givoly, D., C. Hahn, A. Ofer and O. H. Sarig, 1992, "Taxes and Capital Structure: Evidence from Firms' Response to the Tax Reform Act of 1986," *Review of Financial Studies*, 5, 331–355.

Goldstein, R., N. Ju and H. Leland, 2001, "An Ebit-based Model of Dynamic Capital Structure," *Journal of Business*, 74, 483–512.

Gomes, A. R. and G. M. Phillips, 2005, "Why Do Public Firms Issue Private and Public Securities," NBER Working Paper No. W11294.

Goyal, V. K., K. M. Lehn and S. Racic, 2002, "Growth Opportunities and Corporate Debt Policy: The Case of the U.S. Defense Industry," *Journal of Financial Economics*, 64, 35–59.

Graham, J. R., 2003, "Taxes and Corporate Finance: A Review," *Review of Financial Studies*, 16, 1075–1129.

Graham, J. R., 2000, "How Big are the Tax Benefits of Debt?" *Journal of Finance*, 55, 1901–1941.

Graham, J. R. and C. Harvey, 2001, "The Theory and Practice of Corporate Finance: Evidence from the Field," *Journal of Financial Economics*, 60, 187–243.

Graham, J. R. and K. Narasimhan, 2004, "Corporate Survival and Managerial Experiences During the Great Depression," Working Paper, Duke University and University of Pennsylvania.

Graham, J. R. and A. L. Tucker, 2006, "Tax shelters and corporate debt policy," *Journal of Financial Economics*, 81, 563–594.

Green, R. and B. Hollifield, 2003, "The Personal-Tax Advantages of Equity," *Journal of Financial Economics*, 67, 175–216.

Hackbarth, D., 2007, "Managerial Traits and Capital Structure Decisions," *Journal of Financial and Quantitative Analysis, Forthcoming*.

Halov, N. and F. Heider, 2005, "Capital Structure, Risk and Asymmetric Information," Working Paper, New York University and ECB.

Harris, M. and A. Raviv, 1991, "The Theory of Capital Structure," *Journal of Finance*, 46, 297–356.

Haugen, R. A. and L. W. Senbet, 1978, "The Insignificance of Bankruptcy Costs to the Theory of Optimal Capital Structure," *Journal of Finance*, 33, 383–393.

Helwege, J. and N. Liang, 1996, "Is There a Pecking Order? Evidence from a Panel of IPO Firms," *Journal of Financial Economics*, 40, 429–458.

Hennessy, C. A. and T. A. Whited, 2005, "Debt Dynamics," *Journal of Finance*, 60, 1129–1165.

Hirshleifer, J., 1966, "Investment Decision under Uncertainty: Applications of the State-Preference Approach," *Quarterly Journal of Economics*, 80, 252–277.

Hirshleifer, J., 1958, "On the Theory of Optimal Investment Decisions," *Journal of Political Economy*, 66, 329–352.

Holmstrom, B. and S. N. Kaplan, 2001, "Corporate Governance and Takeovers in the U.S.: Making Sense of the 80s and 90s," *Journal of Economic Perspectives*, 121–144.

Hovakimian, A., T. Opler and S. Titman, 2001, "The Debt-Equity Choice," *Journal of Financial and Quantitative Analysis*, 36, 1–24.

Huang, R. and J. R Ritter, 2007, "Testing Theories of Capital Structure and Estimating the Speed of Adjustment," *Journal of Financial and Quantitative Analysis*, Forthcoming.

Jalilvand, A. and R. S. Harris, 1984, "Corporate Behavior in Adjusting to Capital Structure and Dividend Targets: An Econometric Study," *Journal of Finance*, 39, 127–145.

Jensen, M. C., 1986, "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review*, 76, 323–329.

Jensen, M. C. and W. H. Meckling, 1976, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3, 305–360.

Ju, N., R. Parrino, A. M. Poteshman and M. S. Weisbach, 2005, "Horses and Rabbits? Trade-off Theory and Optimal Capital Structure," *Journal of Financial and Quantitative Analysis*, 40, 259–281.

Jung, K., Y. Kim and R. Stulz, 1996, "Timing, Investment Opportunities, Managerial Discretion, and the Security Issue Decision," *Journal of Financial Economics*, 42, 159–185.

Kane, A., A. J. Marcus and R. L. McDonald, 1984, "How Big is the Tax Advantage to Debt?," *Journal of Finance*, 39, 841–853.

Kayhan, A. and S. Titman, 2007, "Firms' Histories and Their Capital Structures," *Journal of Financial Economics*, 83, 1–32.

Kemsley, D. and D. Nissim, 2002, "Valuation of the Debt Tax Shield," *Journal of Finance*, 57, 2045–2073.

Korajczyk, R. A. and A. Levy, 2003, "Capital Structure Choice: Macroeconomic Conditions and Financial Constraints," *Journal of Financial Economics*, 68, 75–109.

Korajczyk, R. A., D. Lucas and R. McDonald, 1990, "Understanding Stock Price Behavior Around the Time of Equity Issues," in R. Glenn Hubbard (Ed.), *Asymmetric Information, Corporate Finance, and Investment*, University of Chicago Press, Chicago.

Kraus, A. and R. H. Litzenberger, 1973, "A State-Preference Model of Optimal Financial Leverage," *Journal of Finance*, 33, 911–922.

Leary, M. T. and M. R. Roberts, 2007, "The Pecking Order, Debt Capacity, and Information Asymmetry," Working Paper, Cornell University and University of Pennsylvania.

Leary, M. T. and M. R. Roberts, 2005, "Do Firms Rebalance their Capital Structures?" *Journal of Finance*, 60, 2575–2619.

Lemmon, M. L. M. R. Roberts, and J. F. Zender, 2007, "Back to the Beginning: Persistence and the Cross-Section of Corporate Capital Structure," *Journal of Finance*, Forthcoming.

Lemmon, M. L. and J. F. Zender, 2004, "Debt Capacity and Tests of Capital Structure Theories," Working Paper, University of Utah and University of Colorado at Boulder.

Lewellen, J. and K. Lewellen, 2006, "Internal Equity, Taxes, and Capital Structure," Working Paper, Dartmouth College.

Little, R. J. A. and D. B. Rubin, 2002, *Statistical Analysis with Missing Data*, 2nd edition John Wiley & Sons, NewYork.

Long, M. S. and I. B. Malitz, 1985, "The Investment-Financing Nexus: Some Empirical Evidence," *Midland Finance Journal*, 53–59.

Mackay, P. and G. M. Phillips, 2005, "How Does Industry Affect Firm Financial Structure?" *Review of Financial Studies*, 18, 1433–1466.

Maksimovic, V. and G. M. Phillips, 1998, "Asset Efficiency and the Reallocation Decisions of Bankrupt Firms," *Journal of Finance*, 53, 1495–1532.

Malmendier, U., G. A. Tate and J. Yan, 2005, "Corporate Financial Policies with Overconfident Managers," Working Paper, Stanford University and University of Pennsylvania.

Marsh, P., 1982, "The Choice between Equity and Debt: An Empirical Study," *Journal of Finance*, 37, 121–144.

Masulis, R. W., 1980ab. "The Effects of Capital Structure Change on Security Prices: A Study of Exchange Offers," *Journal of Financial Economics*, 8, 139–177.

Masulis, R. W. and A. Korwar, 1986, "Seasoned Equity Offerings: An Empirical Investigation," *Journal of Financial Economics*, 15, 91–118.

Mauer, D. C. and A. J. Triantis, 1994, "Interaction of Corporate Financing and Investment Decisions: A Dynamic Framework," *Journal of Finance*, 49, 1253–1277.

Mayer, C. and O. Sussman, 2004, "A New Test of Capital Structure," Working Paper, Oxford University.

McDonald, R. L., 2004, "Portfolio Choice and Corporate Financial Policy When There are Tax-intermediating Dealers," Working Paper, Northwestern University.

Mehrotra, V., W. Mikkelson and M. Partch, 2003, "The Design of Financial Policies in Corporate Spin-offs," *Review of Financial Studies*, 16, 1359–1388.

Mello, A. S. and J. E. Parsons, 1992, "Measuring the Agency Cost of Debt," *Journal of Finance*, 47, 1887–1904.

Mikkelson, W. H. and M. M. Partch, 1986, "Valuation Effects of Security Offerings and the Issuance Process," *Journal of Financial Economics*, 15, 31–60.

Miller, M. H., 1977, "Debt and Taxes," *Journal of Finance*, 23, 261–275.

Minton, B. A. and K. H. Wruck, 2001, "Financial Conservatism: Evidence on Capital Structure from Low Leverage Firms," Working Paper No. 2001–6, University of Ohio.

Modigliani, F. and M. H. Miller, 1963, "Corporate Income Taxes and the Cost of Capital: A Correction," *American Economic Review*, 53, 433–443.

Modigliani, F. and M. H. Miller, 1958, "The Cost of Capital, Corporate Finance and the Theory of Investment," *American Economic Review*, 48, 261–297.

Molina, C. A., 2005, "Are Firms Underleveraged? An Examination of the Effect of Leverage on Default Probabilities," *Journal of Finance*, 60, 1427–1459.

Morellec, E., 2004, "Can Managerial Discretion Explain Observed Leverage Ratios?" *Review of Financial Studies*, 17, 257–294.

Myers, S. C., 2003, "Financing of Corporations," In G. Constantinides, M. Harris, and R. Stulz (Eds.), *Handbook of the Economics of Finance: Corporate Finance*, Elsevier North-Holland.

Myers, S .C., 1984, "The Capital Structure Puzzle," *Journal of Finance*, 39, 575–592.

Myers, S. C., 1977, "Determinants of Corporate Borrowing," *Journal of Financial Economics*, 5, 147–175.

Myers, S. C. and N. S. Majluf, 1984, "Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have," *Journal of Financial Economics*, 13, 187–221.

Noe, T., 1988, "Capital Structure and Signaling Game Equilibria," *Review of Financial Studies*, 1, 321–355.

Parrino, R. and M. Weisbach, 1999, "Measuring Investment Distortions Arising from Stockholder-Bondholder Conflicts," *Journal of Financial Economics*, 53, 3–42.

Petersen, M., 2007, "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," *Review of Financial Studies*, Forthcoming.

Rajan, R. and L. Zingales, 1995, "What Do We Know About Capital Structure: Some Evidence from International Data," *Journal of Finance*, 50, 1421–1460.

Ravid, S. A. and M. Spiegel, 1997, "Optimal Financial Contracts for a Start-up with Unlimited Operating Discretion," *Journal of Financial and Quantitative Analysis*, 32, 269–286.

Ritter, J. R. and I. Welch, 2002, "A Review of Ipo Activity, Pricing, and Allocations," *Journal of Finance*, 57, 1795–1828.

Rubinstein, M., 2003, "Great Moments in Financial Economics: Ii. Modigliani-Miller Theorem," *Journal of Investment Management*, 1, 2.

Schwartz, E., 1959, "Theory of the Capital Structure of the Firm," *Journal of Finance*, 14, 18–39.

Shyam-Sunder, L. and S. C. Myers, 1999, "Testing Static Tradeoff Against Pecking Order Models of Capital Structure," *Journal of Financial Economics*, 51, 219–244.

Smith, C. W. and R. L. Watts, 1992, "The Investment Opportunity Set and Corporate Financing, Dividend and Compensation Policies," *Journal of Financial Economics*, 32, 263–292.

Stiglitz, J. E., 1973, "Taxation, Corporate Financial Policy and the Cost of Capital," *Journal of Public Economics*, 2, 1–34.

Stiglitz, J. E., 1969, "A Re-examination of the Modigliani-Miller Theorem," *American Economic Review*, 59, 784–793.

Strebulaev, I. A., 2007, "Do Tests of Capital Structure Theory Mean What They Say?" *Journal of Finance*, 62, 1747–1787.

Sufi, A., 2007, "Information Asymmetry and Financing Arrangements: Evidence from Syndicated Loans," *Journal of Finance*, 62, 629–668.

Taggart, R. A., Jr. 1985, "Secular Patterns in the Financing of U.S. Corporations," In B. M. Friedman, *Corporate Capital Structures in the United States*, University of Chicago Press, New York.

Taggart, R. A., Jr. 1977, "A Model of Corporate Financing Decisions," *Journal of Finance*, 32, 1467–1484.

Titman, S., 1984, "The Effect of Capital Structure on a Firm's Liquidation Decision," *Journal of Financial Economics*, 13, 137–151.

Titman, S. and S. Tsyplakov, 2007, "A Dynamic Model of Optimal Capital Structure," *Review of Finance*, 11, 401–451.

Titman, S. and R. Wessels, 1988, "The Determinants of Capital Structure Choice," *Journal of Finance*, 43, 1–21.

Tserlukevich, Y., 2006, "Can Real Options Explain Financing Behavior?" Working Paper, Hong Kong University of Science and Technology.

Viswanath, P. V., 1983, "Strategic Considerations, the Pecking Order Hypothesis, and Market Reactions to Equity Financing," *Journal of Financial and Quantitative Analysis*, 28, 213–234.

Warner, J. B., 1977, "Bankruptcy Costs: Some Evidence," *Journal of Finance*, 32, 337–347.

Welch, I., 2004, "Capital Structure and Stock Returns, *Journal of Political Economy*, 112, 106–131.

Weston, J. F., 1955, "Toward Theories of Financial Policy," *Journal of Finance*, 10, 130–143.

Williams, J. B., 1938, *The Theory of Investment Value*, Harvard University Press, Cambridge, MA.

Wright, S., 2004, "Measures of Stock Market Value and Returns for the US. Nonfinancial Corporate Sector, 1900–2002," *Review of Income and Wealth*, 50, 561–584.

Yen, S., 2005, "Growth Opportunities and Governance Structure Choices," Working Paper, City University of Hong Kong.

*Chapter 13*

# CAPITAL STRUCTURE AND CORPORATE STRATEGY

CHRIS PARSONS

*McGill University*

SHERIDAN TITMAN

*University of Texas*

## Contents



*Handbook of Empirical Corporate Finance, Volume 2*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

**Abstract**

In this chapter we review and discuss empirical studies that examine how a firm's financing choice affects its strategic decisions and relationships with its nonfinancial stakeholders, such as its customers or workforce. Generally, high leverage appears to inhibit a firm's ability or willingness to compete aggressively, especially against well-financed competitors. Debt also disciplines the manager–worker relationship, preventing managers from hoarding labor during economic downturns. Many of the studies also indicate that the firm's relationships with its customers can be disrupted by concerns over the firm's long term viability. A second purpose of this study is to highlight and discuss approaches researchers have taken to address endogeneity. Because the firm chooses leverage in advance, most of the studies we consider focus on exogenous shocks—either to the firm's competitive environment or to its leverage ratio. For each study, we describe the particular endogeneity problem and then discuss each author's approach to it, emphasizing differences between approaches when they arise.

## 1. Introduction

The connection between how a firm is financed and its other business choices is one of the most intensely examined issues in corporate finance. When considering this issue, it is convenient to think about a firm's decisions as generally falling into one of two broad categories—investment activity and corporate strategy. While in practice these decisions are linked, thinking about a firm's investment activity and corporate strategy separately allows us to better isolate the effects of capital structure on each.[1] The focus of this chapter is on the interaction between the firm's capital structure choice and corporate strategy issues that go beyond the investment choice. In particular, we consider how capital structure affects a firm's interaction with its nonfinancial stakeholders and competitors.

We define *nonfinancial stakeholders* as parties that have either a direct or indirect interest in the firm's long-term viability. These parties include customers who may be concerned about the quality of the firm's product or may anticipate additional interaction with the firm after an initial purchase, workers who develop firm-specific human capital, and suppliers who may require an investment in relationships with the firm. Each of these parties may demand compensation for the costs they will bear if the firm goes out of business, thus imposing "financial distress" costs on a firm whose capital structure introduces the possibility of bankruptcy. In addition, a firm can also face financial distress costs that arise from the actions of its *competitors*, who may choose to compete more aggressively when the firm is financially weakened.

These theories suggest that financial distress can be costly, and that the potential for incurring these financial distress costs will influence the firm's capital structure choices. Moreover, the theories suggest that a firm's financial condition will influence both its own and its stakeholders' strategic decisions. Although the majority of our discussion focuses on identifying and quantifying how leverage affects these strategic decisions, it is useful to think about this as a key step in determining optimal capital structure. For example, we will later see that firms undertaking leveraged buyouts (LBOs) may, in some circumstances, face more intense competitive pressure from rivals. The standard trade-off theory (discussed in detail by Frank and Goyal in Chapter 12) prescribes that firms take these additional pressures into account when deciding whether or not to follow through with the LBO. Similarly, if an LBO increases competition from a firm's rivals, then will a LBO firm have an incentive to stockpile cash rather pay out dividends?

[1] It has long been recognized that a firm's debt-to-equity mix can influence both its ability to invest and the characteristics of the projects it selects (Jensen and Meckling, 1976; Myers, 1977). When bankruptcy is possible, the incentives of those directing the firm's investment decisions may not lead them to make choices that maximize firm value. Examples of such conflicts of interest between equity holders and creditors are well known in the literature: asset substitution, debt overhang, choice of inferior investments with short payoff horizons, and refusal to liquidate. We refer to these costs as *investment* inefficiencies brought about by the potential for bankruptcy.

The stakeholder/competition theories have been tested in two different ways. The first type of test uses firm characteristics to proxy for the firm's sensitivity to stakeholder concerns, and it examines whether these characteristics are associated with lower leverage ratios. Most notably, Titman and Wessels (1988) find that firms with more unique or specialized products, as measured by high R&D/sales and selling expenses/sales ratios, tend to be less levered. This evidence suggests that less debt is used when financial distress is likely to be particularly costly to customers, suppliers, and workers.

The second type of study, which is the main focus of this review, investigates the extent to which leverage affects the relationships between firms and their stakeholders and competitors. This more recent literature examines how debt affects wages, the level of employment, sales, and product market prices. Although many of these studies present evidence of financial distress costs, it is often difficult to identify the specific sources of these costs. For example, although a financially distressed firm may experience declining sales and market share, these could potentially be caused by aggressive rival firms or by customers wary of doing business with a struggling company. In some cases, the empirical methodology allows us to directly identify which of a firm's specific relationships are impacted by financial distress; in others, the results may only be suggestive.

We begin with a group of studies that specifically focus on how debt affects a firm's relationship with its workers. Studies by Bronars and Deere (1991), Sharpe (1994), and Hanka (1998) focus on slightly different aspects of this relationship. The studies by Sharpe and Hanka primarily demonstrate that debt disciplines management by reducing a firm's ability to "hoard" labor during bad times, thereby increasing the sensitivity of the firm's employment to demand. Bronars and Deere present a model in which debt protects shareholder wealth by reducing the funds potentially available to labor unions; these authors then present empirical evidence supporting the contention that debt is used to deter the threat of unionization.

We then consider papers that address how financial distress can arise from disruptions in the firm–customer relationship. As noted originally by Titman (1984), a firm whose products require future servicing or maintenance may be particularly concerned about how its capital structure influences customer perceptions about its long-term health. We review two studies that contribute to our understanding of how debt affects the firm–customer relationship—Opler and Titman (1994) and Zingales (1998). While each paper presents evidence consistent with a customer-based explanation for sales and market-share declines, both explicitly consider the interaction between capital structure and competition as well.[2]

[2] "A notable omission from the the set of stakeholder-firm relationships we examine herein is that between a firm and its suppliers. Only very recently (since the final draft of this review article and its publication) has a study specifically explored this issue. In a forthcoming paper, Banerjee, Dasgupta, Kim (2007) examine both sides of a relationship where a customer constitutes a significant portion of a supplier's sales (generally 10% or more). The authors find that in such situations, both the supplier and customer maintain low leverage ratios—the former to maintain financial slack in case the relationship deteriorates, and the latter to induce customer-specific investments from the dependent supplier. Importantly, this finding is only present among industries producing durable products, where long-lasting relationships between customers and suppliers are likely to be important."

As already mentioned, a firm can also face financial distress costs through the actions of its *competitors*, since excess debt may make the firm more vulnerable to the predation of its rivals. Most studies that examine this possibility consider shocks to either leverage or competition. Phillips (1995) focuses on sharp increases in leverage in four manufacturing industries, and Chevalier (1995a, 1995b) and Chevalier and Scharfstein (1996) analyze the prices at local supermarkets after LBOs. In each of these studies, debt is shown to influence the prices charged for products, the market shares gained or lost, or the probability of exit or entry into the market. Perhaps more importantly, because of both the nature of the products and the empirical design, the observed changes in sales, market shares, or market presence are likely to be due to interactions with other firms, as opposed to interaction with nonfinancial stakeholders (such as customers).

Complementing these studies that examine competitive responses to dramatic changes in capital structure, Khanna and Tice (2000, 2005) examine competitive responses to competition shocks. The authors investigate how discount department retailers respond when Wal-Mart enters their market, with an eye on what determines the incumbent's ability to compete. Finally, Campello (2003) documents the effects of capital structure on product market competition for a large cross section of industries over a number of years.

A secondary purpose of this survey is to discuss endogeneity problems that arise in empirical corporate finance research in general and to describe how researchers studying capital structure explicitly deal with this problem. As we will see, virtually all empirical studies that attempt to shed light on the connection between capital structure and a firm's corporate strategy potentially suffer from significant endogeneity problems. Indeed, one theme we consistently find in these studies is a careful approach to endogeneity.

We begin our survey of the interaction between capital structure and its effect on real decisions by the firm with a brief review of the classical endogeneity problem in regression, after which we discuss ways in which academic researchers have addressed the issue. In contrast to the work by Li and Prabhala (Chapter 2), who provide a more detailed treatment of econometric techniques that address these endogeneity issues, our discussion is intended to provide economic intuition of the causes and solutions to endogeneity problems. As such, the "solutions" to endogeneity we discuss are often more clever than technically complicated, exploiting opportunities to approach the problem in such a way that mitigates the potentially confounding affects of endogeneity.

## 2. Endogeneity

Endogeneity is probably the most significant problem plaguing researchers in empirical corporate finance. Statistically, endogeneity means that the model's errors are not truly *random*, since they are partially predictable from information contained in the explanatory variables. Practically, endogeneity means that a regression is misspecified in a way that makes identifying a causal effect between two economic variables difficult, if not impossible.

How does endogeneity arise? One way is through reverse causality. For example, consider a regression with firm sales as the dependent variable and potential determinants of sales (including leverage) on the right-hand side. Although it is certainly possible that leverage can influence sales, it is *also* possible that sales can influence leverage. If poorly performing firms raise more debt capital than their better performing counterparts, then a negative relationship will arise, but not for the reason suspected.

Another related type of endogeneity can arise when the researcher is not able to control for important determinants of the dependent variable. For example, suppose that one wished to explain the debt ratios of firms in a cross section (as in Titman and Wessels, 1988). One of the posited determinants of leverage is the uniqueness of the firm's products, one proxy for which is the firm's ratio of research-and-development (R&D) costs to sales. However, another determinant of leverage is a firm's growth opportunities, since these firms may have the incentive and ability to engage in costly asset substitution; more importantly, growth firms often invest heavily in R&D. These observations in tandem highlight why including *only* R&D-to-sales may prove difficult to interpret. If the coefficient on R&D-to-sales is negative, is it because the uniqueness of the firm's products imposes costs on the firm's customers, or is it because the firm's flexibility in investment imposes potential costs on would-be creditors?

The analysis that follows explores endogeneity problems of this sort, along with the solutions various authors offer to mitigate potential bias. In many cases, the authors take advantage of an exogenous shock that is unlikely to be correlated with most (or any) of the potentially endogenously determined variables in the system. In others, careful instruments for the endogenous variable are employed in a reduced-form model that is free of bias. Some authors argue that transformed variables (often in lagged form) remove the potential for an endogenous relationship. As will become clear, there are more often than not several potential sources of endogeneity in a single study; many authors choose to deal directly with the most prominent endogeneity threats, treating the remainder with robustness checks. One feature of endogeneity is that it is often far easier to recognize than to adequately treat. Indeed, the approaches we examine often differ significantly, reflecting the difficulty faced in treating potentially endogenous relationships.

## 3. The determinants of capital structure choice

If a firm's leverage is affected by its relationships with nonfinancial stakeholders and competitors, then cross-sectional differences in debt levels should be observed among firms that differ in their sensitivities to these relationships. For example, firms that experience a deterioration in their stakeholder relationships when they encounter financial distress would, all else equal, be expected to choose lower debt ratios. A primary contribution of Titman and Wessels's (1988) study is to empirically document relations between a firm leverage choice and its attributes, many of which proxy for how sensitive the firm's stakeholders are to its financial distress.

The basic empirical framework is an application of the LISREL system originally developed by Karl Jöreskog and Dag Sörbom.[3] The main advantage of employing this framework is that it allows for debt ratios to be determined by a family of *unobservable* firm attributes, which are specified as linear functions of *observable* proxies. The model includes eight attributes that the authors identify as potentially affecting a firm's leverage ratio: collateral value of assets, nondebt tax shields, growth, uniqueness, industry classification,[4] size, volatility, and profitability. For the eight unobservable attributes, the authors specify fifteen observable variables (obtained mainly from accounting data for each firm), each of which proxies for one or more of the attributes.[5] The analysis can be thought of as proceeding in two steps. First, the authors measure each of the firm's eight attributes by relating them to the observable proxy variables (the measurement model). Second, observed leverage ratios are related to each of these measured attributes (the structural model). Both steps are estimated simultaneously.

The authors measure short-term, long-term, and convertible debt (each divided by either book or market value of equity), regressing each dependent variable separately on the family of measured attributes defined in the measurement model. With regard to short- and long-term debt, only uniqueness proved a statistically significant determinant of leverage in each specification of the model. Firms offering unique products, as measured by the ratios of R&D, selling expense to sales, and labor quit rates exhibit lower debt ratios, whether measured in relation to either book or market values of equity. In addition, although an industry dummy, intended to measure the required service and maintenance associated with a firm's products, was generally statistically insignificant (with one exception), the point estimate of its partial effect was negative in every specification of either short- or long-term debt. The authors interpret the combination of this evidence as supportive of Titman (1984), which predicts lower debt ratios for firms whose liquidation imposes significant costs on its workers, customers, and suppliers.

One of the indicators of uniqueness—how often workers voluntarily left their jobs (quit rates)—reliably predicted debt ratios; firms in industries with high quit rates exhibit high leverage. One interpretation of this evidence is that in industries where workers quit frequently, financial distress is unlikely to be particularly costly to workers. We now proceed to a more recent class of studies that investigates in more detail the interaction between debt and this class of nonfinancial stakeholders.

[3] LISREL is an acronym for *LI* near *S*tructural *REL*ationships. Technically, LISREL is a computer program developed to do covariance structure analysis. Such covariance structural models are widely used in economics, marketing, and psychology Source: *Introducing LISREL, a Guide for the Uninitiated*, Diamantopoulos, A., Siguaw, J. (2002).

[4] The authors isolate with a dummy variable firms that produce heavy equipment and machinery, since these firms are likely to require future servicing and maintenance.

[5] The family of observable proxies includes information from the balance sheet (ratio of intangible assets to total assets, ratio of inventory plus gross plant and equipment to total assets), income statement (depreciation over total assets, R&D-to-sales ratio), and statement of cash flows (ratio of capital expenditures-to-total assets). Many proxies are allowed to proxy for more than one firm attribute.

*3.1. Debt and the firm–worker relationship*

Sharpe (1994) and Hanka (1998) each consider how debt affects a firm's relationship with its employees. Their basic findings are that firms with high leverage pay lower wages, fund pensions less aggressively, and provide less job security to their workers during downturns. Consistent with the preceding discussion regarding endogeneity, both studies carefully consider potentially omitted variables that may bias their estimates.

Sharpe considers this issue explicitly, acknowledging that although employment growth (one of the dependent variables he estimates) should be related to current and expected sales growth, sales growth is an endogenous variable that may also depend on employment. Alternatively, both sales and employment may mutually depend on factors unobserved by the econometrician. Sharpe estimates pooled regressions for multiple firms over time of the form:

$$\begin{aligned}\Delta E_{it} &= (\beta_1 + \beta_3\ \mathrm{LEV}_{i,t-2} + \beta_5\ \mathrm{SIZE}_{\,i,t-2})\Delta S_{i,t+1} \\ &\quad + (\beta_2 + \beta_4\ \mathrm{LEV}_{i,t-2} + \beta_6\ \mathrm{SIZE}_{\,i,t-2})\Delta S_{i,t} \\ &\quad + (\beta_0 + \beta_7\ \mathrm{LEV}_{i,t-2} + \beta_8\ \mathrm{SIZE}_{\,i,t-2}) + u_{i,t}\end{aligned} \tag{1}$$

in which $E$ refers to the number of employees at year end, LEV to book leverage, SIZE to inflation-adjusted capital stock, and $S$ to sales. All changes are divided by their initial levels. The point of estimating (1) is to investigate how leverage affects a firm's sensitivity of employment to its current and future sales. Sharpe's primary interest is to ask whether leverage or size affects a firm's tendency to "hoard labor" during downturns. The signs of $\beta_3$ and $\beta_4$ indicate whether employment changes in highly levered firms are more sensitive to shocks in current and future sales than those of their less levered counterparts. Similarly, coefficients $\beta_5$ and $\beta_6$ tell us whether a firm's size influences how sensitive its employment is to current and future sales shocks.

The preceding specification must address a number of endogeneity issues. One is that sales and employment growth are mutually dependent, since changes in employment can certainly cause changes in sales (think about reducing the size of the sales force). To address this issue, changes in sales, $\Delta S$, are regressed against a set of macroeconomic instruments that are presumably exogenous from the perspective of each firm. These instruments include changes in interest rates, ratios of inventories to sales, growth in industrial production, and the Consumer Price Index (CPI) inflation rate. By effectively asking whether firms with different leverage ratios react differently to changes in the business cycle (which are presumably not predictable by firms), Sharpe is able to isolate the effect of size and leverage on unexpected shocks to a firm's demand.

Sharpe's argument that size is largely exogenous seems reasonable; however, leverage is certainly not exogenous, having been chosen by management simultaneously with the firm's employment level. To address the problem of simultaneity, Sharpe uses lags of both book (alternatively market) leverage and size in his empirical model. The hope is that leverage *levels* chosen at least a year in the past are not correlated with *changes* in current employment, after controlling for current and future sales growth with macroeconomic instruments. Sharpe does *not*, however, assume that using lagged leverage ratios solves the endogeneity problem. Indeed, he acknowledges that firms may select their debt ratios based in part on the costs associated with adjusting their labor forces; that is, firms with more flexible labor forces may be less risky and thus able to choose higher debt ratios. We discuss this matter in more detail later in this chapter.

Sharpe estimates Equation (1), separately for durables and nondurables from 1959 to 1985. While it is indeed the case that leverage increases a firm's employment sensitivity to sales, the impact is more pronounced in the subsample of firms that produce durable goods.[6] In the durable category, a firm with zero debt lays off 8% of its workers in response to a 10% decline in sales, while a firm with 100% debt reduces its workforce by 11.5%. This finding is important in its own right. For example, one direct implication is that a highly levered firm may have difficulty attracting employees and inducing them to build firm-specific human capital if highly levered firms are viewed as providing more risky employment.

Sharpe acknowledges, however, that interpreting the impact of leverage on a firm's sales elasticity of employment should be done carefully. As mentioned previously, all of Sharpe's findings are consistent with firms optimally choosing capital structures that reflect their labor-adjustment costs. For example, as we mentioned earlier, a firm's debt choice may reflect the *specificity* of the human capital it affords its employees. If the knowledge and skills imparted to a firm's workers are transferable to other firms, then layoffs are not particularly costly events; therefore, firms offering fungible human capital may choose high-debt levels. Sharpe's evidence is also consistent with other interpretations, such as that advanced by Jensen (1986 1989). This alternative suggests that debt provides discipline to managers who incur psychological costs from laying off their workers during recessions. Sharpe's finding of asymmetry between hiring and firing (debt increases layoffs during recessions but does not increase hiring during expansions) is consistent with this interpretation.

Hanka motivates his work largely from Sharpe's study, claiming that although the results are suggestive, its design makes it difficult to infer increased layoff risk due to leverage. Again, endogeneity is the culprit, this time from omitted variables. What if, Hanka argues, the observed correlation between debt and employment reductions is due to factors that were not included as controls, such as poor historical performance or low-growth opportunities? Using a set of variables from 1950 to 1993 including wages, funding of pensions, and use of seasonal employees, Hanka augments Sharpe's analysis by explicitly controlling for determinants of these dependent variables. He finds that highly levered firms pay lower wages, are more likely to lay off their employees, and fund pensions less generously, conclusions that largely agree with those advanced by Sharpe.

[6] As Sharpe indicates, "if there is any effect of leverage or size on the cyclicality of employment and sales, it should be more easily detected among firms in industries for which cyclical fluctuations are a more dominant feature of the dynamics" (p. 1062).

Hanka's empirical tests use leverage as an explanatory variable to explain dependent variables that relate to employment. Similar to Sharpe's study, Hanka's tests suffer from two important potential sources of endogeneity: (1) reverse causality: leverage is selected by firms, perhaps taking into account the costs they face of changing the size of their workforces, and (2) omitted variables: employment is determined by a complex set of factors, for which many are difficult to control with proxies.

In some cases, it is possible to identify the likely direction of the endogeneity bias, which can potentially strengthen the results. For example, Hanka mentions that if firms select low-debt levels when facing high costs of financial distress, then high-debt levels should predict fewer employment reductions, since a choice of high leverage indicates a low probability of financial distress and accompanying layoffs. This would bias the results against finding the observed empirical results.[7] Regarding the omitted variables problem, Hanka is particularly careful to attempt to control for potential determinants of employment such as growth opportunities and performance, which may be correlated with leverage.

Hanka presents results from a censored regression of employment layoffs on various sets of controls.[8] In a model without controls for divestitures, performance, and growth opportunities, debt is seen to be positively related to layoffs. A firm that increases its debt from the 10th to the 90th percentile increases its layoffs by 140%. When controls for asset sales are added, this effect is cut in half; when controls from current and prior performance are added, the effect is halved yet again. While controls for operating efficiency and divestitures remove a significant amount of debt's impact on employment reductions, debt is still seen to play a role in a firm's employment policy.

As Hanka argues, however, debt's strong correlation with performance makes it "difficult to be sure that the effects of performance have been completely purged from the results." Hanka addresses this concern by forming portfolios based on performance (ROA) and debt. As he notes, if performance (rather than debt) is driving the results, then employment reductions should increase with declining performance, regardless of leverage. However, Hanka presents evidence that this is clearly *not* the case. While poorly performing firms do indeed lay off more employees than their better performing counterparts, the leverage effect is just as strong—often more so. For the most levered quartiles, firms with both better performance and lower debt exhibit less employment

[7] Sharpe mentions that reverse causality may occur if characteristics related to the dependent variable (cost of employment reductions) influence a firm's choice of a dependent variable (leverage). Hanka acknowledges that if firms facing low costs of reducing their workforces choose higher debt levels, then the resulting estimates may be biased. This type of endogeneity, unlike that mentioned in the paper's body, would bias the magnitude, but not the sign of the resulting estimate. The combination of the above arguments causes Hanka to argue that although "endogeneity may bias the magnitude of the empirical results, [it] cannot easily cause their sign to be opposite that of the true causal relation" (p. 252).

[8] Hanka argues that a Tobit specification—where employment increases are censored at zero—is appropriate, since the goal of the model is to measure debt's impact on a worker's probability of being laid off. Although such a model ignores variation with regard to *increases* in employment, Hanka mentions that the results are still significant without censoring.

reduction; since worse performance should increase employment reductions, it is hard to argue that debt fails to influence a firm's layoff decision, even after controlling for performance. Hanka thus addresses the potential endogeneity from omitted variables through both regression controls, as well as through nonparametric tests that apply dependent sorts on the potentially endogenous variable of interest.

Hanka also presents evidence that firms with higher debt ratios pay lower wages, after controlling for size, industry, changes in employment, and the fraction of assets depreciated (to capture life-cycle effects). A firm that increases its leverage from the 10th to 90th percentile pays about $2300 less annual wage per employee, which is slightly less than 8% of the average of $28,000. In accompanying tests, Hanka also shows that highly levered firms fund pensions less generously and that more levered firms are more likely to rely on seasonal employees.

Hanka's interpretation is that debt "disciplines" managers, forcing them to make choices that may be personally unpleasant. These results, as well as those in Sharpe, are thus consistent with Jensen's (1986) free cash flow theory. One may also be able to develop an explanation for these results based on Myers (1977) debt overhang theory; a highly levered firm may underinvest in its employees when they are financially distressed or financially constrained. Finally, these results are also consistent with the idea that firms with more flexible labor forces can handle higher debt loads. In other words, it may be the nature of the labor force that generates observed capital structures rather than the capital structures influencing employment policy. While both Hanka and Sharpe are aware of this endogeneity problem, there is no apparent way to unambiguously determine the direction of causality in this case.

In the next paper we discuss views the firm–employee relationship through the perspective of labor unions, showing how debt can also alter this dynamic in a way that benefits shareholders. Bronars and Deere (1991) develop a model in which union behavior depends on the firm's capital structure—debt induces unions to act less aggressively.[9] In the first of two model specifications, a labor union faces the choice of either forcing the firm into bankruptcy and then negotiating with creditors or accepting a lower wage. Because creditors are assumed to operate the firm with an efficiency loss, it can be shown that the union's optimal strategy is to accept lower wages and avoid bargaining over a smaller surplus. In a second specification, the union forms and sets its wage simultaneously. When bankruptcy is costly for workers,[10] "the union will moderate its demands in the face of outstanding debt." In either specification, debt shields funds that would otherwise flow from shareholders to workers; the empirical implication is that firms facing a greater threat of unionization use debt more aggressively.

Empirically, this can be written as:

$$DE_{fi} = X_{it}\beta + \gamma\pi_{fi} + \varepsilon_{fi}, \tag{2}$$

[9] See also Dasgupta and Sengupta (1993) and Perotti and Spier (1993).

[10] Bronars and Deere mention that such costs may arise from job loss or organizing another (perhaps unsuccessful) union drive.

where $DE_{fi}$ is firm $f$'s debt-to-equity ratio in industry $i$, $X_{it}$ is a vector of control variables, $\pi_{fi}$ is the (unobserved) probability that firm $f$ is unionized, and $\varepsilon_{fi}$ is an error term. Since whether or not a firm is unionized is a binary variable, Bronars and Deere use the *industry average* for unionization as a proxy for the threat of unionization at the firm level. Thus, two firms with differing union status within the same industry are treated as having identical threats of unionization.

Bronars and Deere face a potential bias due to the way both the threat of unionization and leverage are measured. Consider, for example, the null hypothesis that the threat of unionization has no impact on a firm's choice of leverage—in this case, firm leverage and the threat of unionization should be uncorrelated. However, when unions form unexpectedly, market values of equity decline (to reflect potential wealth transfers from equityholders to the union), which increase the measured values of firm leverage. Thus, if we compare unionization and debt ratios across industries, there will be a positive correlation due to the negative impact on equity value caused by unionization. In this case, the null hypothesis may still be valid, despite positive correlation between average unionization rates and leverage levels.

Bronars and Deere correct for this potential bias in two ways. First, they estimate a model that measures the equity loss when unionization campaigns are successful. An adjustment factor for each industry is calculated, based on that industry's unionization rate and the average fraction of equity lost after successful campaigns.[11] The second adjustment is based on the idea that higher wages earned by union workers come out of the pockets of equityholders. By taking the present value of the union rents and adding them back to the observed equity values, the authors produce estimates of equity values when facing only the *threat* of unionization, not its actual occurrence.

Although regression results using the second (that based on the union wage premium) adjustment yielded an insignificant coefficient on unionization rates, the majority of Bronars and Deere's results are strongly supportive of a positive relation between the threat of unionization and debt ratios.[12] In both regressions using the first adjustment to equity as well as those with an alternative leverage measure that does not require adjustments (debt-to-margin and debt-to-paid in capital), debt ratios are strongly positively related to the unionization rates in their industries.

### *3.2. Debt and the firm–customer relationship*

The remainder of the studies we examine in this chapter show how capital structure affects a firm's sales and/or market share. As we will see, these studies present us with at least

[11] The authors note that because they are adjusting the dependent variable (leverage) by an adjustment factor that is itself a function of one of the regressors (unionization rate), coefficients are downward biased.

[12] In Bronars and Deere's setting, it may be tempting to ask why labor unions do not explicitly bargain over the debt ratio. The reason is that labor laws specifically prohibit unions from negotiating over issues that do not directly influence workers.

two empirical hurdles. The first problem is endogeneity; although we are interested in debt's effect on a firm's sales and market share, we suspect that shocks to sales influence observed debt ratios. Second, even if we can properly infer the correct direction of causality, it may still be difficult to identify *how* the observed changes in performance occur (i.e., through customers or competitors). In this section, we focus on the aspects of two studies (Opler and Titman, 1994; Zingales, 1998) that relate to the firm–customer relationship.

The issues considered in these studies were originally explored by Altman (1984), who attempts to measure both the direct costs (i.e., those paid explicitly by the creditor(s) in the event of reorganization/liquidation such as legal fees) and indirect costs (e.g., losses of sales and foregone profits) of bankruptcy for a sample of 19 firms following bankruptcy filing. Altman interprets the observed loss in sales as evidence of financial distress costs; specifically, he "assumes that the prospect of bankruptcy will often lead to lower than expected earnings." However, Altman admits that his results are also consistent with the alternative interpretation that "lower than expected earnings could cause the management to declare bankruptcy."

Opler and Titman are primarily concerned with measuring financial distress costs in a way that mitigates this inference problem caused by reverse causality. They do so by identifying industries that are economically distressed rather than directly identifying financially distressed firms. They then measure the financial leverage of firms within these economically distressed industries two years prior to the industry's decline, and they assume that the more highly levered firms within these industries are more likely to be financially distressed than their more conservatively financed counterparts in these industries. Thus, if a firm's financial distress affects sales and other performance measures, the more highly levered firms should do worse than their less levered counterparts in these time periods.

Specifically, Opler and Titman examine data for the 20-year period spanning from 1972 to 1991, and they run regressions of the following form:

$$\begin{aligned}\text{Firm performance} = \alpha\ &+ \beta_1 \text{ Log(Sales)} + \beta_2 \text{ Lagged industry-adjusted profitability}\\ &+ \beta_3 \text{ Industry-adjusted investment/assets}\\ &+ \beta_4 \text{ Industry-adjusted asset sale rate}\\ &+ \beta_5 \text{ Distressed industry dummy}\\ &+ \beta_6 \text{ High leverage dummy}\\ &+ \beta_7 \text{ Distressed industry dummy X High leverage dummy} + \varepsilon\end{aligned}$$

in which a "distressed industry" takes a value of one when the industry median sales growth is negative and the median stock return is below $-30\%$. Firm performance is measured with industry-adjusted sales growth, industry-adjusted stock returns, and industry-adjusted operating income. Since financial distress costs are likely to be the highest for firms experiencing both economic (as measured by the distressed industry dummy) and financial distress (as measured by the high leverage dummy), the main interest is in the coefficient of the interaction term, given by $\beta_7$ above.

The coefficient estimates on leverage ($\beta_6$) are negative when the measures of performance are either industry-adjusted stock returns or industry-adjusted sales growth, indicating that highly leveraged firms performed worse than their peers with lower leverage even in *good* times. The interaction term is also statistically significant and negative in each empirical specification, indicating that this effect is magnified in downturns. Specifically, during industry declines, a firm in the most leveraged decile experiences industry-adjusted sales declines of 25% more than its peers in the least leveraged decile. When operating income is the dependent variable, the coefficient on the interaction term is no longer significant.

Throughout the paper, Opler and Titman express concern about three possible sources of endogeneity bias:

- **Self-selection**. First, it is likely that firms with the highest costs of financial distress choose the lowest debt ratios.[13] While it is impossible to correct for this bias in the data, the authors argue that endogeneity of this type would bias the estimates toward zero. In other words, the tests may underestimate the effect of financial distress.
- **Reverse causality**. Although high leverage levels may lead firms to experience poor performance, poor performance may also lead to higher observed leverage levels (either because distressed firms borrow more, or because their market values decline, which increases their leverage ratios). This is potentially the most serious endogeneity problem and is largely addressed by the empirical design itself. First, the study controls for past profitability. In addition, Opler and Titman compare the performance of firms with high and low leverage in industry downturns (which are largely unanticipated when leverage choices are made) and argue that if having a high-debt ratio is costly, then it should be more costly during industry downturns that can cause the highly levered firms to become distressed.
- **Omitted Variables**. Leverage may serve as a proxy for other firm characteristics that are difficult to control for. Suppose, the authors argue, that poorly run firms fail the fastest in industry downturns and that these firms are also highly leveraged. If this is the case, then even if financial distress is not costly, there will be a negative correlation between leverage and performance. Opler and Titman attempt to mitigate this concern by measuring leverage with book values, since its correlation with profitability is lower than when market values are used.

Although the results of these performance regressions suggest that financial distress is costly, Opler and Titman acknowledge that further analysis is warranted. They claim, for example, that sales declines could be evidence of "efficient downsizing" and that

[13] Although firms with different capital structures are likely to be different along some dimensions, there exists a theoretical rationale for identical firms to be financed differently. For example, in Shleifer and Vishny (1992) firms that are ex-ante identical choose different debt ratios.

the stock return evidence may be driven by a pure leverage effect.[14] To investigate this possibility, they ran the same regressions after trifurcating the sample along three dimensions: size, research and development as a percentage of sales, and industry concentration.[15] Consistent with Baxter (1967), Titman (1984), and Maksimovic and Titman (1991), Opler and Titman find that the interaction term is more pronounced (more negative) among firms that invest heavily in research and development. Since these firms are most likely to offer unique products, and those that benefit from a long-term relationship between the customer and firm, the evidence is consistent with sales declines arising from customers wary of doing business with a struggling firm that may not survive to service and maintain its products. This R&D effect is also found with respect to stock returns. Regarding competitor-driven financial distress costs, Opler and Titman report that sales and market values of equity decline in concentrated industries, where the benefits to predation are likely to be the highest. Finally, they examine asset sales, investment rates, and employment growth rates between firms with high leverage and those with low leverage during industry downturns. Neither asset sales nor investment of highly leveraged firms during industry downturns exceeds those of firms with lower leverage. This finding suggests that manager-driven cost-cutting is not likely to be the explanation for the poor performance differential.

Zingales's (1998) study of the trucking industry's deregulation in 1980 also attempts to separate losses driven by customers and those caused by predating rivals. Like many of the other studies we examine, Zingales recognizes the potential for endogeneity to bias results in studies of the interaction between product markets and leverage. In particular, he notes not only that reverse causality may arise if trucking firms are able to adjust their capital structures in anticipation of deregulation, but also that spurious correlation between leverage and performance can arise if unobservable characteristics are not adequately controlled for in the explanatory variables.

The event Zingales examines is the unanticipated[16] deregulation of the U.S. trucking industry between 1978 and 1979. During this time, the Interstate Commerce Commission (ICC) largely reversed its policy regarding new service applications into the industry and liberalized rate-setting practices. With such deregulation, trucking firms began to engage in intense price competition, which decreased their market values relative to the

[14] Consider, for example, two firms that differ only in their leverages. When industry returns are negative, the more highly levered firm will experience comparatively lower returns purely from leverage, even in the absence of financial distress costs.

[15] The authors also investigate asset sales, exit, employment, and investment, in hopes of shedding light on whether the observed sales declines could have arisen from management acting optimally in response to economic distress. The bulk of the evidence (employment being the exception) did not favor a manager-driven explanation.

[16] Because part of the paper's appeal is based on the exogeneity of the shock, Zingales is careful to defend the deregulation event against the criticism that it was anticipated. First, Zingales reviews the *S&P Outlook* newsletter during the years around his study, and he concludes that 1977 was the "watershed" date (his measures of leverage, therefore, were collected prior to this date). Second, the author notes that market values of equity in the trucking sector suffered serious declines from 1978 to 1980, evidence that the deregulation event was largely unanticipated by the market.

preregulation environment. The result was that, for most firms, their ratio of debt to market value of equity (largely represented by trucking firms' operating certificates) was exogenously increased,[17] which provides a natural experiment to examine the effects of capital structure on the product market environment.

The main class of results addresses whether more efficient firms were most likely to have survived the deregulation, regardless of their financing arrangement. The basic model has the following form:

$$Pr\{\text{survival in 1985}\}_i = f(X_i^{1977}, Lev_i^{1977}) + \varepsilon_i$$

where $X_i^{1977}$ refers to a vector of proxies for the level of operating efficiency, $Lev_i^{1977}$ is the net debt-to-capital ratio (calculated as total debt minus cash reserves divided by total debt plus equity), and $\varepsilon_i$ is a mean zero noise term. The elements of $X_i^{1977}$ intended to capture efficiency are the log of sales, fraction of intangible assets, return on sales, proportion of wages over total costs, and nine regional dummies. Since leverage is related to profitability, and since profitability affects survival, a spurious correlation between leverage and exit may arise if determinants of survival (other than leverage) are not included in the empirical model. Return on sales is intended to control for the efficiency of each motor carrier, a control that is crucial in defending a relation between eventual exit and ex-ante leverage. Size is included as another proxy for efficiency (since the largest firms may be the most efficient) and also serves to control for access to financing. The fraction of intangible assets is important because it potentially measures the monopoly rents enjoyed by a carrier prior to deregulation through its operating certificates. Finally, the proportion of a firm's operating expenses dedicated to wages and benefits provides a measure of the firm's sensitivity to union demands. As we have already seen, the threat of unionization is likely to affect a firm's choice of leverage (Bronars and Deere, 1991) and may also eventually affect its probability of survival. Regional dummies are included to account for potential heterogeneity across different geographical areas.

Zingales's regression indicates that highly levered firms, even after accounting for their operating efficiency, are less likely to survive deregulation than their more conservatively financed rivals. This result is robust to several specifications of operating performance and even to the ex-ante probability of default as developed by Altman (1973). More interesting perhaps are the results presented for different segments of the trucking industry, the less-than-truck (LTL) and truckload (TL) segments. The larger LTL segment provides smaller hauls (less than 10,000 lb), relying on large investments in hubs and distributional networks. Because the value of LTL firms depends largely on customer service and on developing relationships with clients, liquidation would likely be very costly for this segment.[18] In contrast, the TL segment provides shipment service

[17] Zingales notes that between 1977 and 1982, the market values of trucking firms fell substantially.

[18] Liquidation of an LTL firm would likely destroy much of the value derived from firm-specific investments made by customers and employees. In contrast to firms with more tangible assets, recovering value from intangible assets in liquidation would likely require additional investments of organizational capital.

with loads greater than 10,000 lb, and is characterized by more intense postregulation competition and easier-to-finance capital investments. Since the value of TL firms is derived mostly from heavy trucks and equipment (as opposed to the comparatively poor collateral value of LTL assets), financial distress should be less costly for TL firms.

Zingales runs the survival model for three subgroups, those that derived (1) less than 30% of the revenues from LTL shipments, (2) between 30 and 70% from LTL, and (3) more than 70% from LTL activity. Interestingly, only in the groups deriving significant revenues from LTL shipments—where service is important—did prior leverage levels negatively affect the probability of surviving the deregulation. LTL firms derive significant value from intangible assets—distributional networks, hubs, and customer relationships—which likely imposes significant costs on these firms in the event of a restructuring. In contrast, the trucks and trailers of TL firms are redeployed with comparative ease, increasing the appeal of workouts for firms in this segment. The differential impacts of firm leverage on liquidation for each segment highlight the important role of customer-driven financial distress. When LTL firms actually face a liquidation/restructuring decision, deterioration of the firm-customer relationship may have caused irreparable loss of value to the point where a workout is infeasible or impossible.

Zingales considers that debt may affect a firm's ability to compete (and therefore, to eventually survive) by either reducing its ability to invest or hampering its competitive position (either because of predation or deterioration in its relationships with nonfinancial stakeholders). Zingales indeed finds some evidence that debt leads to less investment. Specifically, in his examination of trucking firms' investments until 1980 (to prevent bias from liquidating firms dropping out of the sample), he finds that the most highly levered firms invested less than those with less debt. Furthermore, the linkage between investment and leverage was strongest in those firms eventually forced out of the industry; there was no statistically significant relation between debt and investment for those firms that survived the deregulation. While this evidence is important, Zingales notes that this evidence is in itself unable to distinguish between alternative stories that may generate negative correlation between leverage and investment. For example, some highly leveraged firms may have been viewed as viable firms by the market and were supplied the necessary funds to survive the deregulation. Others were not viewed as favorably, were unable to obtain the necessary financing to maintain investment, and were forced to exit the industry.

Zingales also shows that in addition to investing less aggressively, highly leveraged firms began charging lower prices, starting approximately two years after deregulation (1982). A number of explanations are consistent with this evidence: (1) well-financed rivals engage in predation (Telser, 1966; Bolton and Scharfstein, 1990), and (2) they are induced to compete more aggressively (Brander and Lewis, 1986), and (3) customers of highly leveraged firms demand compensation for the possibility of bankruptcy (Titman, 1984).[19] Unfortunately, the fact that trucking companies compete on a national

[19] It should be noted that Chevalier and Scharfstein (1996) and Dasgupta and Titman (1998), discussed later in this chapter, provide models in which more highly levered firms charge higher prices.

scale renders it virtually impossible to define a local market in which to examine prices. Thus, although predation may represent part of the explanation for the price declines after deregulation, Zingales is unable to directly test this hypothesis. Instead, Zingales argues that the distinction between the services provided by LTL and TL carriers allows him to test an implication that applies to only one of the above possible explanations.

Because so much of an LTL carrier's value is related to its customer service, Zingales argues that if price declines for highly leveraged firms are *most* significant in this sector, then this would serve as evidence in favor of customer-driven financial distress. In contrast to the other explanations that should affect the LTL and LT segments relatively equally, price declines attributable to customers wary of doing business with a potentially bankrupt firm should be disproportionately present in the segment where customer service is most important. Although a pool of all firms (both TL and LTL) exhibits the result that highly leveraged firms charge lower prices, Zingales splits the sample by percentage of revenue derived from each segment, showing that firms deriving significant (more than 30%) revenue from LTL activity are almost totally responsible for the observed declines. Zingales interprets this evidence that "leveraged carriers discount their services to compensate consumers for the risk associated with the probability of default of the carrier."

While this interpretation may indeed reflect the most important source of financial distress among LTL firms, one cannot completely eliminate the possibility that predation is playing an important role as well. The reason is that predation is most effective in situations where some type of capital is destroyed (i.e., where actual deterioration of a firm's business occurs, as opposed to temporary depression of prices). Since entry into the TL segment is relatively easy, it is difficult to imagine that predatory pricing would be particularly effective in this segment.

### *3.3. Debt and competition*

Until now, we have focused on how the possibility of financial distress alters a firm's relationship with its customers and employees. In this section, we shift our attention to how capital structure may influence a firm's ability or willingness to compete with its rivals. Most studies rely on natural experiments involving either the firm's product market environment or leverage—and sometimes both. Phillips (1995) and Chevalier (1995a, 1995b) empirically investigate the interaction between product market outcomes and capital structure by examining competitive responses to sharp increases in leverage. A subsequent group of studies by Scharfstein and Chevalier (1996), Khanna and Tice (2000, 2005), and Campello (2003) analyze shocks to competitive environments, exploring how differences in ex-ante capital structure are associated with differential responses and competitive outcomes. In both classes of studies, all these authors seek to investigate how debt influences a firm's position in its competitive environment, whether measured by pricing, market share, or likelihood of surviving.

Phillips's (1995) study examines sharp leverage increases in four industries—fiberglass, tractor trailers, polyethylene, and gypsum—in order to explore the impact on firms' competitive strategies as they relate to production, market share, and price setting. In each of the industries Phillips studied, the largest firm had undergone an LBO, resulting in at least a 25% increase in the debt-to-value ratio. Furthermore, each industry is relatively homogeneous; reducing the likelihood that differential product quality may influence the ensuing results. In nonparametric tests, Phillips shows that in three of the industries (fiberglass, tractor trailers, and polyethylene), the largest firms undertaking leveraged buyouts *lost* market share, as measured by three-year percentage of sales in the industry. In the final industry (gypsum), the industry's leading firm *increased* its market share from 47.7 to 51.1%, despite an increase in its debt-to-value ratio from 35 to 90%.

Phillips's main tests of capital structure's interaction with product markets are conducted at the industry level. To isolate the effect of debt on aggregate supply, Phillips estimates industry-level supply models for each of the four industries. Since the dependent variable in each equation is price, and because price and quantity are simultaneously determined, Phillips estimates each of his supply equations with two-stage simultaneous equations.[20] He also presents a reduced-form equation for each industry in which industry quantity is regressed on the exogenous variables from the price equation. In all specifications, input prices and scales of production are included as controls.

The results from these regressions for each industry indicate that debt influences product prices. In particular, the average debt ratio is positively associated with product prices in every industry except gypsum, where it is significantly negative. Furthermore, Phillips's reduced-form equations indicate that the industry *quantity* is negatively related to industry debt ratio for the same three industries, gypsum (again) representing the only exception.

Chevalier and Scharfstein (1996) and Dasgupta and Titman (1998) each present models where firms compete by setting price (i.e., Bertrand competition) in which debt commits the leveraged firm to behave less aggressively. The results in three of the four industries considered by Phillips (e.g., fiberglass, insulation, and tractor trailer) are consistent with these models. Phillips's findings in the gypsum industry, where the increase in debt led to stronger competition, is more supportive of models by Brander and Lewis (1986) and Maksimovic (1988) where firms compete by setting prices (e.g., Cournot competition). Since gypsum is a commodity with relatively low barriers to entry it is likely that the Cournot assumptions are more applicable. In addition, since overinvestment is likely to be less sustainable in a highly competitive industry, it is less likely that increased leverage will result in reductions in capacity investments.

[20] Phillips's methodology is a two-stage simultaneous equation framework in which instruments are used for price in the demand equation and for quantity in the supply relationship. In the demand equation, input prices (potentially including the price of oil and/or electricity, wages, etc.) are the instruments for product price.

Kovenock and Phillips (1997) add to Phillips (1995) by considering how leverage recapitalization affects individual firm investment and plant closure decisions. In addition, Kovenock and Phillips (1997) recognize the potential endogeneity problem in Phillips (1995), namely, that if firms undertaking LBOs are able to anticipate the effects of their recapitalizations, then any direct effect of debt on a firm's decisions may be obscured.

In their paper studying 10 industries, Kovenock and Phillips (1997) control for the endogeneity of the capital structure decision directly through a two-stage approach. In stage one, they run a logistic regression that explains the firm's decision to recapitalize as a function of industry variables, including capacity utilization and market concentration. In stage two, they examine both the decision to exit the industry and to close plants (in separate regressions), but importantly, they include the *predicted probability* of recapitalization from the first stage.[21] The main result is that high leverage appears to make firms more passive, increasing plant closures and decreasing investment. However, this effect is found only in highly concentrated markets, which the authors interpret as agency problems being "more prevalent in concentrated industries, where the discipline of the market does not weed out nonoptimizing firms (pg. 771)."

Chevalier also takes advantage of the sharp increases in leverage following LBOs in two closely related papers on the supermarket industry. She uses LBOs in grocery store chains to study the effect of changes in capital structure, exploring competitive responses to exogenous shocks in leverage. The primary endogeneity problem is that LBOs may be chosen with their anticipated consequences in mind, which complicates the inferences that can be made regarding the effect of leverage on prices. Suppose, for example, that LBOs are chosen in instances where competition was anticipated to be increasing.[22] Finding that LBOs preceded more competitive price setting would therefore not be surprising, but it would have little to do with the leverage increase itself.

As Chevalier (1995a) notes, "the local-market nature of supermarket competition helps to 'clean out' the endogeneity of the LBO in the study of entry, exit, and expansion." In other words, because the LBO choice is made by the firm at the *national* level, the competitive responses of its rivals in any *local* area play only a minor role in the LBO decision. Therefore, because Chevalier's analysis is conducted at the local level, the sharp increases in leverage brought about by an LBO can be viewed as exogenous, reducing the probability that the resulting estimates suffer from endogeneity bias.

Chevalier (1995a) shows that LBOs "soften" product market competition, as measured by the entry and expansion decisions of rivals of firms that undertook LBOs. Chevalier's (1995b) second paper explores in greater depth one particular dimension in which firms compete—prices. She finds that LBOs have significant impact on the local prices

[21] The authors also use a "high-debt" dummy variable in addition to the predicted probability of recapitalization, findings similar results.

[22] For example, Jensen (1986) suggests that increased debt focuses managerial attention and reduces the incentives to pursue wasteful expenditures. If competition is likely to increase, then perhaps shareholders see a greater need for the disciplining role of debt.

supermarkets charge for their goods; however, the direction of the post-LBO price changes depends on the financial structure of the LBO firm's rivals. When its competitors are less leveraged themselves, LBOs lead to decreased prices, which can be viewed as evidence of opportunistic predation by less financially constrained rivals. When competitors are also highly leveraged, prices rise following LBOs, evidence that supports Chevalier's earlier study showing that LBOs soften product market competition.

Chevalier (1995a) examines data from 85 Metropolitan Statistical Areas (MSAs) obtained from industry publications. To explain the expansion choice, she runs an ordered probit (+1 for expansion, −1 for retrenchment), conditioning on various market controls including changes in population, the Herfindahl Index, and the size of each store, as well as the market share of rival firms that had previously undertaken an LBO. When this model is run only for incumbent firms that were not themselves recapitalized by LBOs, the marginal effects of market leverage strongly suggest that high concentrations of LBOs soften product market competition. For example, if a firm with a 10% market share undertook an LBO, this shock would increase the probability that a non-LBO rival firm would add stores by approximately 6.5%.

As mentioned previously, the main alternative to Chevalier's hypotheses is that LBO decisions are endogenous and may have been driven by firm characteristics that are related to investment or pricing choices. For example, the "weakest" firms may have been LBO targets. Although the experimental design itself largely alleviates this concern, Chevalier presents evidence that LBO firms did not exhibit significant differences in pre-LBO performance. In particular, comparison of accounting data for the 31 publicly traded firms in 1985, including operating margins, market-to-book ratios, and ratio of capital expenditures to assets, indicates almost no evidence that the "types" of firms selected for LBOs were materially different. In addition, Chevalier tests the reactions of stock prices to LBO announcements for their rivals. Since a firm's stock price tends to increase following a rival's LBO, Chevalier concludes that the LBO itself likely softened the expected future competition, rather than the alternative hypothesis.[23] However, it is also possible that the positive stock price responses of the rivals reflected the possibility that they themselves face an increased probability of undergoing an LBO.

Although Chevalier argues that product market competition changes substantially following LBOs, her 1995 study in the *American Economic Review* does little to shed light on the mechanism through which such shifts may occur. For example, do LBO firms compete less aggressively in the pricing arena (i.e., underinvesting in market share)? Does heightened leverage lead to underinvestment in other areas, perhaps eliminating or delaying renovations or store upkeep? Chevalier (1995a) specifically examines the former possibility, asking whether LBOs cause price changes in the supermarket industry. Prices may fall after LBOs, Chevalier argues, if deep-pocketed rivals predate on their more financially constrained rivals. Alternatively, prices may rise when one or more firms in a local market undertakes an LBO, consistent with either leverage-induced

[23] Note that the positive stock price reactions indicate that the LBOs were largely surprises, making ex-ante strategizing in anticipation of the imminent recapitalizations unlikely.

underinvestment in market share (as discussed in Chevalier and Scharfstein, 1996, and Dasgupta and Titman, 1998) or the reversal of underpricing due to agency problems (Jensen, 1989).[24]

Chevalier examines price changes around the time of an LBO, specifically comparing prices from one month prior to the LBO to prices after the LBO (the post-LBO prices are measured from one to six months after the event date). To control for local market conditions that may have influenced supermarket prices, she includes changes in local unemployment as well as price changes in nongrocery items, finding that the coefficients on both are statistically insignificant for all specifications. Of most interest are the coefficients on factors expected to influence either predation by rivals or the willingness of competing firms to accommodate higher prices.

Chevalier finds that the coefficient for the share of supermarket chains in the city that had undertaken LBOs prior to the time window is positive for all specifications and is usually statistically significant. In other words, prices rise when a firm that competes with other highly leveraged rivals undertakes an LBO itself. Importantly, Chevalier runs the test again with data from a time window that completely predates the LBO event (for example, examining price changes from month $-6$ to month $-1$) Under this specification, she finds “absolutely no relationship between the LBO share of rival firms in the city and price change in the period *prior* to the LBO” (p. 1105). This evidence addresses the possibility that price changes occurring after the LBO date may simply be extensions of a preexisting time trend.

In the same empirical model as previously described, Chevalier finds that the coefficient for the store market share of the largest non-LBO chain in the city is negative for all six specifications and is significant at the 1% level for the longer time windows. This suggests that prices are likely to fall in the presence of a single, large, non-LBO competitor, which can be interpreted as evidence of predation. When the regression is run for the period preceding the LBO, the market share of the largest non-LBO chain has no discernible relationship to local price changes. This again suggests that the LBO itself, rather than a preexisting trend or anticipation of the leverage change, represents the reason for the ensuing price change.

Chevalier then extends her analysis to show that price declines following LBOs accomplish the rivals’ presumed goal of driving highly leveraged rivals from the market. With a probit model of exit by LBO firms, Chevalier shows that declines in the grocery price index contributes positively to exit, as would be predicted by a predation explanation.[25] These pieces of evidence in tandem suggest that falling prices after LBOs are most likely the result of predation by more conservatively financed rival firms.

[24] According to Jensen, managers derive utility from large empires, which in the current application, may lead undisciplined agents to inefficiently depress prices to maximize market share rather than profits. To the extent that an LBO aligns incentives of managers and shareholders inefficiently, low prices should rise, reflecting the newfound (and proper) incentives of management.

[25] When the probit model is examined for the time prior to the LBO, no relationship is observed between prices and firm exit, indicating that relatively long periods of declining prices are not alone sufficient to drive firms from the market.

Chevalier and Scharfstein (1996) present a model in which a firm's reliance on external finance alters its incentive to build market share, leading highly leveraged firms to *increase* prices during market downturns. This argument may provide an alternative explanation to Opler and Titman (1994), who show that highly leveraged firms lose sales and market shares to their more conservatively financed rivals (possibly because highly leveraged firms refuse to cut prices) and that this effect is most pronounced during downturns. Debt effectively shortens the firm's horizon by introducing the possibility of liquidation, so that firms relying on external capital have incentives to take actions that boost immediate profits, even at the expense of long-run market share. The main implication is that the output prices of liquidity-constrained firms are predicted to exhibit countercyclicality; they raise prices more (or cut them less) in recessions than their more conservatively financed rivals. Furthermore, the model predicts that even the prices of *unconstrained* rivals are expected to exhibit some—albeit a lesser—degree of countercyclicality, since price markups of the *constrained* firms influence the competitive strategies (in this case, prices) of their rivals.

Chevalier and Scharfstein explore three closely related empirical predictions: (1) more financially constrained firms exhibit more countercyclical pricing, (2) when firms face more financially constrained rivals, their markups should exhibit higher countercyclicality, and (3) industry-average markups should increase when firms within an industry are more financially constrained.

As Chevalier and Scharfstein note, perhaps the most direct of these three hypotheses would be "to relate firm-specific measures of the markup to measures of corporate liquidity." However, as also mentioned, just as leverage is likely to affect prices (through the mechanism proposed), prices are also likely to affect leverage. It is easy to imagine examples in which a shock to a firm's leverage may lead it to raise prices, which may in turn damage its market share and firm value, which may further impact its leverage. To minimize this problem of reverse causality, the authors examine exogenous events that impose liquidity constraints on some firms more than others, investigating whether those facing stricter liquidity constraints raised their prices relative to their less constrained rivals.

Chevalier and Scharfstein use local-market pricing data from the supermarket industry to test their hypothesis that liquidity constraints cause firms to reduce their investment in market share. In one of their tests, average supermarket prices for several cities were regressed against a set of explanatory variables that include each city's sensitivity to oil price shocks as well as the importance of national chains in the local market.[26]

In 1986, the price of oil fell by nearly 50%, inducing severe recessions in several states, including Texas, Louisiana, Oklahoma, Wyoming, and Alaska. Although grocers operating in these states experienced a negative shock, the impact was less severe for national chain stores, whose parents had operations in states relatively insensitive to the oil price spike. The national chain stores in these states could therefore afford to

[26] The authors use an index of grocery prices, which is a weighted average of prices for each city. The data were provided by the American Chamber of Commerce Researchers Association (ACCRA).

capture market share from their rivals by cutting prices deeply in recessions. Examining city-average prices for the six quarters spanning 1985:4–1987:1, the authors find that price declines are most severe in oil-sensitive cities containing a significant national supermarket chain presence. As they argue, the effect of price declines in these cities is quite large. For a given city in an "oil state," an increase of one standard deviation in the fraction of stores owned by national chains from its mean of 0.35 to 0.58 decreases the expected percentage change in the local price index from $-0.020$ to $-0.045$. Chevalier and Scharfstein also present evidence showing that firms that recently did a leveraged buyout increase prices more in severe declines than their less levered counterparts.

The city-level tests address the relation between a firm's capital structure and its investment decisions, largely ignoring the role debt may play in a firm's competitive response to its rivals in the market. Using firm-level pricing data from the first quarter of 1991 to the last quarter of 1992, Chevalier and Scharfstein explore how leverage impacts a firm's pricing, paying particular attention to how the financial position of rival firms influences this decision. As in the city-level tests, whether a store was owned by a firm that undertook an LBO is used as a proxy for being subject to financial constraints. While an LBO dummy may be viewed as endogenous, it is important to remember that the LBO decision is made at the *company* (as opposed to the store) level, such that a given firm's response (or that of its competitors) is not likely to have motivated the recapitalization. In order to test the hypothesis that more financially constrained firms raise prices compared to their less constrained competitors, and that more constrained rivals magnifies this effect, the authors run regressions of the form:

$$\begin{aligned}\Delta\text{Price} = \alpha\ &(\text{LBO}) + b\,(\text{LBO} \times \Delta\text{EMP})\\ &+ c\,(\text{OLBOSHARE}) + d\,(\text{OLBOSHARE} \times \Delta\text{EMP})\\ &+ e\,(\Delta\text{EMP}) + f\,(\Delta\text{WAGE}) + \varepsilon\end{aligned}$$

in which LBO represents a dummy if the parent company had undertaken a leveraged buyout, $\Delta$ EMP is the percentage change in employment in the city's state during the period, OLBOSHARE is the share of stores in the local market owned by an LBO firm, and $\Delta$ WAGE is the percentage change in wages for workers in sales occupations. The dependent variable is the percentage change in price for a firm's price index for a particular city.

Three main results may be gleaned from this regression. First, LBO firms charge higher prices, as indicated by a significantly positive coefficient on the LBO dummy. Since this may reflect increases in *costs* for LBO firms rather than markups, the fact that the coefficient on the LBO $\times$ $\Delta$ EMP interaction term is negative and significant is important. When local markets suffer, as measured by negative employment changes, LBO firms raise prices more than their less financially constrained rivals, which is consistent with the idea that the higher prices are caused by financial constraints rather than higher costs. Also of interest is the coefficient on the local share of LBO firms, OLBOSHARE, which is positive and significant, indicating that leverage causes rivals to increase prices.

Furthermore, the interaction term is negative and significant, indicating that slow economic growth magnifies the effect. In a city with low employment growth of 0.5% (one standard deviation below the mean), an increase in OLBOSHARE by one standard deviation from the mean of 14.9 to 30.0% would lead the non-LBO firm to more than double its price increase from 1.4 to 2.9%.

Campello (2003) builds on Chevalier and Scharfstein's study, first asking "Are markups more countercyclical in highly leveraged industries?" and second, "Does a firm's capital structure affect its ability to build market share, so that competitive outcomes are indeed influenced by a firm's financing mix and the financial condition of its competitors?" While these questions are linked, the empirical approach taken by Campello is quite different, so we consider each question separately.

Campello evaluates Chevalier and Scharfstein's (1996) theory of markup (i.e., prices over marginal costs) countercyclicality, specifically the implication that the possibility of bankruptcy reduces the incentives of firms to invest in building market share in downturns. He starts by extending Bils's (1987) analysis by comparing the cyclicality of markups in industries with different leverage, finding *higher* markup cyclicality in highly leveraged industries.[27] Using data from 20 manufacturing industries segregated by two-digit SIC codes (codes 20–39), Campello runs industry-level (the "i" index refers to industries) regressions of the form:

$$\begin{aligned}\text{Markup}_{i,t} &= \eta + \alpha(-\Delta\ \text{Log(GDP)}_t) + \beta\ \text{Leverag}e_{i,t-1}\\ &\quad + \lambda[\text{Leverage}_{i,t-1} x(-\Delta\ \text{Log(GDP)}_t] + \varepsilon_{i,t}\end{aligned}$$

where GDP is the gross domestic product, and markup and leverage are measured as industry averages. Unlike Chevalier and Scharfstein's (1996) study however, in which the authors' empirical design sidestepped the problem of having to observe marginal costs,[28] Campello constructs a markup measure that incorporates prices, hourly wages, the overtime rate, the number of hours employed (both regular and overtime), the number of workers, and gross output.[29] Campello finds a significantly positive coefficient on the interaction between industry leverage and macroeconomic declines, concluding that "these estimates suggest that negative shocks to demand prompt firms to raise

[27] Bils (1987) empirically documents that while marginal costs are "markedly procyclical," output prices are not as responsive to fluctuations in the health of the economy. The result is that price-cost margins ("markups") are highly *countercyclical*.

[28] Chevalier and Scharfstein note that price increases could arise from either markup or from increases in marginal costs. The empirical design focuses on price changes for both local and national supermarkets in both oil and nonoil states. If prices increase more dramatically for local chains in oil states (as the authors find), then either: (i) markups of local chains are higher than national chains in oil states, or (ii) costs increase for local chains relative to national chains *only in the oil states*, which the authors argue is very unlikely. See Section II of the paper for more details.

[29] To construct the markup series, Campello gathered industry price data from the Bureau of Labor Statistics (BLS) *Producer Price Indexes.* Data on the number of production workers, the weekly average hours, and the average hourly wage were obtained from the BLS *National Employment, Hours and Earnings.*

price-cost margins more (or cut them less) in industries with more externally financed competitors." In response to a 1% decline in gross domestic product (GDP), Campello estimates that a hypothetical "all-debt" industry would experience markups of 42% more than a "zero-debt" industry.

Since Campello takes industry leverage as exogenous, the strength of his conclusions regarding leverage-induced markup countercyclicality depends on his ability to adequately control for the determinants of industry leverage. Specifically, suppose that two industries differ in their abilities to respond to macroeconomic shocks, perhaps because the more flexible industry can efficiently scale down production in the face of recession. Given this, it is likely that the firms in the flexible industry will carry more debt.

To address this potential endogeneity problem, Campello follows Sharpe (1994), splitting the sample based on whether an industry's sales are sensitive to the business cycle. If cross-sectional differences in macroeconomic sensitivities drive Campello's results, then the coefficient on the interaction term should not be significant for firms with low sales-to-GDP sensitivities, yet it is. A second possibility is that debt may be used to finance expansions, so that firms with high utilization rates needing to expand will issue debt when the economy grows.[30] For this reason, Campello explicitly includes lagged industry capacity utilization as a control, finding that the results are unchanged.[31] Campello is careful to control for several other potential factors that potentially drive cross-sectional differences in industry debt ratios, including as controls energy prices (which may affect product costs), industry capacity, the sensitivities of each industry to the business cycle, and industry concentration (which may alter a firm's ability to collude by manipulating prices). Campello argues that his set of controls is sufficient to suggest that the debt level itself, rather than an omitted determinant of the debt level, induces the markup countercyclicality he observes.

The second class of tests Campello runs is conducted at the firm level. Like the industry-level tests, Campello's firm-level tests capitalize on the exogeneity of macroeconomic shocks and examine how the sensitivity of a firm's performance to macroeconomic shocks is influenced by the firm's debt ratio. The empirical methodology proceeds in essentially three steps. First, for each quarter from 1976:1–1996:4, he sorts all manufacturing firms (SIC codes 200–399) into quintiles ranked by book debt-to-long-term assets. Then, for the highest and lowest quintiles, he runs *cross-sectional regressions* of the following form during time $t$:

$$\begin{aligned}\Delta\text{Log(Sales)}_{i,t} = \eta &+ \alpha_K \Delta\text{Log(Sales)}_{i,t-1} + \cdots + \alpha_K \Delta\text{Log(Sales)}_{i,t-4} \\ &+ B_K \Delta\text{Log(PPE)}_{i,t-1} + \cdots + \beta_K \Delta\text{Log(PPE)}_{i,t-4}\end{aligned}$$

30 In the presence of scale economies, decreasing marginal costs may then lead to markdowns for such expanding firms. Therefore, debt (issued when the economy grows) may appear to increase markup countercyclicality (since marginal costs decline, reducing prices) if it is used to finance expansions, but not for the reason posited in Chevalier and Scharfstein (1996).

31 The other control variables are intended to address other potential sources of endogeneity from omitted variables. See Section III of Campello (2002) for more details.

$$+ \Lambda_1 \text{ Log(Assets)}_{i,t-1} + \cdots + \Lambda_4 \text{ Log(Assets)}_{i,t-4} \\ + \delta \text{ Leverage}_{i,t-1} + \varepsilon_{i,t}$$

saving the vector of sales-to-leverage sensitivities $\delta_t$ for both the high- and low-leverage quintiles. All variables are converted to deviations from their industry means. During the final step, Campello runs two time-series regressions (one for each leverage quintile) of the form:

$$\delta_t = \eta + \varphi_1 \Delta\text{Activity}_{t-1} + \cdots + \varphi_4 \Delta\text{Activity}_{t-4} + \gamma \text{ Trend}_t + \varepsilon_t$$

where Activity is one of several proxies for downward macroeconomic shocks. The primary question such a regression addresses is whether or not highly leveraged firms have sales-to-leverage sensitivities that are themselves more sensitive to economic downturns. While Campello finds that in high-debt industries none of the $\Delta$ Activity coefficients are significantly different from zero, the opposite is true in low-debt industries, where the same coefficients are significantly negative in all but one of the specifications. Importantly, the negative impact of debt on sales growth during recessions is limited to industries *in which a firm's rivals have low leverage*. For example, consider two otherwise identical firms operating in a low-leverage industry, except that one is 10% above while the other is 10% below the (low) industry average. After a 1% decline in GDP, the sales growth of the more indebted firm is predicted to be 1.3% lower than the more conservatively financed rival. Were this same test applied in a "high-debt" industry, no differences between the firms' sales growths would be observed.

Both Campello's industry and firm-level evidence can be interpreted as supportive of Chevalier and Scharfstein's theory that cash shortfalls induced by external finance reduce a firm's incentive to invest in market share. It may also be the case however, that debt reduces a firm's ability (rather than willingness) to compete with opportunistic rivals. If less levered firms predate on their more levered rivals during downward demand shocks, then highly leveraged firms in low-debt industries would lose market share in precisely the way Campello documents. In this way, his firm-level tests fail to distinguish between Chevalier and Scharfstein's theory of liquidity-induced underinvestment and Telser's (1966) and Bolton and Scharfstein's (1990) predation models.

Our preceding discussion indicates that macroeconomic changes provide an exogenous source of variation that allows researchers to examine the effect of leverage on competition. An alternative way to examine these issues is to examine how firms respond to exogenous shocks to their competitive environments. Khanna and Tice (2000) examine precisely this issue, studying the rapid nationwide expansion of the discount retailer Wal-Mart during 1975–1996. They focus on how characteristics such as debt, ownership, focus, and profitability lead incumbent firms to react differently to Wal-Mart's expansion into their respective regions.

As in each of the studies we have considered so far, such an investigation requires a careful treatment of endogeneity. In particular, since Wal-Mart's entry into a particular market may be driven by the collective inabilities of incumbents to respond, it may be difficult to infer cause and effect between incumbent characteristics and competitive responses. The concern is that perhaps Wal-Mart chooses to expand into regions with weak competitors and that these firms (perhaps because of a history of poor performance) may have high leverage. While it certainly would not be surprising to find that highly leveraged firms respond to Wal-Mart's entry less aggressively, it may be impossible to tell whether debt itself inhibits the incumbent's response, or whether debt is simply correlated with other characteristics that render incumbents less likely to respond aggressively.

Khanna and Tice convincingly argue that this is *not* the case. Wal-Mart's expansion decisions appear to be driven by its own distributional efficiencies rather than by characteristics of its potential competitors. The observation that incumbent characteristics play only "a relatively minor role" in the expansion decisions allows Khanna and Tice to analyze the effects of capital structure and other incumbent firm characteristics on the reactions to entry by Wal-Mart.

The main classes of tests are variations of an ordered probit in which incumbent actions are ranked by the degree of response aggressiveness to Wal-Mart. The dependent variable is the firm's response, ordered from most aggressive (adding stores, which is assigned a value of $+1$) to least aggressive (reducing the number of stores, assigned a value of $-1$). Although Khanna and Tice analyze the impact of many firm and market characteristics, we focus primarily on marginal effects of capital structure on the incumbent responses. When the sample is restricted to public incumbent firms, high debt-to-asset ratios are associated with less aggressive capital investments, as measured by expansion and retrenchment decisions. In particular, when all controls are evaluated at their mean values, an increase in the debt ratio of 10% decreases the probability of expansion by 2.7% and increases the probability of retrenchment by 3.5%.[32]

Khanna and Tice also study whether or not the incumbent had undergone an LBO influences its response to Wal-Mart. Interestingly, they find that LBO firms mount *more* aggressive responses, which contrasts with the Chevalier evidence we described earlier. Khanna and Tice suggest that although this evidence may indicate that "LBO decisions are different from leverage decisions," they encourage a cautious interpretation due to the small number of LBOs in their sample and a potential endogeneity problem. Specifically, it might be the case that firms with more aggressive management were more likely to undertake LBOs.

Khanna and Tice (2005) expand on their earlier work in which they investigate Wal-Mart's decision to enter a particular market by considering Wal-Mart's location *within* that market. Interestingly, Wal-Mart places its stores closer to rival stores that are less efficient and more highly leveraged, consistent with the idea that leverage weakens a firm's ability to withstand competition.

In other results, Khanna and Tice (2005) present evidence suggesting predation on highly leveraged firms by more conservatively financed rivals. During the period

[32] In addition, the authors find that firms more "focused" on the discount retailing business (as measured by discount retail sales divided by total firm sales), larger firms, and more profitable firms compete more aggressively with Wal-Mart.

1982–1995, they study market-level average prices in various metropolitan areas and consider whether prices charged by discount retailers were different between recessions and normal times. The authors find that higher industry leverage (averaged across all discounters within a given market) implies higher average prices during normal times, but lower prices during recessions. While this evidence suggests that leverage softens competition during downturns, even more compelling is that the price difference between normal times and recessions was greatest in markets with high-debt *dispersion,* suggesting predation as the mechanism driving the observed procyclicality in prices. Consistent with the authors' interpretation of predation, in markets with homogeneous capital structures, prices were no lower in recessions than in normal times.

Further evidence in support of a predation hypothesis comes from regressions examining the sensitivity of exit to price declines. Within the predation framework, price cuts during recessions should drive out financially troubled rivals, a result confirmed in the empirical analysis. In markets where firms have similar capital structures, the extent to which firms exit is not related to price cuts during recessions (even in markets where most firms are highly leveraged). However, in markets where only *some* firms are financially distressed, price cuts increase the probability that a highly leveraged firm is forced out of the market during downturns. Combining this result with those studying the procyclicality of prices between markets with different debt ratio dispersions, Khanna and Tice make a convincing case that predation is likely to be an important influence on prices and survival.

## 4. Conclusion

The studies surveyed in this chapter indicate that a firm's capital structure has a nontrivial effect on its relationships with competitors and nonfinancial stakeholders such as its workers, suppliers, and customers. Generally, the evidence from this literature suggests that debt magnifies the effects of economic downturns and predation, effectively making bad situations even worse for highly leveraged firms. During a recession or downward shock to profitability, a highly leveraged firm can expect to disproportionately lose market share and sales (Opler and Titman, 1994; Zingales, 1998; Campello, 2003), to lay off workers and pay lower wages (Sharpe, 1994; Hanka 1998), and to reduce investment to conserve cash (Chevalier and Scharfstein, 1995; Khanna and Tice, 2000). Debt also appears to render firms more susceptible to predation, as directly suggested by Chevalier's (1995a, 1995b) analysis of supermarket prices and indicated as a possibility in Zingales's (1998) study of deregulation in the trucking industry.

Perhaps just as interesting is what *cannot* be concluded about capital structure's effects on a firm's competitive strategy and relationship with its stakeholders. Since debt ratios are chosen by the firm, cross-sectional differences in capital structure result from cross-sectional differences in firm-level explanatory variables (e.g., the degree to which a firm's customers will be harmed in the event of liquidation). Since many of the determinants of capital structure are either unobservable or do not have good proxies, it is often impossible to distinguish between the direct effect of debt on firm performance and that of an omitted variable(s) that partially determines the debt ratio.

For example, Sharpe (1994) finds that the employment choices of firms with higher leverage ratios are more sensitive to the business cycle. He argues that this finding is consistent with both Titman (1984), whereby firms with less firm-specific human capital (which is presumably destroyed during liquidation) choose higher leverage ratios, as well as with Jensen (1988, 1989), in which debt forces managers to lay off workers even if psychologically costly. A third possibility is that firms simply differ in how easily they can adjust the sizes of their labor forces and that these differences influence debt choices. While it would be interesting to understand whether highly leveraged firms lay off more workers in downturns because of differences in firm-specific human capital, agency problems, or cost structure, in most studies such precision is not possible. Further research will hopefully produce clearer answers to some of these questions.

Other promising areas of research include a consideration of the interaction between corporate governance and the impacts of leverage on a firm's competitive strategy. For example, continuing discussion of Sharpe's (1994) study, it would interesting to understand whether the debt appears to provide more or less discipline between firms that differ in the strength of their corporate governance (e.g., independent corporate versus insider-dominated boards). Since strict corporate governance and leverage may both discipline management, interesting relationships may arise between governance and the effects of leverage on a firm's employment policy.

Also relatively unexplored are the implications for asset pricing that may arise from the impact of capital structure on product market competition. For example, if financial constraints make a firm more vulnerable to predation, and if predation is more likely to occur in economic downturns, then financially constrained firms will have higher betas and hence, higher costs of capital. However, as Chevalier (1995b) shows, this may not hold in every instance, since predation appears to require both a financially vulnerable victim and a well-financed competitor. Do highly leverage firms with deep-pocketed rivals face higher costs of finance than similarly leveraged firms in industries where predation is more difficult?

But perhaps the most important question is whether the effects of capital structure on a firm's relationships with its stakeholders and competitors are significant enough to play an important role in how firms actually determine their capital structures. For example, when a firm undertakes an LBO, is the anticipated response of competitors a first-order effect? Are the concessions made by firms whose customers depend on its long-term viability large enough to merit serious consideration by the firm? These issues were considered by Titman and Wessels (1988) and others in research that predates work examining the effect of debt on strategic choices, which is the focus of this chapter. Perhaps the next step in this research is to determine how the insights of this more recent literature can be used to refine these earlier cross-sectional tests. While we have anecdotal evidence to suggest that these stakeholder issues are serious concerns of management, we have little direct evidence describing how capital structure decisions explicitly take these issues into account.

## References


Altman, Edward I., 1984, "A Further Empirical Investigation of the Bankruptcy Cost Question." *Journal of Finance*, 39 (4), 1067–1089.

Altman, Edward I., 1973, "Predicting Railroad Bankruptcies in America," *Bell Journal of Economics and Management Science*, 4 (1), 184–211.

Banerjee, S., Dasgupta, S., and Kim, Y., 2007. "Buyer-Supplier Relationships and the Stakeholder Theory of Capital Structure," *Journal of Finance*.

Baxter, Nevins D., 1967, "Leverage, Risk of Ruin and the Cost of Capital," *Journal of Finance*, 22 (3), 395–403.

Bils, Mark, 1987, " The Cyclical Behavior of Marginal Cost and Price," *American Economic Review*, 77 (5), 838–855.

Bolton, Patrick and David S. Scharfstein, 1990, "A Theory of Predation Based on Agency Problems in Financial Contracting," *American Economic Review*, 80 (1), 93–106.

Brander, James A. and Tracy R. Lewis, 1986, "Oligopoly and Financial Structure: The Limited Liability Effect," *American Economic Review*, 76 (5), 956–970.

Bronars, Stephen G. and Donald R. Deere, 1991, "The Threat of Unionization, the Use of Debt, and the Preservation of Shareholder Wealth," *Quarterly Journal of Economics*, 106 (1), 231–254.

Campello, Murillo, 2003, "Capital Structure and Product Markets Interactions: Evidence from Business Cycles," *Journal of Financial Economics*, 68 (3), 353–378.

Chevalier, Judith A., 1995a, "Capital Structure and Product-Market Competition: Empirical Evidence from the Supermarket Industry," *Journal of Finance*, 50 (4), 1095–1112.

Chevalier, Judith A., 1995b, "Do LBO Supermarkets Charge More? An Empirical Analysis of the Effects of LBOs on Supermarket Pricing," *American Economic Review*, 85 (3), 415–435.

Chevalier, Judith A. and David S. Scharfstein, 1996, "Liquidity Constraints and the Cyclical Behavior of Markups," *American Economic Review*, 85 (2), Papers and Proceedings of the Hundredth and Seventh Annual Meeting of the American Economic Association Washington, DC, January 6–8, 1995, 390–396.

Cutler, David M. and Lawrence H. Summers, 1988, "The Cost of Conflict Resolution and Financial Distress: Evidence from the Texaco Pennzoil Litigation." *RAND Journal of Economics*, 19 (2), 157–172.

Dasgupta, Sudipto and Kunal Sengupta, 1993, "Sunk Investment, Bargaining and Choice of Capital Structure." *International Economic Review*, 34 (1), 203–220.

Dasgupta, Sudipto and Sheridan Titman, 1998, "Pricing Strategy and Financial Policy," *Review of Financial Studies*, 11 (4), 705–737.

Glazer, Jacob, 1989, "Live and Let Live: Collusion among Oligopolists with Long-term Debt," Working Paper.

Hanka, Gordon, 1988, "Debt and the Terms of Employment," *Journal of Financial Economics 4.*

Jensen, Michael, 1989, "Capital Markets, Organizational Innovation, and Restructuring," Presented at the Board of Governors of the Federal Reserve System, Washington, DC.

Jensen, Michael C., 1998, "Takeovers: Their Causes and Consequences." *Journal of Economic Perspectives*, 2(1), 21–48.

Jensen, Michael C. and Meckling, William H., 1976, "Theory of the Firm: managerial behavior, agency costs, and capital structure," *Journal of Financial Economics*, 3, 305–360.

Jensen, Michael C., 1986, "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review*, 76 (2), Papers and Proceedings of the Ninety-Eighth Annual Meeting of the American Economic Association, 323–329.

Khanna, Naveen and Sheri Tice, 2005, "Pricing, Exit and Location Decisions of Firms: Evidence on the Role of Debt and Operating Efficiency," *Journal of Financial Economics*, 75, 397–427.

Khanna, Naveen and Sheri Tice, 2000, "Strategic Responses of Incumbents to New Entry: The Effect of New Ownership Structure, Capital Structure, and Focus," *Review of Financial Studies,* 13 (3), 749–779.

Kovenock, Dan and Gordon M. Phillips, 1997, "Capital Structure and Product Market Behavior: An Examination of Plant Closing and Investment Decisions,"*Review of Financial Studies*, 10 (3), 767–803.

Maksimovic, Vojislav, 1988, "Capital Structure in Repeated Oligopolies." *RAND Journal of Economics,* 19 (3), 389–407.

Maksimovic, Vojislav and Sheridan Titman, 1991, "Financial Policy and Reputation for Product Quality," *Review of Financial Studies*, 4 (1), 175–200.

Myers, Stewart C., 1977, "The Determinants of Corporate Borrowing," *Journal of Financial Economics,* 5, 146–175.

Opler, Tim C. and Sheridan Titman, 1994, "Financial Distress and Corporate Performance," *Journal of Finance*, 49 (3), Papers and Proceedings of the Fifty-Fourth Annual Meeting of the American Finance Association, Boston, Massachusetts, January 3–5, 1994, 1015–1040.

Perotti, Enrico C. and Kathryn E. Spier, 1993, "Capital Structure as a Bargaining Tool: The Role of Leverage in Contract Renegotiation," *American Economic Review*, 83 (5), 1131–1141.

Phillips, Gordon, 1995, "Increased Debt and Industry Product Markets: An Empirical Analysis," *Journal of Financial Economics,* 37*,* 189–238.

Sharpe, Steven A., 1994, "Financial Market Imperfections, Firm Leverage, and the Cyclicality of Employment," *American Economic Review*, 84 (4), 1060–1074.

Shleifer, Andrei and Robert W. Vishny, 1992, "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," *Journal of Finance*, 47 (4), 1343–1366.

Telser, L. G., 1966, "Cutthroat Competition and the Long Purse." *Journal of Law and Economics*, 9, 259–277.

Titman, Sheridan and Roberto Wessels, 1988, "The Determinants of Capital Structure," *Journal of Finance*, 43 (1), 1–19.

Titman, Sheridan, 1984, "The Effect of Capital Structure on a Firm's Liquidation Decision," *Journal of Financial Economics*, 13 (1), 137–151.

Zingales, Luigi, 1998, "Survival of the Fittest or the Fattest? Exit and Financing in the Trucking Industry," *Journal of Finance*, 53 (3), 905–938.

*Chapter 14*

# BANKRUPTCY AND THE RESOLUTION OF FINANCIAL DISTRESS*

EDITH S. HOTCHKISS

*Boston College*

KOSE JOHN

*New York University*

ROBERT M. MOORADIAN

*Northeastern University*

KARIN S. THORBURN

*Dartmouth College*

## Contents



* We thank the editor, B. Espen Eckbo, for helpful comments and suggestions.

*Handbook of Empirical Corporate Finance, Volume 2*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

## Abstract

This chapter reviews empirical research on the use of private and court-supervised mechanisms for resolving default and reorganizing companies in financial distress. Starting with a simple framework for financial distress and a quick overview of the theoretical research in this area, we proceed to summarize and synthesize the empirical research in the areas of financial distress, asset and debt restructuring, and features of the formal bankruptcy procedures in the United States and around the world. Studies of out-of-court restructurings (workouts and exchange offers), corporate governance issues relating to distressed restructurings, and the magnitude of the costs and the efficiency of bankruptcy reorganizations are among the topics covered.

## Keywords

bankruptcy, financial distress, bankruptcy costs, fire sales, bankruptcy auctions, reorganizations, Chapter 11

## 1. Introduction

Bankruptcy law and related out-of-court mechanisms governing default on debt contracts form one of the essential building blocks of a private economy. The law provides a general structure that helps claimholders resolve unforeseen conflicts arising when the firm defaults on its debt payments. It also determines the allocation of control over the distressed firm to various claimholders and the extent to which market mechanisms are used in resolving financial distress. This in turn affects investors' willingness to provide capital ex ante and thus firms' choice of capital structure and cost of capital.[1]

The design of bankruptcy procedures varies widely across the world. Some countries, like the United States and France, have laws that are favorable to the incumbent management and the continuation of the firm as an ongoing entity. Other countries, like the UK and Sweden, rely on the market in allocating the failing firm's assets. With emerging economies striving to adopt adequate bankruptcy procedures, the relative efficiency of existing procedures has become an important topic for debate.

Furthermore, the use of high leverage in corporate restructuring and the popularity of junk bonds (original issue high-yield bonds) have been important aspects of the U.S. corporate finance scene since the 1980s. Leverage increases are accompanied by increased potential for default and bankruptcy. These structures raise the importance to financial economists, managers, and legal scholars of understanding how firms deal with financial distress.

An active academic literature that examines various aspects of the use of private and court-supervised mechanisms for resolving default has developed over the last two decades. The purpose of this chapter is to summarize and synthesize the empirical research in the areas of financial distress, asset and debt restructuring, and formal bankruptcy procedures in the United States and around the world.[2]

The survey is organized as follows. Section 2 presents a simple conceptual framework for analyzing financial distress that guides the organization of the empirical literature in the subsequent sections. The bulk of the evidence is from the United States, and we turn to the international evidence at the end (Section 8). We review the U.S. evidence in the following order: Evidence on asset restructurings is reviewed in Section 3. Studies of out-of-court debt restructuring (workout and exchange offers) are described in Section 4. Section 5 reviews corporate governance issues related to the restructuring of financially distressed firms. Sections 6 and 7 discuss different aspects of formal bankruptcy proceedings in the United States, in particular the magnitude of costs and the efficiency of the outcome. In Section 8, research on insolvency procedures in other countries is surveyed. Section 9 concludes by offering some comments and suggestions for the direction of future research.

[1] These ex-ante effects are analyzed in the literature on optimal security design and capital structure. See, for example, Allen (1989), Allen and Gale (1994), and Allen and Winton (1995) for security design, and Harris and Raviv (1991) and Part 3 of this book for capital structure. We focus here on the ex-post efficiency of the distress resolution mechanisms.

[2] See Wruck (1990), John and John (1992), John (1993), and Senbet and Seward (1995) for earlier surveys of this literature.

## 2. Theoretical framework

This section presents a simple framework describing financial distress and the mechanisms available to resolve distress. The framework provides an overview of the issues analyzed in the theoretical literature and hence a motivation for the questions examined in the empirical literature.

### *2.1. Restructuring of assets and financial contracts*

The financial contracts of a firm can be broadly categorized into hard and soft contracts. An example of a hard contract is a coupon-paying debt contract that promises periodic payments by the firm to its bondholders. If these payments are not made on time, the firm is in violation of the contract, and bondholders can seek legal recourse to enforce the agreement. Lack of liquidity does not constitute a mitigating circumstance for non-payment. Obligations to suppliers and employees are other examples of hard contracts. In contrast, common stock and preferred stock are examples of soft contracts. Here, even though equityholders have expectations of receiving regular cash payouts from the firm, the level and frequency of these payouts are discretionary policy decisions made by the firm. Specifically, the payouts can be suspended or postponed based on the availability of liquid resources remaining in the firm after satisfying the claims of the hard contracts.

The assets of a firm also have a natural categorization based on liquidity. Cash and marketable securities that can quickly be converted into cash are liquid assets. Long-term investments, such as plant and machinery, which may only produce liquid assets in the future, are considered illiquid or hard assets.

These categorizations of the financing contracts of a firm and its assets form the basis for a straightforward definition of financial distress. A firm is in financial distress at a given point in time when the liquid assets of the firm are not sufficient to meet the current requirements of its hard contracts. Mechanisms for resolving financial distress do so by rectifying the mismatch through restructuring the assets or restructuring the financing contracts, or both. In this survey, we examine the costs of resolving financial distress using either method.

On the asset side, the hard assets can be wholly or partially sold to generate additional cash in order to meet the current obligations. Premature sale of illiquid assets, however, may result in the destruction of going-concern value and involves a cost of liquidation. This cost can be thought of as the difference between the going-concern value of the assets (i.e., the present value of all future cash flows produced by the assets) and the highest value that can be realized if the assets are sold immediately. The cost of liquidation, and hence the cost of the asset restructuring, depends on a variety of factors such as what fraction of the assets needs to be sold and what operational relationship the liquidated assets has to those that are retained. If the assets can be sold as a going-concern package instead of a piecemeal sale of assets, the liquidation costs may be lower. Similarly, if the assets are sold in a competitive auction to a buyer who can use these assets efficiently, liquidation costs may be very low or—if the buyer is a higher-value user than the seller—even positive. In other words, the efficiency of the asset-restructuring channel will depend on the liquidation costs associated with the sale of the required assets.

Shleifer and Vishny (1992) analyze the determinants of liquidation costs related to asset sales in financial distress. They focus on different aspects of market liquidity, including credit constraints in the industry, asset fungibility (the number of distinct uses and users for a particular asset), and participation restrictions (e.g., regulations on foreign acquisitions and antitrust restrictions). In their model, industry outsiders are lower-valuation users of the assets. Shleifer and Vishny argue that the price received in a distressed asset sale may suffer from large discounts if the entire industry is financially distressed and industry insiders are unable to compete for the assets due to liquidity constraints.

An alternative way of dealing with financial distress involves restructuring the financial contracts. One mechanism for this restructuring is to negotiate with creditors and reformulate the terms of hard contracts such that the current obligations are reduced or are deferred to a later date. Another technique is to replace the hard contract with soft securities that have residual rather than fixed payoffs. In general, debt restructuring provides relief from financial distress by replacing the existing debt with a new debt contract that reduces the interest or principal payments, or extends the maturity, or exchanges equity securities for the debt.

An additional financial restructuring mechanism that helps correct the imbalance between current assets and requirements of the hard contracts is to raise current liquidity by issuing additional new claims against future cash flows. Although the original hard contracts are left unaltered, the claim structure of the firm is changed by the new financing undertaken. When the newly issued claims are a softer contract or have longer maturity, the total package of financing becomes less onerous on the firm, resolving financial distress. An infusion of private equity is an example of this type of restructuring.

Both asset restructurings and debt restructurings can be accomplished either through a formal court-adjudicated process or in a voluntary out-of-court workout. The choice of method used to resolve financial distress depends on the relative costs and benefits of each mechanism. For example, in an illiquid secondary market, the costs of asset restructurings are likely to be high, and financial restructuring may constitute a dominant restructuring mechanism. By the same token, if asset restructuring involves asset sales through efficient mechanisms such as auctions, the overall costs of resolving financial distress may be lower.

### 2.2. *Efficiency issues in recontracting*

The efficiency of the mechanisms for resolving distress can be measured by the loss in asset value incurred in the process of the asset and debt restructuring. A number of factors related to the structure of the firm's claims and to the institutional framework governing the process for restructuring contribute to these costs. To understand these factors, it is useful to first consider a simple theoretical setting in which distress can be resolved costlessly.

In this simple setting, a single lender has access to the same information as corporate insiders, and the debt contract is complete; that is, a complete state-contingent set of contracts can be written and are enforced by the legal system. Here, either an initial contract can be designed that imposes the financial restructuring necessary to avoid a premature liquidation of assets, or the contract will be renegotiated costlessly in default in order to avoid suboptimal liquidations. For example, if at any time the firm's current liquidity falls short of the current coupon obligations of the debt contract, the debt contract is renegotiated. In the negotiation, the lender is promised a combination of cash in the current period that the firm can pay without liquidating assets and additional cash flow in the future. The expected value of this combination is equal to the cash flow guaranteed by the old debt contract. Under symmetric information, the lender knows that the restructuring of the debt is such that he or she is indifferent between the new contract and the old one, and will accept the proposed contract. Moreover, the firm is no longer financially distressed under the new contract. In this example, the distress resolution is completely efficient and simply accomplished through a costless restructuring of the debt contract.

In practice, however, contracts are by nature generally incomplete. Neither outside investors nor the court system can verify the detailed information required to enforce many contracts. The current cash available, for instance, may not be observable to outside parties, preventing the enforcement of contracts that are contingent on these cash flows. Moreover, managers may have some latitude to divert a portion of the firm's cash flows according to their personal preferences. Hart and Moore (1998) show that when one cannot contract on cash flows, creditors must be given some rights to liquidate physical assets in order to make borrowing viable. Otherwise, managers would always choose to default strategically and divert available cash to them. Anticipating this situation, creditors would not be willing to lend money to the firm. In contrast, if creditors are given the right to sell assets following nonpayment (default), the threat of liquidation helps deter strategic defaults. To keep the threat credible, suboptimal asset sales may sometimes occur following liquidity-induced defaults.

Financial restructuring can provide a solution to this problem. Mechanisms facilitating debt restructuring will reduce the costs of premature asset sales following liquidity defaults. The same mechanisms, however, will reduce creditors' rights to liquidate assets following a strategic default, encouraging such defaults. The efficiency of the debt-restructuring mechanism ultimately depends on the relative importance of these two effects. Harris and Raviv (1991) and Bolton and Scharfstein (1996) develop related arguments.

In addition to the incomplete contracting problem, asymmetric information between debtors and creditors about the value of the assets—ongoing firm value and liquidation value—can impede a mutually beneficial debt renegotiation. As pointed out by Brown (1989), a private workout is always successful when there is symmetric information between management and a single creditor. Many of the theoretical models in the area examine the effect of incomplete contracting and asymmetric information on the efficiency of contracting, as well as the mechanisms necessary to resolve financial distress arising from a failure to meet the terms of the debt contract.

A third problem in practice is that there are usually multiple creditors with interests that are not congruent. Depending on the nature of the debt contract (private debt vs. public debt, syndicated vs. nonsyndicated debt), it may be difficult to achieve an agreement among creditors. Moreover, each creditor may have incentives to be the first to force a liquidation of the firm's assets in order to guarantee payment in full. It has been argued that one of the central reasons for needing a bankruptcy law is to curb the inefficiencies resulting from this "common pool" problem.

The presence of all these factors will influence the firm's choice of restructuring venue—that is, whether it will recontract privately or will instead choose to enter formal bankruptcy proceedings. We discuss this choice further in Section 2.4, following a brief review of the main ingredients of the formal bankruptcy process in the United States.

### *2.3. Rules and procedures of the U.S. bankruptcy code*

For most firms in the United States, formal bankruptcy practices are governed by the Bankruptcy Reform Act of 1978 and, more recently, the Bankruptcy Reform Act of 2005. Bankruptcy petitions are filed in one of 94 regional bankruptcy courts, often based on the physical location of the company's assets.[3] Corporations generally file for liquidation under Chapter 7 or for reorganization under Chapter 11. Although creditors may initiate an involuntary filing under Chapter 7, management is often successful in converting the case to Chapter 11, allowing an attempt to reorganize. Because management can challenge an involuntary petition, bankruptcy filings are more frequently initiated by management.

For firms filing under Chapter 7, the court appoints a trustee that organizes a sale of the firm's assets. Proceeds from the asset sales are distributed to claimholders according to the absolute priority rule, implying that junior claims do not receive any payment until senior claims are paid in full. Each claimholder's distribution depends on the seniority of his claim and the total amount of proceeds received from the sale of assets.

Filings under Chapter 11 are treated as corporate reorganizations, and the bankrupt firm is expected to continue as a going concern after leaving bankruptcy. Consistent with the objective of reorganization, the major provisions of Chapter 11 are designed to allow the firm to continue operating. In general, incumbent management continues to run the business in Chapter 11. To protect the firm during the reorganization, Chapter 11 imposes an automatic stay that stops all payment of interest and principal to creditors and prevents secured creditors from foreclosing on their collateral. The debtor firm may also obtain debtor-in-possession (DIP) financing, taking the form of a line of credit or new financing for routine business expenses. Firms typically file a motion for authorization of a DIP loan at the same time as the Chapter 11 petition or shortly thereafter. Under Section 364 of the Bankruptcy Code, these post-petition loans are granted a super-seniority status

[3] LoPucki and Whitford (1991) examine the choice of venue for 43 large, publicly traded companies in financial distress. They find that firms often engage in "forum shopping," that is, file in a court where the firm has little physical presence, avoiding courts that appear hostile to extensions of exclusivity or aggressively regulate attorney's fees. See also Eisenberg and LoPucki (1999) for evidence on forum shopping.

that effectively strips seniority covenants from existing debt. This reduces the default risk of the new loan, hence encouraging new lending.

To manage the large number of creditors and equityholders that may be involved in the reorganization, the Bankruptcy Code provides for the appointment of committees to represent the interests of different claimholder classes before the court. Committees normally consist of the seven largest members of a particular class who are willing to serve, and they are empowered to hire legal counsel and other professional help at the expense of the firm. A committee representing unsecured creditors is almost always appointed. Other committees can be appointed at the discretion of the Executive Office for U.S. Trustees or the court to represent other claimholder classes, including stockholders.[4]

In order to emerge from Chapter 11, the bankrupt firm must develop a reorganization plan that restructures and reallocates the financial claims on the firm. Similar claims are grouped into classes depending on the priority and other characteristics of the claims. The plan specifies what each class of claimants will receive in exchange for their pre-bankruptcy claims. The distributions typically consist of a mix of cash, new debt securities, equity, and other distributions.

The reorganization plan may embrace a substantial restructuring of the operations. For example, firms operating in Chapter 11, and particularly those with poor operating performance, undertake significant asset sales. In a successful reorganization plan, the firm must demonstrate to the bankruptcy court that, after emerging from bankruptcy, the firm is unlikely to refile for bankruptcy in the near future, either because of an inappropriate capital structure or because of continued poor operating performance.

The rules under which negotiation of a plan takes place give substantial bargaining power to the filing firm, or debtor. One source of bargaining power is that the debtor has the exclusive right to propose a reorganization plan for the first 120 days following the Chapter 11 filing. Prior to the 2005 Bankruptcy Reform Act, bankruptcy judges had considerable discretion to extend this exclusivity period. If the debtor retains exclusivity, then creditors can only accept or reject a reorganization plan that management proposes. Acceptance of the plan requires an affirmative vote by a majority (two-thirds in value and one-half in number) of the claimholders in each impaired class.[5]

The Bankruptcy Code encourages bargaining among claimholders and promotes achieving agreement over the reorganization plan with limited court intervention. However, if the plan is not approved by each impaired class, the court can unilaterally impose or "cram down" the plan on dissenting classes as long as the plan is "fair and equitable." That is, the market value of the new securities distributed to each class under the plan must be at least equal to what the class would receive in a liquidation of the firm. In practice, cram-downs are extremely rare (Klee, 1979). It is in the joint interest of all classes to avoid a cram-down, because application of the fair and equitable standard requires the

[4] Although firms file in specific bankruptcy courts, various aspects of the administration of the case are overseen by the Executive Office for U.S. Trustees.

[5] An impaired class is one in which the distributions under the reorganization plan are insufficient to meet the terms of the original claims. Equityholders are always presumed to be impaired in bankruptcy.

court to determine the firm's going-concern value in a special hearing. These hearings are considered extremely time-consuming and costly.

Avoidance of cram-down also explains observed deviations from absolute priority, where stockholders or other junior claimants receive some payment under a reorganization plan that provides for less than full payment of senior claims. Since classes that receive nothing under the plan (including stockholders) are considered as objecting to the plan, more senior creditors have an incentive to voluntarily relinquish part of their claim in order to reach an agreement. Empirical studies show that deviations from absolute priority are a common feature of Chapter 11 reorganizations (see Sections 4.1 and 5.1).

The Bankruptcy Reform Act of 2005 enhances the rights of creditors in Chapter 11 reorganizations. Some of the more important changes are restrictions on the use and size of management bonuses and severance payments; limitations of the exclusivity period (for management to propose a reorganization plan) to a maximum of 18 months; extension of the fraudulent conveyance look-back period to two years; and reduction of the time that the debtor has to assume or reject leases.

### *2.4. The choice between private and court-supervised restructuring*

With a single lender, complete contracting, and symmetric information, the efficient method of resolving financial distress would be a private restructuring of the debt contract. In a more realistic setting, however, a costless private workout is not feasible, and the firm must weigh the costs and benefits of a private workout against those of a court-supervised proceeding.

Impediments to reaching a settlement in a private restructuring include information asymmetries that arise between poorly informed outside creditors and better informed managers or insiders of the firm; holdout problems when the firm's debt is held by a large number of diffuse creditors; and various conflicts of interest exacerbated when a firm has multiple layers of creditors. Giammarino (1989) and Mooradian (1994) demonstrate that poorly informed creditors may prefer a more costly bankruptcy alternative when information problems are severe. Carapeto (2005) argues that informational asymmetries could lead to extended bargaining, requiring several plans of reorganization before an agreement is reached.

As proposed by Mooradian (1994), Chapter 11 bankruptcy may serve as a screening device when outsiders cannot observe the economic efficiency of the financially distressed firm. Given the debtor's bargaining power and the associated preservation of equity value in Chapter 11, inefficient firms prefer to restructure in court rather than mimic efficient firms in a private restructuring. The self-selection on the part of inefficient firms reduces the information asymmetry between management and outsiders, thus mitigating the impediment to private restructuring for efficient firms. Alternatively, Hotchkiss and Mooradian (2003) suggest that by submitting a bid for the bankrupt firm, a coalition of management and creditors convey positive information about the value of the firm. This may encourage outsider bidders to enter the auction, hence facilitating an efficient redeployment of the bankrupt firm's assets.[6]

It is possible that Chapter 11 may fail to resolve information asymmetries, leaving creditors uncertain about the viability of the distressed firm. Kahl (2002) claims that with sufficient uncertainty, it may be optimal for creditors to postpone the liquidation decision and gather more information about the firm's survival characteristics. Under this strategy, some inefficient firms will be allowed to emerge from Chapter 11 and, if unsuccessful post-bankruptcy, instead be liquidated at a later date.

Gertner and Scharfstein (1991) focus on the conflicts that arise when there are multiple creditors. In particular, holdout problems can arise when a class of claims, such as public debt, is diffusely held. Under the Trust Indenture Act of 1939, a change in the interest rate, principal amount, or maturity of public debt outside of a formal bankruptcy requires unanimity. As a result, public debtholders cannot coordinate their out-of-court restructuring decision. If the out-of-court restructuring is successful and a more costly bankruptcy is avoided, holdouts are paid according to the original debt contract. The cost is borne entirely by the bondholders who participated in the exchange and accepted a reduction in the value of their claim. Small claimants, such as individual bondholders and trade creditors, may realize that their decision to hold out will not materially affect the outcome of the restructuring offer (Grossman and Hart, 1981), and therefore have few incentives to participate. Thus, even though it may be collectively in the interest of public debtholders to agree to the out-of-court restructuring and avoid bankruptcy, it is likely to be individually rational for bondholders to hold out. Chapter 11 is designed to resolve holdout problems, however, since a majority vote is binding on all members of a creditor class.

Abstracting from information and contracting problems, Haugen and Senbet (1978) suggest that bankruptcy is a capital structure decision that should not be linked to liquidation, which is a capital budgeting decision. If the capital structure problem can be resolved by restructuring the financial claims, then firms will avoid costly bankruptcy procedures and privately agree on a financial restructuring. Haugen and Senbet (1978, 1988) maintain that the costs of such private mechanisms are small and should form an upper bound on the costs of managing financial distress. Similarly, Jensen (1989, 1991) argues that since private restructuring represents an alternative to formal bankruptcy, it pays to avoid bankruptcy when the informal mechanism is cost-efficient. Roe (1983) has made similar arguments.

A complication to the restructuring choice is, however, that a redistribution of the financial claims on the firm may not be independent of the firm's asset restructuring decisions. For a highly leveraged firm in financial distress, different claimholders may have conflicting incentives as to the investment decisions. The issue is that the value of junior claims increases with the riskiness of the firm's assets, while the value of senior claims decreases with risk. At the extreme, a conflict can arise as to whether to liquidate or reorganize the firm. Senior creditors that are first in line may prefer an inefficient

[6] Povel and Sing (2007) warn that outsiders may worry about overpaying when winning against a better informed insider, and suggest that bankruptcy auctions should be biased against insiders.

liquidation that converts the firm's assets into cash and provides senior debtholders with a safe distribution. In contrast, junior creditors or out-of-the-money shareholders may prefer inefficient continuation because it has a potential upside. The models in Bulow and Shoven (1978), White (1980), and Gertner and Scharfstein (1991) show that inefficient liquidation versus reorganization decisions may occur when there are multiple classes of creditors.

Zender (1991) models a distressed restructuring as a means of transferring control from equityholders to debtholders. He argues that the shift in decision making improves the efficiency of investment decisions. If decision making is transferred to the creditor who effectively is the residual claimholder, that is, holds the claim whose value is the most sensitive to a change in firm value, the incentives of the controlling security holder will be aligned with firm value maximization.[7]

It is often not only financial distress—that is, that the hard contract obligations are too large—but also economic distress that leaves the firm unable to pay its debts. Optimally, assets of economically inefficient distressed firms should be moved to higher value uses and users, while economically efficient distressed firms should be allowed to continue to operate.[8] The problem is that economic efficiency or inefficiency may not be readily observable.

Moreover, managers may not voluntarily reveal that a firm is economically inefficient. A manager who has private benefits of control and who is interested in preserving his or her job may seek to continue to operate the firm as an ongoing concern and also when it is efficient to liquidate the firm. Aghion, Hart, and Moore (1992) and White (1996) argue that the incentives to undertake high-risk but negative net present value projects increase when managers expect to get a harsh treatment in bankruptcy, for example by losing his or her job. Eckbo and Thorburn (2003), however, suggest that the manager's desire to continue to run the firm following a successful restructuring may counteract any such incentives to overinvest at the expense of bondholders.

In a perfect world, claimholders of a financially distressed firm would always renegotiate and voluntarily agree to a restructuring of the firm's capital structure. In reality, however, with impediments such as information asymmetries, holdout problems, and conflicting interests, firms sometimes resort to bankruptcy for a court-supervised reorganization. In any restructuring of hard contracts or hard assets, the choice of restructuring venue ultimately affects the cost of the restructuring and the impact it has on the firm's investment decisions.

## 3. Asset restructuring

As outlined in the preceding, one set of mechanisms to deal with financial distress involves restructuring the asset side of the balance sheet in order to generate sufficient

[7] One way of transferring the liquidation versus continuation decision to the marginal claimholder is to sell the bankrupt firm in an auction, where the highest bidder gets to decide over the future use of the assets.

[8] The inefficient bankruptcy outcomes of allowing economically inefficient firms to continue and liquidating economically efficient firms are labeled Type I and Type II errors, respectively, by White (1989).

cash to meet the requirements of the hard contracts. Assets can be sold, either piecemeal or in their entirety, to other firms and new management teams. Asset sales can be done privately or through court-supervised procedures, for example, during bankruptcy reorganization (e.g., Chapter 11 of the Bankruptcy Code) or under a liquidation process (e.g., Chapter 7 of the Bankruptcy Code). Each of these alternatives has different costs attached. The incidence and efficiency of asset restructuring to resolve financial distress will depend on the structure of the bankruptcy system in place. This section describes the empirical evidence on the sale of individual assets by distressed firms in the United States. Studies of sales of entire firms in Chapter 11 are discussed in Section 5, and sales of bankrupt firms in other countries are described in Section 8.

The literature described here broadly addresses the following questions: how frequently do distressed firms sell assets; what determines whether distressed firms will sell assets in or out of bankruptcy court; do asset sales lead to efficient outcomes, in that assets are moved to higher value uses; and do "fire sales" exist, where assets are sold at depressed prices?

### *3.1. Frequency and determinants of asset sales*

Financially distressed firms may face a liquidity shortfall, yet be constrained in their ability to raise external funds to meet their obligations. In this situation, asset sales may serve as an alternative source of funds by which liquidity-constrained firms can generate cash. Consistent with this view, Lang, Poulsen, and Stulz (1995) find that asset sales typically follow a period of poor stock performance. On average, these sales announcements are associated with a positive stock price reaction.

In contrast to the evidence for poorly performing firms, Brown, James, and Mooradian (1994) find insignificant returns to announcements of asset sales for a sample of 62 distressed companies. The announcement returns are, however, significantly lower for sellers who use the proceeds to retire debt than for sellers who use the proceeds for other purposes. Firms using sales proceeds to repay debt are more likely to sell assets in poorly performing industries. Also, the greater the proportion of short-term bank debt, the more likely are the sale proceeds to be paid out to creditors, indicating that creditors may influence the decision to liquidate assets. The asset sales appear to benefit the creditors of the financially distressed firm more than its equityholders, suggesting that creditors may force a premature liquidation of the assets.

Asset sales may also convey information about the financial condition of the seller. Sicherman and Pettway (1992) report lower announcement returns for firms divesting assets following a credit downgrade than for sellers with no such downgrade. Brown, James, and Mooradian (1994) examine the characteristics of distressed firms that sell assets and find that sellers typically have experienced a period of extremely poor operating performance and are in poor financial condition. Moreover, the selling firms tend to be distinguished by multiple divisions or subsidiaries. Leverage has also been found to be a determinant of asset sales. Ofek (1993) and Kruse (2002) show that the probability of asset sales increases in the firm's debt level.

A number of papers more generally document the frequency of asset sales for financially distressed companies. Asquith, Gertner, and Scharfstein (1994), Brown, James, and Mooradian (1994), and Hotchkiss (1993, 1995) all demonstrate a high frequency of asset sales for distressed firms, whether out of court or as part of a Chapter 11 restructuring. For example, Hotchkiss (1995) shows that many firms that successfully emerge from Chapter 11 sell a substantial portion of their assets while in bankruptcy.

Asquith, Gertner, and Scharfstein (1994) find that significant asset sales are an important means of avoiding bankruptcy. They find that only 3 out of 21 companies (14%) that sell over 20% of their assets subsequently file for bankruptcy compared to 49% of firms with small or no asset sales. Firms that sell a large fraction of their assets are more likely to complete a successful debt exchange (62% versus 28%). The proceeds are often used to pay off senior private debt. Moreover, the probability of asset sales decreases with industry leverage, suggesting that asset sales may be limited by industry conditions.

### *3.2. Do "fire sales" of assets exist?*

If creditors exert pressure on firms to inefficiently liquidate assets, the value of the firm declines. Not only should we see negative effects on the value of equity and junior debt claims, but firms should also be observed to sell assets at depressed prices in their "fire sale" attempts to raise cash. As discussed earlier, Shleifer and Vishny (1992) argue that distressed firms are likely to be selling assets at a time when potential buyers for those assets—firms in the same industry—are financially distressed as well, contributing to depressed prices. Their model predicts that distressed sellers will receive lower prices and be more likely to sell to industry outsiders in periods when the industry is financially distressed. Moreover, the more specialized the assets, the greater this fire-sale discount.

Several studies examine these issues and their implications for the efficiency of restructurings. An empirical caveat, however, is that it is almost impossible to know whether prices are low because industry demand is low or because industry insiders are liquidity constrained and unable to pay their full valuation. If industry demand has dropped, a low price simply represents an updated (and efficient) market valuation of the assets. If demand exists but a lack of liquidity prevents potential buyers from bidding aggressively, the discount is a true cost associated with the forced asset sale. Most studies construct a model price to represent the fundamental value of the asset, and they compute the fire-sale discount as the difference between this model price and the actual price. Obviously, any evidence on fire-sale discounts is limited by the quality of the estimate of such fundamental values.

Pulvino (1998, 1999) addresses the question of whether fire sales exist. He shows that financially constrained airlines receive lower prices relative to a model price when selling used aircraft than their unconstrained rivals. He also finds that the conditional prices that bankrupt airlines receive for their used aircraft typically are lower than those received by distressed but nonbankrupt firms. Therefore, not only do distressed sellers receive lower conditional prices, but the bankruptcy status of the seller appears to further influence the outcome. Moreover, when the airline industry is depressed—defined as periods when prices are generally low—capital-constrained airlines are more likely to sell to industry outsiders (financial institutions) than are unconstrained airlines. Overall, the evidence in Pulvino (1998, 1999) is consistent with the Shleifer and Vishny (1992) model.

Two related papers study the impact of asset and industry-level characteristics on asset sales. Ramey and Shapiro (2001) examine individual equipment sales that follow three California plant closures in the aerospace industry. They find that actual transaction prices take place at a discount from estimated replacement costs. This discount is greater for equipment that is more specialized to the aerospace industry and when the buyer is an industry outsider. Kim (1998) investigates the significance of asset liquidity in the contract drilling industry, measured by trading volume and the depth of the buyers' market. She shows that the turnover of illiquid assets drops when the industry is distressed, defined as periods of low crude oil prices and few active rigs. Moreover, sellers of illiquid assets are more financially constrained than sellers of liquid assets and buyers, suggesting that firms avoid selling highly specific assets until it is necessary.[9]

Maksimovic and Phillips (1998) examine whether assets sold by manufacturing firms are redeployed efficiently. Using plant-level data from the U.S. Census Bureau, they track changes in the productivity of assets and operating cash flows for firms entering Chapter 11 and their nonbankrupt industry rivals. Maksimovic and Phillips (1998) show that industry conditions are important in explaining economic decisions such as asset redeployment. Bankrupt firms in high-growth industries are more likely to sell assets than bankrupt firms in declining industries. Furthermore, in high-growth industries, the productivity of the assets sold increases under new ownership. This evidence is consistent with the efficient redeployment of assets to more productive uses and does not support the notion of fire sales in distressed industries. Interestingly, industry conditions are more important than Chapter 11 status in explaining changes in the productivity of assets, regardless of whether they are sold or retained by the firm.

Andrade and Kaplan (1998) also contribute to the body of evidence on asset sales by distressed firms. In a sample of highly leveraged transactions that subsequently became distressed, they find that the total costs of financial distress, measured as the change in the market value of the firm, are independent of the industry's stock performance. Since the market value includes the costs associated with asset sales, their evidence fails to establish that distressed industries force asset sales at greater discounts.

Overall, asset sales appear to be an important component of how firms deal with financial distress. The asset sales are often undertaken in conjunction with a restructuring of the firm's debt contracts. While such asset sales may be costly, because they are so commonly

[9] Asset liquidity can also influence the firm's choice of capital structure. Firms with more liquid assets tend to have higher debt levels and longer maturities; see, for example, Alderson and Betker (1995) and Benmelech, Garmaise, and Moskowitz (2005).

observed, it is conceivable that they still constitute a relatively low-cost mechanism to help resolve financial distress.

## 4. Debt workouts

Debt restructurings can be used to soften the hard contracts that cause financial distress. As outlined in Section 2, the distressed firm may reduce or defer payments on its debt contracts, or replace the debt with soft securities that have residual rather than fixed payoffs. We define a debt restructuring as a transaction in which an existing debt contract is replaced by a new debt contract with a reduction in the required interest or principal payments or an extension of maturity, or exchanged for common stock or securities convertible into common stock. In an out-of-court debt restructuring, claims are renegotiated via a workout or an exchange offer, without resorting to formal bankruptcy proceedings. A workout typically involves renegotiation of bank debt and other privately held claims. Publicly traded debt is restructured through an exchange offer, in which the distressed debt is exchanged for new debt or equity securities. This section surveys the empirical evidence related to different types of debt restructurings.

### *4.1. The choice between out-of-court restructuring and formal bankruptcy*

Many firms first attempt to resolve financial difficulties via a workout or exchange offer. Private mechanisms to restructure a financially distressed firm are expected to be less costly than formal bankruptcy proceedings. The greater are the cost savings, the greater are claimholders' incentives to settle privately. However, as discussed in Section 2.4, there are substantial impediments that hinder private restructurings, including asymmetric information, conflicts of interest among claimants, and holdout problems. When private mechanisms to resolve financial distress fail, the firm enters Chapter 11.

Early empirical work indicates that a substantial fraction of firms fail to successfully restructure out-of-court and file for Chapter 11 bankruptcy. Gilson, John, and Lang (1990) examine 169 financially distressed public companies that experienced extreme stock price declines and for which a debt restructuring is mentioned in the *Wall Street Journal*. Of these distressed firms, 80 (47%) restructure their debt out-of-court, while the remaining 89 firms (53%) fail to privately restructure their debt and subsequently file for Chapter 11. Franks and Torous (1994) investigate 161 firms that are downgraded to CCC or below by Standard and Poor's. They identify equal proportions of firms that complete a distressed exchange offer (76 firms) and firms filing for Chapter 11 (78 firms).

It is possible that the proportion of firms that successfully restructure out-of-court has declined. Altman and Stonberg (2006) track the size of the defaulted public bond and private debt markets. Recently, approximately 60% of defaults are concurrent with a bankruptcy filing, and many more defaulted bonds subsequently enter Chapter 11. This is an increase from the earlier years of Chapter 11 and suggests that private workouts have become relatively less common for distressed firms. One explanation could be legal rulings related to the treatment of claims in the event of a subsequent bankruptcy that discourage out-of-court restructurings relative to bankruptcy (Jensen, 1991).[10]

Following the legal rulings that discourage out-of-court restructurings, prepackaged bankruptcies (prepacks) became more widely used in the early 1990s and now replace some out-of-court restructurings, particularly for firms with public debt (Tashjian, Lease, and McConnell, 1996). Prepacks are a hybrid through which a reorganization plan is negotiated with creditors prior to bankruptcy and filed concurrently with the bankruptcy petition. They are sometimes done in conjunction with an out-of-court exchange offer; if the exchange offer fails to receive sufficient support, the firm can enter Chapter 11 and use votes solicited simultaneously with the exchange offer to confirm a reorganization plan in bankruptcy. Thus, firms filing prepacks can take advantage of certain attractive features of a Chapter 11 filing, such as beneficial tax treatment and voting rules to overcome a holdout problem, without going through long and costly bankruptcy proceedings (Betker, 1995a). Baird and Rasmussen (2003) estimate that one quarter of 93 large-firm Chapter 11 bankruptcies in 2002 were prepackaged bankruptcies.

Gilson, John, and Lang (1990) examine the determinants of firms' choice between formal bankruptcy and out-of-court restructuring. They find that the probability of completing an out-of-court restructuring is higher the greater proportion of the firm's assets that is intangible. The value of intangible assets is more likely to erode in bankruptcy, for example, through asset fire sales or perishing customer demand. Since bankruptcy is relatively more costly for firms with more intangible assets, these firms have greater incentives to preserve value via an out-of-court restructuring.

The study by Gilson et al. (1990) further shows that private workouts are more common when the firm has fewer distinct classes of debt outstanding and a greater proportion of the firm's long-term debt is bank debt. Conflicts of interest among different classes of creditors are more manageable the smaller the number of distinct creditor classes. Moreover, because banks are better informed than public debtholders, reducing potential information asymmetries, it is easier and therefore less costly for firms with banks as dominant creditors to renegotiate their debt. The bank debt is also more likely to be pivotal to the restructuring the greater is the proportion of bank debt, forcing the bank to internalize some of the restructuring costs. In contrast, with a greater proportion of diffusely held debt, such as public debt or trade credit, holdout problems become more severe.

Franks and Torous (1994) compare characteristics of the financial recontracting for firms completing public debt exchange offers and firms entering Chapter 11. They find

[10] One such decision was made in the LTV Corp. bankruptcy case. The bankruptcy case was filed on July 17, 1986 in the Southern District of New York, U.S. Bankruptcy Court. The court ruled on January 32, 1990 that debtholders who had participated in a prior out-of-court restructuring could only make a bankruptcy claim for the new reduced principal amount, while holdouts could claim the original principal amount. This decision discourages creditors from agreeing to reduce the principal of their debt claim in an out-of-court restructuring.

that the firms restructuring out-of-court are more solvent and liquid, and have less negative stock returns prior to the restructuring. Unlike Gilson et al. (1990), however, Franks and Torous do not find that firms restructuring out-of-court have a greater proportion of bank debt. This could be because the firms in their sample are larger and therefore rely less heavily on bank debt or because the bank loans of these firms often are syndicated and hence involve a larger number of banks. James (1995) and Asquith, Gertner, and Scharfstein (1994) also show that the presence of public bonds junior to the bank debt may impede restructurings.

Chatterjee, Dhillon, and Ramirez (1995) show that the firm's level of debt, its short-term liquidity, and the potential for coordination problems among creditors jointly determine the choice of restructuring mechanism. Firms filing for Chapter 11 are characterized by poor operating performance, high leverage, and coordination problems among creditors, whereas firms restructuring out-of-court tend to have relatively strong operating cash flows. They also examine firms filing prepacks and find that they typically have relatively strong operating performance but, in contrast to firms doing workouts, face an immediate liquidity crisis.

Asquith, Gertner, and Scharfstein (1994) provide similar evidence on the relationship between the firm's liability structure and the form of the restructuring. In particular, companies with more secured private debt and those with more complex public debt structures are more likely to enter Chapter 11. The larger fraction of secured debt may indicate a relatively low proportion of intangible assets, and thus less costly bankruptcy proceedings. They also find that 59% of firms whose banks agree to a debt restructuring ultimately enter bankruptcy, which suggests that these firms either did not reduce leverage sufficiently or did not adequately restructure assets to avoid bankruptcy. Altogether, the evidence indicates that conflicts between classes of claimants and holdout problems impede out-of-court restructurings, constrain the structure of out-of-court restructurings, or limit the effectiveness of out-of-court restructurings in the resolution of financial distress.

Although there are substantial impediments or limitations to out-of-court restructurings, the direct restructuring costs are likely to be substantially lower for an out-of-court restructuring than for a court-supervised bankruptcy. Measuring the direct costs of an out-of-court restructuring is often difficult because these costs are typically not reported separately from other nonrestructuring related operating expenses of the distressed firm. For example, although several studies of bank loan restructurings have been made, researchers have been unable to identify the related expenses. The costs can be observed, however, for the restructuring of public debt via a formal exchange offer. Gilson, John, and Lang (1990) document an average cost for 18 exchange offers of 0.6% of the book value of assets. The cost for 29 exchange offers studied by Betker (1997) is somewhat higher, with a mean of 2.5% of the pre-exchange assets (median 2.0%). In addition, out-of-court restructurings take significantly less time than Chapter 11 proceedings, suggesting that various indirect costs may be lower as well.

These estimates are useful in two respects. First, relatively low direct costs may make an out-of-court restructuring desirable relative to formal bankruptcy, particularly for firms with less severe impediments to a privately negotiated solution. Second, in considering a firm's ex-ante optimal leverage, relatively low costs of reorganizing would encourage firms to take advantage of the tax benefits of debt through higher leverage.

The stock market reaction to the announcement of a workout versus a bankruptcy filing corroborates the lower costs of workouts. Chatterjee, Dhillon, and Ramirez (1995) report less negative abnormal returns for announcement of workouts than Chapter 11 filings. Gilson, John and Lang (1990) show that stock returns on the announcement of debt renegotiations are more negative for firms that subsequently file for Chapter 11, suggesting that the market is able to identify firms that are more likely to succeed in restructuring their debt out-of-court.

Another circumstance indicating that there is greater firm value to share in workouts than in bankruptcy is documented by Franks and Torous (1994). They find that senior creditors in workouts are willing to forego a greater share of the value of the reorganized firm in favor of equityholders through deviations from the absolute priority rule. In exchange offers, all creditor classes relinquish some financial consideration to equity (on average 9% of the value of the reorganized firms), while the magnitude of these deviations is much smaller for firms in Chapter 11 (on average 2% of firm value). The fact that senior creditors are willing to give up a greater fraction of the firm to junior claimants in a workout suggests that on average firms attempting workouts may be less severely financially insolvent than bankrupt firms. Alternatively, if senior creditors prefer a smaller fraction of a potentially more valuable firm in a workout than a larger fraction of a potentially less valuable firm in bankruptcy, then this suggests lower overall costs of a workout compared to a bankruptcy.

### *4.2. Characteristics of debt restructurings*

A number of studies have documented various aspects of out-of-court restructurings such as the medium of exchange and debt recovery rates. Asquith, Gertner, and Scharfstein (1994) study the characteristics of private bank debt restructurings. They report that bank lenders respond to financial distress in various ways, including requiring accelerated payments and reducing further lending. Banks also waive covenants but rarely agree to a reduction in the principal amount of their claim. James (1995) expands on these results for a sample of 102 debt restructurings. He shows that banks make concessions only if public debtholders also agree to restructure their claims. In general, banks are more likely to forgive principal and take equity when a smaller fraction of the debt is held by public creditors.

James (1996) demonstrates that bank participation in the workout is important because it facilitates public debt exchange offers. Compared to restructurings in which banks do not participate, exchange offers accompanied by bank concessions have a higher likelihood of succeeding and involve significantly greater reductions in public debt outstanding and less senior debt offered to bondholders. Thus, the characteristics of a firm's debt structure help explain what form of restructuring will be feasible.

Evidence on the characteristics of distressed public debt exchanges is presented by Franks and Torous (1994). They find that a majority of the payments in exchanges of senior public debt are in the form of cash (29%) and new senior debt (38%), whereas the majority of payments in exchanges of junior debt constitutes common stock (67%). They further show that creditor recovery rates tend to be substantially higher in distressed exchange offers than in Chapter 11 reorganizations. Also, relative to Chapter 11 reorganizations, cash is used less extensively and equityholders typically get to retain a larger fraction of the reorganized firm's equity.

Brown, James, and Mooradian (1993) examine how the type of securities offered in a debt restructuring relates to information asymmetries about the firm's prospects. When firms offer equity to private lenders, who tend to be better informed about the firm, and senior debt to public debtholders, this conveys positive information about firm value. In contrast, abnormal announcement returns are negative when private lenders are offered senior debt and public lenders are offered equity.

The participation of investment banks in public debt exchange offers is investigated by Mooradian and Ryan (2005). Firms can chose to conduct a public debt exchange offer without involving an investment bank. Though costly, 61% of the sample firms engage an investment bank as an intermediary in the distressed exchange offers. Mooradian and Ryan show that investment bank participation decreases with the level of commercial bank debt outstanding and increases with bank loan concessions, firm size, number of public debt contracts outstanding, and size of the proposed debt reduction. This suggests that financially distressed firms hire an investment bank to manage their exchange offers when the debt structure is complex and there is a greater need for help in mitigating potential impediments to an out-of-court restructuring. Interestingly, the investment-bank-managed exchange offers involve less senior debt to bondholders, achieve greater debt reduction, and result in better post-restructuring operating performance.

The use of coercive tactics to alleviate holdout problems can be beneficial to the firm. A coercive offer involves a consent agreement to issue a more senior class of debt (which only requires a two-thirds majority vote) combined with an exchange offer replacing the current debt with a more senior debt issue requiring lower interest payments, less principal, or longer maturity. The offer is coercive because if the exchange offer is successful, a creditor holding out ends up with a more junior claim, albeit with more favorable terms. Chatterjee, Dhillon, and Ramirez (1995) report higher completion rates and a higher proportion of bonds tendered or exchanged when exchange offers are coercive, indicating that the coercion helps alleviate the holdout problem. They also show that the equity and debt price reactions to the announcement of the exchange offer indicate that coercion may benefit stockholders without being detrimental to bondholders.

The general conclusion from much of this literature is that absent holdout problems and other coordination problems, private debt restructurings such as exchange offers provide a lower cost restructuring mechanism than formal bankruptcy. Moreover, various characteristics of the financially distressed firm's capital structure and asset composition determine the severity of the impediments to a successful out-of-court restructuring.

## 5. Governance of distressed firms

The governance structure in bankruptcy determines the relative influence of different stakeholders over the process and hence the outcome of the reorganization. Because bankruptcy is a likely event if an out-of-court restructuring fails, the governance structure in bankruptcy also affects the relative power of claimants outside of bankruptcy and is thus influential in shaping any out-of-court restructuring.

Many aspects of a firm's governance are affected when a firm becomes financially distressed. The fiduciary duties of managers and directors, normally owed to the firm's shareholders, expand to include creditors. With conflicting interests between various debtholders and equityholders, corporate executives may be caught in the middle. Both managers and directors typically experience a higher turnover than normal. Also, most significant restructurings lead to large changes in ownership, with creditors often emerging as new owners of the firm. The mechanisms through which the change in control occurs, however, can be quite different from those for nondistressed firms. This section discusses various aspects of governance and their impact on the incentives of managers and other participants in the restructuring process.

### *5.1. Conflicts of interest and the fiduciary duties of managers and directors*

When a corporation is solvent, the managers and directors have fiduciary duties to the corporation and its shareholders. When a company is in financial distress, however, decisions increasing the value of equity may in fact reduce total firm value. Thus, it is no longer clear that the decision making should be left to agents whose incentives are aligned with equity. The courts recognize this problem by extending the fiduciary duties of directors and officers to also include creditors when the firm becomes insolvent (Branch, 2000). This expansion of the fiduciary duties creates potential difficulties in defining managers' responsibilities, however, since shareholders and senior creditors often have opposing interests.

The 1989 bankruptcy of Eastern Airlines, described by Weiss and Wruck (1998), illustrates the potential magnitude of such conflicts. Relying on an offer to purchase the company, Weiss and Wruck estimate the equity going-concern value at the time of filing at approximately \$1.2 billion. Based on the perceived continuation value, creditors and other groups initially supported management's attempts to reorganize. However, even as Eastern continued to experience large operating losses, it was granted the right to use cash available from asset sales to continue operating. Weiss and Wruck estimate a decline in the value of the airline of more than \$2 billion over a 22-month period in bankruptcy. If management were acting solely in shareholders' interests, its best strategy was to continue operating the airline, hoping for a recovery of the business. Given the decline in asset value, however, creditors would have fared better if the cash had been used to pay their claims rather than continue funding unprofitable operations. Reorganization attempts ultimately failed, and Eastern was liquidated under Chapter 7 in 1991.

While Eastern Airlines' bankruptcy provides an extreme example of the tensions between incentives to reorganize versus liquidate, conflicts between different claimholders are manifested in many reorganization cases. Macy's bankruptcy is another example of conflicting interests between stakeholders (Noe and Rebello, 2003). After filing for bankruptcy in 1992, Macy's management embarked on a plan to restructure the operations and close underperforming stores with the objective of ultimately emerging from bankruptcy as an independent company. Negotiations between management, shareholders, and creditors over the reorganization plan remained deadlocked, however. To break the deadlock, Macy's creditors enlisted Federated Department Stores to make a bid for the bankrupt company. Management contested the acquisition and repeatedly sought extension of the exclusivity period to prevent competing reorganization plans. A fraction of the board headed by a bondholder, Laurence Tisch, opposed management's plan. Eventually, Federated and Macy's creditors jointly filed a plan under which Federated gained control of the company, providing no distribution to shareholders.

Until a debt restructuring is completed, the interests of different claimholders regarding the firm's investment decisions can deviate substantially. Chapter 11 provides features that are aimed at balancing such conflicts of interest. The "pro-debtor" provisions of the Bankruptcy Code yield considerable influence to incumbent management over the course of the restructuring and development of the reorganization plan. At the same time, both creditors and the court are granted substantial oversight of the proceedings. Unsecured creditors typically are represented by a committee, giving them influence over the negotiation process. The appointment of other committees, however, is more uncertain. Betker (1995b), for example, documents the formation of an equity committee in one-third of his sample of 75 large Chapter 11 cases.

To speed up the confirmation of a reorganization plan, preventing further deterioration of asset values, senior creditors may agree to a side-payment to junior creditors and equityholders. Such side-payments show up as deviations from the absolute priority rule. The priority of claims is violated for three-quarters of the Chapter 11 cases in Franks and Torous (1989), Eberhart, Moore, and Roenfeldt (1990), and Weiss (1990). For a more recent sample of Chapter 11 filings, Bris, Welch, and Zhu (2006) find violations of the absolute priority rule in only 12% of the cases. The much lower incidence of deviations from the priority of claims could partly be explained by a smaller firm size in Bris et al. (2006) and thus a less complex proceeding, and partly by a change in the view and enforcement of creditor rights. This trend is, however, corroborated by Bharath, Panchapegesan, and Werner (2007), who examine 531 large firms that filed for Chapter 11 between 1991 and 2005. While 26% of the bankruptcy cases in the 1990s involve deviations from absolute priority, such deviations are recorded for only 9% of the cases after year 2000.

Similarly, for a sample of 153 large corporate Chapter 11 filings in 2001, Ayotte and Morrison (2007) report that very few reorganization plans (6% or less) violate absolute priority rules by distributing any value to equityholders.[11] They argue

[11] This measure does not account for distributions to equityholders of warrants, which are usually the right to buy out the creditors at the face value of their claims.

that governance in Chapter 11 has shifted to emphasize creditor control and creditor conflict. Senior lenders exercise control through pre- and post-petition lines of credit, which limit the debtor's access to financing and impose strict requirements on business activity. Three-quarters of their sample firms obtain DIP financing, typically secured by a lien on all corporate assets. The vast majority of loans contain covenants imposing line-item budgets, profitability targets, and deadlines for reorganization plans. If these covenants are violated, the lender is generally free to seize collateral unilaterally, without seeking court approval. Ayotte and Morrison (2007) further document that junior lenders use claims trading, committees, and other tactics to gain control over the reorganization process. Acting through the unsecured creditors committee, junior creditors file objections in over half of the sample cases. In almost as many cases, DIP lenders object to actions proposed or taken by incumbent management. Amendments to the U.S. Bankruptcy Code effective October 2005 have further increased creditor influence in Chapter 11.

In sum, when a firm becomes financially distressed, the residual claim often shifts from equityholders to creditors. This creates conflicts of interest regarding the firm's investment and continuation decisions that have an important effect on bankruptcy outcome.

### 5.2. *Management and board changes*

Critics of Chapter 11 bankruptcy suggest that the process protects bad managers from being removed. Bradley and Rosenzweig (1992) argue, on the one hand, that bankruptcy law allows management to go relatively unpunished, retaining control over corporate assets, even when their own actions helped to render the firm insolvent. On the other hand, operating decisions of healthy firms will be affected by an increased likelihood that managers are replaced in the event of financial distress. For example, managers may be reluctant to undertake highly profitable (positive net present value) but also highly risky investments if they are likely to be fired should the investment fail.

Several academic papers examine whether financial distress is costly to managers in the sense that they are likely to lose their jobs. Gilson (1989) examines the turnover of managers carrying the title of CEO, chairman, and president over a four-year period beginning two years prior to bankruptcy filing or debt restructuring. For 69 firms filing for bankruptcy, 71% of managers are replaced over the four years. This turnover rate is significantly higher than that of financially distressed firms that successfully restructure their debt out of court. None of the executives who lose their position are employed by another publicly traded firm over a three-year period following their departure, suggesting that the personal costs are significant.

Other studies of management replacement rates for failing firms show similarly high turnover. Betker (1995b) reports a 91% turnover of CEOs in office two years prior to filing by the time the firm emerges from bankruptcy. In comparison, Weisbach (1988) and Warner, Watts, and Wruck (1988) document substantially lower CEO turnover rates for nondistressed firms. Moreover, both studies show that management turnover increases as firm performance deteriorates. In a more recent study, Ayotte and Morrison (2007) find that 70% of CEOs are replaced within two years of a bankruptcy filing.

While the turnover of managers is abnormally high for distressed firms in general, certain bankruptcy courts (e.g., Delaware) have been alleged to maintain relatively strong pro-debtor biases. LoPucki (2004) argues that managers choose to file for bankruptcy in such districts, where they expect to receive favorable rulings that help them retain control of the reorganization process. The documented high turnover of managers, however, runs counter to the notion that they are overly protected by the process. Gilson (1989), Betker (1995b), and Hotchkiss (1995) show that although a significant fraction of managers is able to stay in place until a plan is proposed, it is unlikely that they still remain when the firm emerges from bankruptcy.

Financial distress also leads to significant changes in the membership and composition of boards. Distressed firms require a substantial commitment of time and attention from managers and directors to address the firm's operating problems and develop a restructuring plan. Some directors may resign in anticipation of the firm's problems and the implications for the board. Such concerns can potentially make it difficult to recruit new outside directors. Countering the problems with a shrinking board is that certain parties, such as large creditors or outsiders investing in the distressed firm, may seek board seats to protect their interests in the restructuring.

Gilson (1990) finds that although average board size declines for distressed firms, replacement directors often possess some special skill or interest in managing troubled companies (for example, investment bankers or workout specialists). On average, only 46% of the board members prior to financial distress are still present two years after a reorganization or debt restructuring. Hotchkiss and Mooradian (1997) show that "vulture" investors are frequently active in the governance of firms defaulting on their public debt. These investors join the board of directors for 28% of the firms they study, often maintaining these positions for at least one year after emergence from bankruptcy.

In summary, the literature documents the increase in top management turnover rates as firms become financially distressed, suggesting large personal costs for incumbent managers. Director turnover is also high, often resulting in new restructuring specialists joining the board.

### *5.3. Management compensation in financial distress*

Compensation contracts are a common means to align managers' incentives. In financial distress, the compensation policy is often an integral part of the firm's overall restructuring strategy, for example, through providing incentives that facilitate negotiations with creditors or encourage a speedy resolution. Once in bankruptcy, contracts with key employees are subject to the approval of the bankruptcy court.

Gilson and Vetsuypens (1993) examine the compensation contracts of managers that are in place as the firm enters financial distress and the contracts of the managers replacing them. They find that managers who retain their position through a debt restructuring often take a substantial cut in salary and bonus. Replacement CEOs who were previous employees of the firm earn a median of 35% less than their predecessors. In contrast, the median outside replacement CEO earns 35% more than the manager he or she replaces.

The compensation of CEOs of emerging firms exhibits high sensitivity to the post-bankruptcy stock performance (Gilson and Vetsuypens, 2003). For a sample of 63 Chapter 11 cases, Gilson, Hotchkiss, and Ruback (2000) show that half of the managers receive stock and options in the reorganized firm. Stock-based incentive compensation, however, may be associated with a downward bias in cash flows projected for the reorganized firm. A low reorganization value can create a windfall for managers if the option exercise price is set to that low value or the number of shares that managers receive increases with a lower initial stock price. Nevertheless, the form of the compensation contract for managers of the reorganized firm will affect management's efforts in developing a reorganization plan.

A common approach in financial distress is to tie management compensation to the successful resolution of the firm's bankruptcy or debt restructuring, or to the recovery of certain creditor groups. Gilson and Vetsuypens (2003) describe cases in which the CEO is granted a substantial salary increase as a reward for successfully bringing the firm through its financial restructuring or in which part of the CEO's compensation is deferred until the financial restructuring is completed. They further observe cases in which the CEO incentives are tied to the value of creditor claims, for example, by awarding claims with similar characteristics as those held by creditors, or paying a bonus based on the amount of cash creditors receive under the reorganization plan or as a result of asset sales.

Another prevalent practice that has been criticized is the granting of generous retention plans to certain executives and key employees for remaining with the company during the course of the bankruptcy reorganization. Such key employee retention plans (KERPs) led to widespread controversies since they were often accompanied by massive layoffs and wage concessions, and they are now severely limited by the 2005 amendments to the U.S. Bankruptcy Code. Two recent court rulings, however, circumvent these limitations by allowing the debtors to use bonus compensation plans to provide adequate financial incentives to management during the reorganization.[12]

The repricing of executive stock options for firms that have performed poorly has also received much attention. Repricing refers to the practice of lowering the strike price of previously issued employee stock options, typically following a significant stock price decline. Although repricing may reward management following a period of poor performance, it may also be necessary in order to restore appropriate incentives for management.[13] Chidambaran and Prabhala (2003) show that a majority of the repriced options have a new vesting period or exercise restrictions related to continued employment. This suggests that repricing may be useful in the motivation and retention of keyemployees.

[12] In re Global Home Products, LLC 1 and In re Nellson Nutraceutical, 2.

[13] See Acharya, John, and Sundaram (2000) for a theoretical analysis of the trade-off between reducing current-period incentives and restoring continuation incentives that determine the optimality of repricing options.

Repricing has more recently been replaced by a practice known as rescission. In a rescission, shares received by the employee from exercise of the options are returned to the company in exchange for a refund of the strike price. Similar to repricing, this practice has been criticized as symptomatic of poor governance, yet it may be necessary to restore incentive structures.

Overall, CEO salaries tend to decline when their firms become financially distressed. The distressed firms, however, often put in place new management compensation contracts that increase the sensitivity of pay either to a successful resolution of the restructuring or to post-bankruptcy equity performance. Stock or option grants in the emerging firms risk leading to a downward bias in the valuation on which a reorganization plan is based.

### *5.4. Changes in ownership and control*

A distressed debt restructuring typically results in a substantial change in the ownership of the firm. The primary reason is that the poor performance has eroded the equity value, so that shareholders often receive little or no equity in the reorganized firm. Much of the reorganized firm's stock is distributed to a subset of existing creditors, who become the new owners of the firm.

In Gilson's (1990) study of 61 firms filing for bankruptcy, on average 80% of the common stock in the reorganized firm is distributed to creditors. The distribution of stock in exchange for pre-petition debt claims can frequently lead to a change in control. Federal and state banking laws provide U.S. banks with authority to hold common stock received in loan restructurings. For three-quarters of all 111 financially distressed firms in Gilson's (1990) sample, bank lenders and other creditors receive significant blocks of voting stock in the restructured firm. Banks receive on average 36% of the firm's common stock and frequently appoint representatives to the board of directors. James (1995) studies 102 distressed bank debt restructurings and finds that banks take equity positions in 31% of the transactions. Moreover, the banks typically maintain a substantial equity stake for at least two years following the restructuring.

Although asset sales are common, early studies of ownership changes of firms in Chapter 11 detect relatively few acquisitions of the bankrupt firm as a whole by other operating companies (Hotchkiss and Mooradian, 1998). A possible explanation is that Chapter 11, by allowing incumbent management to retain control, discourages potential acquirers. Furthermore, industry rivals may be distressed and lack the financial strength to bid for the bankrupt firm.

Hotchkiss and Mooradian (1998) examine a sample of 55 acquisitions of firms in Chapter 11 by other public companies. The bidding firm is often in the same industry and frequently has some prior relationship (such as a previous asset purchase) with the target. One-third of the transactions they examine involve multiple bidders, indicating substantial competition for the bankrupt targets. Transactions prices, however, are significantly lower than those paid for nonbankrupt firms matched on size and industry. More recently, as sales of businesses through Section 363 of the bankruptcy code have become more common, M&A activity involving bankrupt targets is observed more frequently (Baird and Rasmussen, 2003).

Along with the increase in takeover activity in Chapter 11, changes in control through claims trading have also become more commonplace. The market for trading claims of distressed firms has grown dramatically since the early 1990s. This market provides banks and other creditors with an opportunity to exit the process earlier, with new investors taking the place of existing creditors in the negotiation of a restructuring plan. A common strategy for an investor who specifically seeks control of a distressed company is to purchase a large block of debt. With a stake sufficiently large to block a reorganization plan, the investor gains influence over the course of the restructuring. Depending on the final negotiated terms of the plan, the stake potentially can be converted into a controlling ownership position. The debt security that will ultimately be exchanged for equity is commonly referred to as the "fulcrum" security. Examining a sample of 288 firms defaulting on their debt between 1980 and 1993, Hotchkiss and Mooradian (1997) find that vulture investors become blockholders (owning more than 5% of the reorganized firm's stock) for half of the sample firms and gain control of 16% of the firms. Some investors have developed a reputation for using this strategy to gain control of firms in bankruptcy, and as a result, they manage a portfolio of reorganized firms (Apollo Advisors, for example).

Equity infusions in the reorganized firm can also shift control to a new investor. Gilson, Hotchkiss, and Ruback (2000) find such equity investments for 12 of the 63 firms (19%) in their sample, resulting in the investors owning a median of 54% of the reorganized firm's stock. The activity of these investors, together with high management and board turnover, contributes to significant changes in the governance of distressed firms.

## 6. Bankruptcy costs

A restructuring can be costly because of asymmetric information, coordination problems among creditors, and conflicting interests of different claimholders. Distressed firms incur direct expenses for lawyers, accountants, financial advisers, and other turnaround professionals. In addition, over the course of a distressed restructuring, the firm may pursue a suboptimal investment policy or inefficiently liquidate assets due to insufficient liquidity and limited ability to obtain new financing. Indirect costs of financial distress include unobservable opportunity costs, such as lost sales driven by the firm's deteriorating financial condition and lack of management attention on the business itself. This section reviews estimates of the different costs related to financial distress and bankruptcy.

### 6.1. Direct costs

Studies estimating the direct costs for firms reorganizing in Chapter 11 are listed in Table 1 (Altman and Hotchkiss, 2006). The sample-size weighted average direct cost across the seven studies of Chapter 11 is 6.5% of the book value of assets. Since there is no single

Table 1

Estimates of direct costs of formal bankruptcy proceedings in the United States

| Study | Sample | Time period | Estimated costs |
|---|---|---|---|
| **Traditional Chapter 11 cases:** | | | |
| Warner (1977) | 11 bankrupt railroads; estimated mean market value $50 million at filing. | 1933–1955 | Mean 4% of market value of firm one year prior to default. |
| Altman (1984) | 19 Chapter 11 cases; mean assets $110 million before filing. | 1974–1978 | Mean 4% (median 1.7%) of firm value just prior to bankruptcy for 12 retailers; 9.8% (6.4%) for 7 industrial firms. |
| Weiss (1990) | 37 cases from 7 bankruptcy courts; average total assets before filing $230 million. | 1980–1986 | Mean 3.1% (median 2.6%) of firm value prior to filing. |
| Betker (1997) | 75 cases; mean assets FYE before restructuring $675 million. | 1986–1993 | Mean 3.9% (median 3.4%). |
| Lubben (2000) | 22 cases; median assets $50 million. | 1994 | Mean 2.5%. |
| LoPucki and Doherty (2004) | 48 cases from Delaware and Southern District of NY; mean assets at filing $480 million. | 1998–2002 | Mean 1.4% of assets at beginning of case. |
| Bris, Welch, and Zhu (2006) | 225 cases from Arizona and Southern District of NY; mean pre-bankruptcy assets $19.8 million. | 1995–2001 | Mean 9.5%, median 2%. |
| **Prepackaged bankruptcies:** | | | |
| Betker (1997) | 48 prepackaged Chapter 11 cases; mean assets FYE before restructuring $675 million. | 1986–1993 | Mean 2.8% (median 2.4%) of pre-bankruptcy total assets. |
| Tashjian, Lease, and McConnell (1996) | 39 prepackaged Chapter 11 cases; mean book value assets FYE before filing $570 million. | 1986–1993 | Mean 1.8%, median 1.4% of book value of assets at fiscal year-end preceding filing. |
| **Chapter 7 cases and liquidations:** | | | |
| Ang, Chua, and McConnell (1982) | 86 liquidations, Western District of Oklahoma; estimated mean pre-bankruptcy assets $615,516. | 1963–1979 | Mean 7.5% (median 1.7%) of total liquidating value of assets. |
| Lawless and Ferris (1997) | 98 Chapter 7 cases from 6 bankruptcy courts; median total assets $107,603. | 1991–1995 | Mean 6.1% (median 1.1%) of total assets at filing. |
| Bris, Welch, and Zhu (2006) | 61 Arizona and S.D.N.Y. Chapter 7 cases; mean pre-bankruptcy assets $501,866. | 1995–2001 | Mean 8.1%, median 2.5% of pre-bankruptcy assets. |

*Source*: Altman & Hotchkiss (2006), p. 95.

source of comprehensive information for Chapter 11 cases, studies make use of court documents collected from one or more of the federal bankruptcy courts. The studies cover a wide variety of firms, including everything from large railroads (Warner, 1977) to relatively small firms (Lawless and Ferris, 1997). The range of estimates of direct costs is therefore quite wide, with means ranging from 1 to 10% and medians from 2 to 6%.

Researchers generally interpret these numbers as evidence of relatively low direct costs, particularly in relation to the potential tax benefits of using debt. The direct costs also appear to have a fixed component, explaining why a Chapter 11 reorganization may not be feasible for some smaller firms. For large public companies in bankruptcy, the mean professional fees as a percentage of pre-filing assets ranges from 1 to 3% (Lubben, 2000; Weiss, 1990). Though relatively small on a percentage basis, the dollar amount of fees in large public bankruptcy cases can be significant.

Firms undertaking prepackaged bankruptcies seek agreement among claimholders on terms of the financial restructuring prior to filing. Prepackaged bankruptcies generally allow firms to exit bankruptcy within months and are therefore expected to have lower direct costs than a lengthier bankruptcy proceeding. Betker (1997) finds direct costs for prepackaged bankruptcies of on average 2.8% of the pre-bankruptcy total assets. This cost estimate includes all the pre-bankruptcy expenses of informal bondholder committees and banks, where most of the costs are incurred and for which the bankrupt firm routinely pays. Tashjian, Lease, and McConnell (1996) show that direct costs for prepacks average 1.8% of the book value of pre-filing assets and 1.6% for the subsample of cases that are pre-voted. Thus, the costs of prepackaged bankruptcies appear to fall somewhere between those observed for traditional Chapter 11 cases and those documented by Gilson, John, and Lang (1990) for out-of-court exchange offers.

While most attention is devoted to the costs of Chapter 11 proceedings, a few studies examine the costs of liquidations under Chapter 7. Bris, Welch, and Zhu (2006) document bankruptcy expenses of on average 8.1% (median 2.5%) of pre-bankruptcy assets for a sample of 61 smaller nonpublic firms. Based on their estimates of the post-bankruptcy remaining value, however, the bankruptcy fees exceed the value of the entire estate in two-thirds of the cases. Lawless and Ferris (1997) find that the fees in Chapter 7 on average amount to 6.1% of total assets.

Bankruptcy costs are likely to increase with the time that the firm spends in bankruptcy. Franks and Torous (1989) report that the average bankruptcy takes 2.7 years for 14 firms filing after the 1978 Bankruptcy Code took effect. The average time from filing of the bankruptcy petition to resolution is 2.5 years in Weiss (1990) and 2.2 years in Franks and Torous (1994). For the sample in Bris, Welsh, and Zhu (2006), which is both more recent and contains smaller firms, the average Chapter 11 proceeding lasts 2.3 years (median 2.4 years). They show that the length of the bankruptcy procedure is independent of firm size but varies with the specific judge overseeing the case. The duration of Chapter 11 reorganization is further found to decrease with the operating profitability of the industry

(Denis and Rodgers, 2007). Bharath, Panchapegesan, and Werner (2007) show that the time to resolution in Chapter 11 has declined and on average is 16 months in the 2000–2005 period. Morrison (2007) finds a median duration of 8 months for the 36 small business Chapter 11 cases in 1998 that emerge as going concerns.[14]

The relatively low direct costs of exchange offers discussed in Section 4.2, as well as the increasing use of prepackaged bankruptcies, suggest that cost savings can be significant for firms that successfully restructure without entering a traditional Chapter 11 procedure.

### *6.2. Indirect costs*

The magnitude of indirect costs relative to direct costs, and therefore their importance to theories of debt structure and reorganization, can be large. Indirect costs, however, are unobservable and therefore more challenging to estimate empirically.

One of the first attempts to study indirect costs is Altman (1984). Altman compares expected profits to actual profits over the three years prior to bankruptcy (years $-3$ to $-1$) for a sample of 19 firms entering Chapter 11. Expected profits are based either on a comparison of each firm's sales and profit margin to industry levels prior to year $-3$ or on security analyst estimates. He finds that the indirect costs, that is, the difference in profits, average 10% of firm value just prior to bankruptcy. The combined direct and indirect costs are on average 17% of firm value. It is, however, impossible to distinguish whether the decline in profits is a result of the financial distress itself (and therefore is an indirect cost) or a result of the same economic factors that caused financial distress in the first place.

Opler and Titman (1994) address this causality problem by selecting firms in industries that experience economic distress, defined as declining industry sales and median stock returns below $-30\%$. They find that firms with higher leverage ratios prior to the onset of industry economic distress experience a greater decline in market share and operating profits, consistent with the notion that there are significant indirect costs of financial distress.

Subsequent studies recognize that in order to provide specific estimates of indirect costs, it is useful to separate the effects of financial versus economic distress. Andrade and Kaplan (1998) examine 31 firms that become distressed subsequent to a highly leveraged transaction. Given the high ex-ante leverage of these firms, they are largely financially distressed but not economically distressed, allowing an observation of the costs of "pure" financial distress. Andrade and Kaplan (1998) report that the distressed firms cut capital expenditures, sell assets, and delay restructuring or filing for Chapter 11 in a way that appears to be costly. Based on changes in firm market value over time, they estimate the net costs of financial distress to range from 10 to 20% of firm value. In addition, they find that these costs are concentrated in the period after the firm becomes distressed, but

[14] See also Flynn (1989), Gilson, John, and Lang (1990), Hotchkiss (1995), and Betker (1997) for evidence on the length of Chapter 11 proceedings.

before it enters Chapter 11, suggesting that the indirect costs are not caused by Chapter 11 itself.[15]

As discussed in Section 3.2, Maksimovic and Phillips (1998) show that industry conditions are much more important than bankruptcy status to explain the productivity, asset sales, and closure decisions of Chapter 11 firms. Similar to Andrade and Kaplan (1998), this indicates that few real economic costs are attributable to Chapter 11 and that bankruptcy status is marginal to indirect costs. Pulvino (1999), in contrast, finds that bankrupt airlines sell aircrafts at prices that generally are lower than those received by distressed but nonbankrupt firms, implying that bankruptcy status could influence these costs.

A bankruptcy filing may convey negative private information about industrywide business conditions. Studying the effect of the bankruptcy announcements of 59 failing firms, Lang and Stulz (1992) find a 1% price decline in a value-weighted portfolio of competitor stock. The effect is greater for relatively highly leveraged industries. In contrast, competitors in concentrated industries with low leverage experience positive announcement returns, perhaps because the exit creates a windfall for the surviving rival firms. The negative stock price reaction of industry rivals to the bankruptcy announcement of large firms is confirmed by Ferris, Jayaraman, and Makhija (1997). They further show that competitors who subsequently file for bankruptcy experience the greatest decline in equity value. Haensly, Theis, and Swanson (2001), in contrast, find insignificant announcement returns for industry rivals. The stock returns are negative, however, in industries with relatively high leverage. Hertzel, Li, Officer, and Rodgers (2007) show that distress related to bankruptcy filing also is associated with negative and singificant stock price effects for suppliers, in particular when intra-industry contagion is severe.

Debt recovery rates, defined as the bankruptcy payoff to creditors as a fraction of the face value of their claims, reflect the value of the distressed firm's assets net of all direct and indirect costs. Franks and Torous (1994) report total recovery rates of on average 51% for 37 Chapter 11 cases. Bris, Welch, and Zhu (2006) document average recovery rates in Chapter 11 of 69% (median 79%). A caveat with these recovery rate estimates, however, is that a majority of the distributions are in the form of new claims valued at face value. For a subsample of 12 firms in Frank and Torous with available market values for all claims in the reorganized firm, the median recovery rate is a lower 41%.

Although the evidence is mixed with respect to whether indirect costs are largely incurred during the period of financial distress prior to bankruptcy or while in formal bankruptcy, such costs appear to be of greater magnitude than the direct bankruptcy costs. Thus, firms with potentially large opportunity costs of operating in financial distress are more likely to choose lower debt levels ex ante and, once in financial distress,

[15] Kaplan (1989, 1994) provides an illustration of the indirect costs of financial distress in the context of Campeau's acquisition of Federated. See also Cutler and Summer's (1988) analysis of the Texaco-Pennzoil litigation.

select a restructuring mechanism that resolves the financial distress both faster and more fully.

## 7. The success of chapter 11 reorganization

One measure of a "successful" restructuring is that a consensual agreement between claimants is ultimately reached, putting in place a modified set of financial contracts and/or liquidating all or a portion of the firm's assets to meet its obligations. In terms of Chapter 11, however, "success" implies that the firm is able to reorganize rather than liquidate. If the Bankruptcy Code is structured such that some inefficient firms are allowed to reorganize (i.e., their estimated going-concern value is less than their unobserved liquidation value at the time of reorganization), researchers need to consider the performance of the firm some time after it has emerged to ultimately argue whether the restructuring has been successful. In this section, we focus on these aspects of Chapter 11 restructurings, rather than on outcomes of private restructurings, as this literature relates to the important debate over the efficient design of a bankruptcy code.

### *7.1. Outcomes of chapter 11 filings*

The Executive Office for U.S. Trustees provides statistics for confirmation rates of Chapter 11 cases in the United States. It is clear from their statistics that many firms entering Chapter 11 ultimately are not successful in having a plan of reorganization confirmed; for the years 1990 through 2003, confirmation rates do not exceed 45% in any single year. The national average confirmation rate for this time period is only 29% of cases. Furthermore, many of the plans that are confirmed are "liquidating Chapter 11" plans, providing an alternative mechanism for liquidation other than the Chapter 7 process. The large number of cases that do not reach confirmation are ultimately closed with no remaining value, or converted to a Chapter 7 case.[16]

For the subset of Chapter 11 cases successfully confirming a plan, the disposition of the firm's assets can still vary in significant ways. Unfortunately, information is generally not available for nonpublic companies, and for public firms it must be compiled from various sources including news services. Hotchkiss and Mooradian (2004) examine 1770 public companies that filed for Chapter 11 between 1979 and 2002. A publicly cited resolution of the outcome by June 2004 is available for some 1400 cases (80%). The remaining cases are either still in bankruptcy as of 2004 or have likely ended in liquidation. The bankrupt firm emerges as a public company (44% of cases) or a deregistered private company (27%), is liquidated (21%, including conversions to Chapter 7), or merges with another operating company (8%). Similar proportions are

[16] Statistics are available at http://www.usdoj.gov/ust. Statistics include all states except Alabama and North Carolina.

reported by Hotchkiss (1995) for the subset of firms filing prior to 1988. Using data from two courts (Arizona and the Southern District of New York) where bankruptcy documents are available electronically, Bris, Welch, and Zhu (2006) find that 52% of 150 firms reorganized under Chapter 11 firms continue as independent companies.[17]

A smaller number of firms merge with another operating company while in bankruptcy. Hotchkiss and Mooradian (1998) show that the combined cash flows of the merged company increase by more than is observed for similar nonbankrupt transactions, suggesting that these mergers represent a successful restructuring outcome. For smaller firms, acquisitions are more common. White (1984) finds that in a sample of 64 small corporations in Chapter 11, 23% of the firms are sold as going concerns, 47% adopt reorganization plans, and the remaining 30% are eventually liquidated. Examining 95 relatively small corporate bankruptcy filings in Chicago during 1998, Morrison (2007) reports that 9 firms (9%) are sold as a going concern and another 27 firms (28%) exit as a reorganized entity, while 28 firms (29%) are shut down in bankruptcy and the remaining 31 firms (33%) exit Chapter 11 without a new capital structure, typically followed by a subsequent liquidation.[18]

Several studies have examined factors influencing the probability that a firm successfully emerges from Chapter 11. Hotchkiss (1993) shows that firm size, measured by pre-petition assets, is the utmost important characteristic determining whether a firm will be successfully reorganized rather than liquidated. Many of the emerging firms have considerably downsized while in bankruptcy. She suggests that the ability to divest assets and use the proceeds to fund the remaining operations is critical to the firm's survival in Chapter 11. Similarly, Denis and Rodgers (2007) provide documentation that firms with significant reductions in assets and liabilities in bankruptcy are more likely to emerge as going concerns. If asset prices are temporarily depressed by low industry demand, a liquidation or sale may be relatively costly to the creditors of defaulted firms. Acharya, Bharath, and Srinivasan (2007) show that most distressed firms emerge as restructured entities during periods of industry distress, possibly as a way of avoiding costly asset fire sales.

Carapeto (1999) and Dahiya, John, Puri, and Ramirez (2003) argue that access to DIP financing is an important factor in a successful reorganization. The availability of DIP financing is particularly important to firms in desperate need of fresh working capital, such as retailers whose suppliers might otherwise discontinue business. Using a sample of 538 public companies in Chapter 11, Dahiya et al. (2003) show that the probability of emerging as a reorganized entity is higher for firms receiving DIP financing. The benefits from DIP financing are further documented by Chatterjee, Dhillon, and Ramirez (2004), who report significantly positive abnormal stock and bond returns at the announcement of DIP loans.

Another factor that could affect the reorganization is the individual judge's interpretation and application of the bankruptcy law. Chang and Schoar (2006) find significant

[17] For evidence on the outcome of public firm Chapter 11 reorganization, see also Weiss (1990), LoPucki and Whitford (1993), Denis and Rodgers (2007), and Kalay, Singhal, and Tashjian (2007).

[18] See also Flynn (1989) for the outcome of small-firm bankruptcies.

differences across judges in the propensity to grant or deny creditor motions (e.g., to dismiss a case, lift an automatic stay, extend the exclusivity period, and use cash collateral). Their evidence is consistent with Hotchkiss (1995): she uses a dummy variable to indicate cases filed in the Southern District of New York, which at the time had a reputation for pro-debtor rulings favoring management attempts to reorganize (Weiss, 1990; LoPucki and Whitford, 1991). Cases filed in this district have a somewhat higher probability of subsequently entering a second bankruptcy or distressed restructuring.

Baird and Rasmussen (2003) argue that modern Chapter 11 practices are quite different from those observed a decade ago, with creditor control now being a dominant theme. They examine the 93 public large firms that completed their Chapter 11 reorganization in 2002. Of these, 52 (or 56% of the sample) are sales under Section 363 of the Bankruptcy Code or as part of a reorganization plan. Of the remaining cases, two-thirds (26 firms) reach an agreement with creditors prior to filing a prepackaged bankruptcy and one-third (15 firms) are reorganized in a traditional Chapter 11 proceeding.

Although the use of Chapter 11 may have changed over time, it is still true that large public firms are more likely to survive Chapter 11 as a going concern, while small firms have a higher probability of liquidation.

### *7.2. Post-bankruptcy performance*

In choosing a restructuring mechanism, firms consider both the cost of the restructuring itself and the extent to which the restructuring is able to resolve the financial difficulties. Distressed firms with plenty of intangible assets, and thus high indirect costs of bankruptcy, are more likely to choose a restructuring mechanism that minimizes the chance of a subsequent bankruptcy filing. In other words, these firms may choose to incur the immediate high costs of a comprehensive restructuring as long as it leads to greater debt reduction and a superior post-restructuring operating performance.

Conflicts of interests may further explain why firms fail to fully correct corporate investment policy in a restructuring. Incumbent managers are more likely to push for a continuation of the operations that preserves their private benefits of control rather than a more comprehensive restructuring involving the sale of a substantial part of the firm's assets. Management looking out for the interests of equityholders may also choose to file for Chapter 11 in order to take advantage of the bargaining power allocated to equity and the preservation of shareholder value.

If financial distress is not fully resolved for firms reorganizing in bankruptcy or if Chapter 11 suffers from economically important biases toward continuation of unprofitable firms, poor investment decisions will be reflected in the post-bankruptcy performance of emerging firms. Hotchkiss (1995) examines the operating performance of firms that emerge as public companies from Chapter 11 by 1989. Over 40% of the firms continue to experience operating losses in the first three years following bankruptcy. Accounting ratios such as return on assets and profit margins are substantially lower than for industry rivals. In the first year after emerging from Chapter 11, almost 75% of the sample firms have a lower operating performance (EBITDA/sales) than that of nonbankrupt firms in the same industry. Hotchkiss and Mooradian (2004) find similar results for a more recent time period. More than two-thirds of their sample firms underperform industry peers for up to five years following bankruptcy, and over 18% of the firms have negative operating income in the year following emergence.

Maksimovic and Phillips (1998) examine changes in the asset composition for firms that survive Chapter 11. By tracking the productivity of individual plants, regardless of whether these plants are sold or closed down, they are able to avoid the impact of survivorship bias, since they can examine asset performance even if the original owner of the assets is liquidated or emerges from Chapter 11 as a private company. They show that plants that are retained by bankrupt firms have lower productivity compared to the assets that are sold off, suggesting that firms in bankruptcy retain their least profitable assets. Thus, the performance changes may partially be a result of asset sales and closures, and not of changes in the efficiency of the retained assets.

In a recent paper, Kalay, Singhal, and Tashjian (2007) study changes in the operating performance of 113 firms that reorganized in Chapter 11 in the 1990s. The failed firms experienced significant profitability improvements while in bankruptcy in absolute terms as well as compared to industry rivals, suggesting that the reorganization may provide net benefits to the distressed firms. The performance improvements are smaller for firms with complex debt structure (more classes of debt) and greater for firms with higher prefiling debt ratios, possibly because the automatic stay on debt payments is particularly valuable to these firms.

An alternative to examining accounting-based performance measures is provided by Alderson and Betker (1999), who estimate the return that could have been earned by liquidating the firm's remaining assets and investing the proceeds in a portfolio of securities. Alderson and Betker (1999) compare the market value of 89 firms five years after emerging from bankruptcy (including all cash distributions to claimholders) to an estimated value if the assets would have been liquidated at emergence. The annualized return is then compared to the return of the S&P index over the same time period. They find that reorganized firms on average neither underperform nor overperform the S&P index. One interpretation of this study is that based on cash flow returns, emerging firms perform at par with the market overall, ignoring any differences in systematic risk.

Measures of operating profitability after emergence are likely to be strongly related to stock price performance as well. However, studies of stock price performance largely address the efficiency of pricing the securities at emergence, rather than the efficiency of the decision to reorganize. Still, these studies provide yet another view of post-bankruptcy success, in particular from an investment point of view. One difficulty in interpreting studies of emerging firm stock returns, however, is that only a fraction of firms that emerge relists their stock. For example, only 60% of the emerging firms studied by Hotchkiss (1995) relist their stock on NYSE, AMEX, or NASDAQ post-bankruptcy. If the worst firms are systematically unable to relist their stock, studies of post-bankruptcy stock performance may be biased to reflect the better performing firms.

The most comprehensive study of post-bankruptcy stock price performance to date is that of Eberhart, Altman, and Aggarwal (1999), who examine the equity performance of 131 firms emerging from Chapter 11 by 1993. They report large positive excess stock returns over the 200 trading days following emergence using different benchmarks. Compared to the return of a portfolio of nonbankrupt firms matched on industry and size, the average cumulative abnormal return (ACAR) is 25% (median 6%). Using the market model, the ACAR of the reorganized firms over the same period is 139% (median 5% to 7%). In sum, emerging firms exhibit large positive and significant abnormal stock returns in the first year post-bankruptcy. For a smaller sample but over a much longer time interval (five years subsequent to distress), Goyal, Kahl, and Torous (2003) document average abnormal returns close to zero using a value-weighted reference portfolio, but highly negative abnormal returns (−51%) using a size and book-to-market reference portfolio.

As a whole, the research suggests that a considerable portion of firms emerging from bankruptcy continue to perform poorly based on various performance measures. Underperformance may be related to firms that insufficiently reduced their debt burden with the restructuring, or that failed to undertake sufficient asset restructuring, enabling them to implement a feasible reorganization plan. The ultimate measure of success, therefore, is whether the firm is able to subsequently avoid another distressed restructuring or bankruptcy.

A number of studies have documented the incidence of repeated failures of distressed firms. The high rate of subsequent failures occurs despite the Chapter 11 requirements that the company must demonstrate the feasibility of the reorganization plan before it can be confirmed.[19] Among the earliest, LoPucki and Whitford (1993) report that 32% of 43 large Chapter 11 cases confirmed by March 1988 reenter Chapter 11 within four years. Hotchkiss (1995) shows that one-third of the emerging firms in her sample need to again restructure either through a private workout, a second bankruptcy, or an out-of-court liquidation. Gilson (1997) reports a failure rate of 25% for 108 distressed firms that recontracted with creditors in Chapter 11 or out of court. More recent statistics for the incidence of "Chapter 22" filings show that this pattern continues.

The high rate of subsequent failures has several potential explanations. One possibility is that firms have not sufficiently reduced their debt in the restructuring. Gilson (1997) finds that firms remain highly leveraged after emerging from Chapter 11, though less so than firms completing an out-of-court restructuring. Firms emerging from bankruptcy have a median ratio of long-term debt to total capitalization of 47%, and three-quarters of the firms are more highly levered than their industry rivals. Another explanation is that management is overly optimistic about the prospects for the reorganized firm. Hotchkiss (1995) shows that the continued involvement of incumbent management in the restructuring process increases the probability of post-bankruptcy failure. Finally, it has been

[19] According to Ÿ1129(a)(11) of the Bankruptcy Code, the reorganization plan must be feasible. The statute specifically requires the bankruptcy judge to find that approval of the reorganization plan "is not likely to be followed by the liquidation or the need for further financial reorganization of the debtor."

suggested that the pro-debtor orientation of the Bankruptcy Code and the courts permit inefficient firms to reorganize.[20] It is likely that all these factors combined play a role in the high failure rate of firms reorganized under Chapter 11.

The governance structure of the reorganized firm, however, appears to have an important relationship to post-bankruptcy success. Hotchkiss and Mooradian (1997) find that when a vulture investor remains active in the governance of the firm post-Chapter 11, the fraction of firms experiencing operating losses is a mere 8%. Improvements in performance relative to pre-default levels are greater when a distressed investor joins the board, becomes the CEO or chairman, or gains control of the firm. When there is evidence of vulture involvement but this investor subsequently is passive in the restructured company, performance appears no better than for those firms with no evidence of vulture involvement. Thus, the continued presence of distressed investors in the governance of the restructured firm is strongly related to different measures of post-bankruptcy success.

To sum up, a majority of large public firms emerge from Chapter 11 as independent companies, while small private firms are more likely to be liquidated in bankruptcy. Surviving firms frequently exhibit poor operating performance and frequently default on their debt again. Nevertheless, stock returns of surviving firms exceed various benchmarks in the first year following bankruptcy, raising the possibility that the market initially undervalues some reorganized firms.

## 8. International evidence

Bankruptcy laws vary considerably across the world. All countries provide liquidation procedures, where control over the firm shifts to creditors and assets are sold piecemeal or as a going concern. There are, however, major differences in the provisions for court-supervised reorganization—that is, a court settlement that permits the firm to continue as an ongoing concern while the financial claims are restructured. Some countries offer few alternatives to a sale of the distressed firm's assets. Other codes provide substantial shelter for incumbent management and equityholders, typically favoring a continuation of the operations. The degree to which the company's business is protected from creditors also varies. In some bankruptcy systems, the existing debt contracts are stayed and new debt receives super-priority status. Under other codes, secured claimholders have the right to seize collateral, potentially thwarting a continuation of the business.[21]

Although there is substantial variation, two distinct systems stand out: reorganization and auction codes. A reorganization code provides strong provisions for a court-supervised renegotiation of the firm's capital structure. Creditors have limited influence over the bankrupt firm, and incumbent management is typically allowed to continue to

[20] See, for example, Bradley and Rosenzweig (1992). The Southern District of New York and Delaware have been mentioned in this context.

[21] For a specific situation related to the enforcement of a debt contract against a hotel, Djankov, Hart, McLiesh, and Shleifer (2006) show that the contract is enforced more efficiently in countries with higher per capita income and quality of contract enforcement, and predicts debt market development.

run the operations. Chapter 11 of the U.S. Bankruptcy Code is a prominent example of a reorganization code. An auction code, in contrast, mandates a public sale of the bankrupt firm's assets.[22] The bidder offering the highest price decides whether the firm is liquidated piecemeal or survives as a going concern. Creditor interests are at the forefront, and the fate of management is determined by the buyer in the auction. As discussed later in this section, the Swedish bankruptcy code is a good example of an auction code.

Proponents of reorganization codes point to the perceived weaknesses of auction codes. There are concerns that the markets for distressed firms' assets are illiquid, forcing fire sales at depressed prices and perhaps producing a suboptimal allocation of assets. Moreover, bidding costs may be prohibitive due to uncertainty about the distressed firm's prospects, deterring potential bidders from entering the auction (Aghion, Hart, and Moore, 1992). It has also been suggested that managers, dreading the uncertainty about their position that the auction implies, may delay filing and engage in value-destroying, risk-shifting activities in an attempt to entirely avoid bankruptcy (White, 1996; Hart, 2000). In contrast, managers may be encouraged to file promptly under a management-friendly reorganization code, hence preserving firm value and increasing the likelihood of a successful reorganization.

Obviously, reorganization codes embrace a different set of inefficiencies. While an auction makes use of the market, the reorganization code uses negotiations to determine the value and future use of the bankrupt firm's assets. Since a restructuring of the capital structure entails the distribution of new financial claims, the negotiations also involve how much and what type of securities the various creditors will receive. Reaching one negotiated solution covering all these aspects can be a lengthy and costly procedure for the distressed firm. The auction, on the other hand, separates these decisions and thus provides a speedier resolution.

Another potential issue associated with a reorganization code is the substantial control rights given to incumbent management, effectively removing the residual claimholders (creditors) from the decision making. While this approach may encourage management to file without delay, it also opens the way for decisions that benefit self-interested managers. It is possible that the default is a result of managerial incompetence. Allowing the incumbent managers to retain control of the firm may delay a necessary change in management or prevent closure of the operations when a piecemeal liquidation of the assets is optimal.[23] In contrast, in the auction, the highest bidder who has its own money at stake determines whether the firm will continue to operate as a going concern or whether the assets are to be redeployed.

The total costs imposed by the bankruptcy code determine claimholders' incentives to voluntarily restructure the claims outside of the formal bankruptcy procedure. Claessens

[22] Mandatory auctions are often referred to as liquidations. In this context, however, liquidation simply implies that the assets are redeployed through a sale. This may or may not imply a termination of the operations.

[23] Franks and Loranth (2005) suggest that lack of court oversight and poorly designed trustee compensation contracts lead to inefficient continuation of poorly performing firms under the Hungarian reorganization code.

and Klapper (2005) find that the bankruptcy filing rate generally is higher in countries with an efficient judicial system. Moreover, controlling for judicial efficiency, bankruptcy tends to be used more frequently in countries where the insolvency procedures give creditors more rights.[24] Thus, when comparing outcomes under different bankruptcy codes, one should keep in mind the caveat that a distinct set of firms may file for bankruptcy under each code. [25]

The magnitude of the potential inefficiencies in different bankruptcy systems is an empirical question. Nevertheless, evidence on bankruptcy reorganization outside the United States is sparse. In the following section, we review evidence on the restructuring of distressed firms in the UK, Sweden, France, Germany, and Japan.

### *8.1. The United Kingdom: receivership*

UK companies have access to several court-supervised procedures. In the dominant procedure, *Receivership*, a secured creditor appoints a receiver representing the interests of this creditor. The receiver realizes the security and, after deducting his expenses and paying any higher priority claims, uses the proceeds to pay off the appointing creditor. If the claim is secured by floating charge collateral, an administrative receiver gets full control over the firm and can reorganize the firm or sell assets without permission from other creditors or the court.[26] There is no automatic stay of debt claims. Creditors secured with fixed liens on particular assets have the right to repossess their collateral, even if the assets are vital for the firm's operations. Any excess balance is distributed to remaining claimholders according to the absolute priority of their claims. Unsecured creditors have little influence over the procedure.

The UK also provides two court-administered reorganization procedures, *Administration* and, for small firms, *Company Voluntary Arrangements* (CVAs), which give the firm temporary relief from its creditors. A secured creditor can veto these procedures, however, and instead appoint a receiver. Thus, in practice, the court can appoint an administrator that represents all creditors only in the absence of secured creditors. Reformed UK insolvency procedures took effect in 2003. The new UK law cuts back the rights of creditors secured by floating charge, including that to appoint an administrative receiver. Holders of floating charge claims issued prior to September 15, 2003, however, retain the same rights as before. Overall, UK insolvency procedures are considered to be creditor-oriented.

In general, the UK receivership code provides little protection of the operations. The liquidation decision is typically left to secured creditors, who lack incentives to generate

[24] See also Claessens, Djankov, and Klapper (2003) for an examination of the use of bankruptcy in East Asia.

[25] The design of the bankruptcy code may also have other ex-ante effects. Acharya and Subramanian (2007) suggest, for example, that firms generate more patents in economies with weaker creditor rights.

[26] The collateral of a floating charge claim includes inventory, accounts receivables, working capital, and intangible assets.

proceeds above the value of their claim. Franks, Nyborg, and Torous (1996) propose that the allocation of control rights to secured creditors leads to underinvestment and excessive termination of economically viable firms. Franks and Nyborg (1996), however, argue that premature liquidation can be avoided if the creditor appointing the receiver has large private benefits of control associated with the survival of the bankrupt firm.

Interestingly, the evidence indicates that a large fraction of distressed UK firms indeed survive as ongoing concerns. Franks and Sussman (2005) examine 542 small-to-medium-sized financially distressed UK firms that are transferred to their bank's workout unit. They report that 60% of sample firms filing for the UK receivership code continue to operate as going concerns after bankruptcy. In a sample of UK firms filing for administrative receivership, Kaiser (1996) finds that almost half are sold as going concerns. Similarly, Davydenko and Franks (2006) show that 43% of small UK firms that default on their debt are liquidated piecemeal.

Franks, Nyborg, and Torous (1996) suggest that the UK receivership code is speedy, which would imply low direct costs. Nevertheless, Franks and Sussman (2005) report direct costs averaging 33% of asset values. They note that a lack of competition among receivers may explain the high costs and point to much lower costs (mean 14%) when the Royal Bank of Scotland recently required receivers to tender for their appointments. We are not aware of specific data on the duration of the UK bankruptcy procedure. Nevertheless, the firms in Franks and Sussman (2005) spend on average 7.5 months in the bank's workout unit, and the median length of reorganization is 1.4 years for a subset of the defaulted firms in Davydenko and Franks (2006).

Secured creditors seem to fare relatively well in the UK procedure, as expected. Franks and Sussman (2005) document average bank recovery rates of 75%, with a median of 100%. Nearly all of the firms' assets are pledged as collateral to the bank. Interestingly, banks tend to liquidate collateral at prices close to the face value of the secured claim, possibly because secured creditors have few incentives to generate additional proceeds for junior claimants. Similarly, Davydenko and Franks (2006) report an average bank recovery rate of 69% (median 82%).

Since secured creditors fare relatively well in formal bankruptcy, one would predict voluntary workouts to be relatively rare in the UK. Davydenko and Franks find that 75% of small firms that default on their debt enter formal bankruptcy, with the remaining 25% of firms reorganizing out of court. When large distressed companies issue new equity, however, UK banks appear quite willing to make concessions out of court. Franks and Sanzhar (2006) show that banks make concessions for one-third of 111 financially distressed, publicly traded UK firms that issue new equity. These concessions include forgiveness of principal, debt for equity swaps, and provisions for new loans. Concessions are offered to firms with higher leverage and greater debt impairment, representing situations where the expected wealth transfer to debtholders is relatively large.

Acharya, John, and Sundaram (2005) contend that the allocation of control rights in bankruptcy determines the impact of asset specificity on the firm's optimal capital structure. On one hand, when assets are specific to the industry, liquidation values may be low and a forced liquidation relatively costly to the firm. On the other hand, when assets are nonspecific, the costs from inefficiently continuing the firm may be high. Thus, firms with high asset specificity will choose a lower debt level under a creditor-friendly system, which is prone to inefficient liquidations, than under a reorganization-oriented code. In contrast, firms with low asset specificity will choose a lower debt level under a debtor-oriented code, which risks allowing excessive continuation. Contrasting firms in the UK and the United States—classified as having creditor-friendly and debtor-friendly bankruptcy systems, respectively—Acharya, John, and Sundaram (2005) find variations in debt ratios consistent with their predictions.[27]

Overall, the weak protection of the firm's operations and the strong rights allocated to secured creditors in UK bankruptcy may raise concerns of excessive liquidations. Nevertheless, firm survival and recovery rates in the UK compare well to the U.S. Chapter 11. Thus, the strong creditor orientation in formal bankruptcy does not appear to be detrimental to the restructuring of distressed UK firms.

### *8.2. Sweden: auctions*

In Sweden, bankruptcy is resolved through a mandatory auction. The proceeding is run by a court-appointed trustee with fiduciary responsibility to all creditors. This trustee organizes the sale of the firm in an auction. The winning bidder determines whether the firm is liquidated piecemeal or continues to operate as an ongoing concern. Payment must be in cash, and creditors are paid strictly according to the absolute priority of their claims.

The trustee typically retains the incumbent management team to run the operations of the firm in bankruptcy. In contrast to the UK, the Swedish code restricts the liquidation rights of creditors. Debt payments are stayed, and collateral cannot be repossessed. Moreover, trade credits and other debt raised while in bankruptcy get super-priority. These provisions help protect the operations until the firm is auctioned off.

Swedish insolvency law provides a forum for renegotiation of unsecured debt called composition (*ackord*). Secured debt and priority claims (taxes and wages) must be offered full repayment, and junior creditors at least 25% of their claim. These high thresholds make composition unfeasible for the vast majority of distressed firms. A new reorganization law was enacted in 1996, but Buttwill and Wihlborg (2004) argue that the new law shares many of the weaknesses of the old composition procedures and is rarely used. Thus, in Sweden, court-supervised renegotiation of the firm's debt contracts is effectively not an alternative to auction bankruptcy.

Thorburn (2000) examines a sample of 263 small, private Swedish firms filing for bankruptcy between 1991 and 1998. Her evidence counters widespread fears that bankruptcy auctions tend to excessively force liquidation. She demonstrates that three-quarters of firms continue as a going concern under the buyer's reign, with the remaining

[27] See also Vig (2007), who argues that a recent change in bankruptcy law in India that strengthens the rights of secured creditors has had an important influence on capital structure and the use of secured financing.

one-quarter of firms being liquidated piecemeal.[28] The probability for a going concern sale increases in the fraction of intangible assets, perhaps because these assets generate little value in a piecemeal liquidation. To gauge the quality of the continuation decision, Eckbo and Thorburn (2003) examine the operating profitability of the Swedish firms emerging from bankruptcy. They show that auctioned firms perform at par with industry competitors for several years, also when the incumbent CEO retains control. This contrasts to the evidence in Hotchkiss (2005) that firms emerging from U.S. Chapter 11 tend to underperform their industry rivals.

Thorburn (2000) also estimates the costs of Swedish bankruptcy proceedings. She reports direct costs of on average 6% of pre-filing book value of assets, with an average of 4% for the one-third largest firms in her sample.[29] When measured as a fraction of the market value of assets in bankruptcy, costs average 19%, with a median of 13%. The direct costs decrease with firm size and increase with measures of industry distress, suggesting that trustees may increase their sales effort in periods when auction demand is relatively low. Importantly, the auction is speedy, with an average time from filing to sale of the assets of only two months, implying relatively low indirect costs.[30]

The value of the assets remaining at the end of the bankruptcy process reflects all the different costs imposed on the financially distressed firm. This value is split between the firm's creditors. The higher the total costs of bankruptcy, the lower are creditor recovery rates. In Swedish bankruptcy, creditors' claims are paid with the cash generated in the auction. Thorburn (2000) reports average recovery rates of 35% (median 34%). Recovery rates are higher in going-concern sales (mean 39%) than in piecemeal liquidations (mean 27%). Secured creditors receive on average 69% (median 83%).[31]

A potential issue with a creditor-oriented code is that it may encourage management to delay filing and undertake value-reducing risk-shifting investments in an effort to stay out of bankruptcy. Eckbo and Thorburn (2003) argue, however, that managerial incentives to preserve private benefits of control may counteract potential risk-shifting incentives. Specifically, managers may implement a conservative value-maximizing investment policy for the financially distressed firm in an attempt to increase the joint likelihood that the firm survives as a going concern and that current management gets rehired by the buyer in the auction. For the sample of Swedish small-firm bankruptcy filings, Eckbo and Thorburn (2003) show that the probability that the incumbent manager continues to run the auctioned firm increases in a measure for the private benefits of control. They

[28] Prior to 1993, Finnish bankruptcy also mandated a sale of the firm. In a sample of 72 small firms filing under the old Finnish code, Ravid and Sundgren (1998) find that only 29% of the firms are sold as a going concern.

[29] Ravid and Sundgren (1998) report average direct costs of 8% of pre-filing book value of assets for the small firm bankruptcies in Finland.

[30] Note that while the firm's operations are auctioned off quickly, the bankruptcy proceeding continues and last on average around three years.

[31] Ravid and Sundgren (1998) find average recovery rates of 34% in going-concern sales and 36% in piecemeal liquidations in Finnish bankruptcy.

also find that bidders screen managers on quality in the rehiring decision and that CEOs suffer large income declines conditional on bankruptcy filing (about 40% relative to CEOs of nonbankrupt companies). Their evidence supports the notion that managers' drive to retain control of the operations conditional on default may counterbalance ex-ante incentives to risk shift.

Most European countries hold directors and managers personally liable and impose civil and criminal penalties if they fail to file in a timely manner or to inform creditors when the firm becomes insolvent. To the extent that these laws are enforced, such penalties may help trigger prompt action, further offsetting potential tendencies to delay filing.

A common objection to auctioning firms in bankruptcy is the concern that markets for distressed firms' assets are illiquid, forcing fire sales at depressed prices. Stromberg (2000) suggests, however, that salebacks may help avoid costly asset fire sales in periods of industry distress. He shows that the probability that the old owner buys back the firm in the bankruptcy auction increases with industry leverage and operating performance, and decreases with the proportion of nonspecific assets.

Eckbo and Thorburn (2008) model the participation in the auction of a secured creditor with an impaired debt claim on the bankrupt firm. The more impaired the secured claim, the greater incentive the creditor has to provide financing to rival bidders and encourage aggressive bidding, thus increasing the expected recovery. Eckbo and Thorburn show that the bankrupt firm's bank frequently enhances auction liquidity by providing bid financing. The premiums paid by the winning bidder decrease with an estimate of the secured creditor's expected recovery in the event of piecemeal liquidation, consistent with the predicted bidding behavior.

In a companion paper, Eckbo and Thorburn (2007) test the implications of industry distress using prices paid and debt recovery rates in the bankruptcy auctions. They estimate fundamental values of the auctioned assets in a cross-sectional model and examine how industry liquidity factors (leverage and interest coverage ratios) affect the standardized residuals from the price regression. There is some evidence of fire-sales discounts in piecemeal liquidations, but not when the bankrupt firm is acquired as a going concern. Neither industry-wide distress nor the industry affiliations of the buyer affect the prices in going-concern sales. Eckbo and Thorburn (2007) further show that bids often are structured as leveraged buyouts, which relaxes liquidity constraints and reduces bidder underinvestment incentives in the presence of debt overhang. It is possible that distressed industry insiders overcome liquidity constraints by using LBO financing. Eckbo and Thorburn (2007) also find evidence that prices are lower in prepackaged filings than in other going-concern sales, suggesting that prepacks may help preempt excessive liquidation when the auction is expected to be illiquid. Liquidation preemption seems to be a risky strategy, however, as prepackaged bankruptcies have much higher refiling rates than firms sold in a regular auction.

Overall, the evidence on Swedish bankruptcy filings suggests that mandatory auctions provide a relatively efficient mechanism for restructuring financially distressed firms.

Survival rates, direct costs, and recovery rates compare well with extant evidence from the United States and the UK. Moreover, there is no evidence that firm value is destroyed because of distorted ex-ante incentives to risk-shift. While bankruptcy auctions risk forcing asset sales at depressed prices, the evidence suggests that the incentives of secured creditors and old owners combined with opportunities for LBO financing help increase auction demand, effectively counteracting fire-sales tendencies.

### *8.3. France: weak creditor rights*

France provides very strong protection of distressed businesses through its formal reorganization procedure, *Redressement Judiciare*.[32] The objectives of the procedure are, in order of priority, to continue the firm's operations, to maintain employment, and to pay back creditors. A court-appointed administrator oversees the reorganization. Debtholders are restricted from directly participating in the restructuring process. They are represented by a court officer and can raise their concerns only through this court-appointed creditor representative. Employees, however, may appoint their own representative.

The administrator evaluates the prospects for reorganization and presents a reorganization plan to the court. Creditors cannot reject the court's decision, nor does confirmation of a reorganization plan or sale of collateral require approval of secured creditors. Creditors are offered new, altered claims in place of their old impaired debt claims. Although the court cannot force creditors to write down their claims, it can redefine the terms of the loan, including maturity. Thus, in practice, creditors often prefer to accept a write-down with timely repayment to a promised repayment in full in an uncertain distant future.

Debt payments are stayed during the bankruptcy process, and the administrator can raise new super-priority financing without creditor approval. If the firm is sold, the court can choose a lower bid that provides better prospects for continued operations and employment. Moreover, government and employee claims have first priority to proceeds generated in a sale of collateral, effectively forcing a deviation from absolute priority rules.[33]

The French code, with its explicit objective to maintain operations and preserve jobs, has a predisposition to allow continuation of inefficient firms. Nevertheless, the evidence indicates that relatively few firms survive bankruptcy in France. Kaiser (1996) reports that only 15% of filing firms continue to operate as a going concern after bankruptcy reorganization. In a broader sample comprised of bankruptcy filings and voluntary workouts, Davydenko and Franks (2006) find that 62% of French firms are liquidated piecemeal, which is a higher fraction than in the UK. Despite the poor odds for survival, they show that a vast majority (87%) of firms that default on their debt enter formal bankruptcy.

[32] French insolvency law also provides a separate proceeding for liquidation (*Liquidation Judiciare*) and a rarely used procedure for renegotiation of debt contracts prior to default (*Reglément Amiable*).

[33] Certain types of collateral, such as receivables and guarantees, are exempt from this rule.

The low survival rates in France translate into relatively low creditor recovery rates. Davydenko and Franks (2006) document an average bank recovery rate in French proceedings of 47% (median 39%), which is much lower than the recovery rates reported for UK banks. The median reorganization takes three years. Moreover, French banks take more collateral than bank lenders in the UK and Germany, possibly reflecting the poor standing of banks in French bankruptcy.

Overall, although bankruptcy law in France is set up to promote firm survival, the actual result seems to be the opposite. Firm survival rates and creditor recovery in France compare poorly with evidence from the UK and the United States. It is possible that the costs associated with the extremely creditor-hostile French insolvency procedures ultimately are borne by the distressed firms and their claimholders. Or perhaps French firms restructure their debt prior to default in order to entirely avoid reaching the point where they are subject to insolvency laws, leaving only the lemons to the bankruptcy procedure.

### *8.4. Germany: bank-driven reorganizations*

The German 1999 reorganization procedure, *Insolvenzordnung*, gives the financially distressed firm three months to engineer a reorganization plan under the supervision of a court-appointed administrator.[34] This plan outlines the financial and asset restructuring of the firm, including a potential sale of the firm as a going concern. The reorganization plan must receive creditor approval before it can be implemented. Creditors vote with a simple majority rule. Similar to the United States, the court may cram down a plan on a dissenting class of creditors as long as the plan leaves the class better off than would be the case with a piecemeal liquidation of the assets. Creditor claims are stayed during the three-month reorganization period. The firm can raise new debt financing with super-priority subject to creditor approval.

The evidence on distressed firms in Germany primarily dates from the period before the new reorganization code took effect. Davydenko and Franks (2006), for example, examine firms that defaulted on their debt between 1984 and 2003. They document that 57% of distressed German firms are liquidated piecemeal, which is higher than liquidation ratios reported for Sweden and the UK and lower than liquidation ratios in France. The median duration of the reorganization procedure in Germany is 3.8 years, and banks recover on average 59% (median 61%) of their claims.

An important impediment to out-of-court agreements is holdout problems among dispersed creditors. In Germany, the debt is typically concentrated with a house bank that often also has an equity interest. As a result, one should expect coordination failures to be relatively rare in Germany. According to Kaiser (1996), most German firms with a chance of survival are reorganized in an out-of-court workout. Davydenko and Franks (2006), however, find that 78% of the distressed firms in their sample enter formal

[34] Prior to 1999, German insolvency law offered an auction liquidation process (*Konkursordnung*) and a rarely used procedure for the reorganization of unsecured claims (*Vergleichsordnung*).

bankruptcy, with the remaining 22% of sample firms working things out with creditors informally.

Elsas and Krahnen (2002) study the role of lending relationships for 75 financially distressed German firms initiating private workouts. They find that house banks and banks holding a secured claim are more likely to participate in a voluntary restructuring. Brunner and Krahnen (2004) show that German bank lenders often coordinate their reorganization efforts by forming a bank pool when medium-sized firms become financially distressed. They report that banks strike a formal contractual pool arrangement for 45% of the distressed firms, and the probability of bank pool formation increases with the number of bank relationships and the degree of distress.

While the German procedure has some resemblance to the U.S. Chapter 11, it imposes a strict three-month limit on the reorganization. This period risks being too short to allow a firm with complex operations and capital structure to carefully develop a reorganization plan. The evidence on the new reorganization procedure, however, is at this point insufficient for us to draw any conclusions about how well the new German code works.

### *8.5. Japan: keiretsu banks*

Japan's bankruptcy code has historically been oriented toward a liquidation of the filing firm. Managers typically lost their jobs, and creditors controlled the outcome of the bankruptcy proceeding. Over the last decade, however, Japan has undertaken a series of revisions of its insolvency procedures aimed at strengthening the provisions for restructuring financially distressed firms as ongoing concerns.

A prominent feature of the Japanese business environment is industrial groups called *keiretsus*. At the core of a keiretsu are banks, which finance much of the industrial operations, both as creditors and equityholders of the firms affiliated with the group. Hoshi, Kashyap, and Scharfstein (1990) examine the role of a keiretsu affiliation for a sample of 125 publicly traded firms that become financially distressed. They find that distressed firms associated with a keiretsu invest more and sell more than nonkeiretsu firms in the years following the onset of financial distress. This suggests that keiretsu banks help relax financial constraints, possibly mitigating the costs of financial distress.[35] Helwege and Packer (2003) study the role of keiretsu banks for the outcome of bankruptcy for 172 troubled Japanese firms. They report that the probability of liquidation is higher for firms affiliated with keiretsu banks than for nonkeiretsu firms, controlling for firm size. However, since there is no discernible difference in the profitability of the liquidated firms, they conclude that there is no evidence that keiretsu banks force excessive liquidations.

In sum, Japan has traditionally provided creditor-oriented insolvency procedures often dominated by large keiretsu banks. There is insufficient evidence at this point, however, to determine whether financial ties with keiretsu banks help or are detrimental to the reorganization of distressed firms.

[35] Claessens, Djankov, and Klapper (2003) show that financially distressed firms in East Asia are less likely to file for bankruptcy if they are owned by banks or affiliated with a business group.

## 9. Conclusion

This chapter surveys the body of empirical research that focuses on the use of private and court-supervised mechanisms for resolving default by restructuring companies in financial distress. We organize and synthesize this literature in the context of a simple model of financial distress. After a quick overview of the theoretical issues, we identify some main themes to anchor the empirical research in the areas of financial distress, asset and debt restructuring, and the formal bankruptcy procedures in the United States and abroad. Studies of out-of-court restructurings (workouts and exchange offers), corporate governance issues related to distressed restructurings, the magnitude of costs and outcomes of bankruptcy reorganizations, and the relative efficiency of bankruptcy codes in different countries are among the topics surveyed.

It is customary (as we have done in this survey) to make a distinction between two types of systems for resolving default: one in which the business is sold to a third party, possibly through an auction; and another in which the firm is reorganized under the current claimholders. Although these two philosophies of resolving default—liquidation and reorganization—have been viewed as entirely different approaches (we have discussed their relative merits in Section 8), claimholders in the U.S. Chapter 11 reorganization system are increasingly relying on the market to mimic solutions provided by an auction (liquidation) system. Most of the research on firms reorganizing under the U.S. Bankruptcy Code, however, dates from the 1980s and early 1990s. The peak in default rates in 2002, combined with creeping changes in insolvency practices and an escalation in the enforcement of creditor rights, has caused a growing demand for new research that can help us understand the process that governs the restructuring of financially distressed firms in the current environment.

The active trading of distressed debt at all priority levels combined with the participation of sophisticated investors is significantly affecting Chapter 11 mechanisms. The extensive trading in distressed debt has led to high turnover in the identity of the creditors of companies in financial distress. Nontraditional investors, such as private equity investors and hedge funds, have increased their role in these markets and therefore as creditors of troubled firms. Based on their estimate of the value of the business and the legal priority of the various claims, many strategic investors acquire the fulcrum class of claims, that is, the securities where they expect that the equity value will reside after the reorganization is completed. Taking a private equity perspective on their investment, these investors seek to become owners of the enterprise, fix it, and then sell it at an optimal time. In this manner, the multiple creditors basically replicate the characteristics of a third-party sale, although the restructuring process is that of a conventional reorganization. Moreover, creditors frequently require, and courts are more willing to approve of, an outright sale of major assets of the distressed firm, either through an auction under Section 363 of the U.S. Bankruptcy Code or as part of the reorganization plan. The overall effect of all these changes on the efficiency of the U.S. bankruptcy procedures yet remains to be documented and analyzed.

With emerging economies searching for an optimal bankruptcy system and the opportunity for the European Union to harmonize its insolvency rules, the efficiency of various systems across the industrial economies has received increasing attention. The differences in the insolvency codes of the advanced economies and the differential degree of creditor rights available in the legal systems of those economies have been the focus of policy makers in many countries. It is also recognized that changes in insolvency codes and their enforcement could have important influence on how firms access capital as well as on the efficiency of investment in the economy. It is evident from our review of the research on insolvency procedures outside the United States that a lot still remains to be done in this area. One important and still mainly unanswered question is how different institutional characteristics of individual countries interact with their respective insolvency rules. For example, to what degree is the success of Swedish bankruptcy auctions tied to the dominant role of banks in this economy?

The argument has often been made that the direct costs of bankruptcy seem too small to justify the relatively low leverage ratios that we typically observe. A response to this observation has been that leverage and financial distress might have other indirect costs that need to be taken into consideration. Our understanding of the nature and magnitude of these indirect costs of financial distress is still very preliminary. In some sense, the indirect costs of bankruptcy arise from the value lost from investments that optimally should but in practice are not undertaken (an opportunity cost). Finding reliable measures of such unobservable phenomena is very difficult and requires clever empirical strategies.

In designing bankruptcy systems, it is important to consider their effect on a variety of issues, including capital structure choices, investment incentives, and risk choices that arise from the law and its implementation. For obvious reasons, most of the existing literature has focused on the ex-post efficiency of the mechanisms for resolving default, that is, on events following the onset of financial distress. In order to assess the optimality of various mechanisms for resolving default, however, we also need to consider their ex-ante efficiency. The international evidence plays an important role in the search for an optimal bankruptcy system.

## References

Acharya, Viral V., Sreedhar T. Bharath and Anand Srinivasan, 2007, "Does Industry-Wide Distress Affect Defaulted Firms? Evidence from Creditor Recoveries," *Journal of Financial Economics*, 85, 787–821.

Acharya, Viral V., Kose John and Rangarajan K. Sundaram, 2005, "Cross-Country Variations in Capital Structures: The Role of Bankruptcy Codes," CEPR Discussion Papers 4916.

Acharya, Viral V., Kose John and Rangarajan K. Sundaram, 2000, "On the Optimality of Resetting Executive Stock Options," *Journal of Financial Economics*, 57, 65–101.

Acharya, Viral V. and Krishnamurthy Subramanian, 2007, "Bankruptcy Codes and Innovation," Working Paper, London Business School.

Aghion, Philippe, Oliver Hart and John Moore, 1992, "The Economics of Bankruptcy Reform," *Journal of Law, Economics and Organization*, 8, 523–546.

Alderson, Michael J. and Brian L. Betker, 1999, "Assessing Postbankruptcy Performance: An Analysis of Reorganized Firms' Cash Flows," *Financial Management*, 28, 68–82.

Alderson, Michael J. and Brian L. Betker, 1995, "Liquidation Costs and Capital Structure," *Journal of Financial Economics*, 39, 45–69.

Allen, Franklin, 1989, "The Changing Nature of Debt and Equity; A Financial Perspective," Conference Series; [Proceedings], *Federal Reserve Bank of Boston*, 12–48.

Allen, Franklin and Douglas Gale, 1994, "Optimal Security Design," *Financial Innovation and Risk Sharing*, 157–197.

Allen, Franklin and Andrew Winton, 1995, "Corporate Financial Structure, Incentives, and Optimal Contracting," In R. A. Jarrow, V. Maksimovic, and W. T. Ziemba (Eds.), *Handbooks in Operations Research and Management Science*, Vol. 9: *Finance*, Elsevier/North Holland, New York.

Altman, Edward I., 1984, "A Further Empirical Investigation of the Bankruptcy Cost Question," *Journal of Finance*, 39, 1067–1089.

Altman, Edward I. and Edith S. Hotchkiss, 2006, *Corporate Financial Distress and Bankruptcy*, 3rd edition, John Wiley & Sons, Hoboken, NJ.

Altman, Edward I. and William Stonberg, 2006, "The Market in Defaulted Bonds and Bank Loans," *Journal of Portfolio Management*, 32, 93–106.

Andrade, Gregor and Steven N. Kaplan, 1998, "How Costly Is Financial (not Economic) Distress? Evidence from Highly Leveraged Transactions that Became Distressed," *Journal of Finance*, 53, 1443–1493.

Ang, James S., Jess H. Chua and John J. McConnell, 1982, "The Administrative Costs of Corporate Bankruptcy: A Note," *Journal of Finance*, 37, 219–226.

Asquith, Paul, Robert Gertner and David Scharfstein, 1994, "Anatomy of Financial Distress: An Examination of Junk Bond Issuers," *Quarterly Journal of Economics*, 109, 625–658.

Ayotte, Kenneth M. and Edward R. Morrison, 2007, "Creditor Control and Conflict in Chapter 11," Working Paper, Columbia Law School.

Baird, Douglas G. and Robert K. Rasmussen, 2003, "Chapter 11 at Twilight," *Stanford Law Review*, 56, 673–699.

Benmelech, Efraim, Mark J. Garmaise and Tobias J. Moskowitz, 2005, "Do Liquidation Values Affect Financial Contracts? Evidence from Commercial Loan Contracts and Zoning Regulation," *Quarterly Journal of Economics*, 120, 1121–1254.

Betker, Brian L., 1997, "The Administrative Costs of Debt Restructurings: Some Recent Evidence," *Financial Management*, 26, 56–68.

Betker, Brian L., 1995a, "An Empirical Examination of Prepackaged Bankruptcy," *Financial Management*, 24, 3–18.

Betker, Brian L., 1995b, "Management's Incentives, Equity's Bargaining Power, and Deviations from Absolute Priority in Chapter 11 Bankruptcies," *Journal of Business*, 68, 161–183.

Bharath, Sreedar T., Venky Panchapegesan and Ingrid Werner, 2007, "The Changing Nature of Chapter 11," Working Paper, University of Michigan.

Bolton, Patrick and David S. Scharfstein, 1996, "Optimal Debt Structure and the Number of Creditors," *Journal of Political Economy*, 104, 1–25.

Bradley, Michael and Michael Rosenzweig, 1992, "The Untenable Case for Chapter 11," *Yale Law Journal*, 101, 1043–1095.

Branch, Ben S., 2000, "Fiduciary Duty: Shareholders versus Creditors" *Financial Practice and Education*, 10, 8–13.

Bris, Arturo, Ivo Welch and Ning Zhu, 2006, "The Costs of Bankruptcy: Chapter 7 Liquidation versus Chapter 11 Reorganization," *Journal of Finance*, 61, 1253–1303.

Brown, David T., 1989, "Claimholder Incentive Conflicts in Reorganization: The Role of Bankruptcy Law," *Review of Financial Studies*, 2, 109–123.

Brown, David T., Christopher M. James and Robert M. Mooradian, 1994, "Asset Sales by Financially Distressed Firms," *Journal of Corporate Finance*, 1, 233–257.

Brown, David T., Christopher M. James and Robert M. Mooradian, 1993, "The Information Content of Exchange Offers made by Distressed Firms," *Journal of Financial Economics*, 33, 93–118.

Brunner, Antje and Jan P. Krahnen, 2004, "Multiple Lenders and Corporate Distress: Evidence on Debt Restructuring," CEPR Discussion Papers 4287.

Bulow, Jeremy I. and John B. Shoven, 1978, "The Bankruptcy Decision," *Bell Journal of Economics and Management Science*, 9, 437–456.

Buttwill, Klas and Clas Wihlborg, 2004, "The Efficiency of the Bankruptcy Process. An International Comparison," Working Paper, Copenhagen Business School.

Carapeto, Maria, 2005, "Bankruptcy Bargaining with Outside Options and Strategic Delay," *Journal of Corporate Finance*, 11, 736–746.

Carapeto, Maria, 1999, "Does Debtor-in-Possession Financing Add Value?" IFA Working Paper No. 294–1999, Cass Business School Research Paper, England.

Chang, Tom and Antoinette Schoar, 2006, "Judge Specific Differences in Chapter 11 and Firm Outcomes," Working Paper, MIT Sloan School of Management.

Chatterjee, Sris, Upinder S. Dhillon and Gabriel G. Ramirez, 2004, "Debtor-in-Possession Financing," *Journal of Banking and Finance*, 28, 3097–3111.

Chatterjee, Sris, Upinder S. Dhillon and Gabriel G. Ramirez, 1995, "Coercive Tender and Exchange Offers in Distressed High-yield Debt Restructurings: An Empirical Analysis," *Journal of Financial Economics*, 38, 333–360.

Chidambaran, Nenmara K. and Nagpurnanand R. Prabhala, 2003, "Executive Stock Option Repricing, Internal Governance Mechanisms, and Top Management Turnover," *Journal of Financial Economics*, 69, 153–189.

Claessens, Stijn, Simeon Djankov and Leora F. Klapper, 2003, "Resolution of Corporate Distress in East Asia," *Journal of Empirical Finance*, 10, 199–216.

Claessens, Stijn and Leora F. Klapper, 2005, "Bankruptcy around the World: Explanations of Its Relative Use," *American Law and Economics Review*, 7, 253–283.

Cutler, David M. and Lawrence H. Summers, 1988, "The Costs of Conflict Resolution and Financial Distress: Evidence from the Texaco-Pennzoil Litigation," *RAND Journal of Economics*, 19, 157–172.

Dahiya, Sandeep, Kose John, Manju Puri and Gabriel Ramirez, 2003, "Debtor-in-Possession Financing and Bankruptcy Resolution: Empirical Evidence," *Journal of Financial Economics*, 69, 259–280.

Davydenko, Sergei A. and Julian R. Franks, 2006, "Do Bankruptcy Codes Matter? A Study of Defaults in France, Germany, and the UK," ECGI Finance Working Paper No. 89.

Denis, Diane K. and Kimberly J. Rodgers, 2007, "Chapter 11: Duration, Outcome and Post-Reorganization Performance," *Journal of Financial and Quantitative Analysis*, 42, 101–118.

Djankov, Simeon, Oliver Hart, Caralee McLiesh and Andrei Shleifer, 2006, "Debt Enforcement around the World," NBER Working Papers: 12807.

Eberhart Allan C., Edward I. Altman and Reena Aggarwal, 1999, "The Equity Performance of Firms Emerging from Bankruptcy," *Journal of Finance*, 54, 1855–1868.

Eberhart Allan C., William T. Moore and Rodney L. Roenfeldt, 1990, "Security Pricing and Deviations from the Absolute Priority Rule in Bankruptcy Proceedings," *Journal of Finance*, 45, 1457–1469.

Eckbo, B. Espen and Karin S. Thorburn, 2008, "Creditor Incentives and Bidding in Bankruptcy Auctions," *Journal of Corporate Finance*, forthcoming.

Eckbo, B. Espen and Karin S. Thorburn, 2007, "Automatic Bankruptcy Auctions and Fire-Sales," *Journal of Financial Economics*, forthcoming.

Eckbo, B. Espen and Karin S. Thorburn, 2003, "Control Benefits and CEO Discipline in Automatic Bankruptcy Auctions," *Journal of Financial Economics*, 69, 227–258.

Eisenberg, Theodore and Lynn M. LoPucki, 1999, "Shopping for Judges: An Empirical Analysis of Venue Choice in Large Chapter 11 Reorganization," *Cornell Law Review*, 84, 967–1003.

Elsas, Ralf and Jan P. Krahnen, 2002, "Collateral, Relationship Lending, and Financial Distress: An Empirical Study on Financial Contracting," Working Paper, Frankfurt.

Ferris, Stephen R., Narayanan Jayaraman and Anil K. Makhija, 1997, "The Response of Competitors to Announcements of Bankruptcy: An Empirical Examination of Contagion and Competitive Effects," *Journal of Corporate Finance*, 3, 367–395.

Flynn, E., 1989, "Statistical Analysis of Chapter 11," Unpublished Study, Administrative Office of the U.S. Courts, Washington, DC.

Franks, Julian R. and Gyongyi Loranth, 2005, "A Study of Inefficient Going Concerns in Bankruptcy," CEPR Discussion Paper No. 5035.

Franks, Julian R, and Kjell G. Nyborg, 1996, "Control Rights, Debt Structure and the Loss of Private Benefits: The Case of the UK Insolvency Code," *Review of Financial Studies*, 9, 1165–1210.

Franks, Julian R., Kjell G. Nyborg and Walter Torous, 1996, "A Comparison of US, UK, and German Insolvency Codes," *Financial Management*, 25, 19–30.

Franks, Julian R. and Sergey V. Sanzhar, 2006, "Evidence on Debt Overhang from Distressed Equity Issues?" Working Paper, London Business School.

Franks, Julian R. and Oren Sussman, 2005, "Financial Distress and Bank Restructuring of Small to Medium Size UK Companies," *Review of Finance*, 9, 65–96.

Franks, Julian R. and Walter N. Torous, 1994, "A Comparison of Financial Recontracting in Distressed Exchanges and Chapter 11 Reorganizations," *Journal of Financial Economics*, 35, 349–370.

Franks, Julian R. and Walter N. Torous, 1989, "An Empirical Investigation of U.S. Firms in Reorganization," *Journal of Finance*, 44, 747–769.

Gertner, Robert and David Scharfstein, 1991, "A Theory of Workouts and the Effects of Reorganization Law," *Journal of Finance*, 46, 1189–1222.

Giammarino, Ronald M., 1989, "The Resolution of Financial Distress," *Review of Financial Studies*, 2, 25–47.

Gilson, Stuart C., 1997, "Transactions Costs and Capital Structure Choice: Evidence from Financially Distressed Firms," *Journal of Finance*, 52, 161–196.

Gilson, Stuart C., 1990, "Bankruptcy, Boards, Banks, and Blockholders: Evidence on Changes in Corporate Ownership and Control When Firms Default," *Journal of Financial Economics*, 26, 355–387.

Gilson, Stuart C., 1989, "Management Turnover and Financial Distress," *Journal of Financial Economics*, 25, 241–262.

Gilson, Stuart C., Edith S. Hotchkiss and Richard S. Ruback, 2000, "Valuation of Bankrupt Firms," *Review of Financial Studies*, 13, 43–74.

Gilson, Stuart C., Kose John and Larry H.P. Lang, 1990, "Troubled Debt Restructurings: An Empirical Study of Private Reorganization of Firms in Default," *Journal of Financial Economics*, 27, 315–353.

Gilson, Stuart C. and Michael R. Vetsuypens, 1993, "CEO Compensation in Financially Distressed Firms: An Empirical Analysis," *Journal of Finance*, 43, 425–458.

Goyal, Amit, Matthias Kahl and Walter N. Torous, 2003, "The Long-Run Stock Performance of Financially Distressed Firms: An Empirical Investigation," Working Paper, University of North Carolina.

Grossman, Sanford J. and Oliver D. Hart, 1981, "The Allocational Role of Takeover Bids in Situations of Asymmetric Information (in the Allocation Role of the Stock Market)," *Journal of Finance*, 36, 253–270.

Haensly, Paul J., John Theis and Zane Swanson, 2001, "Reassessment of Contagion and Competitive Intra-Industry Effects of Bankruptcy Announcements," *Quarterly Journal of Business & Economics*, 40, 45–63.

Harris, Milton and Artur Raviv, 1991, "The Theory of Capital Structure," *Journal of Finance*, 46, 297–355.

Hart, Oliver, 2000. "Different Approaches to Bankruptcy," Harvard Institute of Economic Research Working Papers 1903, Harvard—Institute of Economic Research.

Hart, Oliver and John Moore, 1998, "Default and Renegotiation: A Dynamic Model of Debt," *Quarterly Journal of Economics*, 113, 1–41.

Haugen, Robert A. and Lemma W. Senbet, 1988, "Bankruptcy and Agency Costs: Their Significance to the Theory of Optimal Capital Structure," *Journal of Financial & Quantitative Analysis*, 23, 27–38.

Haugen, Robert A. and Lemma W. Senbet, 1978, "The Insignificance of Bankruptcy Costs to the Theory of Optimal Capital Structure," *Journal of Finance*, 33, 383–393.

Helwege, Jean and Frank Packer, 2003, "The Decision to Go Public: Evidence from Mandatory SEC Filings of Private Firms," Working Paper, Ohio State University.

Hertzel, Michael G., Zhi Li, Micah S. Officer and Kimberly Rodgers, 2007, "Inter-firm Linkages and the Wealth Effects of Financial Distress along the Supply Chain," *Journal of Financial Economics*, forthcoming.

Hoshi, Takeo, Anil Kashyap and David Scharfstein, 1990, "The Role of Banks in Reducing the Costs of Financial Distress in Japan," *Journal of Financial Economics*, 27, 67–88.

Hotchkiss, Edith S., 1995, "Postbankruptcy Performance and Management Turnover," *Journal of Finance*, 50, 3–21.

Hotchkiss, Edith S., 1993, "Investment Decisions under Chapter 11 Bankruptcy," Doctoral dissertation, New York University.

Hotchkiss, Edith S. and Robert M. Mooradian, 2004, "Postbankruptcy Performance of Public Companies," Working Paper, Boston College and Northeastern University.

Hotchkiss, Edith S. and Robert M. Mooradian, 2003, "Auctions in Bankruptcy," *Journal of Corporate Finance*, 9, 555–574.

Hotchkiss, Edith S. and Robert M. Mooradian, 1998, "Acquisitions as a Means of Restructuring Firms in Chapter 11," *Journal of Financial Intermediation*, 7, 240–262.

Hotchkiss, Edith S. and Robert M. Mooradian, 1997, "Vulture Investors and the Market for Control of Distressed Firms," *Journal of Financial Economics*, 43, 401–432.

James, Christopher, 1996, "Bank Debt Restructurings and the Composition of Exchange Offers in Financial Distress," *Journal of Finance*, 51, 711–727.

James, Christopher, 1995, "When Do Banks Take Equity in Debt Restructurings?" *Review of Financial Studies*, 8, 1209–1234.

Jensen, Michael C., 1991, "Corporate Control and the Politics of Finance," *Journal of Applied Corporate Finance*, 4, 13–33.

Jensen, Michael C., 1989, "Active Investors, LBOs, and the Privatization of Bankruptcy," Journal of Applied Corporate Finance, 2, 35–44.

John, Kose, 1993, "Managing Financial Stress and Valuing Distressed Securities: A Survey and a Research Agenda," *Financial Management*, 22, 60–78.

John, Kose and Teresa John, 1992, "Coping with Financial Distress: A Survey," *Financial Markets, Institutions and Instruments*, 1, 63–78.

Kahl, Matthias, 2002, "Economic Distress, Financial Distress, and Dynamic Liquidation," *Journal of Finance*, 57, 135–168.

Kaiser, Kevin M.J., 1996, "European Bankruptcy Laws: Implications for Corporations Facing Financial Distress," *Financial Management*, 25, 67–85.

Kalay, Avner, Rajeev Singhal and Elizabeth Tashjian, 2007 "Is Chapter 11 Costly?," *Journal of Financial Economics*, 84, 772–796.

Kaplan, Steven N., 1994, "Campeau's Acquisition of Federated: Post Bankruptcy Results," *Journal of Financial Economics*, 35, 123–136.

Kaplan, Steven N., 1989, "Campeau's Acquisition of Federated: Value Destroyed or Value Added," *Journal of Financial Economics*, 25, 191–212.

Kim, Chyhe Esther, 1998, "The Effects of Asset Liquidity: Evidence from the Contract Drill Industry," *Journal of Financial Intermediation*, 7, 151–176.

Klee, 1979, "All You Ever Wanted to Know about Cram Down under the New Bankruptcy Code," *American Bankruptcy Law Journal*, 53, 133–171.

Kruse, Timothy A., 2002, "Asset Liquidity and the Determinants of Asset Sales by Poorly Performing Firms," *Financial Management*, 31, 107–129.

Lang, Larry, Annette Poulsen and Rene Stulz, 1995, "Asset Sales, Firm Performance, and the Agency Costs of Managerial Discretion," *Journal of Financial Economics*, 37, 3–37.

Lang, Larry H. P. and Rene M. Stulz, 1992, "Contagion and Competitive Intra-Industry Effects of Bankruptcy Announcements: An Empirical Analysis," *Journal of Financial Economics*, 32, 45–60.

Lawless, Robert M. and Stephen P. Ferris, 1997, “Professional Fees and Other Direct Costs in Chapter 7 Business Liquidations,” *Washington University Law Quarterly*, 75, 1207–1236.

LoPucki, Lynn M., 2004, “A Team Production Theory of Bankruptcy Reorganization,” *Vanderbilt Law Review*, 57, 741–782.

LoPucki, Lynn M. and Joseph W. Doherty, 2004, “The Determinants of Professional Fees in Large Bankruptcy Reorganization Cases,” *Journal of Empirical Legal Studies*, 1, 111–141.

LoPucki, Lynn M. and William C. Whitford, 1993, “Patterns in the Bankruptcy Reorganization of Large, Publicly Held Companies,” *Cornell Law Review*, 78, 597–618.

LoPucki, Lynn M. and William C. Whitford, 1991, “Venue Choice and Forum Shopping in the Bankruptcy Reorganization of Large Publicly Held Companies,” *Wisconsin Law Review*, 1, 1–63.

Lubben, Stephen J., 2000, “The Direct Costs of Corporate Reorganization: An Empirical Examination of Professional Fees in Large Chapter 11 Cases,” *American Bankruptcy Law Journal*, 509, 508–552.

Maksimovic, V. and G. M. Phillips, 1998, “Asset Efficiency and the Reallocation Decisions of Bankrupt Firms,” *Journal of Finance*, 53, 1495–1532.

Mooradian, Robert M., 1994, “The Effect of Bankruptcy Protection on Investment: Chapter 11 as a Screening Device,” *Journal of Finance*, 49, 1403–1430.

Mooradian, Robert M. and Harley E. Ryan, Jr., 2005, “Out-of-Court Restructurings and the Resolution of Financial Distress: Section 3(a)(9) Compared to Investment-Bank-Managed Exchange Offers,” *Journal of Business*, 78, 1593–1624.

Morrison, Edward R., 2007, “Bankruptcy Decisionmaking: An Empirical Study of Continuation Bias in Small Business Bankruptcies,” *Journal of Law and Economics*, forthcoming.

Noe Thomas H. and Michael J. Rebello, 2003, “Reputation and the Market for Distressed-Firm Debt,” *Journal of Financial and Quantitative Analysis*, 38, 503–522.

Ofek, Eli, 1993, “Capital Structure and Firm Response to Poor Performance,” *Journal of Financial Economics*, 34, 3–30.

Opler, Tim C. and Sheridan Titman, 1994, “Financial Distress and Corporate Performance,” *Journal of Finance*, 49, 1015–1040.

Povel, Paul and Rajdeep Singh, 2007, “Sale-Backs in Bankruptcy,” *Journal of Law, Economics and Organization*.

Pulvino, Todd C., 1999, “Effects of Bankruptcy Court Protection on Asset Sales,” *Journal of Financial Economics*, 52, 151–186.

Pulvino, Todd C., 1998, “Do Asset Fire Sales Exist? An Empirical Investigation of Commercial Aircraft Transactions,” *Journal of Finance*, 53, 939–978.

Ramey, Valerie A. and Matthew D. Shapiro, 2001, “Displaced Capital: A Study of Aerospace Plant Closings,” *Journal of Political Economy*, 109, 958–992.

Ravid, Abraham S. and Stefan Sundgren, 1998, “The Comparative Efficiency of Small-Firm Bankruptcies: A Study of the US and Finnish Bankruptcy Codes,” *Financial Management*, 27, 28–40.

Roe, Mark J., 1983, “Bankruptcy and Debt: A New Model for Corporate Reorganization,” *Columbia Law Review*, 83, 527–602.

Senbet, L. and J. Seward, 1995, “Financial Distress, Bankruptcy and Reorganization,” in Jarrow, Maksimovic, and Ziemba (Eds.), Vol. 9: *Finance*, Elsevier, New York, Chapter 28.

Shleifer, Andrei and Robert W. Vishny, 1992, “Liquidation Values and Debt Capacity: A Market Equilibrium Approach,” *Journal of Finance*, 47, 1343–1366.

Sicherman, Neil W. and Richard H. Pettway, 1992, “Wealth Effects for Buyers and Sellers of the Same Divested Assets,” *Financial Management*, 21, 119–128.

Stromberg, P., 2000, “Conflicts of Interest and Market Illiquidity in Bankruptcy Auctions: Theory and Tests,” *Journal of Finance*, 55, 2641–2692.

Tashjian, Elizabeth, Ronald C. Lease and John J. McConnell, 1996, “Prepacks: An Empirical Analysis of Prepackaged Bankruptcies,” *Journal of Financial Economics*, 40, 135–162.

Thorburn, Karin S., 2000, “Bankruptcy Auctions: Costs, Debt Recovery, and Firm Survival,” *Journal of Financial Economics*, 58, 337–368.

Vig, Vikrant, 2007, "Access to Collateral and Corporate Debt Structure: Evidence from a Natural Experiment," Working Paper, London Business School.

Warner, Jerold B., 1977, "Bankruptcy Costs: Some Evidence," *Journal of Finance*, 32, 337–347.

Warner, Jerold, Ross L. Watts and Karen H. Wruck, 1988, "Stock Prices and Top Management changes," *Journal of Financial Economics*, 20, 461–492.

Weisbach, Michael S., 1988, "Outside Directors and CEO Turnover," *Journal of Financial Economics*, 20, 431–460.

Weiss, Lawrence A., 1990, "Bankruptcy Costs and Violation of Claims Priority," *Journal of Financial Economics*, 27, 285–314.

Weiss, Lawrence A. and Karen Hopper Wruck, 1998, "Information Problems, Conflicts of Interest, and Asset Stripping: Chapter 11's Failure in the Case of Eastern Airlines," *Journal of Financial Economics*, 48, 55–97.

White, Michelle J., 1996, "The Costs of Corporate Bankruptcy: A US-European Comparison," in J. Bhandari and L. Weiss (Eds.), *Corporate Bankruptcy: Economic and Legal Perspectives*, Cambridge University Press, Cambridge, UK, 467–500.

White, Michelle J., 1989, "The Corporate Bankruptcy Decision," *Journal of Economic Perspectives*, 3, 129–151.

White, Michelle J., 1984, "Bankruptcy Liquidation and Reorganization," In Dennis Logue and James Seward (Eds.), *Handbook of Modern Finance*, Warren, Gorham & Lamont, Boston, E7-1-E7-44.

White, Michelle J., 1980, "Public Policy toward Bankruptcy: Me-First and Other Priority Rules," *Bell Journal of Economics*, 11, 550–564.

Wruck, Karen Hopper, 1990, "Financial Distress, Reorganization, and Organizational Efficiency," *Journal of Financial Economics*, 27, 419–444.

Zender, Jaime F., 1991, "Optimal Financial Instruments," *Journal of Finance*, 46, 1645–1663.

*Chapter 15*

# CORPORATE TAKEOVERS*

SANDRA BETTON

*John Molson School of Business*
*Concordia University*

B. ESPEN ECKBO

*Tuck School of Business*
*Dartmouth College*

KARIN S. THORBURN

*Tuck School of Business*
*Dartmouth College*

## Contents



* Surveying the vast area of corporate takeovers is a daunting task, and we have undoubtedly missed many interesting contributions. We apologize to those who feel their research has been left out or improperly characterized, and welcome reactions and comments. Some of the material in Section 3 is also found in Eckbo (2008).

*Handbook of Empirical Corporate Finance, Volume 2*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

## Abstract

This chapter surveys the recent empirical literature and adds to the evidence on takeover bids for U.S. targets, 1980–2005. The availability of machine readable transaction databases have allowed empirical tests based on unprecedented sample sizes and detail. We review both aggregate takeover activity and the takeover process itself as it evolves from the initial bid through the final contest outcome. The evidence includes determinants of strategic choices such as the takeover method (merger v. tender offer), the size of opening bids and bid jumps, the payment method, toehold acquisition, the response to target defensive tactics, and regulatory intervention (antitrust), and it offers links to executive compensation. The data provides fertile grounds for tests of everything ranging from signaling theories under asymmetric information to strategic competition in product markets and to issues of agency and control. The evidence is supportive of neoclassical merger theories. For example, regulatory and technological changes, and shocks to aggregate liquidity, appear to drive out market-to-book ratios as fundamental drivers of merger waves. Despite the market boom in the second half of the 1990s, the proportion of all-stock offers in more than 13,000 merger bids did not change from the first half of the decade. While some bidders experience large losses (particularly in the years 1999 and 2000), combined value-weighted announcement-period returns to bidders and targets are significantly positive on average. Long-run post-takeover abnormal stock returns are not significantly different from zero when using a performance measure that replicates a feasible portfolio trading strategy. There are unresolved econometric issues of endogeneity and self-selection.

## Keywords

takeover, merger, tender offer, auction, offer premium, bidder gains, toeholds, markups, hostility, executive compensation, arbitrage, announcement return, long-run performance, monopoly, antitrust

## 1. Introduction

Few economic phenomena attract as much public attention and empirical research as the various forms of transactions in what Manne (1965) dubbed "the market for corporate control." Corporate takeovers are among the largest investments that a company ever will undertake, thus providing a unique window into the value implications of important managerial decisions and bid strategies, and into the complex set of contractual devices and procedures that have evolved to enable the deals to go through. Empirical research in this area has focused on a wide range of topics including the impact of statutory and regulatory restrictions on the acquisition process (disclosure and target defenses), strategic bidding behavior (preemption, markup pricing, bid jumps, toeholds, payment method choice, hostility), short- and long-run abnormal stock returns to bidders and targets (size and division of takeover gains), and the origin and competitive effects of corporate combinations (efficiency, market power, and antitrust policy). In this survey, we review empirical research on each of these and related topics.

The structure of our survey differs from most earlier empirical reviews, where the focus tends to be on the final bid in completed takeovers.[1] We follow the approach begun by Betton and Eckbo (2000) and examine the entire takeover process as it evolves from the first bid through bid revision(s) and toward the final outcome (success or failure). This more detailed focus on the takeover process is also found in more recent publications.[2] We provide new empirical updates in some areas, using takeovers found in the Thomson Financial SDC database for the period 1980–2005. One limitation of the survey is that we do not discuss the general interplay between the market for corporate control, ownership structure and corporate governance (with the exception of hostile bids).[3] We also limit the review to empirical studies of takeovers of U.S. target firms.[4] Takeovers by financial buyers such as leveraged buyouts (LBOs) are surveyed in Eckbo and Thorburn (2008a), Chapter 16 of this volume.

Throughout, we use the term *takeover* generically for any acquisition of corporate control through the purchase of the voting stock of the target firm, regardless of whether the bid is in the form of a merger agreement or a tender offer. Moreover, in our vernacular, the first observed bid for a specific target starts a takeover "contest" whether or not subsequent bids actually materialize. All initial bids start a contest in the sense of attracting potential competition from rival bidders and/or incumbent target management. This is true even after signing a merger agreement as director fiduciary duties require the target board to evaluate competing offers all the way until target shareholders have voted to accept the agreement (the fiduciary out). Also, we know from the data that a friendly

[1] Jensen and Ruback (1983), Jarrell, Brickley, and Netter (1988), Eckbo (1988), Andrade, Mitchell, and Stafford (2001), Martynova and Renneboog (2005, 2007).

[2] Bhagat, Dong, Hirshleifer, and Noah (2005), Boone and Mulherin (2007b), Betton, Eckbo, and Thorburn (2007). See also the survey by Burkart and Panunzi (2006).

[3] Research on corporate ownership structure, managerial private benefits of control, shareholder activism and voting, etc., is surveyed in Becht, Bolton, and Roll (2003), Dyck and Zingales (2004), and Adams and Ferreira (2007).

[4] See Martynova and Renneboog (2006) for the European takeover market.

merger negotiation is not a guarantee against the risk of turning the takeover process into an open auction for the target. The contest perspective helps us understand why initially friendly merger bids are sometimes followed by tender offers and vice versa, why we sometimes observe bid revisions even in the absence of rival bidders, why target hostility emerges even when the initial bidder appears to be friendly, and why the auction for the target sometimes fails altogether (no bidder wins).

We begin in Section 2, "Takeover activity," with a brief discussion of takeover waves, followed by a detailed description of the initial bids in an unprecedented sample consisting of more than 35,000 takeover contests for U.S. public targets over the period 1980–2005. The description includes initial deal values, degree of actual competition (single-bid versus multiple-bid contests), success rates, the deal form (merger versus tender offer), payment method (cash, stock, or a mix), target attitude (hostile v. neutral or friendly), product market connection (horizontal v. nonhorizontal), public versus private status of the bidder and the target, time to second bid, and total contest duration. We also characterize the actual institutional environment in which firms are sold, including rules governing tender offers and various contractual innovations designed to support merger negotiations. Moreover, this section comments on the determinants of the choice between merger and tender offer, and it discusses the impact of mandatory disclosure rules on premiums in tender offers.

We then move to Section 3, "Bidding strategies." In theory, a complex set of factors determine the design of optimal bids.[5] These include auction design, the nature of bidder valuations, the private information environment, target ownership structure, and bidding costs. A key empirical challenge is to establish whether there is evidence of strategic bidding and/or signaling effects in the data. As the first mover in the takeover game, the initial bidder is in a unique position, so strategic bidding behavior is likely to be most evident in the first bid. Thus, our empirical analysis is structured around the actions of the first bidder making a control-offer for the target.

We begin Section 3 with a brief description of the classical free-rider model of Grossman and Hart (1980b) and of the standard auction setup in models with a single seller. This helps frame some of the subsequent empirical test results. We then review empirical work on strategic decisions, including the initial bidder's choice between merger and tender offer, the payment method, pre-bid acquisition of target shares in the market (toehold bidding), markup pricing following a pre-bid target stock price runup, takeover defenses, and acquisitions of formally bankrupt targets. This section focuses on how the various actions affect the initial and final offer premium.

In the first part of Section 4, "Takeover gains," we discuss estimates of the announcement effect of takeovers on the wealth of bidder and target shareholders. In their review of the empirical evidence from the 1960s and 1970s, Jensen and Ruback (1983) conclude that the average sum of the deal-related stock market gains to bidders and targets is significantly positive. Subsequent surveys have also made this conclusion (Jarrell, Brickley, and Netter, 1988; Andrade, Mitchell, and Stafford, 2001). On the other hand,

[5] For surveys of takeover theories, see Spatt (1989), Hirshleifer (1995), Burkart and Panunzi (2006), and Dasguptha and Hansen (2007).

as pointed out by Roll (1986) and strongly emphasized in Moeller, Schlingemann, and Stulz (2004), bidder deal-related abnormal returns are often negative. Drawing on Betton, Eckbo, and Thorburn (2008c), we show that the value-weighted sum of announcement-induced three-day abnormal stock return to bidders and targets is significantly positive. This conclusion holds for the entire sample period 1980–2005 as well as for each of the five-year subperiods. We also discuss the large bidder dollar losses from the period 1998–2001 that are the central focus of Moeller, Schlingemann, and Stulz (2004).

In the second part of Section 4, we review and update estimates of abnormal stock returns to merged firms over the five-year period following successful completion of the takeover. We show that post-merger performance is on average negative if one benchmarks the returns with the returns to nonmerging firms matched on size and book-to-market ratio. However, the abnormal performance is insignificantly different from zero when using standard asset pricing benchmarks. These conflicting inferences concerning long-run performance produced by the matched-firm technique and the "Jensen's alpha" (regression) procedure is reminiscent of the debate in the literature on security offerings.[6]

In Section 5, "Bondholders, executives, and arbitrageurs," we review empirical studies of the wealth implications of mergers for bondholders, for bidder and target executives and directors, and for arbitrageurs. Issues for bondholders include the potential for a wealth transfer from stockholders to bondholders as a result of the coinsurance effect of takeovers, and protection against event-risk. For executives, a key issue is the disciplinary role of the market for corporate control, and whether undertaking value-decreasing takeovers is costly in terms of increased turnover and/or reduced compensation. Merger (risk) arbitrage is an investment strategy that tries to profit from the spread between the offer price and the target stock price while the offer is outstanding. It is essentially a bet on the likelihood that the proposed transaction closes. Research documents the determinants of the arbitrage spreads, trading volumes, the role of transaction costs in establishing these positions, and the returns to arbitrage activity.

Finally, in Section 6, "Takeovers, competition, and antitrust," we broaden the focus to the industry of the bidder and target firms. The key empirical issue centers on the extent to which mergers are driven by opportunities for creating market power. While the potential for market power is most obvious for horizontal combinations (as recognized by the antitrust authorities), vertical mergers may generate buying power vis-à-vis suppliers. We review empirical tests employing estimates of abnormal stock returns to the industry rivals of the merging firms. These estimates show that mergers tend to cause a wealth effect throughout the industry of the target firm. One consistent interpretation is that synergy gains generated by takeovers represent quasi-rents from scarce resources owned throughout the target industry. The alternative hypothesis — that the industry wealth effect represents the present value of monopoly rents from collusive behavior—is consistently rejected by the empirical studies. We end this section with a brief discussion of implications for antitrust policy.

The survey concludes in Section 7 with a summary of the key findings and some directions for future research.

[6] See the reviews by Ritter (2003) and Eckbo, Masulis, and Norli (2007).

## 2. Takeover activity

### 2.1. Merger waves

A merger wave is a clustering in time of successful takeover bids at the industry- or economy-wide level. This is shown in Figure 1 for U.S. publicly traded firms over the period 1926–2006. The figure plots the annual fraction of all firms on the University of Chicago's Center for Research in Security Prices (CRSP) database in January of each year which delists from the stock exchange due to merger during the year. Looking back, aggregate takeover activity appears to occur in distinct waves—peaks of heavy activity followed by troughs of relatively few transactions.

Merger activity tends to be greatest in periods of general economic expansion. This is hardly surprising as external expansion through takeovers is just one of the available corporate growth strategies. As seen in Figure 1, aggregate takeover activity was relatively high in the late 1960s, throughout the 1980s, and again in the late 1990s. These waves are typically labeled the conglomerate merger wave of the 1960s, the refocusing wave of the 1980s, and the global wave or strategic merger wave of the 1990s.[7]

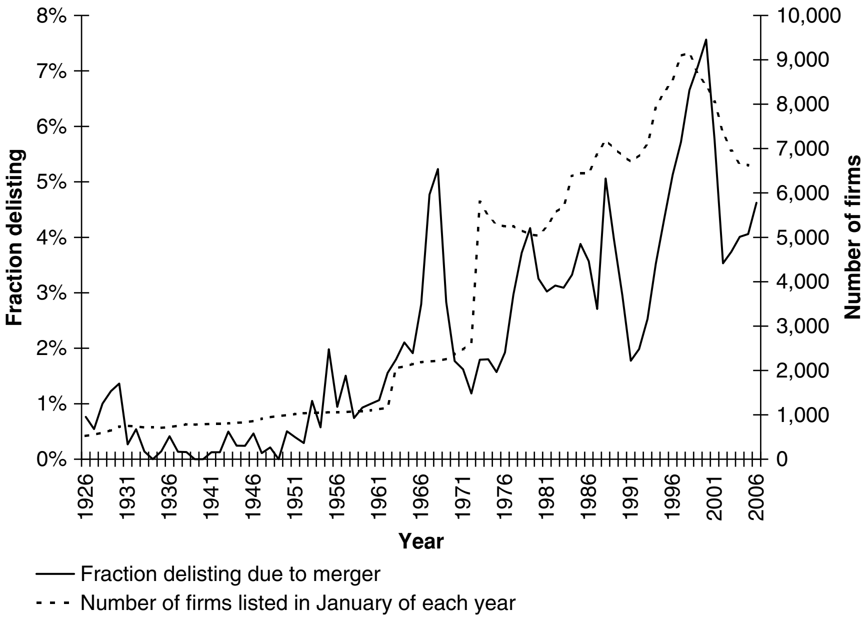



Fig. 1. Annual fraction of all publicly traded (CRSP) firms in January of each year which delists due to merger during the year, 1926–2006.

[7] The merger wave of the late 1890s and early 1900s (not shown in Figure 1) has been referred to as the "Great merger wave" (O'Brien, 1988) or the monopolization wave (Stigler, 1950).

These labels indicate the character of the typical merger within the wave. Thus, a majority of the mergers in the 1960s were between firms operating in unrelated industries (conglomerate mergers). It is possible that the internal capital market created through conglomerate merger may have reduced financing costs for unrelated corporate entities.[8] On the other hand, since conglomerates tend to reduce (diversify) the risk of managerial human capital and to create "business empires" perhaps valued excessively by CEOs, the conglomerate wave may also reflect an agency problem. The agency view is strengthened by the fact that executive compensation showed little sensitivity to firm performance at the time (Jensen and Murphy, 1990). Thus, value-reducing diversifying mergers may have had little consequence for CEOs, leading to excessive conglomeration. However, estimates of abnormal stock returns around the conglomerate takeovers of the 1960s do not indicate that these investments were on average detrimental to shareholder wealth.[9]

The merger wave of the 1980s includes a number of mergers designed either to downsize or to specialize operations. Some of these corrected excessive conglomeration, others responded to excess capacity created by the 1970s recession (following the creation of the OPEC oil cartel), while yet others responded to the important advances in information and communication technologies (Jensen, 1986, 1993). The 1980s also experienced the largest number of hostile bids in U.S. history. The subsequent spread of strong takeover defenses in the late 1980s halted the use of hostile bids, and the late 1990s saw a "friendly" merger wave, with a primary focus on mergers with global strategic partners.

A complex set of factors are at play in any given merger wave. For example, merger waves may be affected by changes in legal and regulatory regimes. Shleifer and Vishny (1991) suggest that the demand for conglomerate mergers in the 1960s may have been triggered by the stricter antitrust laws enacted in the early 1950s.[10] While this may have had an effect in the United States, it is interesting that countries with lax antitrust laws (Canada, Germany, and France) also experienced diversification waves in the 1960s (Matsusaka, 1996). Industry-specific deregulations may also create merger waves, such as deregulations of the airline industry in 1970s (Spiller, 1983; Slovin, Sushka, and Hudson, 1991) and of the utility industry in 1992 (Jovanovic and Rousseau, 2004; Becher, Mulherin, and Walkling, 2008).

The perhaps most compelling theory of merger waves rests on the technological link between firms in the same industry. A merger implementing a new technological innovation may, as news of the innovation spreads, induce follow-on takeovers among industry rivals for these to remain competitive. This argument goes back at least to Coase (1937), who suggests that scale-increasing technological change is an important driver of merger

[8] Hubbard and Palia (1999), Maksimovic and Phillips (2002). Maksimovic and Phillips (2007) reviews internal capital markets, while Eckbo and Thorburn (2008a) reviews breakup transactions that may follow excessive conglomeration.

[9] Loderer and Martin (1990), Matsusaka (1993), Akbulut and Matsusaka (2003).

[10] One important antitrust development was the 1950 Celler-Kefauver amendment of the 1914 Clayton Act. See Section 6.

activity. Jensen (1993) draws parallels between merger activity and the technological changes driving the great industrial revolutions of the nineteenth and twentieth centuries. Gort (1969) and Jovanovic and Rousseau (2002) use the related-technology notion to build theories of resource reallocations based on valuation discrepancies and Tobin's Q. Rhodes-Kropf and Robinson (2008) propose a search theory where bidders and targets match up based on the degree of complementarity of their resources.

There is substantial evidence of industry-clustering of mergers.[11] Andrade and Stafford (2004) find that mergers play both an expansionary and a contractionary role in industry restructurings. During the 1970s and 1980s, excess capacity tended to drive industry consolidation through merger, while peak capacity utilization triggered industry expansion through nonmerger investment (internal expansion). This phenomenon appears to have reversed itself in the 1990s, as industries with strong growth prospects, high profitability, and near capacity also experienced the most intense merger activity. Maksimovic and Phillips (2001) use performance improvements at the plant level to support the neoclassical reallocation theory of merger waves. Maksimovic, Phillips and Prabhala (2008) show that, for mergers in manufacturing industries, the acquirer on average closes or sells about half of the target firm's plants. Moreover, a simple neoclassical model of production helps predict the choice of which target plants to sell/close. The plants that are kept are often restructured, resulting in productivity increases. Servaes and Tamayo (2007) find that industry peers respond by financing and investment policies when another firm in the industry is the subject of a hostile takeover attempt, suggesting that firms in the same industry are linked by both technology and resource complementarities.

The fact that merger waves are correlated with economic expansions and high stock market valuations has also spurred the development of models in which merger waves result from market overvaluation and managerial timing. The idea is that bull markets may lead bidders with overvalued stock as currency to purchase the assets of undervalued (or less overvalued) targets. In Shleifer and Vishny (2003), target managements accept overpriced bidder stock as they are assumed to have a short time horizon. In Rhodes-Kropf and Viswanathan (2004), target management accepts more bids from overvalued bidders during market valuation peaks because they overestimate synergies during these periods. In both models, the bidder gets away with selling overpriced stock.

Eckbo, Giammarino, and Heinkel (1990) present a rational expectations model of the payment method in takeovers with two-sided information asymmetry (neither the bidder nor the target knows the true value of the shares of the other), in which the fraction of the deal paid in cash signals the bidder's true value. In equilibrium, the target receives correctly priced bidder stock as part of the payment. Their analysis suggests that the pooling equilibrium proposed by Shleifer and Vishny (2003) is sensitive to the possibility of mixed offers. As shown in Figure 7 below, mixed offers represent a substantial portion

[11] Mitchell and Mulherin (1996), Mulherin and Boone (2000), Andrade, Mitchell, and Stafford (2001), Maksimovic and Phillips (2001), Andrade and Stafford (2004), and Harford (2005).

of all takeovers: during the period 1980 through 2005, there were nearly as many mixed cash-stock offers as there were all-stock bids. Moreover, despite the market boom in the second half of the 1990s, the relative proportions of all-cash, all-stock, and mixed cash-stock offers in more than 15,000 merger bids did not change from the first half of the decade. Also, during the 1996–2000 period with peak market valuations, the sum of all-cash and mixed cash-stock bids in mergers equals the number of all-stock merger bids.

Rhodes-Kropf, Robinson and Viswanathan (2005), Ang and Cheng (2006) and Dong, Hirshleifer, Richardson, and Teoh (2006) find that merger waves coincide with high market-to-book (M/B) ratios. One argument is that the M/B ratio is a reliable proxy for market overvaluation and that investor misvaluations tend to drive merger waves. High market valuations may be a fundamental driver of merger waves as bidders attempt to sell overpriced stock to targets (and succeed). Rhodes-Kropf and Viswanathan (2004) present an interesting model in which *rational* (Bayesian) managers accept too many all-stock merger bids when the stock market booms and too few when the market is low. They assume that the market's pricing error has two components, one economywide and another that is firm-specific. When receiving a bid, the target attempts to filter out the marketwide error component. The Bayesian update puts some weight on there being high synergies in the merger, so when the marketwide overvaluation is high, the target is more likely to accept the offer. In other words, bids tend to look better in the eyes of the target when the market is overvalued.

Harford (2005) contrasts these predictions with a neoclassical argument in which the driver of merger waves is market liquidity. That is, under the neoclassical view, market liquidity is the fundamental driver of *both* M/B ratios and merger waves.[12] Harford (2005) constructs a measure of aggregate capital liquidity based on interest rate (default) spreads and uses this measure in a "horse race" with M/B ratios in predicting industry merger waves. He finds that waves are preceded by deregulatory events and high capital liquidity. More importantly, he shows that the capital liquidity variable eliminates the ability of M/B ratios to predict industry merger waves. He concludes that aggregate merger waves are caused by the clustering of shock-driven industry merger waves, not by attempts to time the market.

Patterns of merger waves notwithstanding, predicting individual target firms with any accuracy has proven difficult.[13] Probability estimates are sensitive to the choice of size and type of control sample. Firm size consistently predicts targets across most studies, while results are mixed for other commonly used variables, including factors capturing growth, leverage, market-to-book ratios, and ownership structure.

[12] For example, Shleifer and Vishny (1992) argue that merger waves tend to occur in booms because increases in cash flows simultaneously raise fundamental values and relax financial constraints, bringing market values closer to fundamental values. Harford (1999) shows that firms that have built up large cash reserves are more prone to acquire other firms.

[13] Hasbrouck (1985), Palepu (1986), Mork, Shleifer, and Vishny (1988), Mikkelson and Partch (1989), Ambrose and Megginson (1992), Shivdasani (1993), Comment and Schwert (1995), Cremers, Nair, and John (2008).

### *2.2. Takeover contests, 1980–2005*

As discussed in Section 2.3, after signing a merger agreement, the target board is normally required to consider new outside offers until target shareholders have given final approval of the takeover (the so-called fiduciary out clause). This means that no bidder can expect to lock up the target through negotiations but must be prepared for potential competition. All initial bidders, whether the initial bid is in the form of a merger or a tender offer, face this potential competition. We therefore refer to all initial bids as initiating a *control contest* whether or not multiple bids actually emerge ex-post.

The "contest tree" in Figure 2 shows the potential outcomes of any initial bid. In the first round of the contest, one of three outcomes will occur: (1) the bid is accepted by the target and the contest ends; (2) the bid is rejected and the contest ends; or (3) the bid is followed by one or more rival bids and/or bid revisions before the contest ends. After two or more rounds of bidding, one of three final outcomes will occur: (4) the initial bidder wins control; (5) a rival bidder wins control; or (6) no bidder wins control (the target remains independent). Later in this chapter, we use this contest-tree structure to organize successive bids for the same target and to describe recent bidding activity.

#### *2.2.1. Initial bidders and offer characteristics*

We collect bids from the Thomson Financial SDC mergers and acquisitions database. SDC provides records of individual bids based on information in the news and Securities and Exchange (SEC) filings by the bidder and target firms. As shown by Boone and Mulherin (2007b), targets increasingly initiate takeovers through a process where they privately solicit several potentially interested bidders and select a negotiating partner

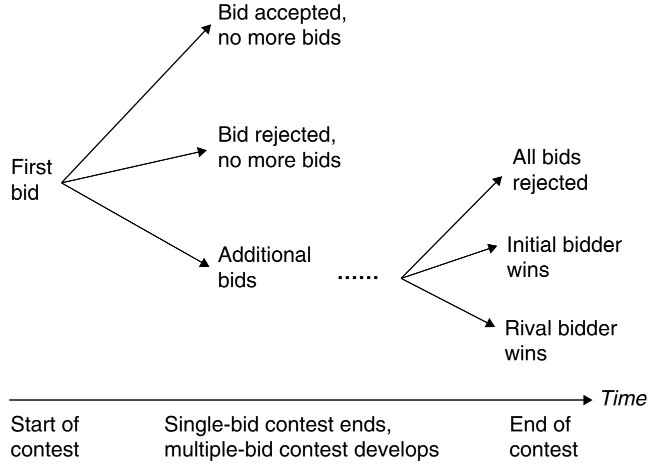



Fig. 2. Takeover contest structure and outcomes.

among the respondents. The initial bidder identified by the SDC may well have emerged from such a process. However, we follow standard practice and use the first *official* (public) bid for the target to start the contest.

The bids are by U.S. or foreign bidders for a U.S. public or private target announced between January 1980 and December 2006. We start by downloading all mergers (SDC deal form M), acquisition of majority interest (AM), acquisition of partial interest (AP), and acquisition of remaining interest (AR).[14] This results in a total of 70,548 deals (bids). We then use the SDC tender flag to identify which of the bids are tender offers and control-block trades.[15] Next, we organize the 70,548 bids into control contests, where a target is identified using the CUSIP number. A control bid is defined as a merger or acquisition (tender offer) of majority interest where the bidder holds less than 50% of the target shares at announcement.[16] The control contest begins with the first control bid for a given target and continues until 126 trading days have passed without any additional offer (including acquisitions of minority interests). Each time an additional offer for the target is identified, the 126 trading day search window rolls forward.

A control bid is successful if SDC's deal status field states "completed." For successful contests, the formal contest ending date is the earlier of SDC's effective conclusion date and target delisting date. Unsuccessful contests (no bid is successful) end with the offer date of the last control bid or partial acquisition plus 126 trading days (given that there were no more bids in the 126-day period).[17] This selection process produces a total of 35,727 contests. Control contests may be single-bid, multiple-bid but single bidder, or multiple bidder. A multiple-bid contest occurs either because there are multiple bidders or because the initial bidder submits a bid revision. Bid revisions are shown on SDC as a difference between the initial and final offer price within one SDC deal entry. For multiple-bidder contests, the identity of the successful bidder is determined by comparing the CUSIP of the successful bidder with the CUSIP of the initial control bidder. If they are the same, then the initial bidder is successful; otherwise a rival bidder is successful.

Tables 1 through 3 and Figures 3 through 6 describe the central characteristics of the total sample of 35,727 initial bids and their outcomes. Table 1 shows how the total sample is split between initial merger bids (28,994), tender offers (4,500), and control-block trades (2,224). Panel A of Figure 3 shows the annual distribution of the initial merger bids and tender offers, confirming the peak activity periods also shown earlier

[14] We exclude all transactions classified as exchange offers, acquisition of assets, acquisition of certain assets, buybacks, recaps, and acquisition (of stock).

[15] This identification proceeds as follows: If the tender flag is "no" and the deal form is a merger, then the deal is a merger. If the tender flag is "no" and the deal form is "acquisition of majority interest" *and* the effective date of the deal equals the announcement date, then the deal is classified as a control-block trade. If the tender flag is "yes", or if the tender flag is "no" and it is not a block trade, then the deal is a tender offer.

[16] If information on the bidder's prior ownership in the target is missing from SDC, we assume that the prior shareholding is zero.

[17] We removed a single contest due to missing target name, 23 contests due to multiple successful bids, and 36 contests where the target was a Prudential-Bache fund.

Table 1

Total number of takeover contests and characteristics of the initial control bid, 1980–2005.

Control bids (mergers and tender offers) and their characteristics are from SDC. Control contests begin with the first control bid for a company and continue until 126 trading days have passed without any offer (including acquisitions of minority interests). Each time an offer for the target is identified, the 126 trading day search window rolls forward. Multiple-bid contests occur when there are either multiple control bidders or the initial bidder revises the bid. Successful offers are identified as completed by the SDC status variable. Initial deal values provided by SDC for the first control offer in the contest are restated in constant 2000 dollars using the Consumer Price Index obtained from the Bureau of Labor Statistics (Series Id: CUUR0000SA0).

| | | Number of contests | | | Number of successful contests | | | | Initial deal values | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Multiple bids | | | Multiple bids | | | (millions of 2000 constant $ ) | | | |
| | | | | | | | Multiple bidders | | | | | |
| | All contests | Single bid | Single bidder | Multiple bidders | Single bid | Single bidder | Initial bidder | Rival bidder | Number | Mean | Stdev | Median |
| **All Contests** | 35,727 | 33,836 | 671 | 1,220 | 26,012 | 476 | 732 | 646 | 21,476 | 419.43 | 2,610.1 | 36.39 |
| **Initial Deal Form** | | | | | | | | | | | | |
| Merger bid | 28,994 | 27,663 | 481 | 850 | 21,087 | 349 | 481 | 462 | 17,367 | 426.33 | 2,838.2 | 34.57 |
| Tender offer | 4,509 | 3,955 | 188 | 366 | 2,707 | 125 | 246 | 183 | 3,267 | 481.36 | 1,384.3 | 79.01 |
| Block trade | 2,224 | 2,218 | 2 | 4 | 2,218 | 2 | 5 | 1 | 842 | 36.95 | 93.1 | 8.45 |

*(Continued)*

Table 1 *(Continued)*

| | Number of contests | | | Number of successful contests | | | | Initial deal values (millions of 2000 constant $ ) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Multiple bids | | | Multiple bids | | | | | | |
| | | | | | | Multiple bidders | | | | | |
| | All contests | Single bid | Single bidder | Multiple bidders | Single bid | Single bidder | Initial bidder | Rival bidder | Number | Mean | Stdev | Median |
| **Initial Payment Method** | | | | | | | | | | | | |
| All-Cash | 4,798 | 4,114 | 268 | 416 | 3,424 | 160 | 266 | 209 | 4,798 | 309.54 | 1,130.1 | 54.82 |
| All-Stock | 6,983 | 6,699 | 161 | 123 | 5,708 | 146 | 181 | 62 | 6,983 | 492.89 | 3,747.1 | 33.56 |
| Mixed Cash-Sec. | 6,995 | 6,516 | 172 | 307 | 5,672 | 133 | 221 | 167 | 6,995 | 538.41 | 2,410.9 | 38.56 |
| No payment data | 16,951 | 16,507 | 70 | 374 | 11,208 | 37 | 64 | 208 | 2,700 | 116.50 | 621.3 | 19.45 |
| **Initial Attitude** | | | | | | | | | | | | |
| Hostile | 590 | 207 | 153 | 230 | 35 | 72 | 126 | 133 | 560 | 1611.8 | 4,055.0 | 301.06 |
| Unsolicited | 435 | 272 | 38 | 125 | 3 | 5 | 6 | 84 | 362 | 608.92 | 2,729.9 | 103.19 |
| Friendly | 34,702 | 33,357 | 480 | 865 | 25,974 | 399 | 600 | 429 | 20,554 | 383.61 | 2,549.4 | 33.94 |
| **Initial offer** | | | | | | | | | | | | |
| horizontal | 10,452 | 10,043 | 154 | 255 | 7,953 | 130 | 198 | 128 | 6,080 | 562.48 | 3,172.0 | 40.17 |

Table 2

Bidder and target public status and initial bidder nationality in takeover contests, 1980–2005.

Control bids (mergers and tender offers) and their characteristics are from SDC. Control contests begin with the first control bid for a company and continue until 126 trading days have passed without any offer (including acquisitions of minority interests). Each time an offer for the target is identified, the 126 trading day search window rolls forward. Multiple bid contests occur when there are either multiple control bidders or the initial bidder revises the bid. Successful offers are identified as completed by the SDC status variable. Initial deal values provided by SDC for the first control offer in the contest are restated in constant 2000 dollars using the Consumer Price Index obtained from the Bureau of Labor Statistics (Series Id: CUUR0000SA0). Bidder nationality and the public status of target and initial bidder is from SDC. "Other bidder" status includes unknown (268), joint-venture (115), individual (54), mutual (23) and government (19).

| | Number of contests | | | | Number of successful contests | | | | Initial deal values (millions of 2000 constant $ ) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Multiple bids | | | Multiple bids | | | | | | |
| | | | | | | | Multiple bidders | | | | | |
| | All contests | Single bid | Single bidder | Multiple bidders | Single bid | Single bidder | Initial bidder | Rival bidder | Number | Mean | Stdev | Median |
| **Public target and status of initial bidder** | | | | | | | | | | | | |
| Public bidder | 8,259 | 7,364 | 397 | 498 | 5,822 | 321 | 449 | 248 | 7,088 | 957.38 | 4,337.4 | 115.75 |
| Private bidder | 3,656 | 3,012 | 180 | 464 | 1,710 | 93 | 162 | 263 | 2,424 | 388.07 | 1,523.9 | 61.65 |
| Other bidder | 1,270 | 1,125 | 47 | 98 | 841 | 32 | 55 | 52 | 986 | 502.83 | 1,806.4 | 84.87 |
| **Private target and status of initial bidder** | | | | | | | | | | | | |
| Public bidder | 15,799 | 15,675 | 29 | 95 | 12,467 | 23 | 45 | 51 | 9,269 | 66.28 | 233.7 | 16.06 |
| Private bidder | 4,482 | 4,429 | 5 | 48 | 3,413 | 4 | 15 | 27 | 940 | 111.41 | 323.1 | 22.92 |
| Other bidder | 2,261 | 2,231 | 13 | 17 | 1,759 | 3 | 6 | 5 | 769 | 86.26 | 416.5 | 16.84 |
| **Nation of initial bidder** | | | | | | | | | | | | |
| United States | 31,845 | 30,184 | 613 | 1,048 | 23,399 | 431 | 648 | 555 | 19,249 | 404.56 | 2,632.0 | 35.97 |
| Canada | 1,044 | 1,011 | 10 | 23 | 688 | 5 | 9 | 11 | 590 | 222.09 | 933.4 | 11.13 |
| United Kingdom | 716 | 681 | 5 | 30 | 574 | 4 | 18 | 14 | 514 | 701.45 | 3,364.2 | 41.43 |
| Other international | 2,122 | 1,960 | 43 | 119 | 1,351 | 36 | 57 | 66 | 1,123 | 649.00 | 2,409.7 | 77.88 |

Table 3

Distribution of the time to completion of control contests for successful U.S. target firms, classified by the type of initial offer and the public status of the bidder and target firms. Total sample of 25,166 successful targets, 1980–2005.

Control contests begin with the first control bid for a company and continue until 126 trading days have passed without any offer (including acquisitions of minority interests). Each time an offer for the target is identified, the 126 trading day search window rolls forward. The table reports the number of trading days from the date of the initial control bid to the effective merger date reported by the SDC. The effective date is the date target shareholders approve the merger agreement.

| Public status | | No of | Trading days from initial control bid to effective date | | | |
|---|---|---|---|---|---|---|
| | | | | | Quartiles | |
| Target | Bidder | Obs. | Mean | Median | Lowest | Highest |
| **Entire sample** | | 25,166 | 64.62 | 42 | 0 | 100 |
| **Merger** | | 2,2030 | 62.42 | 39 | 0 | 100 |
| Public | Public | 5,147 | 107.92 | 96 | 63 | 136 |
| Public | Private | 1,766 | 97.84 | 86 | 42 | 136 |
| Private | Public | 11,131 | 48.42 | 19 | 0 | 73 |
| Private | Private | 3,986 | 27.09 | 0 | 0 | 28 |
| **Tender** | | 3,136 | 80.06 | 52 | 30 | 98 |
| Public | Public | 1,257 | 71.44 | 49 | 31 | 85 |
| Public | Private | 1,030 | 97.8 | 67 | 34 | 123 |
| Private | Public | 533 | 73.61 | 43 | 21 | 84 |
| Private | Private | 316 | 67.38 | 41 | 19 | 92 |

in Figure 1. The number of merger bids exceeds the number of tender offers by a factor of at least three in every sample year and by a factor of seven for the total period. The relative frequency of tender offers peaked in the second half of the 1980s.

The SDC deal value, converted to constant 2000 dollars using the Consumer Price Index of the Bureau of Labor Statistics (Series Id: CUUR0000SA0), averages $436 million for initial merger bids, and $480 million for initial tender offers.[18] The distribution of deal values is highly skewed, with a median of only $35 million for mergers and $79 million for tender offers, respectively. The annual deal values plotted in Panel A of Figure 3 show that tender offers have somewhat greater deal values in the first half of the sample period, and that merger bids have slightly greater deal values than tender offers in the years 1998–2000.

Table 1 also provides information on the initial bidder's choice of payment method, the target's reaction to the initial bid, and the product-market relationship between the initial bidder and the target. SDC provides payment information for 53% of the

[18] SDC deal values are available for 17,367 of the merger bids and for 3267 of the tender offers.

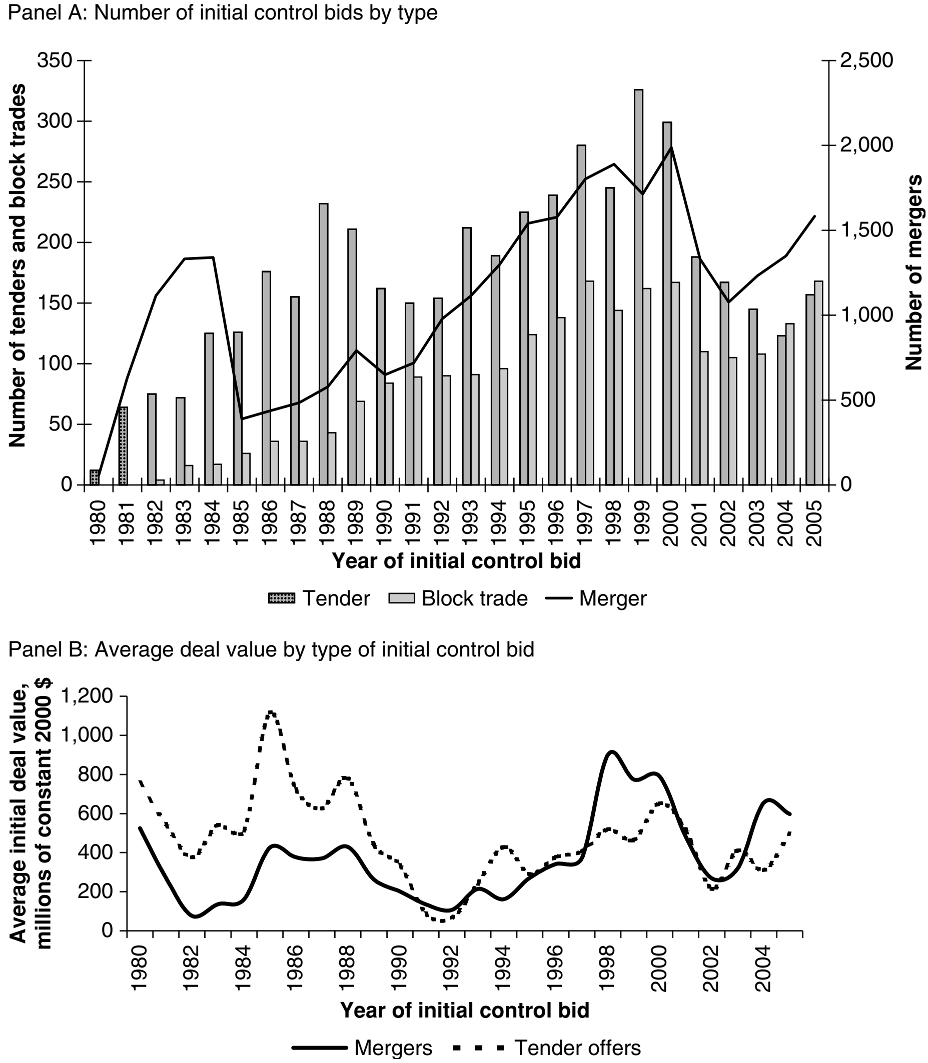



Fig. 3. Initial control bids for U.S. targets, 1980–2005: merger, tender offer, or block trade.

sample bids. Of these, 26% (4798) are classified as all-cash bids, in 37% the method of payment is all-stock, and for 37% the bidder pays with a mix of cash, bidder stock, and/or other (typically debt) securities. In terms of average deal size, mixed and all-stock offers have similar sizes ($538 and $493 million, respectively), while all-cash bids are somewhat lower with $310 million. SDC classifies 590 initial bids as hostile and another 435 bids as unsolicited. All other bids are grouped here as

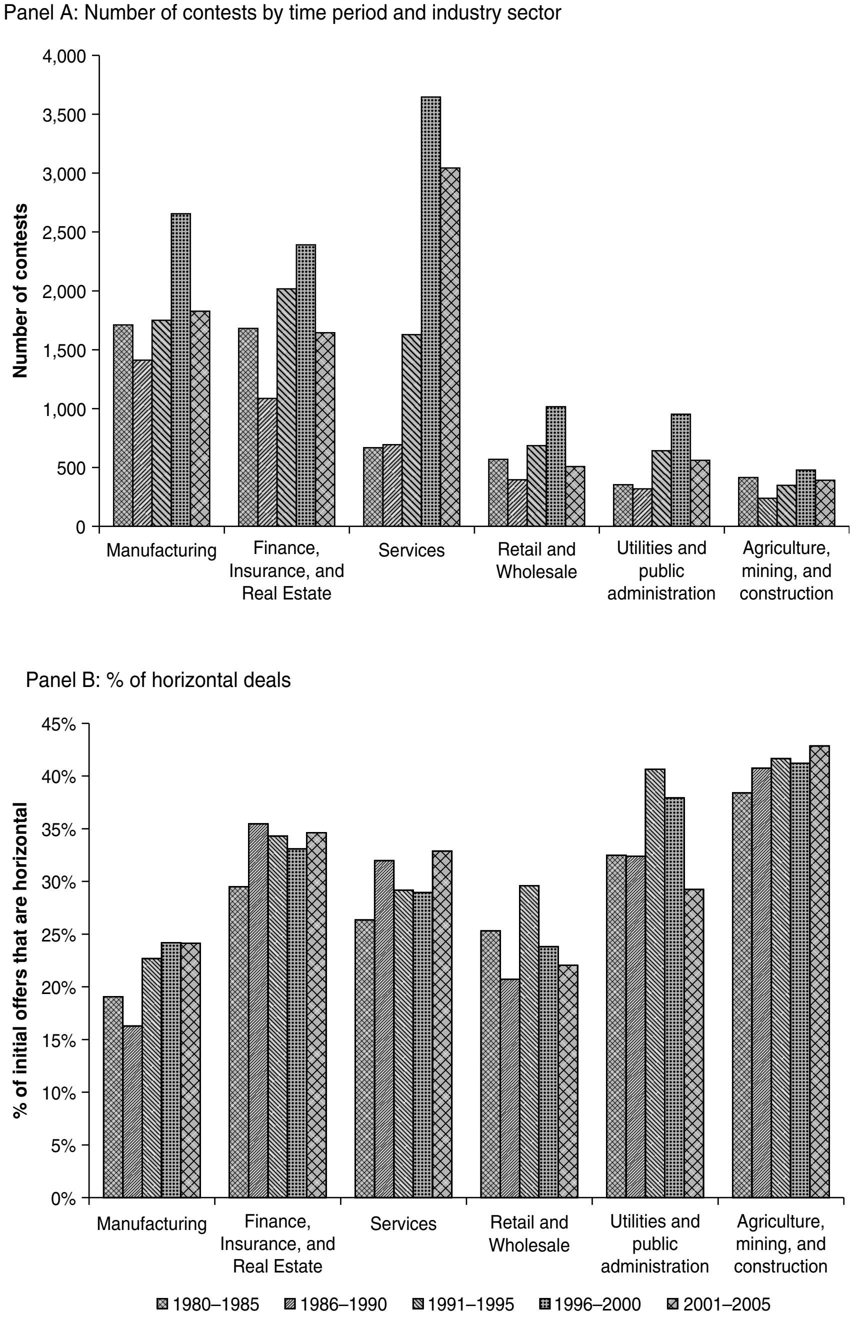
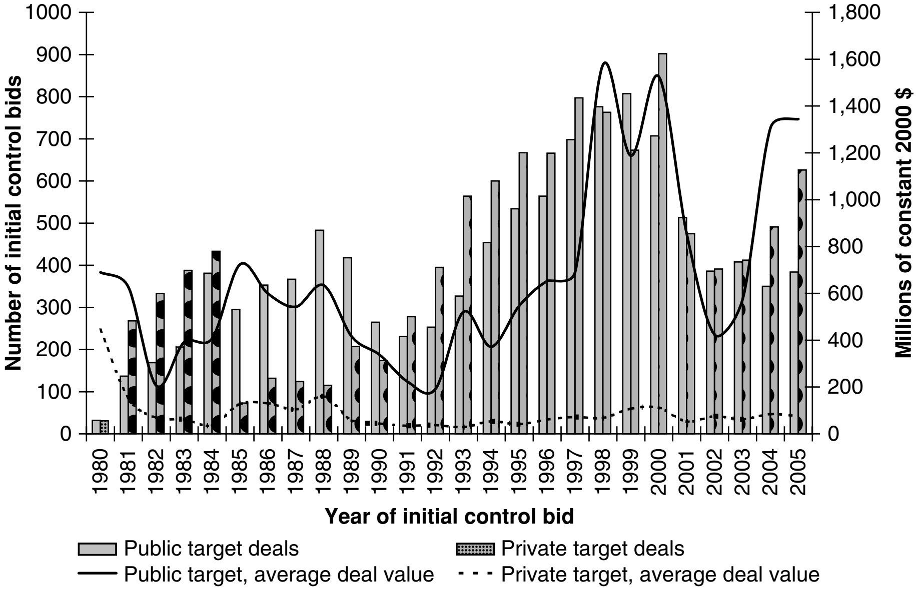



Fig. 4. Initial control bids for U.S. targets, 1980–2005, by 2-digit SIC target industry sector and 4-digit SIC horizontal within sector.

Panel A: Deal values by target public status




Panel B: Deal values by initial bidder public status

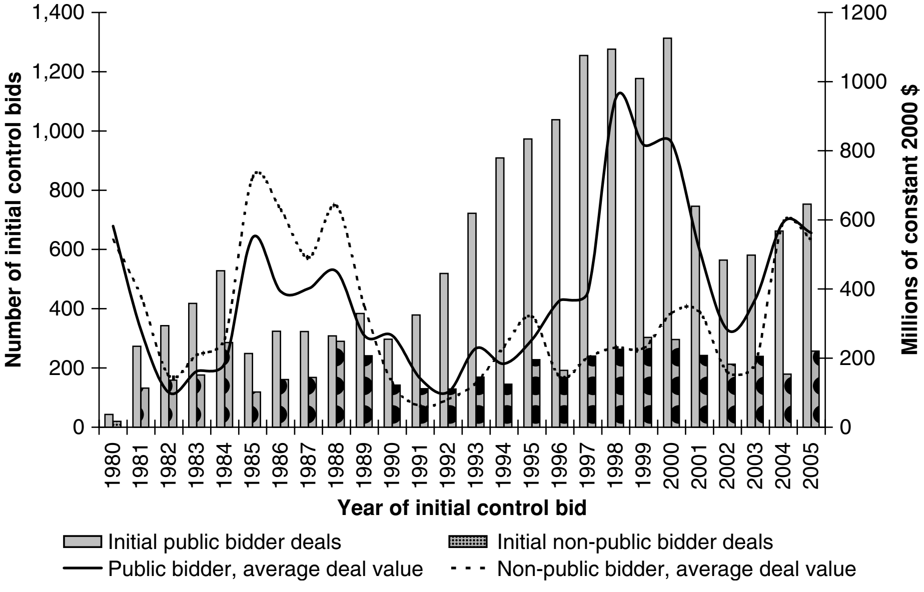



Fig. 5. Initial control bids for U.S. targets, by public status of bidder and target, 1980–2005.

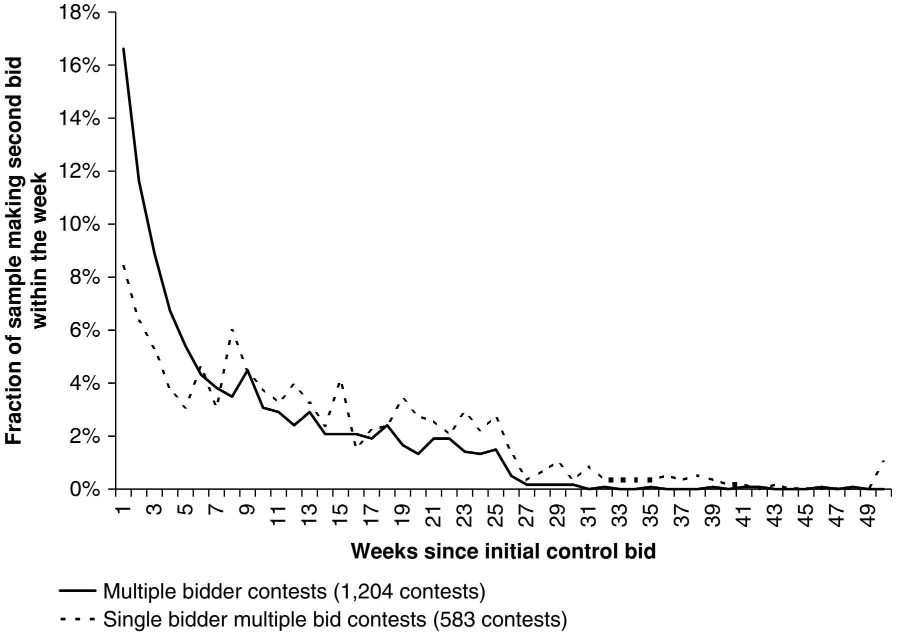



Fig. 6. Weeks from first to second bid in 1,787 contests with multiple bids for U.S. targets, 1980–2005. For the 1,204 contests with multiple bidders, the time from the initial to the second bid averages 5.7 calendar weeks (40 trading days) with a median of 3.7 weeks. For the 583 contests with a single bidder making multiple bids, the average time to the first bid revision is 9 weeks (63 trading days) with a median of 7.6 weeks. Under the 1968 Williams Act, any given tender offer must be open for at least 20 days, and a new bid extends the minimum period accordingly.

friendly—including bids for which SDC does not provide a classification. The hostile bids are by far the largest in terms of size, with an average deal size of $1,612 million versus $609 million for unsolicited offers and $384 million for the average friendly deal.

The last panel in Table 1 shows that 10,452 or 29% of all bids are horizontal (defined as the initial bidder and the target operate in the same four-digit SIC industry). With an average deal value of $562 million, the typical horizontal bid is somewhat larger than the sample average deal size. Figure 4 complements the industry information by listing the total number of bids (Panel A) and the fraction of horizontal initial bids (Panel B) by broad industry sectors and by time period. The industry sectors are Manufacturing; Finance, Insurance, and Real Estate; Services; Retail and Wholesale; Utilities and Public Administration; and Agriculture, Mining and Construction. The two first sectors (Manufacturing, and Finance, Insurance, and Real Estate) are by far the most takeover-intensive sectors in every one of the five five-year subperiods covering the total sample. The only exception is that Services experienced a peak takeover-intensity during 1996–2005. The percentage of the takeover bids that are horizontal tends to be somewhat greater for the least takeover-intensive sectors such as Utilities and Public Administration, and Agriculture, Mining and Construction.

Table 2 and Figure 5 list the sample according to the public status of the target and initial bidder. Of the total sample of 35,727 initial bidders, 67% (24,058) are publicly traded. There are a total of 13,185 publicly traded targets, of which 8,259 receive initial bids from a public bidder. Not surprisingly, these are also the largest deals, with an average of $957 million in constant 2000 dollars (median $116 million). The largest single group is public bidders initiating a contest for a private target, with a total of 15,799 initial bids (44% of the sample). These deal values are typically small, with an average deal value of $66 million (median $16 million). There is also a group of 4,482 private bidder/private targets, comprising 13% of the total database and with an average deal value of $114 million (median $23 million).

Panel A of Figure 5 plots the number and total deal value (in constant 2000 dollars) for public and private target deals over the sample period, while Panel B repeats the plot based on the bidder being either public or private. The number of deals with public targets (Panel A) and with public bidders (Panel B) both increase sharply in the second half of the 1990s. The average deal values when the target is private (Panel A) is small and stable over the entire sample period. Deal values for private bidders (Panel B) are also relatively low, but fluctuate over time in direct proportion to the number of public targets in this group.

Recall that our sampling procedure requires the target but not the bidder to be a U.S. firm. The last panel of Table 2 shows how the bidders split according to nationality. A total of 3,882 or 11% of the total sample of initial bidders are domiciled outside of the United States. Of these, 1,044 bidders are from Canada, 716 from the United Kingdom, and the remaining 2,122 are from a variety of other nations. Interestingly, contests initiated by a foreign bidder are on average large, with a mean of $701 million (median 41 million) when the bidder is from the UK, and $649 million (median $78 million) when the bidder is from the group of "other" countries.

### *2.2.2. Duration, time to second bid, and success rates*

Recall that, starting with the initial offer, we identify the final bid in the contest when 126 trading days have passed without any new offer. Table 3 provides information on the duration of the 25,166 successful contests initiated by a merger or a tender offer. Duration is measured from the date of the initial bid to the effective date of the takeover. The effective date is the day of target shareholder approval of the deal. Given the stringent disclosure rules governing public offer, it is important to separate public from private firms. In the group where both the initial bidder and the target are public, the duration averages 108 trading days (median 96) when the initial bid is a merger offer and 71 days (median 49) when the initial bid is a tender offer. This confirms the conventional view that tender offers are quicker than merger negotiations.

These results are comparable to Betton and Eckbo (2000), who report contest durations for 1,353 tender offer contest, 1971–1990. Of these contests, 62% are single bid with an average duration of 40 trading days (median 29 and highest quartile 52 days). For the multibid contests, the average (median) duration is 70 (51) days. Thus, there is very little change in duration from the 1980s. Also, Table 3 shows clearly that a public target slows down the the takeover process, whether or not the initial bid is a merger or a tender offer. Contests have the shortest duration when both firms are private: 27 days (median 0) for mergers and 67 days (median 41) for tender offers.[19]

Figure 6 shows the distribution of the number of weeks from the initial to the second bid in 1,787 of the 1,891 multibid contests in our sample (Table 1). In general, the expected time to arrival of a second bid depends on the cost to rival bidders of becoming informed of their own valuation of the target, as well as the time it takes to file a formal offer. For some rival bidders, the initial bid may have been largely anticipated based on general industry developments or prior rumors of the target being in play. However, in general, the observed time to the second bid sheds some light on the likelihood that rival bidders have ready access to the resources required to generate takeover gains.

For the 1,204 contests with multiple bidders, the time from the initial to the second bid averages 5.7 calendar weeks (29 trading days), with a median of 3.7 weeks. For the 583 contests with a single bidder making multiple bids, the average time to the first bid revision is 9 weeks (45 trading days) with a median of 7.6 weeks.[20] Thus, the time to the second bid is, on average, shorter when a rival bidder enters than when the second bid represents a bid revision by the initial bidder. These findings are comparable to those in Betton and Eckbo (2000), who report a mean of two weeks (14 trading days) and a highest quartile of 6 week days from the first to the second bid for their sample of 518 multibid tender offer contests.

Several studies provide estimates of the probability that the target will be successfully acquired by *some* bidder (the initial or a rival) following takeover bids. Given our contest focus (Figure 2), we are particularly interested in the probability that the *initial* bidder wins (possible after multiple bid rounds). Betton, Eckbo, and Thorburn (2007) estimate this probability using 7,470 initial merger bids and tender offers. They find that this probability is higher when the initial bidder has a toehold in the target and when the initial bid is all-cash (rather than all-stock or mixed cash-stock), when the bid is a tender offer (rather than merger), and when the bidder is a public company. The probability is also increasing in the pre-bid target stock price runup (the average cumulative target abnormal return from day −42 through day −2 relative to the initial offer day), when the target is traded on the NYSE or the Amex, and when the bidder and target are horizontally related in product markets. Finally, the probability that the initial bidder wins the contest

[19] A contest duration of zero results when the initial offer is announced and accepted on the same day. This is possible in some private deals, provided bidder shareholders do not need to vote on a share issue to pay for the target, and provided the target vote is quick due, say, to high shareownership concentration.

[20] Under the 1968 Williams Act, any given tender offer must be open for at least 20 days, and a new bid extends the minimum period accordingly.

is lower if the target has a poison pill and if the target reaction is hostile. The negative impact of the presence of a poison pill is interesting, for it suggests that pills deter some bids. We return to this issue in Section 3.5.

Finally, Table 1 implies that the probability that all bids fail in a contest is 23% when the contest is initiated by merger and 28% when the initial bid is a tender offer. Thus, as noted by Betton, Eckbo, and Thorburn (2007) as well, merger negotiations are risky for the initial bidder. They are particularly risky when the initial bidder is private. As shown in Table 2, the probability that all bids fail is as high as 40% when the initial bidder is private and the target is public and the bidder approaches with a merger offer.

## *2.3. Merger negotiation v. public tender offer*

### *2.3.1. Merger agreement and deal protection devices*

A merger agreement is the result of negotiations between the bidder and target management teams. The agreement sets out how the bidder will settle any noncash portion of the merger payment. Frequently used contingent payment forms include stock swaps (discussed extensively in Section 3.2), collars, and clawbacks and earnouts.[21] Contingent payment forms allow bidder and target shareholders to share the risk that the target and/or bidder shares are overvalued ex ante. Both parties typically supply fairness opinions as part of the due diligence process.[22]

Whenever the bidder pays the target in the form of bidder stock, the merger agreement specifies the exchange ratio (the number of bidder shares to be exchanged for each target share). A collar provision provides for changes in the exchange ratio should the level of the bidder's stock price change before the effective date of the merger. This helps insulate target stockholders from volatility in the bidder's stock price. Collar bids may have floors and caps (or both), which define a range of bidder stock prices within which the exchange ratio is held fixed, and outside of which the exchange ratio is adjusted up or down. Thus, floors and caps guarantee the target a minimum and maximum payment.

The total payment to target shareholders may also be split between an upfront payment and additional future payments that are contingent upon some observable measure of performance (earnouts, often over a three-year period). This helps close the deal when the bidder is particularly uncertain about the true ability of the target to generate cash flow. It provides target managers with an incentive to remain with the firm over the earnout period, which may be important to the bidder. The downside is that the earnout may distort the incentives of target managers (an emphasis on short-term over longer-term cash flows), and it may induce the new controlling shareholder (the bidder) to manipulate earnings in order to lower the earnout payment. Thus, earnouts are not for everyone.

Merger negotiations protect the negotiating parties against opportunistic behavior while bargaining takes place. Before negotiations start, the parties sign agreements

[21] Officer (2004), Officer (2006), Kohers and Ang (2000), Cain, Denis, and Denis (2005).

[22] Kisgen, Qian, and Song (2006), Makhija and Narayanan (2007), Chen and Sami (2006).

covering confidentiality, standstill, and nonsolicitation. The confidentiality agreement allows the target board to negotiate a sale of the firm without having to publicly disclose the proceedings, and it permits the target to open its books to the bidder. The standstill commits the bidder not to purchase target shares in the market during negotiations, while nonsolicitation ensures that neither the bidder nor the target tries to hire key employees away from the other firm. It is also common for the bidder to obtain tender agreements from target insiders, under which these insiders forsake the right to tender to a rival bidder (Bargeron, 2005).

Delaware case law suggests that a merger agreement must include a fiduciary out clause enabling the target board to agree to a superior proposal if one is forthcoming from a third party.[23] As a result, the target board cannot give its negotiating partner exclusive rights to negotiate a control transfer: it must remain open to other bidders along the way. The resulting potential for bidder competition (after the merger agreement has been signed but before the shareholder vote) has given rise to target termination agreements, starting in the mid-1980s. A termination agreement provides the bidder with compensation in the form of a fixed fee (breakup fee) or an option to purchase target shares or assets at a discount (lockup option) should the target withdraw from the agreement (Burch, 2001; Officer, 2003; Bates and Lemmon, 2003; Boone and Mulherin, 2007a).[24] As discussed in Section 3.3, the value of a target termination agreement may be substantial, and it may affect the initial bidder's optimal toehold strategy.

When merger negotiations close, the bidder seeks SEC approval for any share issue required in the deal, and a merger prospectus is worked out. Writing the prospectus typically takes from 30 to 90 days, so the target shareholder vote is typically scheduled three to six months following the signing of the initial merger proposal.[25] The New York Stock Exchange requires that the shareholders of the bidder firm must also be allowed to vote on the merger if the agreement calls for the bidder to increase the number of shares outstanding by at least 20% in order to pay for the target.

### *2.3.2. Mandatory disclosure and tender offer premiums*

In contrast to the merger process, a public tender offer is relatively quick. A tender offer is an offer made by the bidder directly to target shareholders to purchase target shares. The offer specifies the price per target share, the method of payment (cash, securities, or

[23] *Omnicare Inc. v. NCS Healthcare Inc.,* 818 A.2d 914 (Del. 2003). Delaware law is important as approximately 60% of all publicly traded companies in the United States are incorporated in the state of Delaware. Moreover, decisions in the Delaware Supreme Court tend to set a precedence for court decisions in other states.

[24] The Delaware court views termination fees anywhere in the range of 2 to 5% of the transaction value as reasonable. Termination agreements sometimes allow a reduction in the breakup fee if the target strikes a competing deal within a 30/45-day time frame. There are also cases where the deal includes a bidder termination agreement.

[25] During this waiting period, the bidder also performs a due diligence on key assumptions behind the merger agreement. If the bidder receives 90% of the target shares in a prior tender offer, the bidder can force a merger without calling for a vote among the remaining minority target shareholders (so-called short-form merger).

a mix of the two), whether the offer is restricted to less than 100% of the target shares, conditions for accepting tendered shares (e.g., all or nothing or any or all), and how long the offer is outstanding. The 1968 Williams Act, the main federal law governing public tender offers, requires an orderly auction mechanism: the tender offer must be open for a minimum of 20 business days; competing bid and material bid revisions automatically extend the offer period by 10 days; target shareholders may withdraw all tendered shares for any reason (typically in response to a higher bid) within 15 days; and the bidder must purchase target shares on a pro rata basis among those who tendered their shares.[26]

The 1968 Williams Act also requires public information disclosure.[27] These provisions of the Act were in part a response to perceived takeover abuses in the 1960s, such as "creeping takeovers" and "Saturday night raids" where the bidder quickly gained control of the target shares using all-cash purchases in the market and privately from blockholders. While the stated intention of the Act is to protect target shareholders, a concern for potential bidders is that the mandatory disclosure rules also act to increase the ability of potential rival bidders to compete for the target. As pointed out by Grossman and Hart (1980a) and Jarrell and Bradley (1980), an active market for corporate control presupposes that initial bidders expect to have an advantage over potential rivals when search costs are sunk. Mandatory disclosure rules that increase expected competition among bidders possibly raise offer premiums and therefore deter some bids.[28]

Did the disclosure provisions of the Williams Act raise tender offer premiums? Jarrell and Bradley (1980) examine this issue and find that the average cash tender offer premium increased from 32% to nearly 53% following passage of the Act in 1968. Consistent with higher premium costs, Schipper and Thompson (1983) present evidence indicating that a sample of frequent acquirers earned significantly negative abnormal returns over the months surrounding announcements of the introduction of the Williams Act. Also, Asquith, Bruner, and Mullins (1983), Loderer and Martin (1990), and others report that gains to bidder firms in mergers are on average lower after 1968.

Nathan and O'Keefe (1989) find that the premium increase after introduction of the Williams Act is not restricted to cash tender offers: Cash *mergers* experienced an increase in the average premium from 30% to 67%, while security exchange mergers saw the

[26] Note that, contrary to takeover regulations in many Western countries (Berglof and Burkart, 2003), the Williams Act does not include a mandatory bid rule. A mandatory bid rule requires the bidder to proceed with an offer for 100% of the target shares after acquiring a certain stake in the target (Burkart and Panunzi, 2003). Mandatory bid rules do, however, exist in certain states, including Pennsylvania and Maine. The mandatory bid price varies with jurisdiction but is typically a function of the price(s) the bidder paid for the initial stake.

[27] A tender offer is disclosed through a 14D filing with the SEC. Also, regardless of any plans to acquire the target, an investor purchasing 5% or more of the target shares must file Form 13D with the SEC within a 10-day period. The 13D includes statements concerning the purchaser's intention with respect to the eventual purchase of control. Antifraud provisions were added to the Williams Act in 1970 to back up these disclosure requirements.

[28] However, severe penalties on the release of false (or misleading) information may benefit some bidder firms by making their otherwise voluntarily disclosed information more credible (Eckbo and Langohr, 1989). This positive effect is greater the lower the correlation between rival bidders' private valuations of the target (i.e., the more unique the bidder's contribution to total synergy creation).

average premium increase from 30 to 54%. They also show that the majority of the increase in the average offer premium takes place after 1972. This delay is puzzling and raises the question of whether the premium increase is due to the Williams Act or to some other economic phenomenon.

The Williams Act introduced both disclosure rules and a minimum 20-day offer period. Providing rival bidders with time to respond to the initial bid (the 20-day wait period) is obviously key to increased competition. Thus, studies of the Williams Act cannot isolate the premium impact of the disclosure rules. Specifically, these studies do not answer the fundamental question of whether the introduction of disclosure rules affects offer premiums in an environment where rival bidders already have time to respond.

Eckbo and Langohr (1989) provide evidence on this question using a different institutional setting. In 1970 France introduced mandatory disclosure rules for public tender offers—much like those in the Williams Act. The difference is that France had already established a minimum (four-week) tender offer period much earlier (in 1966). Eckbo and Langohr (1989) find that the average offer premium in successful cash tender offers increased from 34 percent to nearly 61 percent after the 1970 disclosure regulations. Since the minimum tender offer period remained at one month throughout their sample period, this indicates that disclosure requirements *alone* can cause a substantial increase in average offer premiums. Eckbo and Langohr (1989) also study a contemporaneous control sample of privately negotiated controlling-block trades, exempt from the 1970 disclosure regulations. Premiums in these alternative control acquisitions did not increase subsequent to the 1970 regulations.

### *2.3.3. Determinants of the merger choice*

What are some of the determinants of the choice between merger negotiations and a public tender offer? From the bidder's point of view, two immediate advantages of the tender offer process is speed of execution (supported empirically by Table 3) and the fact that it does not require prior approval by—or even prior contact with—target management. Thus, the tender offer is an option for bidders who believe the target will refuse to negotiate ex ante, or should negotiations break down ex-post.[29] Also, many tender offers involve prior contact and even negotiations with the target management (Comment and Jarrell, 1987). Negotiated tender offers may help resolve bargaining issues (e.g., difference of opinions on what constitutes a reasonable bid price), and the arm's length transaction implied by a public tender offer helps protect target managements against charges ex-post that they "sold out" to the bidder.

As discussed in Section 3.5, the target takeover defenses developed in the 1980s, in particular the poison pill, have significantly raised the cost to the bidder of launching a hostile tender offer. This is evidenced by a substantial decline in the frequency of hostile bids over the past 20 years. In today's legal environment, it is likely that virtually

[29] Berkovitch and Khanna (1991), Aktas, deBodt, and Roll (2007), and Betton, Eckbo, and Thorburn (2007) present models in which a tender offer (auction) is an explicit outside option in merger negotiations.

*all* bidders (also those who intend to replace incumbent target management) prefer to approach the target management with a proposal to negotiate. Again, an initially friendly approach preserves the option of making a hostile tender offer down the line. Moreover, a significant benefit of a friendly cooperative approach is that it gives the bidder access to the target books, a crucial factor in pricing.

Systematic empirical evidence on the choice of merger versus tender offer is only beginning to emerge. Kohers, Kohers, and Kohers (2007) study 2,610 completed mergers and 795 successful tender offers for the period 1984–1999. They find that the probability of a tender offer is more likely when the form of payment is all-cash, when the target is defensive, and has high institutional ownership, and when there are multiple bidders. The tender offer form is less likely between two "glamor" companies (i.e., when the bidder and target have low book-to-market (B/M) ratios), and for deals after the 1980s.

Betton, Eckbo, and Thorburn (2008a) study the *initial* bidder's choice between merger and tender offer for 4,618 merger bids and 1,638 tender offers for public U.S. target firms from 1980 through 2002. They separate public bidders (3,119) from private bidders (1,438) and test for differences in their choices. They show that bidder and target B/M values drive the merger choice only when these ratios exceed the median B/M of the respective industry rivals. Public bidders are significantly less likely to select merger over tender offer when the B/M values of the target or of the bidder exceed their respective industry medians. For private bidders, however, this glamor effect does not exist: private bidders are *more* likely to select merger over tender offer when the target's B/M exceeds its industry median (data on private bidders' B/M values are not available). In the 1980s, public bidders were less likely to choose merger, while private bidders were more likely to select this acquisition form. While the target's asset size and target hostility both reduce a public bidder's likelihood of selecting a merger, these factors do not influence the choice of private bidders. Moreover, the greater the concentration of the target's industry, the less likely both public and private bidders are to select merger over tender offer.

## 3. Bidding strategies

### *3.1. Modeling the takeover process*

Before reviewing the empirical evidence on various bidding strategies, it is instructive to briefly characterize the two most common theoretical settings used to model takeover bids. This in turn helps us understand the various empirical hypotheses and their relevance for actual takeover activity.

#### *3.1.1. Free riders and post-offer dilution*

An early workhorse in the theoretical takeover literature is the free-rider model of Grossman and Hart (1980b) and Bradley (1980). They analyze the incentives of dispersed, noncooperative target shareholders to accept a tender offer from a single bidder and the resulting inefficiency of the takeover market. To illustrate, suppose the target's pre-offer (stand-alone) share price is equal to zero and that it is common knowledge that the post-takeover share price will equal $v > 0$. The value-increase $v$ may be thought of as synergy gains resulting from the bidder taking control of the target. The bidder makes a conditional unrestricted bid $b$ for 50% of the target shares (sufficient to transfer control of the target to the bidder).[30] A risk-neutral target shareholder tenders only if the offer price exceeds the expected value of her share if she retains it:

$$\text{Tender if } b \geq \Pr(\text{Success}|i \text{ Retain})v \tag{1}$$

where Pr(.) denotes the probability that the offer succeeds given that the shareholder does not tender.[31]

By inspection of Equation (1), the target shareholder is more willing to tender the lower is the post-takeover value $v$, and the more she believes that retaining reduces the takeover's probability of success. As the number of target shareholder becomes larger, however, the probability that any single shareholder is pivotal for the outcome of the bid becomes arbitrarily small. For such shareholders, the tender criterion in (1) reduces to:

$$\text{Tender if } b \geq v \tag{2}$$

Since the bidder has no economic incentive to make the bid in Equation (2), these shareholders are in effect free-riding on a decision by others to tender. Of course, if all shareholders behave this way, the takeover opportunity never materializes.[32]

Making *every* target shareholder pivotal by a conditional and restricted offer for 100% is unlikely to help. Because the bidder gains control after receiving 50% of the shares, refusing to purchase those shares if she is one share short of 100% is not credible. Also, allowing the bidder to be better informed than target shareholders (about $v$) does not solve the problem. Individual target shareholders now demand an offer price $b \geq E(v|\text{Offer})$ in order to tender, where the right-hand side is the expected valuation of the bidder given that he makes an offer. An offer below this expectation leads target shareholders to infer that $b < v$ and therefore to retain their shares. In this case, there does not exist a rational expectations (perfect Bayesian) equilibrium in which the bidder expects to make a profit from the takeover.[33]

[30] "Conditional" means no shares will be purchased if less than 50% are tendered. "Unrestricted" means any or all tendered shares above 50% will be purchased.

[31] We are ignoring taxes. For example, when $b$ is paid in cash, the offer may trigger a capital gains tax liability.

[32] Just as the free-rider problem can discourage value-increasing bids, value-reducing bids—bids where the post-takeover value of the target is less than its pre-offer value—may be encouraged due to a "pressure-to-tender" problem (Bebchuk, 1985): Conditional on the offer succeeding, tendering may dominate retaining and receiving an even lower value. Thus, paradoxically, there may be "pressure-to-tender" when the bidder is value-reducing. The root cause of this result is, as above, that each target shareholder bases the tendering decision on a comparison between $b$ and $v$, ignoring the pre-takeover value.

[33] Hirshleifer and Titman (1990) prove the existence of a separating equilibrium in which the offer price fully reveals $v$.

There are a number of ways to mitigate the free-rider problem so that the bidder gains on the acquired target shares. Two frequently mentioned mechanisms are post-takeover dilution (Grossman and Hart, 1980b) and pre-takeover toehold acquisition (Shleifer and Vishny, 1986b). Post-takeover dilution reduces the "back-end" value of the takeover and may be enforced through a two-tiered tender offer. The first tier is a bid $b$ while the back end is a minority buyout (enforced by the bidder after acquiring control in the front end) at a lower value $v_d < v$. Alternatively, if fair price rules prevent the minority buyout to take place at a price below the front-end price, the bidder may resort to self-dealing ("asset tunneling"), which is harmful to minority shareholders after the takeover. Examples of such dilution techniques are asset sales at prices below market value, transfer pricing favorable to the majority shareholder, excessive compensation schemes, and so on. These schemes create a wedge between the post-takeover share value to the acquirer and minority shareholders and enable the acquirer to make a profit. Although such transfers may enhance the ex ante efficiency of the takeover market, they are controversial and legally difficult to enforce ex-post. [34]

A firm contemplating making a bid for the target may also decide to purchase target shares—a toehold—in the market at the pre-bid (no-information) target share price. The implications of such toehold acquisitions for optimal bidding are discussed in detail later in this chapter. In the context of the free-rider problem, the important point is that the toehold bidder may gain on the toehold while making zero profits on the shares acquired in the formal takeover bid. Let $\delta$ denote the fraction of the target post-takeover value that may be diluted ex-post, and $\alpha$ the fraction of the target shares held by the bidder prior to the offer, respectively. The bidder makes the conditional unrestricted offer of

$$b^* = (1 - \delta)v \tag{3}$$

which yields a bidder profit of

$$v - (1 - \alpha)b^* = \alpha v + (1 - \alpha)(v - b^*) = \alpha v + (1 - \alpha)\delta v \tag{4}$$

The first term, $\alpha v$, is the gain on the toehold shares, while the second term is the profits on the shares purchased in the takeover. The second term, $(1 - \alpha)\delta v$, shows that dilution is costly for the bidder in that it also reduces the value of the bidder's toehold shares. Thus, the larger the initial stake $\alpha$, the lower the controlling shareholder's incentive to dilute ex-post. In other words, a corporate insider with a larger equity stake is more prone to act in the outside (minority) shareholders' interest (Jensen and Meckling, 1976; Burkart, Gromb, and Panunzi, 1998).

What is the empirical relevance of the free-rider problem in corporate takeovers? The most direct way to evaluate this question is to look at the frequency of (pivotal)

[34] Djankov, La Porta, Lopez-de-Silanes, and Shleifer (2007) survey the opportunities for corporate insiders around the world to dilute minority shareholder value through self-dealings deemed *legal* under a country's corporate laws. Under the European Takeover Directive (article 14), member-states may grant acquirers a squeeze-out right, that is, the right to compel post-takeover minority shareholders to sell their shares after the acquirer has purchased 90% of the target shares.

blockholders in corporate shareownership structures. A large blockholder likely accounts for the possibility that her tendering decision affects the probability that the offer will succeed. In this case, shareholders are willing to tender at a price lower than indicated by expression (1) above (Bagnoli and Lipman, 1988; Holmstrom and Nalebuff, 1992).[35]

The evidence on corporate ownership structures around the world suggests that the existence or one or more large blockholder is the rule rather than the exception.[36] In the United States and elsewhere, small and midsized publicly traded companies typically have one or more large shareholder (defined as a minimum 5% holding).[37] In large-cap firms, individual (or family) blockholdings are less frequent in the United States; however, large blocks held by financial institutions such as pension funds are common for our large firms. As highlighted by Holderness (2006), the evidence challenges the view—originating with Berle and Means (1932)—that U.S. ownership is largely dispersed, and it suggests that free-rider problems in takeovers may be a rarity.[38]

A more indirect way to evaluate the empirical relevance of free-rider problems is to examine characteristics of observed takeover bids. For example, the unequal distribution of takeover gains between target and bidder firms—with most, if not all, of the total gains typically accruing to the target—is often cited in support of the existence of the free-rider problem (Hirshleifer, 1995; Burkart and Panunzi, 2006). However, as discussed in Section 4, there are a number of alternative and plausible reasons for the observed uneven distribution of takeover gains. Moreover, toehold bidding—perhaps the most obvious

[35] In Holmstrom and Nalebuff (1992), there are $N$ target shareholders of equal size, and the bidder needs $K$ of these to tender in order to acquire control. They show that there exists a mixed strategy equilibrium where the takeover succeeds and the bidder makes a positive expected profit. In this equilibrium, individual target shareholders tender with a probability $p = K/N$. Expected profits go to zero when $N$ becomes large.

[36] Following the early international evidence of La Porta, Lopez-de-Silanes, and Shleifer (1999), detailed information on corporate ownership structures has appeared for East Asia (Claessens, Djankov, and Lang, 2000), Western Europe (Faccio and Lang, 2002; Franks, Mayer, and Rossi, 2005), and the United States (Holderness, Kroszner, and Sheehan, 1999; Holderness, 2006; Helwege, Pirinsky, and Stulz, 2007; Dlugosz, Fahlenbrach, Gompers, and Metrick, 2006).

[37] The definition of a block varies in the literature from 5% to 20%. Note that a relatively small block may become pivotal depending on the ownership distribution of the remaining shares. A natural empirical measure of "pivotal" is the Shapley transformation of the block (Shapley, 1953). The Shapley value is the probability that the block will be pivotal, computed using all possible shareholder coalitions (with the block) in which the coalition determines the voting outcome. See, for example, Eckbo and Verma (1994) and Zingales (1994) for applications in corporate finance.

[38] Holderness (2006) studies a random sample of 10% of the firms trading on the NYSE, Amex, and NASDAQ. Large shareholders (which include institutional holdings) on average own 39% of the voting power of the common stock. Moreover, 96% of the firms have at least one 5%+ blockholder, and the average holding of the largest blockholder is 26%. Holderness also reports that 89% of the firms in the S&P 500 Index have large blockholders. Thus, free-rider problems are unlikely. Whether the evidence also challenges the seriousness of the Berle-Means warnings of agency costs associated with delegated management in public firms is, of course, a different issue. It is possible that a large block held by a financial institution (as opposed to an individual investor) carries with it serious agency problems when seen from the point of view of the firm's individual shareholders.

way to mitigate expected free-rider problems—is extremely rare in control-oriented acquisitions (Betton, Eckbo, and Thorburn, 2007).

### 3.1.2. *Auction with single seller*

A second workhorse in the theoretical literature on takeover bidding is the competitive auction. Here, the bidder faces a single seller in the form of a large target shareholder or a target management with sufficient authority to commit to selling in the auction. As noted by Dasgupta and Hansen (2007), auction theory plays an important prescriptive role: to inform a company's board or regulators about the impact of selling processes or rules on shareholder wealth, efficiency, and welfare. They also note that, for such prescriptions to be useful, the auction model must reasonably mimic the actual takeover bidding environment. One important characteristic of any auction is the seller's commitment to stick to the rules of the game. For auction-theoretic results to apply, the seller must be trying to secure the best price for the firm's shareholders by committing to a selling mechanism.[39] As noted earlier, since a publicly traded target's board of directors has a fiduciary obligation to accept the highest offer (provided the board has placed the target "in play"), a takeover is arguably much like an auction even if the target initially negotiates a merger agreement.

The typical assumption is of an open, ascending (English) auction with zero entry and bidding costs, and where the winning bidder pays the second-highest bid.[40] Bidder valuations $v$ (synergies) are private knowledge, but the seller knows the probability distribution function over $v$, $G(v)$. Since bidders tend to have different skill levels in terms of managing the target assets, it is often assumed that the valuations $v$ are uncorrelated across bidders—a "private value." Alternatively, bidder valuations may be correlated—a "common value" environment that requires bidders to shave their bids in anticipation of the "winners curse."[41]

It is also commonly assumed that the bidder's outside option is status quo. That is, the payoff to the bidder is zero when losing the auction. This assumption is effectively relaxed when the bidder has a toehold[42] or a target termination agreement, or when the takeover is a response to changes in industry competition (Morellec and Zhdanov, 2005;

[39] For example, in a first-price auction, in which bidders optimally shave their bids, the seller must be able to commit not to allow further bid revisions by the losing bidder (who, after losing, may want to submit a bid higher than the winning bid).

[40] With zero entry and bidding costs, optimal bid increments are infinitesimal, so the winning bidder pays the second highest price whether or not the auction is defined formally at a first-price or second-price auction.

[41] In a common-value setting, bidders receive private and noisy signals as to the true (common) value of the target. Bidding the full value of the signal would cause the bidder with the largest positive signal error to win and overpay (the "curse"). Optimal bidding takes this possibility into account by reducing the bid to the point where the expected value of the bid *conditional on winning* is nonnegative. Thus, testing for the presence of a winner's curse is equivalent to testing the hypothesis that bidders are irrational (cannot compute). See Boone and Mulherin (2008) for same evidence inconsistent with this hypothesis. In a private value setting, bidders know their true valuations and thus do not face a winner's curse.

[42] Burkart (1995), Bulow, Huang, and Klemperer (1999), Betton, Eckbo, and Thorburn (2007).

Akdogu, 2007b; Molnar, 2008). The toehold provides a positive payoff when the toehold bidder loses to a rival (who purchases the toehold). A termination contract also pays off when the bidder loses and no other bidder wins (the target remains independent). Also, a worsening of the competitive industry equilibrium can place the unsuccessful bidder at a competitive disadvantage vis-à-vis the winner.

### 3.2. *The payment method choice*

As discussed earlier (Table 1), the payment method in takeovers includes all-stock payment, various debt securities, mixes of securities and cash, and all-cash payment.[43] As Table 1 shows for the total sample, 26% of the initial bidders use the all-cash method while the groups of all-stock and mixed offers each cover 37% of the initial bids. Figure 7 plots the fraction of all initial bids that are in the form of each of these three payment methods over the 1980–2005 period. Use of the various payment methods clearly differs between merger bids (Panel A) and tender offers (Panel B): the majority of tender offers use all-cash or a mix of cash and stock, while the majority of merger bids are in the form of all-stock (with the exception of the 1980–1985 period when 90% of the initial merger bids offered a mix of cash and securities).

Notice that in the two subperiods 1990–1995 and 1996–2000, the percentage of all-stock offers in initial merger bids was approximately 55% in *both* periods. This means that (1) nearly half of the initial merger bids in the 1990s used some cash as payment, and (2) the percentage of all-stock merger bids remained unaffected by the significant runup in overall market valuations in the 1996–2000 period.

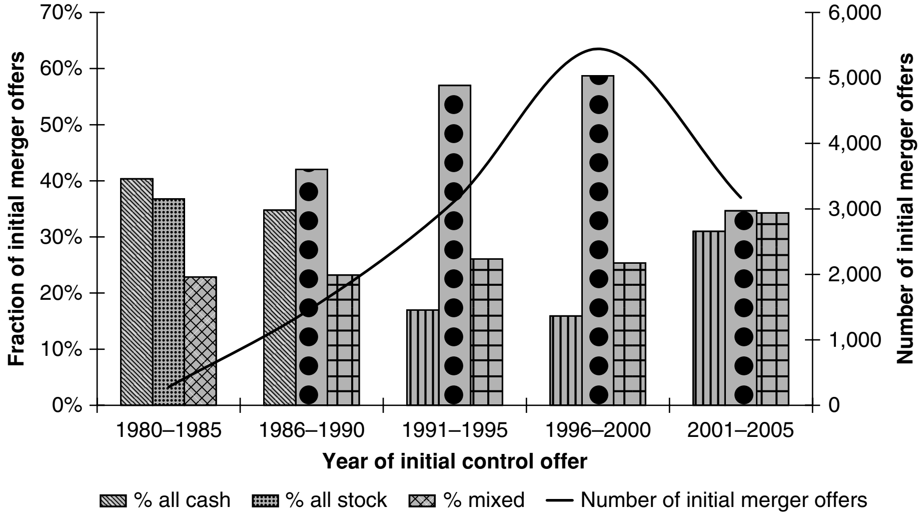
Table 4 summarizes a number of economic hypotheses and related empirical evidence concerning the choice of payment method. The associated empirical evidence is a combination of determinants of the probability of a specific payment method choice (e.g., all-cash versus all-stock), and announcement-induced abnormal stock returns as a function of the payment method. The hypotheses deal with tax effects, deal financing costs under asymmetric information, agency and corporate control motives, and behavioral arguments. These hypotheses are not necessarily mutually exclusive, so a given payment choice may reflect elements of several theories.

#### 3.2.1. *Taxes*

The U.S. Internal Revenue Code (IRC) requires target shareholders to immediately pay capital gains taxes in an all-cash purchase. If the merger qualifies as a tax-free reorganization under Section 368 of IRC, for example by using all-stock as method of payment, target shareholder capital gains taxes are deferred until the shares received in the deal are sold. Mixed cash-stock offers are treated as either all-cash bids or the stock part is treated as an all-stock bid depending on the cash portion and other characteristics of

[43] The cash amount is typically financed using accumulated retained earnings (financial slack) or debt issues prior to the takeover.

Panel A: Distribution of mergers by time period and method of payment




Panel B: Distribution of tender offers by time period by method of payment

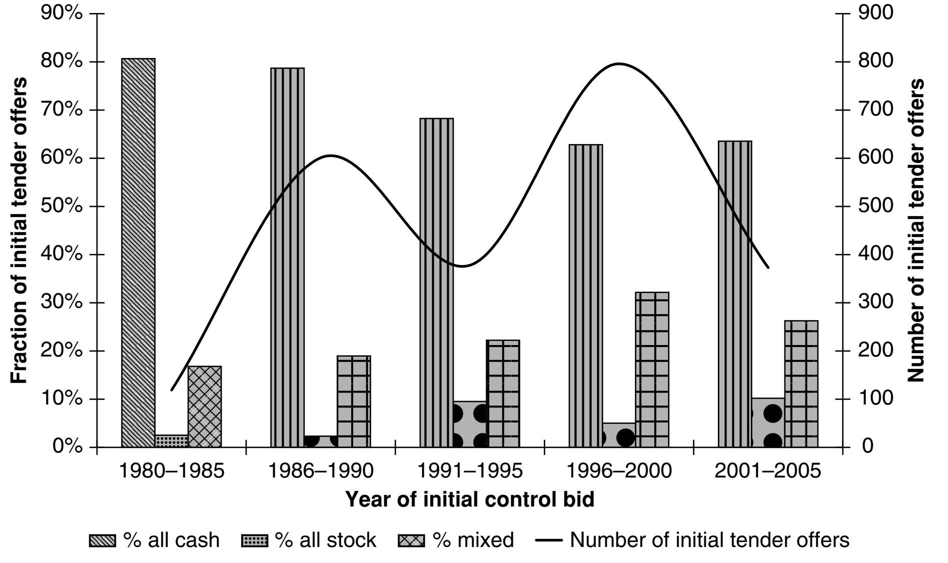



Fig. 7. The initial control bidder's use of all-cash, all-stock, and mixed cash-stock as method of payment. Total sample of 13,503 merger bids and 2,678 tender offers with SDC information on payment method. U.S. targets, 1980–2005.

Table 4

Selected hypotheses and U.S. evidence concerning the choice of payment method in takeovers.

| Theories | Hypotheses | Evidence |
|---|---|---|
| **A. Taxes and the payment method** | | |
| U.S. Internal Revenue Code, Section 368 governing statutory merger.<br><br>Gilson, Scholes, and Wolfson (1988) | H1: *Cash deals may be relatively costly as the implied capital gains tax penalty forces higher target premiums*.<br><br>In a cash-for-stock deal, target shareholders pay capital gains tax immediately (the deal is taxable if target shareholders receive less than fifty percent of the deal in bidder stock). The buyer steps up the tax basis with the takeover premium. In a Stock deal, however, target capital gains taxes are deferred until shares received are sold. No step-up of tax basis for buyer. Buyer in stock deal may make a Section 338 election to be treated as cash deal (Bruner, 2004). | Carleton, Guilkey, Harris, and Stewart (1983): Probability of stock offer increases in bidder's market-to-book ratio.<br><br>Huang and Walkling (1987), Hayn (1989): Target announcement returns in U.S. deals higher for taxable than tax-deferred transactions.<br><br>Franks, Harris, and Mayer (1988): Reach similar conclusion for control-oriented takeovers in U.K.. However, the all-cash premium effect is present also *before* the introduction on capital gains taxes.<br><br>Eckbo and Langohr (1989): Find higher target premiums in all-cash tenders offers for control as well as for minority buyouts in France.<br><br>Brown and Ryngaert (1991): Find empirical support for their proposition that stock should not be used by bidders selecting taxable offers. |
| **B. The payment method choice motivated by asymmetric information** | | |
| Myers and Majluf (1984) | H2: *One-sided information asymmetry: Investor concern with adverse selection produces a negative market reaction to the news of a stock deal.* | Travlos (1987); Asquith, Bruner, and Mullins (1987); Servaes (1991); Brown and Ryngaert (1991); Smith and Kim (1994); Martin (1996); Emery and Switzer (1999); Heron and Lie (2004); Schlingemann (2004) and many others show that bidder announcement-induced abnormal stock returns are on average negative in all-stock offers for public targets.<br><br>However, bidder announcement returns are non-negative in all-stock offers for private targets (Chang, 1998; Fuller, Netter, and Stegemoller, 2002; Moeller, Schlingemannm, and Stulz, 2004; Bradley and Sundaram, 2006; Officer, Poulsen, and Stegemoller, 2007). |

*(Continued)*

Table 4 *(Continued)*

| Theories | Hypotheses | Evidence |
|---|---|---|
| Hansen (1987)<br><br>Fishman (1989) | H3: *Two-sided information asymmetry: Paying with securities induce targets to make more efficient accept/reject decisions than with cash. Stock offers are less likely when (i) the bidder has a relatively large total equity size, and (ii) when the target undervalues the bidder's shares.*<br><br>The value of a stock offer is contingent on the true values of both the bidder and the target. A cash offer that undervalues the target will be rejected, while an equivalent stock offer may be accepted because the stock offer will rise in value ex-post. This ex-post price effect is smaller the smaller the size of bidder's total equity relative to the target's.<br><br>The more the target undervalues the bidder's stock, the more costly a given stock offer, and the more likely the bidder is to use cash. | Hansen (1987): Probability of stock offer increases in bidder's asset size as well as in the size of its liabilities.<br><br>Chemmanur and Paeglis (2003): Probability of stock offer increases in a measure of market mispricing of bidder shares and falls as the dispersion of analyst forecast of bidder earnings increases.<br><br>Betton and Eckbo (2000): Probability that the target accepts the initial bid in tender offer contests is lower for stock offers than for cash bids.<br><br>Travlos (1987): Bidder's announcement-induced abnormal stock returns lower for stock offers than for cash bids. |
| Eckbo, Giammarino, and Heinkel (1990)<br><br>Berkovitch and Narayanan (1990) | H4: *Two-sided information asymmetry where bidders in equilibrium choose a mix of cash and stock. There exists a fully revealing separating equilibrium in which the greater the proportion of the deal paid in cash, the greater the true value of the bidder.*<br><br>In Eckbo, Giammarino, and Heinkel (1990), target adverse selection pushes the bidder towards using stock as payment method, while target undervaluation of bidder shares pushes the bidder towards cash. The market uses the proportion of the deal paid in cash to separate low-value from high-value bidders. In equilibrium, bidder announcement-induced abnormal stock returns are an increasing and convex function of the cash portion of the deal.<br><br>In Berkovitch and Narayanan (1990), the bidder's choice of cash-stock mix affects target returns as well. Greater potential bidder competition raises the optimal amount of cash, and with actual competition all but the lowest type make all-cash offers. | Eckbo, Giammarino, and Heinkel (1990); Eckbo, and Thorburn (2000): The average announcement-induced abnormal stock returns to bidders are highest for all-cash deals, lowest for all-stock deals, with mixed cash-stock deals in between.<br><br>Eckbo, Giammarino, and Heinkel (1990): In cross-sectional regressions, bidder announcement-induced abnormal stock returns are increasing in the cash portion of the deal as predicted. However, the data rejects convexity.<br><br>Betton, Eckbo, and Thorburn (2008c): Shows frequent use of mixed cash-stock offers in tender offers (see also Figure 7). Moreover, there is evidence that multiple bids raise the use of cash, however, the amount of stock used in competitive contests remains significant. |

*(Continued)*

Table 4 *(Continued)*

| Theories | Hypotheses | Evidence |
|---|---|---|
| **C. Capital structure and corporate control motives for the payment method choice** | | |
| Ross (1977)<br><br>Jensen (1986)<br><br>Harris and Raviv (1988)<br><br>Stulz (1988) | H5: *The payment method is selected as part of a broader capital structure choice. Moreover, some bidder managements select (possibly debt financed) cash over stock as payment method in order to avoid diluting their private benefits of control in the merged firm.*<br><br>In Ross (1977) increased leverage raises expected managerial-specific bankruptcy costs. In Jensen (1986), paying with cash drains free-cash flow and reduces agency costs. In Harris and Raviv (1988); Stulz (1988), managers act to protect private benefits of control. | Capital structure: The cash portion of the bid must be financed internally or by a previous security issue. Schlingemann (2004); Toffanin (2005) find a link between the market reaction to takeover announcements and financing decision in the previous year.<br><br>Yook (2003) find greater bidder gains in all-cash offers when the takeover causes downgrading of the merged firm's debt (due to increased leverage). The results are consistent with agency costs of free cash flow (Jensen, 1986).<br><br>Control: Amihud, Lev, and Travlos (1990); Martin (1996); Ghosh and Ruland (1998) find that bidder management shareholdings in the U.S. have negative effect on stock financing.<br><br>Studying European mergers, Faccio and Masulis (2005) find that corporate control incentives to choose cash are particularly strong when in bidder firms with relatively concentrated shareownership structures. Martynova and Renneboog (2006) finds a link between the quality of a country's corporate governance system and the market reaction to stock as payment form. |
| **D. Behavioral motives for the payment method choice** | | |
| Shleifer and Vishny (2003)<br><br>Rhodes-Kropf and Viswanathan (2004) | H6: *Bidders are able to sell overpriced stock to less overpriced targets*<br><br>Bidders attempt to cash in on a temporary market overvaluation of their stocks. In Shleifer and Vishny (2003) they succeed because targets have "short time horizon". In Rhodes-Kropf and Viswanathan (2004) they succeed because targets accept more bids from overvalued bidders during market valuation peaks because they tend overestimate synergies during these periods. | The propensity of all-stock offers increases with M/B ratios (Rhodes-Kropf, Robinson, and Viswanathan, 2005; Aug and Cheng, 2006; Dong, Hirshleifer, Richardson, and Teoh, 2006). This supports the behavioral argument provided M/B is a fundamental driver of takeovers.<br><br>Harford (2005): A macroeconomic measure of capital liquidity (interest rate spreads) drives merger activity and drives out M/B as a predictor of merger activity. This is is inconsistent with the behavioral argument. |

*(Continued)*

Table 4 *(Continued)*

| Theories | Hypotheses | Evidence |
|---|---|---|
| | | Betton, Eckbo, and Thorburn (2008c): There are nearly as many mixed cash-stock offers as all-stock offers, also in the recent period of high market valuations and peak merger activity (1996–2000). Mixed offers are an enigma in the model of Shleifer and Vishny (2003). The fact that the substantial market runup prior to year 2000 did not induce a greater use of all-stock offers as a proportion of all merger bids is inconsistent with the behavioral argument. |

the deal. There is a carry-over of the tax basis in the target to the acquiring company, unless a 338 election is made. Under a 338 election, there is a step-up of the tax-basis of the target assets to the price paid in the takeover (Bruner, 2004). Such elections imply a capital gains tax in the target, and are used only in rare circumstances such as when there are substantial target net operating losses (NOLs) due to expire, or when the target is a subsidiary.

Given these differences in the tax treatment, there is little doubt that taxes play an important role in the bidder's choice of payment method. The more difficult empirical issue is whether the bidder in all-cash offers must pay target shareholders a compensation up front both for the realization of a potential capital gains tax penalty and for the value of the target's unused tax benefits. This depends, of course, on the relative bargaining power of the bidder and the target and is therefore transaction specific. For example, targets that have low-cost substitute ways of capitalizing on unused tax benefits will force bidders to pay for these in the deal (Gilson, Scholes, and Wolfson, 1988).

Hypothesis H1 in Table 4 holds that targets will receive higher offer premiums in all-cash bids than in all-stock offers, where the difference is compensation for the capital gains tax penalty inherent in the cash bid. Early studies that classify takeover premiums according to the payment method include Huang and Walkling (1987) and Hayn (1989) on U.S. data, and Franks, Harris, and Mayer (1988) and Eckbo and Langohr (1989) on acquisitions in the UK and France, respectively. This evidence shows that takeover premiums are indeed significantly greater in all-cash deals than in all-stock offers, which is consistent with H1. Also, Brown and Ryngaert (1991) find empirical support for their hypothesis that stocks are less likely to be found in taxable offers (offers where less than 50% of the offer is to be paid in bidder stock).

On the other hand, Franks, Harris, and Mayer (1988) show that takeover premiums in the UK were greater in cash deals even *before* the introduction of capital gains taxes. Moreover, Eckbo and Langohr (1989) argue that for a tax compensation to induce tendering behavior, it must be included in the value of the option to tender (as opposed to keeping) the target shares. They approximate this option value with the difference between the offer price and the expected post-offer target share price, and they find that this difference is indistinguishable across all-stock and all-cash offers. They also show that the larger total premium in all-cash offers carries over to minority buyouts that convey few if any bidder tax benefits (as the two firms are already consolidated for accounting purposes). This evidence does not support the view that the larger takeover premiums observed in all-cash deals are driven by the tax hypothesis H1.

### *3.2.2. Information asymmetries*

Hypotheses H2–H4 in Panel B of Table 4 suggest that the payment method choice may be economically important—and give rise to premium effects—even in the absence of taxes. When the bidder and target are asymmetrically informed about the true value of their respective shares, the payment method may cause information revelation and affect both the division of synergy gains and the probability that the offer is successful. Hypothesis H2 is motivated by the adverse selection argument of Myers and Majluf (1984) and the associated financing "pecking order" suggested by Myers (1984). H2 focuses on the implication for the market's reaction to the all-stock vs. all-cash announcement: Equity issues to relatively uninformed target shareholders may cause a negative market reaction as investors hedge against the possibility that the bidder's stock is overpriced.

There is substantial empirical evidence that seasoned equity offerings (SEO) are on average met with a negative market reaction (approximately $-3\%$)—even when the SEOs are fully underwritten by reputable investment banks. This is consistent with the hypothesis that outside investors are somewhat nervous that the typical equity issue may be overpriced — despite the substantial due diligence effort and reputational risk exposure of underwriters. The evidence on takeovers indicates that all-equity acquisition announcements also tend to cause a statistically significant (approximately) 1% price bidder price drop when the target is a public company.[44] However, bidder announcement returns are nonnegative (or even positive) in all-stock offers for *private* targets.[45]

Hansen (1987), Fishman (1989), and Eckbo, Giammarino, and Heinkel (1990) provide theoretical analyses that also incorporate adverse selection but where the bidder's choice of payment method is modeled explicitly. An important insight of Hansen (1987) is that ex-post means of payments such as stock can increase the seller's revenue beyond what cash payments can do.[46] This point is easily illustrated using our second-price, independent private value auction with two bidders ($v_1 > v_2$). If bidder 1 (B1) wins with

[44] Travlos (1987), Asquith, Bruner, and Mullins (1987), Servaes (1991), Brown and Ryngaert (1991), Martin (1996), Emery and Switzer (1999), Heron and Lie (2004), and Schlingemann (2004). Because the level of communication between bidder and target management teams in merger negotiations is greater than that between underwriters and the market in SEOs, the potential for adverse selection is also smaller, thus the smaller price drop in all-equity bids than in SEOs.

[45] Chang (1998), Fuller, Netter, and Stegemoller (2002), Moeller, Schlingemann, and Stulz (2004), Bradley and Sundaram (2006), Officer, Poulsen, and Stegemoller (2007). Faccio, McConnell, and Stolin (2006) find a similar positive bidder announcement effect of all-stock offers in Europe.

an all-cash offer, the target receives $v_2$ (the second price). Alternatively, with all-stock as the payment method, the bidder offers the target the ownership fraction $z_i$ in the merged firm. Suppose B1 and B2 have the same stand-alone value $v$. The optimal bid is the fraction $z_i$, which satisfies

$$(v + v_i)(1 - z_i) = v \tag{5}$$

or $z_i = v_i/(v + v_i)$. This leaves each bidder with a post-acquisition value equal to the pre-acquisition (stand-alone) value. If B1 wins, the target receives

$$z_2(v + v_1) = \frac{v + v_1}{v + v_2} v_2 > v_2 \tag{6}$$

since $v_1 \geq v_2$. In other words, the all-stock offer extracts a higher revenue from the winning bidder than does the all-cash bid, resulting in more efficient sell/don't sell decisions by the target.[47]

Another insight is that all-stock payment may increase the expected deal value for the bidder if there is little or no uncertainty concerning the true bidder value. Consider a single bidder B who has all the bargaining power. Denote B's with-synergy value as $v_B \equiv v + v_i$. Assume that $v_B$ is known to everyone and that B only knows the probability distribution over the true target value, $v_T \in [\underline{v}_T, \bar{v}_T]$, where $\underline{v}_T < \bar{v}_T$. Moreover, suppose B's strategy is to ensure bid success.[48] The all-cash offer is therefore $c = \bar{v}_T$. This means that B expects to overpay for the target by the amount $\bar{v}_T - E(v_T|\, accept)$, where the expectation is conditional on the target accepting the bid. The corresponding all-stock offer solves $z(v_B + \bar{v}_T) = \bar{v}_T$, or $z = \bar{v}_T/(v_B + \bar{v}_T)$. The expected overpayment cost is now

$$z[v_B + E(v_T|\, accept)] - E(v_T|\, accept) = \frac{v_B}{v_B + \bar{v}_T}[\bar{v}_T - E(v_T|\, accept)] \tag{7}$$

Since $v_B/(v_B + \bar{v}_T) < 1$, the expected overpayment cost of securities is less than that of cash, reflecting the contingent nature of stock as payment form (payment in shares causes the target to share in the overpayment ex-post). Cash, on the other hand, precommits the bidder to a target value ex ante.

If we also allow $v_B$ to be private information (two-sided information asymmetry), then the above preference for a stock offer is reversed provided the bidder shares are sufficiently undervalued by the target. With two-sided information asymmetry, let $\hat{v}_B$ denote target beliefs about bidder value. In this case, the all-stock offer which guarantees

[46] See also Hansen (1985) and DeMarzo, Kremer, and Skrzypacz (2005), and Dasgupta and Hansen (2007) for a review.

[47] In Fishman (1989), the alternative to cash is a debt instrument secured in the *target's* asset. This also eliminates target uncertainty about the true value of the bidder's payment for all-security offers and leads to efficient target accept/reject decisions.

[48] This bid strategy is maintained in the model of Eckbo, Giammarino, and Heinkel (1990). In Hansen (1987), high-value bidders separate themselves by lowering their all-stock offers $z$, which is costly as it reduces the probability that the target will accept. The signaling cost is the reduction in the bidder's expected synergy gains from a reduction in $z$.

success solves $z(\hat{v}_B + \bar{v}_T) = \bar{v}_T$, and the difference between the expected overpayment cost of an all-stock and an all-cash offer becomes

$$\bar{v}_T \frac{(v_B - \hat{v}_B) - (\bar{v}_T - E(v_T|\,accept))}{v_B + E(v_T|\,accept)} \tag{8}$$

which is positive or negative depending on whether the target undervalues ($v_B - \hat{v}_B > 0$) or overvalues ($v_B - \hat{v}_B < 0$) the bidder shares, respectively. Consistent with this, Chemmanur and Paeglis (2003) find that the probability of a stock offer falls when measures of bidder share underpricing increase.

As discussed earlier (see Figure 7), mixed cash-stock offers are pervasive across the entire sample period. Eckbo, Giammarino, and Heinkel (1990) and Berkovitch and Narayanan (1990) model equilibrium mixed offers.[49] In the separating equilibrium of Eckbo, Giammarino, and Heinkel (1990), bidder types are separated by the fraction of the total target payment that is paid in cash. Consistent with a separating equilibrium, Eckbo, Giammarino, and Heinkel (1990) and Eckbo and Thorburn (2000) find that abnormal announcement returns are, on average, highest in all-cash offers and lowest in all-stock deals, with mixed offers in between.[50]

Eckbo, Giammarino, and Heinkel (1990) present cross-sectional regressions tests of their signaling model. To illustrate, let $\gamma_j$ denote the announcement-induced bidder abnormal return. The separating equilibrium implies that

$$\gamma_j = h_j\left(\frac{c_j}{v_T}\right), \quad h'_j, h''_j > 0, \tag{9}$$

where $c_j$ is the cash payment, $v_T$ is the average pre-bid target value, and the superscripts $h'_j$ and $h''_j$ denote first and second derivatives, respectively. That is, in the separating equilibrium, the market reaction to the takeover announcement is an increasing and convex function of the cash portion of the deal. The cross-sectional regression tests confirm the "increasing" part, but fails to identify a significant second derivative (convexity). Additional empirical tests are required to sort out why convexity fails.

### 3.2.3. *Capital structure and control motives*

Under hypothesis H5 in Panel C of Table 4, the payment method is selected as part of a broader capital structure choice. Moreover, some bidder managements select cash over stock to avoid diluting private benefits of control. Attempts to link the payment method choice to financing sources for the cash portion of the bid are only starting to emerge. For example, Yook (2003) finds greater bidder gains in all-cash offers when the takeover causes downgrading of the merged firm's debt (due to increased leverage). He interprets this as consistent with the free-cash flow argument of Jensen (1986).

49 In Hansen (1987) and Fishman (1989), bidders select between all-stock and all-cash offers but do not mix the two.

50 These two studies use mergers in Canada where offering less than 50% of the deal in cash does not trigger capital gains taxes. In the United States, the tax code confounds the analysis as it in of itself discourages mixed offers where the cash portion exceeds 50% (Brown and Ryngaert, 1991).

Schlingemann (2004) and Toffanin (2005) examine whether the market reaction to the payment method choice is a function of the type of cash financing. While the market is aware of any pre-bid public security issues, the acquisition bid announcement possibly resolves uncertainty regarding use of the issue proceeds. If this resolution is economically important, the source of financing for the cash portion of the bid will affect the market reaction to the takeover attempt. The empirical results indicate a prior-cash-financing-source component in acquisition announcement returns.

Schlingemann (2004) reports that, after controlling for the form of payment, financing decisions during the year before a takeover play an important role in explaining the cross section of bidder gains. Bidder announcement period abnormal returns are positively and significantly related to the amount of ex-ante equity financing. This relation is particularly strong for high $q$ firms. He further reports a negative and significant relation between bidder gains and free cash flow. This relation is particularly strong for firms classified as having poor investment opportunities. The amount of debt financing before a takeover announcement is not significantly related to bidder gains. Interestingly. Toffanin (2005) finds that the well-known positive market reaction to all-cash bids requires the cash to have been financed either using internal funds (retained earnings) or borrowing. All-cash acquisitions financed by a prior equity issue earn zero or negative abnormal returns.

Early theories incorporating private benefits of control in the contexts of takeovers and capital structure choice are Stulz (1988) and Harris and Raviv (1988). In our context, an all-cash offer preserves the bidder's control position, while an all-stock offer may significantly dilute this position (e.g., a merger of equals). The potential for control dilution may therefore drive the use of cash. Several empirical papers examine the payment method choice from this angle. For example, Amihud, Lev, and Travlos (1990), Martin (1996), and Ghosh and Ruland (1998) all find that bidder management shareholdings in the United States have negative effects on stock financing. Similarly, studying European mergers, Faccio and Masulis (2005) find that corporate control incentives to choose cash are particularly strong in bidder firms with relatively concentrated shareownership structures. Overall, corporate control motives are likely to play a role in some all-cash mergers. Martynova and Renneboog (2006), who also examine acquisitions in Europe, find a link between the quality of a country's corporate governance system and the market reaction to stock as payment form. All-stock offers are more likely in countries with greater levels of shareholder rights protection.

### *3.2.4. Behavioral arguments for all-stock*

The hypothesis here is that bidders are able to sell overpriced stock to less overpriced targets (H6). We discussed this hypothesis in Section 2.1 on merger waves and so will provide only a summary here. In the model of Shleifer and Vishny (2003), bidders succeed in selling overpriced stock to target managers with a short time horizon. In Rhodes-Kropf and Viswanathan (2004), bidders succeed as targets (rationally) accept more bids from overvalued bidders during market valuation peaks because they tend to overestimate synergies during these periods. Empirically, the propensity to select all-stock offers increases with M/B ratios. If one views the M/B ratio as a proxy for stock overvaluation, then this empirical regularity supports the behavioral argument for all-stock selections.[51] On the other hand, Harford (2005) finds that a macroeconomic measure of capital liquidity (interest rate spreads) drives merger activity and drives out M/B as a predictor of merger activity. This finding reduces the likelihood that market overvaluation systematically drives the bidder's selection of all-stock as the payment method.

Earlier we reported that there are nearly as many mixed cash-stock offers as all-stock offers, even in the recent period of high market valuations and peak merger activity (1996–2000). Because mixing cash and stock increases the ability of undervalued bidders to separate out from the pool of overvalued bidders (Eckbo, Giammarino, and Heinkel, 1990), the substantial presence of mixed offers undermines the pooling equilibrium of Shleifer and Vishny (2003). Also, our finding in Figure 7 that the substantial market runup prior to year 2000 did not induce greater use of all-stock offers as a proportion of all merger bids further undermines the behavioral argument. In sum, while some bidders undoubtedly get away with selling overpriced stock to their targets, additional research is needed to systematically contrast behavioral to rational theories of the payment method choice in takeovers.

### *3.3. Toehold bidding*

In this section, we first discuss optimal bids when the initial bidder has a toehold and has also negotiated a termination agreement. We then review the empirical evidence on toehold bidding.

#### *3.3.1. Optimal bids*

We use a standard auction setting with two risk-neutral bidders. The bidders have private valuations that are independent and identical distributed (i.i.d.) with distribution and density functions $G(v)$ and $g(v)$, respectively. The initial bidder (B1) has toehold $\alpha \in [0,\ 0.5)$ acquired at the normalized pre-takeover target share price of zero. B1 has negotiated a merger agreement with the target management that includes a termination fee $t \in (0,\ v)$. A rival bidder (B2) challenges the agreement and forces an open auction. The termination fee is paid by B2 if B2 wins, or by the target if neither B1 nor B2 wins (the target remains independent). The no-bidder-wins outcome occurs with an exogenous probability $\theta$.[52]

Since the termination fee represents a claim of $t$ on the target, the fee reduces B2's private valuation to $v_2 - t$. B2's optimal bid is therefore $b_2^* = v_2 - t$: bidding less risks foregoing a profitable takeover, while bidding more risks overpaying for the target. Given

[51] Rhodes-Kropf, Robinson, and Viswanathan (2005), Ang and Cheng (2006), Dong, Hirshleifer, Richardson, and Teoh (2006).

[52] The probability $\theta$ captures exogenous factors that may derail merger negotiations or cause all bidders to abandon a takeover auction. For example, the market may revise upwards its estimate of the target's stand-alone value during the contest, causing the takeover to be unprofitable for both B1 and B2. Betton, Eckbo, and Thorburn (2007) reports that close to 30% of takeover contests end up in the no-bidder-wins state. This issue is discussed further below.

B2's optimal bid, and noting that the net termination fee paid to B1 if B2 wins is $(1-\alpha)t$, B1's expected profits from bidding $b$ is

$$E(\Pi) = \{(v)G(b+t) - (1-\alpha)\int_t^{b+t}(v_2 - t)g(v_2)dv_2 + (t+\alpha b)[1 - G(b+t)]\}(1-\theta) + t(1-\alpha)\theta \tag{10}$$

The right-hand side is the sum of four components. The first three (inside the curly bracket) are, respectively, B1's expected private value, the expected payment for the target, and the expected value from selling the toehold $\alpha$ and receiving $t$ when B2 wins the auction. The fourth term is the expected payoff when no bidder wins. Using Equation (10), the first-order condition for profit maximization, $\partial E(\Pi)/\partial b = 0$, implies an optimal bid for B1 of [53]

$$b_1^* = v_1 - t + \alpha h(b_1^*) \tag{11}$$

where $h(b_1^*) \equiv \frac{1-G(b_1^*)}{g(b_1^*)}$. Notice the following from Equation (11):

- The toehold induces overbidding, that is, a bid greater than the private valuation $v_1$. This means that B1 may win even if B2 is the higher-valuation bidder (when $v_1 < v_2 < b_1^*$).
- The effect of the termination fee is to induce underbidding. For example, a bidder with zero toehold and a termination agreement walks away from the target when rival bids exceed $v_1 - t$ (quitting means receiving $t$ while continued bidding implies an expected profit of less than $t$).
- Since B1's optimal bid is increasing in the toehold, the probability that B1 wins the auction is also increasing in the toehold. This gives economic content to the frequently heard notion among practitioners that toehold bidding is "aggressive" toward the target.
- When $\alpha = 1$, the optimal bid $b_1^*$ is equivalent to the optimal reserve price by a monopolist seller in a take-it-or-leave-it offer (Eckbo and Thorburn, 2008b).

Bulow, Huang, and Klemperer (1999) and Dasgupta and Tsui (2003) examine toehold bidding in a pure common-value setting where both B1 and B2 have toeholds but of unequal size (asymmetric toeholds). Toehold bidding also induces overbidding in a common-value setting, and these researchers show that holding B1's toehold constant, B2's probability of winning goes to zero as B2's toehold becomes arbitrarily small. Even small differences in toeholds can produce significant benefits for the bidder with the greater toehold. Moreover, the expected winning sales price is decreasing in the difference between the toeholds of B1 and B2. This suggests an incentive on the part of the target to sell a toehold to B2—and for B2 to purchase a toehold—in order to even the playing field. Consistent with this, Betton and Eckbo (2000) find that when a rival bidder enters a takeover contest with a toehold, the toehold size is on average roughly the same size as that of the initial bidder (approximately 5%).

[53] To ensure uniqueness, $G(v)$ must be twice continuously differentiable and satisfy the monotonicity condition $\partial(1 - G(v))/\partial g(v) \geq 0$.

### *3.3.2. The toehold puzzle*

A priori, there is a compelling case for acquiring a toehold prior to initiating a takeover bid. The toehold not only reduces the number of shares that must be purchased at the full takeover premium, but it may also be sold at an even greater premium should a rival bidder enter the contest and win the target. This expected toehold gain raises the bidder's valuation of the target, which in turn helps overcome free-rider problems and makes the toehold bidder a more aggressive competitor in the presence of rivals. Early empirical research supports the existence of toehold benefits. Walkling (1985), Jennings and Mazzeo (1993), and Betton and Eckbo (2000) show that toehold bidding increases the probability of winning the target. Consistent with entry deterrence effects of toeholds, Betton and Eckbo (2000) also find that toeholds are associated with lower offer premiums in winning bids.

However, toehold bidding has in fact been declining dramatically over the past two decades and is now surprisingly rare. This decline is apparent in Figure 8, which plots toehold data from Betton, Eckbo, and Thorburn (2007). The toeholds in Figure 8 include target shares held by the bidder long term as well as shares purchased within six months of the actual offer date (short-term toeholds). Betton, Eckbo, and Thorburn (2007) report a sample-wide toehold frequency of 13%. Moreover, the sample-wide frequency of short-term toeholds—defined as target shares purchased within six months of the offer—is only 2%. In sum, toehold benefits notwithstanding, toeholds acquired as part of an active bidding strategy are almost nonexistent.

Presumably, rational bidders avoid toeholds as a response to large toehold costs. Several potential sources of toehold costs have been suggested in the literature, ranging from mandatory information disclosure and market illiquidity to costs associated with target management resistance to the takeover. Consider first the argument that mandatory disclosure rules make toeholds too costly because they reveal the bidder's intentions early in the takeover process. As discussed above, toehold purchases of 5% or more have triggered mandatory disclosure requirements (13d filings with the SEC) since the 1968 Williams Act. Also, under the 1976 Hart-Scott-Rodino Antitrust Improvements Act, share acquisitions exceeding a certain threshold ($60 million in 2007) trigger notification to the antitrust agencies.

As shown in Figure 8, however, toehold bidding was relatively common in the early 1980s. The passage of disclosure rules in the 1970s cannot explain this time-series pattern. Also, the decline in toehold bidding has occurred despite a steady increase in market liquidity over the entire sample period.[54] Furthermore, Betton, Eckbo, and Thorburn (2007) report that the average toehold size (when positive) is as large as 20%, and 13% for short-term toeholds. It is difficult to explain the observed bimodal toehold distribution (centered on either zero or large toeholds) by appealing to general market illiquidity.

[54] Small toeholds, for which concerns with liquidity and disclosure are unimportant, can also have significant investment value as they retain many of the strategic benefits of larger ones. Toehold benefits arise as long as the toehold is greater than that of the rival bidder (Bulow, Huang, and Klemperer, 1999; Dasgupta and Tsui, 2004).

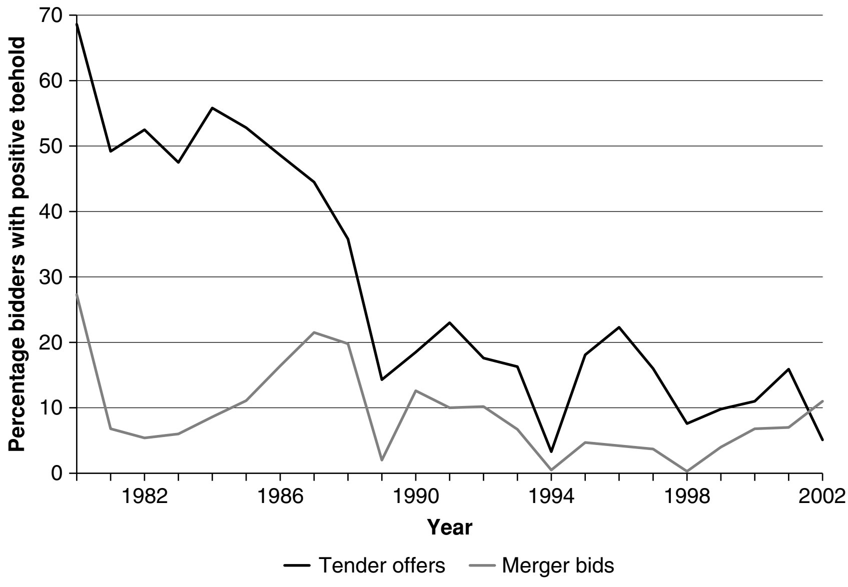



Fig. 8. Annual percentage of initial control bidders with a positive toehold in the target, classified by the type of the initial bid. U.S. public targets. Data source: Betton, Eckbo, and Thorburn (2007).

Goldman and Qian (2005) point to a toehold cost when entrenched target management successfully thwarts the takeover bid. In their model, entrenched target managements may resist a bidder in order to retain the private benefits of control. The degree of target entrenchment is unknown ex ante and, in equilibrium, is signaled ex-post through the size of the bidder's toehold in successfully resisted offers. Successful resistance causes the target share price to drop, and the price drop is greater the greater the bidder's toehold. Bidders trade off expected toehold benefits (greater success probability) with expected toehold costs (greater target price decline when the bid fails), causing some bidders to select small or even zero toeholds. However, the evidence in Betton, Eckbo, and Thorburn (2007) rejects the predicted negative correlation between the sizes of bidder toeholds and target price declines conditional on all bids failing. The potential for a toehold loss in the event that all bids fail following target resistance does not appear to explain the toehold puzzle.

Betton, Eckbo, and Thorburn (2007) develop and test a model in which toehold costs arise endogenously. The takeover game starts with the initial bidder approaching the target with an invitation to negotiate a merger. In line with the fiduciary out requirement discussed earlier, a merger agreement is always followed by a period during which the target board is required to consider any rival bids (until the shareholder vote). The expected outcome of this open auction period determines the outcome of merger negotiations. Since a toehold affects the expected auction outcome (recall the optimal bid in Equation (11)), it also affects the willingness of entrenched target managements to accept the bidder's invitation to negotiate. If the target management rejects negotiations, the bidder foregoes the benefit of the termination agreement and incurs resistance costs during the takeover process.

These toehold-induced bidder costs make it optimal for some bidders to approach the target without a toehold. That is, the expected toehold cost creates a *toehold threshold* (a minimum toehold size), below which the optimal toehold is zero. Betton, Eckbo, and Thorburn (2007) show that the toehold threshold averages 9% in the data, which is consistent with the observed bimodal distribution of observed toeholds (centered on zero or large toeholds). That is, some bidders find that the toehold threshold is too costly to purchase in the market (e.g., due to market illiquidity) and select zero toehold. The key model prediction is that the likelihood of toehold bidding decreases in the toehold threshold estimate (the expected opportunity loss of the termination agreement), which the empirical evidence supports.

The threshold model is also consistent with another stylized fact: toeholds are much more common in hostile than in friendly takeovers. While 11% of initial bidders have toehold when the target is friendly, 50% of the initial bidders in hostile contests have toeholds. The threshold theory suggests that one should observe toehold bidding when the opportunity cost of the toehold is relatively low. A special case is when the opportunity cost is *zero*, which occurs whenever the target's optimal resistance strategy is independent of the toehold. That is, if target management is expected to resist regardless of toeholds, acquiring a toehold is always optimal.[55] Thus, the toehold threshold model predicts a higher toehold frequency in hostile bids, and it is consistent with the observed decline in the frequency of toehold bidding over the 1990s (Figure 8). This decline coincides with a general reduction in hostile bids due to a widespread adoption of strong takeover defenses such as poison pills.

Finally, in the absence of synergistic opportunities with the target ($v = 0$), the owner of a toehold may contemplate making a (false) bid in an attempt to put the target in play. The idea is to try to sell the toehold to a potential rival bidder or (anonymously) to an unwitting market anticipating a successful takeover. Bagnoli and Lipman (1996) present a model with a single bidder selling the toehold shares to individual noise traders through a market maker before calling off the takeover bid. While charges of price manipulation go back at least to the greenmail episodes of the late 1970s, systematic empirical evidence on the feasibility of this type of price manipulation is virtually nonexistent. The context of hostile bids is potentially interesting since hostility may induce the target to produce a white knight committed to purchase the toehold.

### *3.4. Bid jumps and markup pricing*

In this section we examine evidence on the size of bid jumps in multiple-bid contests and investigate how pre-bid target runups affect the initial and final offer prices. Also,

[55] Similarly, toehold bidding occurs when the target's optimal strategy is to never resist.

an interesting question is whether target runups and markup pricing deter toehold acquisitions by the initial bidder.

### 3.4.1. Preemption and bid jumps

As indicated earlier, the high premiums observed in takeovers are consistent with the hypothesis that takeover benefits are partly common to several potential bidders. This is likely when takeover benefits emanate, for example, from replacing inefficient target management or using voting control to extract value from ex-post minority shareholders in the merged firm. These and other forms of bidder–target complementarities often do not require specialized resources owned by a single potential bidder firm. As a result, the first bidder is concerned that the initial bid will alert potential rivals to a profit opportunity. The empirical issue is whether this possibility affects observed bid strategies.

Fishman (1988) analyzes this issue assuming that bidders must pay an investigation cost to identify their respective private valuations of the target. If both bidders enter (so that both investigation costs are sunk), an open English auction with costless bidding ensues and produces the "ratchet" solution $min[v_1,\ v_2]$ (Hirshleifer, 1995). However, there exists an initial bid that deters the second bidder from paying the investigation cost and entering the auction. The high initial (all-cash) bid signals that the initial bidder has a relatively high private valuation for the target, which reduces rival bidders' expected value of winning. For a sufficiently large investigation cost, the expected value is negative and the rival does not enter.

Testing preemption arguments is difficult since one obviously cannot observe deterred bids nor bidder private valuations in observed bids. One must look to auxiliary or related predictions, and the following four categories of results seem relevant. First, entry is rapid when it occurs: the average number of trading days between the first and second control bid is 40 in our sample (Figure 6) and 15 days in Betton and Eckbo (2000). This suggests that the rival bidder's investigation process required to establish its own valuation of the target is not very time-consuming in these cases. Also, some rivals may have completed much of the evaluation prior to the initial bid. Observing the initial bid event may produce a sufficient target valuation estimate to make a bid.

Second, auction outcomes are sensitive to bidder asymmetries. One important form of bidder asymmetry is the size of bidder toeholds. Even small toehold differences can have a large impact on entry and competition. Empirically, Betton and Eckbo (2000) find that when a rival bidder enters a takeover contest with a positive toehold, the toehold size is on average of roughly the same size as that of the initial bidder (approximately 5%). It is as if the rival bidder realizes the initial bidder's toehold advantage and wants to neutralize it upon entry.

Third, both Betton and Eckbo (2000) and Betton, Eckbo, and Thorburn (2007) report that the average offer premium in single-bidder successful tender offer contests (the first node in Figure 2) is slightly higher than the average *initial* offer premium in contests that developed into multiple bids. This is consistent with the argument that the premiums in single-bid successful contests are preemptive in the sense of Fishman (1988). However, the premium effect is weak: the probability of rival bidder entry appears unaffected by the initial offer premium (Betton, Eckbo, and Thorburn, 2008b).

Fourth, Betton and Eckbo (2000) report evidence of significant bid jumps throughout the tender offer contests. For example, the average jump from the initial to the second bid price in the contest is 10%, implying a 31% change in the initial offer premium. The jump from the first to the final bid average 14% (a 65% revision in the initial offer premium), and the average bid jump throughout the entire contest, is 5% (average premium increments of 17%). The evidence of significant bid jumps throughout the contest is consistent with the presence of bidding costs. This in turn supports the notion in Fishman (1988) that initial bidders may strategically raise the first bid in an attempt to deter competition.[56]

### *3.4.2. Runups and markups*

We now turn to the markup pricing phenomenon first documented by Schwert (1996). Initial takeover bids are typically preceded by substantial target stock price runups. The runup reflects takeover rumors generated from various public sources, such as Schedule 13(d) filings with SEC disclosing stake purchases of 5% or more in the target, media speculations, and street talk. The conventional view is that runups reflect takeover rumors based on information that is already known to the bidder.[57] If this view is correct, the runup anticipates an already planned offer premium and does not require a premium revision before the offer is made.

This is not the only possible scenario, however. Schwert (1996) begins his paper with the following question:

> Suppose that you are planning to bid for control of a company and, before you can announce the offer, the price of the target firm's stock begins to rise coincident with unusually high trading volume. You have not been buying the target company's stock, and there is no reliable evidence to show who has been buying. Do you go forward with the offer exactly as you had planned? Or do you take into account the recent movement in the target's stock price and adjust your bidding strategy? (pp. 153–154).

Bidders need a plan for how to react to the runup before making the initial bid. Moreover, such a plan requires an understanding of the true nature of the pre-bid target runup. For example, it is possible that the target runup represents an increase in the target's fundamental (stand-alone) value, in which case the target management may demand a higher price. If so, the bidder may be forced to mark up the offer price to reflect the higher target stock price on the day before the offer is made.

To examine the extent of markup pricing, Schwert (1996) writes the total offer premium as $Premium \equiv Markup + Runup$, where *Runup* is the cumulative target abnormal stock return from day $-42$ through day $-1$ relative to the first bid for the

[56] See also Hirshleifer and Ping (1990) and Daniel and Hirshleifer (2008) for discussions of the implication of bidding costs for optimal bidding strategies.

[57] Jarrell and Poulsen (1989) and King and Padalko (2005) conclude that runups are primarily a result of public information. Meulbroek (1992) and Schwert (1996) find greater target runups in cases where the SEC subsequently alleges insider trading.

target (day 0), and *Markup* is the cumulative abnormal target stock return from day 0 through day 126 (or until delisting, whatever comes first). He then estimates the coefficient $b$ in the following cross-sectional regression:

$$Premium_i = a + bRunup_i + u_i \tag{12}$$

where $u$ is an error term. With a sample of 1,814 mergers and tender offers from the period 1975–1991, Schwert finds a statistically significant $b = 1.13$ for the total sample (with a $t$-value of 2.88 for the null hypothesis of $b = 1$). In other words, in the total sample, a dollar runup in the target stock price raises the total offer premium by approximately a dollar. Under the more conventional view of the runup, *Markup* and *Runup* are substitutes (predicting $b = 0$ in regression (12)), which Schwert's evidence rejects.

Schwert's estimate of the markup is impacted by events occurring *after* the initial offer, such as the entry of rival bidders and bid revisions by the initial bidder, target management resistance, and ultimate target shareholder voting outcomes. Betton, Eckbo, and Thorburn (2008b) use the initial offer price $p_{initial}$ to measure the initial markup directly as $Markup = ln(p_{initial}/p_{-1})$, where $p_{-1}$ is the target share price on the day prior to the initial bid. The runup is measured as $Runup = ln(p_{-1}/p_{-42})$. With a sample of six thousand initial takeover bids for U.S. public targets from the period 1980–2002, they estimate the coefficient $b'$ in the following regression,

$$Markup_i = a' + b' Runup_i + cX + u_i \tag{13}$$

where $X$ is a set of bidder- and target-specific deal characteristics. Betton, Eckbo, and Thorburn find that $b' = -0.18$ for the total sample ($t$-value of $-15.44$). Thus, in the cross section of bids, a dollar increase in the target runup is associated with an increase in the average initial offer price by \$0.82.[58] They also show that the degree of substitution between the markup and the runup is greater when the bidder purchases a target toehold in the runup period, and they conclude that target runups are an unlikely explanation for the sparsity of toehold purchases by initial bidders in the runup period.

Is markup pricing costly in the sense of reducing bidder synergies? To examine this issue, Betton, Eckbo, and Thorburn (2008b) estimate the following cross-sectional regression with bidder takeover-induced abnormal stock returns, BCAR, as dependent variable:

$$\mathrm{BCAR}_i = a_b + b_b Runup_i + c_b X_i + u_i \tag{14}$$

where *Runup* is the *target* runup (as before). The coefficient $b_b$ is positive and highly significant in a sample exceeding 4,000 public bidders. That is, greater target runups are simultaneously associated with markup pricing and greater bidder synergies from the takeover.

Since target synergies are also (obviously) increasing in target runups, the positive estimate of $b_b$ means that the runup is a proxy for *total* synergies in the cross section.

[58] If one changes the dependent variable in Equation (13) to the total initial offer premium premium, $ln(p_{initial}/p_{-42})$, the slope coefficient changes to $1 + b' = 0.82$.

This finding affects the interpretation of the coefficients $b$ and $b'$ in Equations (12) and (13). To illustrate, suppose takeover rumors allow market investors to not only identify the target but also to distinguish targets with high and low expected total synergies. Moreover, suppose competition always forces bidders to grant target shareholders (in the form of a takeover premium) a fixed portion of the total synergies. Bidders expecting the takeover to be profitable now also expect a high pre-bid runup, and mark up the initial offer price ex ante (before the runup). This also produces a markup that is independent of the runup ex-post ($b = 1$), although there are no actual bid revisions following the runup. Ultimately, distinguishing between this total synergy hypothesis and Schwert (1996)'s ex-post markup proposition requires evidence on actual offer price changes made by the initial bidder during the runup period. However, either scenario is consistent with a positive association with target runups and bidder takeover gains.

### 3.5. *Takeover defenses*

In this section, we briefly characterize the legal basis for target takeover defenses, and then we examine the empirical evidence on the shareholder wealth effects of antitakeover measures, in particular poison pills, classified boards, and defensive payouts (greenmail). Figure 9 shows the annual frequency of the sample of 1,052 unfriendly (unsolicited and outright hostile) initial bids previously listed in Table 1.

Since target hostility may simply represent posturing to improve the target's bargaining position, several definitions of hostility exist (Schwert, 2000). The SDC definition probably casts a relatively wide net, as all it ensures is that (1) the bidder (and not the target) initiates the takeover and (2) the target board is initially unprepared and/or unwilling to enter into merger negotiations. Specifically, the SDC classification does *not* necessarily mean that the target is dead set against negotiations, nor does it mean that it is going to implement defensive tactics. However, target defensive actions are more likely in this sample than in cases where initial bids are classified by SDC as solicited or friendly. Notice also that the SDC definition allows a hostile initial bid to be in the form of either a merger or a tender offer (although, as shown in Figure 9, unfriendly initial bids are typically in the form of a tender offer). An example of an initially hostile merger bid is a "bear hug," in which the bidder invites the target to negotiate while reminding the target board that the bidder is likely to pursue a tender offer should the board refuse negotiations.

As shown in Figure 9, the fraction of bids that are unfriendly is relatively high throughout the 1980s and then drops sharply after 1989. Comment and Schwert (1995) analyze the drop in hostility, which is closely associated with the spread of takeover defenses and the development of state antitakeover statutes (control share and business combinations laws). Given this close association, it is natural to view the drop as being caused by increased managerial entrenchment afforded by strong takeover defenses. Comment and Schwert (1995), however, argue that the emergence of takeover defenses played only

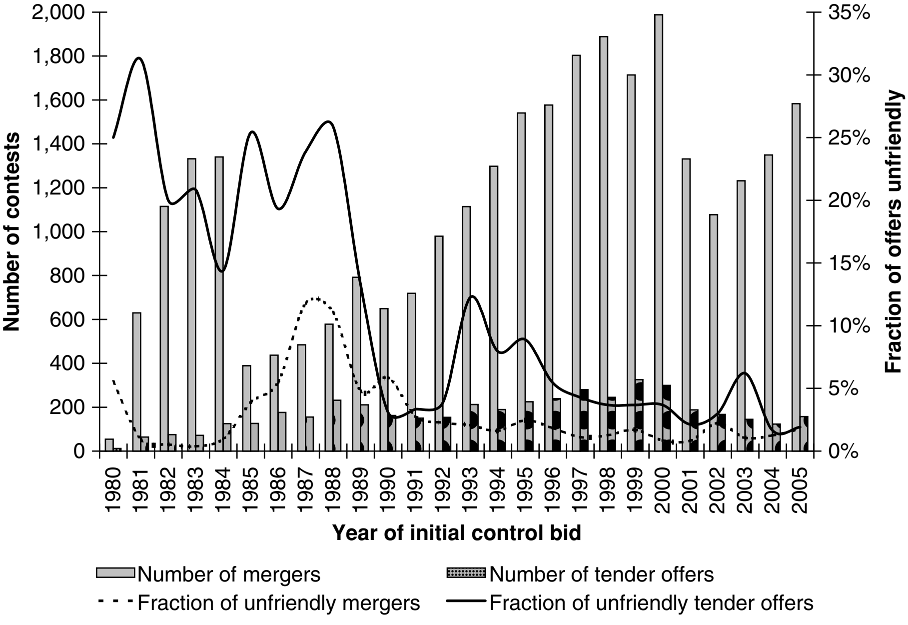



Fig. 9. Target attitude toward the initial control bidder in mergers and tender offers. The attitude is shown as "unfriendly" ($N = 1,025$) if the SDC lists the offer as either "hostile" or "unsolicited." All other offers are shown as "friendly" ($N = 34,702$). U.S. targets, 1980–2005.

a minor role in ending the 1980s merger wave. They point instead to the development of a general economic recession beginning in 1989, which caused a collapse in net new lending to the nonfinancial sector by commercial banks from $33 billion in 1989 to $2 billion in 1990. Commercial banks were the dominant providers of bridge or transaction financing for large, cash acquisitions at the time. Takeover activity was also generally reduced as a result of a drop in the availability of long-term and subordinated financing, in part due to government intervention in the junk bonds in 1989.[59]

While the overall credit crunch undoubtedly slowed the economy and reduced takeover activity, there is also little doubt that the sharp reduction in *unfriendly* takeovers in large part reflects the legal certification and spread of strong antitakeover measures. Indeed, Jensen (1993) argues that the regulatory attack on the junk bond market around 1989

[59] "In August 1989, Congress passed the Financial Institutions Reform, Recovery and Enforcement Act (FIRREA), which penalized savings and loans for holding junk bonds and mandated their sale, while regulators issued guidelines barring commercial bank participation in highly leveraged transactions (including all acquisition loans that raised liabilities to 75% of assets, or doubled the debt ratio while raising it to 50% of assets). The junk bond market crashed in September 1989" (Comment and Schwert, 1995, p. 9).

in of itself may be understood as a broadly organized defensive tactic against unwanted takeovers.[60] While the combination of a poison pill and staggered board is not viewed as a draconian defense in the eyes of the law (see below), there can be no question that these measures when used in combination effectively bar or seriously delay a hostile bid. As discussed in the following, however, the overall degree of deterrence remains unclear from the empirical literature.

### *3.5.1. Legal basis for defensive measures*

In this section, we summarize certain aspects of the highly complex case law governing takeover defenses.[61] We focus on Delaware case law since a majority of U.S. public companies (and more than 60% of the Fortune 500 firms) are incorporated in the state of Delaware.

Delaware case law sanctions the right of a board to "just say no" to an unsolicited takeover bid and to defend itself against that bid if necessary to remain an independent corporation. The case law rests on director fiduciary duties and the judicially developed principle referred to as the business judgement rule. Director fiduciary duties include duty of care and duty of loyalty. Duty of care is typically satisfied as long as the board examines fairness opinions of a bid and spends a minimum amount of board time discussing the value of the proposed deal.[62] Duty of loyalty is typically satisfied as long as the proposed deal does not imply a personal benefit for directors. Moreover, the presence of a majority of independent directors is viewed as a strong indication of the satisfaction of duty of loyalty.

The business judgment rule presumes, when director action is challenged, that the director of a corporation acted on an informed basis, in good faith, and in the best interest of the company. If the board is found to have acted this way, a court will not substitute its judgment for that of the board, and the court is inclined to find some rational purpose for the board action. In the context of a takeover bid, the board may determine in good faith that the continuing independence of the corporation is in the long-term best interests of the corporation and its stockholders. The board "is under no obligation, in

[60] The critical view many business leaders had of the junk bond market is illustrated by the sentiment expressed by J. Richard Munro, chairman and CEO of Time, Inc., in a speech in 1989: "Notwithstanding television ads to the contrary, junk bonds are designed as the currency of 'casino economics' . . . they've been used not to create new plants or jobs or products but to do the opposite: to dismantle existing companies so the players can make their profit. This isn't the Seventh Cavalry coming to the rescue. It's a scalping party" (Munro, 1989, p. 472).

[61] We have benefited greatly from conversations with John G. Gorman, partner in the law firm Luse Gorman Pomerenk & Schtik, P.C. (Washington D.C.). For comprehensive reviews of federal and state rules governing corporate control changes, see, for example, Wasserstein (2000), Lipton and Steinberger (2004), and Gaughan (2007).

[62] The standard for determining breach of the duty of care is generally considered to be gross negligence. *Smith v. Van Gorkam*, 488 A.2d 858 (Del. 1985).

the abstract, to submit to an external summons to the auction block or otherwise transfer control of corporate assets."[63]

A board may even be legally required to oppose an offer that it believes is not in the best interest of the corporation and its stockholders.[64] The board is not obligated to accept an offer simply because it represents a premium over a current market price. Refusal of such an offer is not *prima facie* evidence of a breach of fiduciary duty,[65] *except* when a sale of control of the corporation has been decided. If a determination is made to enter into a sale of control transaction, the fiduciary duties of the directors are enhanced and the directors have an obligation to seek the transaction offering the best value (which may mean highest price) reasonably available to stockholders — the so-called Revlon-duties.[66]

Case law sanctions a wide range of target defensive mechanisms against an unsolicited bid. However, the courts have noted that given the "omnipresent-specter that a board may be acting primarily in its own interests, there is an enhanced duty which calls for judicial examination at the threshold before the protections of the business judgement rule may be conferred."[67] This modified business judgment rule requires that the board initially establishes that (i) it had reasonable grounds for believing there was a danger to corporate policy and effectiveness, and (ii) the measure adopted in response to the threat was reasonable in relation to the threat posed.[68]

If the board's defensive response is not draconian (i.e., it is neither coercive nor preclusive) but within the range of reasonableness given the perceived threat, the board is protected by the modified business judgment rule. The following excerpt from the Unitrin decision, the leading case on a board of directors' ability to use defensive measures to prevent a hostile takeover, illustrates the court's mind-set:[69]

> Proper and proportionate defensive responses are intended and permitted to thwart perceived threats. When a corporation is not for sale, the board of directors is the defender of the metaphorical medieval corporate bastion and the protector of the corporation's shareholders. The fact that a defensive action must not be coercive or preclusive does not prevent a board from responding defensively before a bidder is at the corporate gate . . .. Thus, continuing with the medieval metaphor, if a board reasonably

[63] *Paramount Communications, Inc. v. Time, Inc.*, 571 A.2d 1140 (Del. 1990).

[64] *Gilbert v. El Paso Co.*, 575 A.2d 1131 (Del. 1990).

[65] *Pogostin v. Rice*, 480 A.2d 619 (Del. 1984).

[66] See *Revlon, Inc. v. MacAndrews and Forbes Holdings, Inc.*, 506 A.2d 182 (Del. 1986), and *Paramount Communications, Inc. v. QVC Network, Inc.,* 637 A.2d 34 (Del. 1994).

[67] *Unocal Corp. v. Mesa Petroleum Co.*, 493 A.2d 946 (Del. 1985).

[68] The burden of proving reasonable grounds as to the danger to corporate policy and effectiveness can be met by showing good faith and reasonable investigation. A board's ability to show the reasonableness of the response adopted is enhanced when a majority of the board consists of outside, independent directors, or when the actions taken precede an actual threatened change in control. An "inadequate" offer price can be a reasonably perceived threat. The concern that shareholders may be ignorant of the true value of the company may be considered by the board, and the interests of long-term shareholders versus short-term speculators (such as arbitrageurs) may be taken into account.

[69] *Unitrin, Inc. v. American General Corp.*, 651 A.2d 1361 (Del. 1995).

> perceives that a threat is on the horizon, it has broad authority to respond with a panoply of individual or combined defensive precautions, e.g., staffing the barbican, raising the drawbridge, and lowering the portcullis. Stated more directly, depending upon the circumstances, the board may respond to a reasonably perceived threat by adopting individually or sometimes in combination: advance notice by-laws, supermajority voting provisions, shareholder rights plans, repurchase programs, etc.

A defense that is deemed preclusive because it frustrates, impedes, or disenfranchises a shareholder vote will be held to the so-called Blasius standard of compelling justification[70] and is unlikely to be upheld.[71] For example, a stock repurchase designed primarily to preclude a third party from winning a proxy contest for the selection of directors may not pass the Blasius standard. Also, defensive measures have not fared well in court when the defense has involved a transaction in which existing management will have an equity interest or where the purpose is to favor one party over another.

### *3.5.2. Defenses and offer premiums*

Following the reference made by Manne (1965) to the "external" and "internal" market for corporate control, several authors have similarly categorized antitakeover provisions.[72] The external control market involves takeover bids and specific target responses, while the internal market involves general board actions and shareholder voting. Examples of internal antitakeover provisions are classified (staggered) board (directors are divided into separate classes—typically three—and elected to overlapping terms), unequal voting rights (e.g., two classes of common stock, one with zero voting rights), and various restrictions on shareholder rights to amend company charter and bylaws, to act by written consent, and to call special meetings. Examples of external antitakeover provisions include antigreenmail provisions (prohibition on paying greenmail—the targeted repurchase of a single shareholder's stockholding at a premium), supermajority requirements to approve a merger, blank check preferred stock (used to implement a poison pill), fair price provisions (requires a large shareholder to pay a minimum price set by formula for all shares acquired in the back end of a two-tiered acquisition), and poison pills or shareholder rights plans.

The development of the poison pill is tied directly to the history of greenmail. Following several occurrences of greenmail payments during the late 1970s and early 1980s, Unocal made what turned out to be a landmark decision to reverse the greenmail payment. In 1983, Mr. T. Boone Pickens Jr. and his Mesa Partners II, who held 13.6% of Unocal's

70 *Blasius Indus., Inc. v. Atlas Corp.*, 564 A.2d 651, 660 (Del. 1988).

71 *MM Companies v. Liquid Audio, Inc.*, 813 A.2d 1118 (Del. 2003). In this case, the court invalidated the board's decision to add two new directors to prevent the acquirer from obtaining board control at the subsequent shareholder meeting. The Blasius standard "is to be applied sparingly, and only in circumstances in which self-interested or faithless fiduciaries act to deprive stockholders of a full and fair opportunity to participate and to thwart what appears to be the will of a majority of stockholders" (MONY Group, Inc. Shareholder Litigation, Del. (class action that was settled)).

72 Jarrell and Poulsen (1987), Danielsen and Karpoff (1998), Gompers, Ishi, and Metrick (2003), Moussawi (2004). Bebchuk, Coates, and Subramanian (2002) summarize case law concerning antitakeover provisions.

stock, made an $8.1 billion takeover bid for Unocal. The offer was for $54 a share in cash for 37% of Unocals stock and $54 a share in junior securities for the rest. Unocal's board responded by offering to exchange $72 a share in senior securities for 50.1% of the company's total shares, but barred the Mesa group from participating in the stock repurchase. Delaware Supreme Court upheld Unocal's right to undertake the targeted repurchase.[73]

Attorney Martin Lipton of Wachtell, Lipton, Rosen & Katz was a key legal strategist working for Unocal. Subsequently, Mr. Lipton's law firm proceeded to develop the "shareholder rights plan"—popularly referred to as the posion pill—which is an ongoing commitment to trigger what in essence is a reverse greenmail payment.[74] When adopting the poison pill, the corporation issues to its stockholders (usually by means of a dividend) certain rights to purchase stock. The rights are out of the money (the exercise price exceeds the then market price) and not exercisable until a triggering event. The triggering event is that someone acquires a certain percentage (e.g., 15%) of the firm's voting shares. Pending their exercise, the rights may be redeemed for a nominal value by the board. If triggered, the rights give each holder, other than the stockholder who triggered the pill, the right to purchase shares of the issuing corporation (flip-in) or of the acquirer (flip-over) at a deep discount (e.g., 50%) to the market price. The board may offer pills without prior stockholder approval, and the pills may be issued after having received a hostile bid ("morning after" pill).[75]

In 1985, the Delaware Supreme Court upheld Household International's adoption of a shareholder rights plan as reasonable under the Unocal standard, even though the company did not face a hostile threat.[76] Subsequently, Delaware has upheld the right of a board to refuse to redeem a pill in the face of an all-cash, noncoercive tender offer, even though a majority of the company's stockholders had tendered their shares to the bidder.[77] On the other hand, Delaware courts have invalidated the so-called dead-hand poison pill, which attempted to provide that only incumbent directors could redeem the rights, thus preventing newly elected directors from unwinding the pill.[78] This is an important decision, as one (though costly) way to circumvent the pill is to launch a proxy contest simultaneously with the hostile offer, in the hope of winning enough board seats to have the board rescind the pill and let the offer go through.

The combination of a hostile bid and a proxy contest does not work if the target board is classified or staggered. For example, if only one-third of the board is up for election, the hostile bidder cannot win the majority needed to rescind the pill. Indeed,

[73] *Unocal Corp. v. Mesa Petroleum Co*, 493 A.2d 946 (Del. 1985).

[74] Mr. Lipton's law firm became a dominant supplier of poison pills to U.S. public companies thereafter.

[75] Pill adoption does not require a shareholder vote since it is akin to a dividend payment. Recently, there has been a growing demand from large institutional shareholders such as pension funds to allow shareholders to vote on pill adoptions.

[76] *Moran v. Household International, Inc.* 500 A.2d 1346 (Del. 1985).

[77] *Moore Corp., Ltd. v. Wallace Computer Services, Inc.*, 907 F. Supp. 1545 (D. Del. 1995).

[78] *Quickturn Design Systems, Inc. v. Shapiro* 721 A.2d 1281 (Del. 1998). This version of the pill had been upheld under Georgia Law, but also invalidated under New York law.

as argued by Bebchuk, Coates, and Subramanian (2002), Bebchuk, Cohen, and Ferrell (2004), and Bebchuk and Cohen (2005), board classification may in and of itself constitute an antitakeover device. Bebchuk and Cohen (2005) examine the cross-sectional relationship between board classification and firm value, and find that board classification is negatively correlated with industry-adjusted Tobin's Q. Also, Masulis, Wand, and Xie (2007) find that acquisition announcement-period stock returns are significantly lower for bidders with staggered boards, possibly because board classification reduces forced board turnover and quality. On the other hand, Bates, Becher, and Lemmon (2008) find that board classification does not reduce the probability that a firm, once it is targeted, is ultimately acquired. Moreover, targets with classified boards appear to extract premiums equivalent to those of single-class boards. However, they do find that board classification is associated with a small reduction in the probability of receiving a takeover bid. Rose (2008) also concludes that the presence of staggered boards has more of a detrimental impact on firm value when management is relatively entrenched.

The ambiguities in interpreting the overall consequences for shareholders of a defensive measure such as a staggered board are also present in the debate over the poison pill defense. There is substantial empirical evidence that targets that have adopted poison pills receive offer premiums that are, on average, indistinguishable from offer premiums received by nonpill targets.[79] This evidence is consistent with the following four alternative hypotheses:

H1 Poison pills are irrelevant for determining final takeover premiums.
H2 Poison pills convey bargaining power, which increases the final takeover premium relative to what the premium would have been for the same target without a pill.
H3 Poison pills convey bargaining power that is used to benefit target management at the expense of target shareholders.
H4 Poison pills provide bargaining power, but "shadow" pills are as effective as adopted pills.

Hypotheses H2–H4 maintain that pills do convey bargaining power but that a comparison of offer premiums in samples of firms with or without pills is difficult from an econometric point of view. Pill adoptions are voluntary, which raises complex issues of endogeneity (H2). Controlling for self-selection is difficult because the marginal effect of a poison pill depends on the firm's entire governance system, including executive compensation (H3).[80] Also, in order to isolate the true premium effects of pills, empirical work relies on the existence of two samples, one representing "poison" and the

[79] Comment and Schwert (1995), Field and Karpoff (2002), Heron and Lie (2006), Betton, Eckbo, and Thorburn (2007).

[80] Compensation effects of takeovers are discussed in Section 5.2. Heron and Lie (2006) find that the targets of hostile bids are more likely to adopt poison pills when they have classified boards, suggesting that the two antitakeover devices are interdependent.

other "placebo" effects (Comment and Schwert, 1995). This sampling is difficult, if not impossible, if, as in H4, *all* firms effectively have ready access to the pill (Coates, 2000).

H1 maintains that pills may simply be ineffective and therefore irrelevant for final offer premiums. At first blush, H1 seems to be rejected by the fact that no bidder (to our knowledge) has yet triggered a pill. However, why trigger an ineffective pill if the trigger itself is costly—also to target management? Consider the failed 1996 takeover attempt by U.S. Surgical Corporation of medical device maker Circon Corporation. Exercising the Circon pill would have required Circon shareholders to pay approximately $800 million in cash into a company with a pre-takeover total equity value of $150 million. In return for this massive (and expensive) cash infusion, Circon shareholders would lose a 70% takeover premium and stood to gain only $10 million from the resulting dilution of U.S. Surgical's shareholding in Circon. In general, a pill with this structure may lack credibility and therefore have little effect on bargaining outcomes.[81]

Moreover, the definition of target "hostility" used in the literature probably captures many targets that are ready to negotiate with or without the pill (Schwert, 2000). Bidders that are able to look beyond the pill and determine whether negotiations are possible (based on observable target characteristics or on the bidder's own ability to persuade a hostile target management) may reach a final bargaining outcome that is largely indistinguishable from that observed in samples of ex-ante "friendly" targets. Empirical evidence shows that the probability of receiving a bid (and ultimate bid success) is either unaffected or slightly lower for targets with strong antitakeover defenses.[82]

Finally, several studies estimate the valuation effects of antitakeover charter amendments (which require a shareholder vote), with data primarily from the 1980s. An advantage of studying charter amendments is that the market reaction isolates the net present value of the expected impact of the antitakeover measures on all future takeover activity. A disadvantage, however, is that the lengthy process toward a vote at the shareholder meeting leaks information and leads the market to partially anticipate the event, thus reducing power to register significant changes in market expectations. There is also some controversy over which event date is the most appropriate: the shareholder meeting date, the proxy announcement date, or the proxy mailing date (Bhagat and Jefferis, 1991). Also, as with studies of poison pills, it is important but difficult to properly account for the endogenous nature of the amendment choice, as it is part of the amending firm's entire governance system.[83]

Since the amendments must pass a shareholder vote, a natural null hypothesis is that these serve the interests of shareholders. Under this hypothesis, a takeover amendment

81 In the Circon case, Circon chairman and CEO Richard Auhll appeared to be protecting large private benefits of control. Using information on SDC, approximately half of all pills are cash pills (the exercise price is paid in cash rather than by an exchange of securities.

82 Comment and Schwert (1995), Field and Karpoff (2002), Heron and Lie (2006).

83 Malekzadeh, McWilliams, and Sen (1998), Bhagat and Jefferis (2002).

increases the expected future takeover premium (the change in the probability of a takeover times the change in the premium conditional on a takeover). For example, the amendments may help resolve a target shareholder coordination (holdout) problem and increase the expected takeover price, especially in a two-tier tender offer setting (DeAngelo and Rice, 1983). Or, in the context of optimal contracting, the amendment decreases the expected future takeover premium in return for greater managerial incentives to invest in firm-specific human capital (Shleifer and Vishny, 1989). The main alternative hypothesis is that the amendments further entrench incumbent management (with insufficient offsetting benefits) and that the voting mechanism is unable to prevent the management proposal from passing.[84]

Early studies of share price effects of fair price amendments, classified boards, supermajority requirements, and other "shark repellents" adopted by publicly traded firms find a zero or small negative market reaction. These include DeAngelo and Rice (1983), Linn and McConnell (1983), and Jarrell and Poulsen (1987). Fair-price amendments (the bulk of the sample amendments) are met with an insignificant market reaction, while board classification elicits significantly negative abnormal stock returns. Jarrell and Poulsen (1987) also find that the amendments having the most negative effects are adopted by firms with the lowest percentage of institutional holdings and the greatest percentage of insider holdings. Malatesta and Walkling (1988) distinguish between takeover defenses that do or do not require shareholder approval, and conclude that defenses that are not ratified by a shareholder vote reduce shareholder wealth. Ryngaert (1988) also finds evidence that poison pill adoptions reduce shareholder value, as do news of court decisions upholding poison pill defenses. More recent studies of the market reaction to antitakeover amendments tend to confirm the conclusions of this literature, also after providing a more detailed picture of the interaction with the adopting firm's corporate governance and ownership structures.[85]

Finally, a number of papers examine the valuation effects of greenmail—the precursor to the poison pill—and antigreenmail charter amendments. As indicated above, greenmail refers to an arrangement in which a company repurchases the stock held by a large shareholder, usually at a substantial above-market price. In return, the large stockholder signs a standstill agreement committing not to purchase additional target shares or launch a control bid for typically a 10-year period. Bradley and Wakeman (1983), Dann and DeAngelo (1983), and Mikkelson and Ruback (1991) find that the announcement of greenmail transactions are associated with significantly negative abnormal stock returns of approximately $-2\%$. Mikkelson and Ruback (1991) find that the market reaction is negative only if the stockholder signs a standstill agreement or if it aborts a control

[84] It is well understood that a vote may not necessarily safeguard shareholder interests. See, for example, Austen-Smith and O'Brien (1986), Jarrell and Poulsen (1987), Brickley, Lease, and Smith (1988), and Gordon and Pound (1993) for evidence of voting on antitakeover amendments. Since the 1980s, increasing institutional shareholder activism has made it more difficult for incumbents to secure shareholder support for defensive measures.

[85] Mahoney and Mahoney (1993), McWilliams and Sen (1997), Sundaramurthy, Mahoney, and Mahoney (1997).

contest.[86] Mikkelson and Ruback (1985) show that the total abnormal stock return from the initial 13D filing by the toehold investor until he receives the greenmail payment is significantly positive. In other words, the greenmail payment and standstill do not eliminate all gains from having had a significant blockholder in the firm's ownership structure.

Eckbo (1990b) and Bhagat and Jefferis (1991) present evidence on antigreenmail charter amendments. The typical amendment prohibits the firm from repurchasing some or all of the common (voting) stock of an *interested shareholder*—normally defined as a shareholder who owns 5% or more of the outstanding common stock and who acquired this ownership position within the past two to three years. Virtually all firms retain the option to pay greenmail as long as (1) two-thirds or more of the *disinterested shareholders* approve of the action, or (2) if the shares are repurchased at a fair price, usually defined as an average of the stock's trading prices over the 90 days immediately preceding the share repurchase.

The market reaction to greenmail prohibitions represents the value of the *option* to pay greenmail in the future. If the option value is negative, the market reaction to the amendment will be positive. The option value is negative if the sum of the repurchase premium (greenmail payment), the marginal increase in agency costs from successfully rebuffing future hostile takeover bids, and the increased likelihood of receiving purely extortive bids (bids designed exclusively to generate greenmail payments) exceed the expected benefits. Thus, evidence of a positive market reaction to the greenmail prohibition would support the widely held view that greenmail payments harm nonselling shareholders.[87]

For a subsample where the antigreenmail amendment is proposed by itself (without other simultaneous antitakeover proposals), Eckbo (1990b) has found the average market reaction to the charter amendments to be weakly negative. The market reacts negatively to the greenmail prohibition if the value of the unrestricted option to pay greenmail is positive. However, cross-sectional regressions further indicate that the market reaction is strongly positive when the firm has experienced a recent stock price runup along with takeover rumors. Eckbo (1990b) concludes that the option to pay greenmail is costly when the firm likely has been identified as a target, in which case the antigreenmail amendment removes a possible barrier to the pending takeover.

### *3.6. Targets in bankruptcy*

In this section, we consider evidence on the acquisitions of target firms that have filed for bankruptcy. Since bankruptcy law alters the bargaining position of the target, one

[86] They also report that even with the standstill, 40% of the firms paying greenmail experience a control change within the subsequent three years.

[87] It is not, however, the only possible interpretation. In the asymmetric information model of Shleifer and Vishny (1986a), greenmail payments increase the intrinsic value of the firm while at the same time causing the firm's stock price to fall.

expects the outcome for bidders to be different from that for out-of-court acquisitions. We begin with targets in Chapter 11 of the U.S. Bankruptcy Code, where a decision to put the bankrupt firm up for sale is driven jointly by incumbent management and creditor committee votes. We then consider targets sold in the automatic auction bankruptcy system in Sweden. This code essentially eliminates the target's bargaining opportunities and relies on bidder competition to maximize debt recovery and an efficient reallocation of the target assets.

### *3.6.1. Chapter 11 targets*

Beginning with U.S. bankruptcies, there is growing use of market-based mechanisms to lower the costs of traditional Chapter 11 proceedings. These include prepackaged bankruptcies with a reorganization plan in place at filing (Betker, 1995; Lease, McConnell, and Tashjian, 1996), acquisition of distressed debt by "vulture" investors in order to make voting more efficient (Hotchkiss and Mooradian, 1997), and voluntary sales in Chapter 11 (Hotchkiss and Mooradian, 1998; Maksimovic and Phillips, 1998). Baird and Rasmussen (2003) report that more than half of all large Chapter 11 cases resolved in 2002 used the auction mechanism in one form or another, and that another quarter were prepacks.

Hotchkiss and Mooradian (1998) study acquisitions of targets in Chapter 11. There are two ways in which a firm in Chapter 11 can sell substantially all of its assets: through a Section 363 (of the U.S. Bankruptcy Code) sale or as part of a confirmed reorganization plan. Under a Section 363 sale, management must first obtain an offer and then notify the court, which in turn notifies creditors. The Bankruptcy Code invalidates no-shop agreements and allows creditors to retain advisers at the expense of the debtor firm to search for competing buyers. If there are several potential buyers, the court holds an auction.

Chapter 11 grants the incumbent management exclusive rights within a limited time period (rolling six months) to propose a reorganization plan. As a consequence, hostile acquisitions are difficult and the targets will be more likely for firms whose management has already been replaced or for which managerial private benefits of control are small. It is also possible that management is willing to put the target up for sale when it has private information that the target assets are of relatively low quality. Furthermore, since acquisition bids are subject to creditor approval (just as for any other reorganization plan), complex debt structure makes it more difficult to generate the necessary votes. Thus, targets are also likely to have relatively simple capital structures.

Hotchkiss and Mooradian (1998) start with 1,200 public companies that filed for Chapter 11 between October 1970 and December 1992. Using SEC and Compustat information, they identify 339 firms that reorganized as independent public companies and 111 firms that were acquired by another operating company. Of these, 55 acquirers are publicly traded firms. Target firms spend a median time in bankruptcy of 14 months, compared to 17 months for independently reorganized firms. They find little evidence that acquired firms have unusually simple capital structures (although they tend to have less public debt) or that incumbent management is particularly entrenched. Acquirers tend to be firms in the same industry as the target and have some prior relationship with the target such as an ownership stake. Of the 55 takeovers, 18 transactions have multiple bidders.

Hotchkiss and Mooradian (1998) also report that the bankrupt targets on average are purchased at a 45% discount relative to prices paid for nonbankrupt targets in the same industry. However, they do not consider this as evidence of allocative inefficiency: "Although the transactions are at discount prices, the high proportion of acquirers operating in the same industry as the target, as well as the competitive bidding environment, does not support the conclusion that acquisitions in bankruptcy are sales to lower value users" (p. 243). This conclusion is further supported by their finding that the postmerger cash flow performance of firms combined with bankrupt targets is better than that reported by Hotchkiss (1995) for firms emerging from Chapter 11. Finally, there is evidence of positive and significant abnormal stock returns to both bidders and bankrupt targets for the days surrounding the announcement of the acquisition.

### *3.6.2. Bankruptcy auctions and fire sales*

Next, we consider bankruptcies in Sweden's mandatory auction system. Here, a firm filing for bankruptcy is turned over to a court-appointed trustee who puts the firm up for sale in an auction. This mandatory auction system has an attractive simplicity. All debt claims are stayed during the auction period and the bids determine whether the firm will be continued as a going concern or liquidated piecemeal. A going-concern sale takes place by merging the assets and operations of the auctioned firm into the bidder firm, or into an empty corporate shell—much like a leveraged buyout transaction. Payment must be in cash, allowing the auction proceeds to be distributed to creditors strictly according to absolute priority.

As surveyed by Hotchkiss, John, Mooradian, and Thorburn (2008), bid premiums observed in the mandatory auction bankruptcy system in Sweden provide an important empirical perspective on the viability auctions as a mechanism for resolving bankruptcy. Proponents of the market-oriented auction system point to costs associated with conflicts of interests and excessive continuation of operations due to managerial control over the restructuring process in Chapter 11.[88] These costs most likely explain the trend toward increased use of market-based mechanisms in the United States. On the other hand, opponents of an auction-based system argue that the time pressure of an auction system is costly as it may cause excessive liquidation and fire sales of economically viable firms when potential bidders in the auction are themselves financially constrained.

A series of papers study the Swedish auction system using a sample of 260 auctioned firms.[89] The average auctioned firm has \$5 million in sales and assets of \$2 million

[88] See, for example, Baird (1986), Bebchuk (1988), Jensen (1989), Aghion, Hart, and Moore (1992), Bebchuk and Chang (1992), Bradley and Rosenzweig (1992), and Baird (1993). Hotchkiss (1995) finds that firms emerging from Chapter 11 tend to underperform their industry rivals which is consistent with excessive continuation.

[89] Thorburn (2000), Strömberg (2000), Eckbo and Thorburn (2003, 2007, 2008b).

($8 million and $4 million, respectively, in 2007 dollars), and it has an average of 45 employees.[90] Thorburn (2000) reports that the auctions are quick—lasting an average of two months—and relatively cost-efficient. Moreover, three-quarters of the filing firms survive the auction as a going concern, which is similar to the survival rate of Chapter 11 cases. In going-concern sales, the buyer typically rehires lower-level employees. Top management fares less well: Eckbo and Thorburn (2003) find that while the buyer rehires the old management to run the restructured company in about one-half of the going-concern sales, the old management typically experiences a median wealth decline of $-47\%$ relative to managers of nonbankrupt firms. They argue that this expected personal bankruptcy cost, along with the loss of private benefits of control, counteract shareholder risk-shifting incentives when the firm is in severe financial distress (Jensen and Meckling, 1976). That is, if the CEO's objective includes being rehired by the buyer in the auction, she may implement a relatively conservative investment policy to preserve the possibility of a going-concern sale in the auction.

Does the auction mechanism induce an efficient reallocation of the resources of the bankrupt firm? First, Eckbo and Thorburn (2003) show that firms sold as going-concerns typically perform at par with industry rivals. Second, Eckbo and Thorburn (2007) fail to find auction fire-sale discounts in going-concern sales. That is, the auction produces auction premiums (and post-bankruptcy operating performance) in going-concern sales that are independent of fire-sale conditions such as industrywide financial distress, industry leverage, and whether or not the buyer is an industry insider or outsider.

Third, Eckbo and Thorburn (2007) find that prepackaged auctions (where the buyer has been identified prior to filing) tend to produce prices consistent with the hypothesis that the contracting parties are concerned with preempting piecemeal liquidation. Strömberg (2000) shows that salebacks to the previous owner-manager tend to increase during periods of high industry financial distress, which further helps preempt liquidation. Eckbo and Thorburn (2007) document that prices paid in salebacks are as high as prices in non-saleback going-concern transactions, which fails to support arguments that salebacks carry an inherent conflict of interest with junior creditors.

### *3.6.3. Testing for auction overbidding*

Eckbo and Thorburn (2008b) develop and test the argument that creditor incentives may induce auction overbidding. Recall from Section 3.3 that toehold bidding raises the optimal bid above the bidder's own private valuation of the target, for example, as shown in Equation (11). In the sample of Swedish bankruptcies, the main creditor is always a single bank. Thus, the toehold analogy is that the bankruptcy event effectively creates an instant "creditor toehold" of $\alpha = 1$ when the creditor's debt is impaired at filing. The question is whether the existence of this creditor toehold leads to overbidding in the

[90] A majority of Chapter 11 filings are also by small private firms: Chang and Schoar (2007) report average sales of $2 million and 22 employees in a large and representative sample of Chapter 11 filings between 1989 and 2003. Bris, Welch, and Zhu (2006) report that the median firm filing for Chapter 11 has assets of $1 million.

auction. Given the importance of toehold bidding in the takeover literature, we outline the main test procedure and results below.

Swedish bank regulations prevent the bank from bidding directly in the auction. However, Eckbo and Thorburn (2008b) report that the bank often finances the winning bidder and uses this observation to motivate the following proposition: Bank financing allows the bank to induce bidder 1 to submit a bid $b_c^*$ that involves overbidding and is jointly optimal for both parties.[91] As in Section 3.3, overbidding forces a wealth transfer from bidder 2 to the bank-bidder coalition when bidder 2 wins the auction. This rent transfer raises auction revenue and the bank's expected debt recovery rate.

Suppose the bank forms a coalition with bidder 1. Continuing the notation from Section 3.3, the coalition's optimal bid is as follows:

$$b_c^* = \begin{cases} v_1 + h(b_c) & if \quad v_1 \leq f - h(b_c) \qquad (\textit{unconstrained overbidding}) \\ f & if \quad f - h(b_c) < v_1 < f \quad (\textit{constrained overbidding}) \\ v_1 & if \quad v_1 \geq f, \qquad\qquad\quad (\textit{no overbidding}) \end{cases} \tag{15}$$

where $f$ is the face value of the bank's debt claim. Note that the unconstrained overbidding price is identical to the bid in Equation (11) but with $\alpha = 1$ and a termination fee $t = 0$. A value of $\alpha = 1$ follows because the bank, being the secured creditor with an impaired debt claim, is effectively the seller of the auctioned firm. Thus, the bank has a creditor toehold equal to one. As shown by Hotchkiss and Mooradian (2003) as well (in the context of Chapter 11 sales), a creditor toehold induces overbidding in exactly the same manner as a bidder toehold outside of bankruptcy.

What makes this overbidding theory testable is the constraining effect of the bank-debt face value $f$.[92] To illustrate, let $l$ denote the piecemeal liquidation value of the bankrupt firm, and suppose $l$ is public knowledge at the beginning of the auction. Since $l$ is the sum of the value of the firm's assets if sold individually, it constitutes a price floor in the auction of the firm as a going-concern. Let $r \equiv l/f \in [0, 1]$ denote the bank's debt recovery if the firm is liquidated piecemeal. $r$ is a measure of the bank's debt impairment: low values of $r$ indicate that the bank's debt is highly impaired. For low values of $r$, the bank-bidder coalition fully overbids (unconstrained overbidding). However, as the value of $r$ increases, the amount of overbidding becomes constrained by $f$: the coalition optimally overbids *only* to the extent that overbidding does not benefit junior creditors. If the valuation of the bank's coalition partner is such that $v_1 > f$, the bank will receive full debt recovery even without overbidding, so the optimal coalition bid is simply $b_c^* = v_1$.

[91] The bank may induce the bidder to bear the expected overpayment cost by granting a lower interest on the loan. Eckbo and Thorburn (2008b) show that there exists a positive transfer from the bank to bidder 1 which makes coalition formation incentive compatible for both parties.

[92] This testable restriction does not exist for takeovers outside of bankruptcy. Extant empirical evidence on toehold-induced overbidding is therefore indirect. For example, theory implies that overbidding increases the probability of winning, which is supported by studies of corporate takeover bids with equity toeholds (Betton and Eckbo, 2000).

Eckbo and Thorburn (2008b) prove that the greater the liquidation recovery rate $r$, the lower is the incentive to overbid and, in turn, the lower is the expected premium paid by the winning bidder. They use a professional estimate of the piecemeal liquidation value $l$, published by the bankruptcy trustee at the beginning of the auction.[93] They find that when the firm is sold as a going-concern, final auction premiums are higher the lower is the liquidation recovery rate, as predicted by overbidding. Equally important, in subsamples where the theory implies zero overbidding incentive, the cross-sectional regressions reject overbidding. That is, final auction premiums are unaffected by the liquidation recovery rate when the auction leads to the target being liquidated piecemeal (in which case the going-concern premium is zero), or when the bank's collateral exceeds the face value ($l > f$) so the bank's debt is not impaired.

Overbidding results in allocative inefficiency whenever the bank-bidder coalition wins against a higher-valuation bidder. To examine this possibility, Eckbo and Thorburn (2008b) estimate the post-bankruptcy operating performance of firms sold as going-concerns conditional on the bank-bidder coalition having large overbidding incentives and winning the auction. While this is the most powerful subsample to look for ex-post allocative inefficiency, they show that the post-bankruptcy operating performance in this subsample is at par with or exceeds that of industry rivals. Overall, they conclude from this that the bank's coalition partner tends to be efficient in terms of restructuring and operating the bankrupt firm's asset.

### 3.7. *Offer premium summary*

Reflecting restrictions on the availability of actual offer prices, the bulk of the empirical studies on takeovers are content to use target cumulative abnormal stock returns around the takeover bid as a proxy for the actual offer premium. Obviously, target abnormal stock returns present noisy estimates of offer premiums because they incorporate the probability of bid failure and competition at the initial offer date, and they must be estimated over a long event window to capture the final premium. Thus, it is difficult to properly sort out how bidders determine offer premiums unless one employs offer price data directly.

Several papers study offer prices directly. Bradley (1980) provides the first systematic offer price analysis in the context of public tender offers. Walkling (1985) uses offer premiums to predict tender offer success. Eckbo and Langohr (1989) examine the effect of disclosure rules and method of payment (cash versus stock) on tender offer premiums. Betton and Eckbo (2000) examine bid jumps and offer premium determinants in tender offers. Officer (2003, 2004), Bates and Lemmon (2003), and Bargeron (2005) examine the premium effects of deal protection devices such as termination and share-tendering agreements. Betton, Eckbo, and Thorburn (2007) study the premium effects of toehold

[93] Bidders appear to rely on this estimate as well: when the auction does lead to piecemeal liquidation, the average price paid by the winning bidder is close to (on average 8% above) the trustee's estimate. In contrast, when the bankrupt firm is purchased as a going-concern, the average auction premium more than doubles the trustee's piecemeal liquidation value estimate.

bidding, while Betton, Eckbo, and Thorburn (2008b) are the first to estimate the effect of target runups on markups in initial and final offer prices. Chatterjee, John, and Yan (2008) study the effect of divergence of opinion on bid prices, while Levi, Li, and Zhang (2008) examine whether CEO and director gender affect takeover premiums.

Table 5 shows the cross-sectional determinants of both the initial and final offer premiums. The offer price data used for this table is from Betton, Eckbo, and Thorburn (2008b),

Table 5

Determinants of the offer premium in 4,889 control contests for U.S. public targets, 1980–2002.

The table shows OLS estimates of the initial and final offer premium, defined as $ln(p_{initial}/p_{-42})$ and $ln(p_{final}/p_{-42})$, respectively, where $p_{initial}$ is the initial offer price, $p_{final}$ is the final offer price in the contest, and $p_{-42}$ is the stock price on day $-42$ adjusted for splits and dividends. Amihud liquidity is the average value of $|R_i|/(p_i S_i)$ over days $i \in \{-250, -42\}$, where $R_i$ is the % holding period return, $p_i$ is the closing price and $S_i$ is the number of shares traded. Industry is the four-digit SIC code in CRSP. The sample is 4,889 control bids 1980–2002 for U.S. targets with a stock price $\geq$ \$1 and a market capitalization $\geq$ \$10 million. $p$-values are in parenthesis. The data is from Betton, Eckbo, and Thorhurn (2007).

| | Initial offer premium | | Final offer premium | |
|---|---|---|---|---|
| Mean | 0.43 | | 0.48 | |
| Median | 0.37 | | 0.39 | |
| St. dev. | 0.46 | | 0.47 | |
| Intercept | 0.296 | 0.256 | 0.296 | 0.254 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| **A: Target characteristics** | | | | |
| Size: ln of target market capitalization on day −42 | −0.030 | −0.027 | −0.030 | −0.027 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Target book-to-market > industry median | 0.025 | 0.029 | 0.024 | 0.029 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Target runup: $ln(p_{-1}/p_{-42})$ | 0.808 | 0.811 | 0.808 | 0.811 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Amihud liquidity | 8.55 | 13.29 | 8.71 | 13.46 |
| | (0.311) | (0.114) | (0.302) | (0.110) |
| Poison pill dummy | −0.016 | 0.000 | −0.016 | −0.001 |
| | (0.606) | (0.990) | (0.604) | (0.987) |
| **B: Bidder characteristics** | | | | |
| Positive toehold (vs. zero toehold) | −0.023 | −0.025 | −0.023 | −0.025 |
| | (0.032) | (0.024) | (0.032) | (0.024) |
| Acquirer public (vs. private) | 0.015 | 0.023 | 0.015 | 0.023 |
| | (0.079) | (0.008) | (0.072) | (0.007) |
| Horizontal takeover (same industry) | −0.004 | −0.004 | −0.004 | −0.004 |
| | (0.608) | (0.664) | (0.618) | (0.673) |

(*Continued*)

Table 5 (*Continued*)

| | | | | |
|---|---|---|---|---|
| **C: Deal characteristics** | | | | |
| Tender offer (vs. merger) | −0.061<br>(0.000) | −0.066<br>(0.000) | −0.061<br>(0.000) | −0.066<br>(0.000) |
| All cash consideration (vs. stock or mixed) | 0.019<br>(0.017) | 0.021<br>(0.012) | 0.019<br>(0.017) | 0.021<br>(0.012) |
| Hostile target response (vs. friendly or neutral) | 0.020<br>(0.185) | 0.020<br>(0.185) | 0.019<br>(0.216) | 0.019<br>(0.216) |
| Multiple bidders (vs. single-bidder contest) | | | 0.009<br>(0.497) | 0.008<br>(0.501) |
| Announced in 1980–1989 (vs. 1990–2002) | −0.016<br>(0.056) | | −0.017<br>(0.050) | |
| Year fixed effects | no | yes | no | yes |
| Adjusted $R^2$ | 0.424 | 0.436 | 0.423 | 0.436 |
| F-value | 300.3<br>(0.000) | 115.3<br>(0.000) | 277.2<br>(0.000) | 111.9<br>(0.000) |

and covers a total of 4,889 targets. The premiums are defined as $ln(p_{initial}/p_{-42})$ and $ln(p_{final}/p_{-42})$, respectively, where $p_{initial}$ is the initial offer price, $p_{final}$ is the final offer price in the contest, and $p_{-42}$ is the stock price on day −42 adjusted for splits and dividends. The sample is restricted to targets in the period 1980–2002 with a stock price $\geq$ \$1 and a market capitalization $\geq$ \$10 million. As shown in the first two rows of the table, the mean (median) value of the initial offer premium is 43% (37%), which increases to 48% (39%) by the time of the final bid.

The explanatory variables, which are grouped into target characteristics, bidder characteristics, and deal characteristics, cover the types of decisions discussed throughout Section 3. We alternately use a time dummy for offers taking place in the early sample period (1980–1989) and year fixed effects. Notice also that the information in these variables is known at the time the offer premium was set. We include the variable hostile target response as a determinant of the initial offer premium because we believe this information is basically known at the outset. However, the variable multiple bidders obviously is not and is included as a determinant of the final offer premium only.

Not surprisingly (given the relative paucity of multiple-bid contests in the total sample of 4,889), the explanatory variables have similar coefficients and level of significance for both the initial and final offer premiums. In the order of the discussion of this section, the initial and final offer premiums are

1. significantly *higher* when the bidder is a public company and significantly *lower* if the initial bid is a tender offer (Section 2.3).

2. significantly *greater* when the method of payment is all cash (Section 3.2).
3. significantly *lower* when the bidder has a positive toehold (Section 3.3).
4. significantly *greater* the greater the target runup $ln(p_{-1}/p_{-42})$ prior to the initial bid (Section 3.4). [94]
5. *unaffected* by either the presence of a target poison pill or target hostility to the initial bid (Section 3.5).

Table 5 further shows that the initial and final offer premiums are decreasing in target total equity capitalization on day −42, and they are greater if the target's book-to-market ratio exceeds the industry median B/M (i.e., if the target has few growth options relative to industry rivals). Offer premiums are unaffected by target stock liquidity, by the presence of multiple bidders, and by whether the bidder and target are horizontally related in product markets. Finally, offer premiums have increased from the 1980s.[95]

Officer (2003) and Bates and Lemmon (2003) show that offer premiums are significantly greater when the SDC indicates the existence of a target termination agreement, while Bargeron (2005) finds lower premiums in the presence of a target board/management tender agreement. Moeller (2005) presents evidence indicating that powerful entrenched target CEOs reduce takeover premiums. Chatterjee, John, and Yan (2008) find that takeover premiums are larger the greater the disagreement between the earnings forecasts of financial analysts following the target. Levi, Li, and Zhang (2008) use RiskMetrics Group data on board structure and find that bid premiums are affected by the gender composition of the board. Specifically, bid premiums are lower when the bidder CEO is female, and the higher the target board's proportion of female directors (provided that the female directors are independent appointees).

Several of the variables used to explain the offer premium are themselves endogenous choice variables (payment method, toehold, hostility, termination agreements, bidder's public status), Some of the reported effects appear robust to endogeneity.[96] One variable that does *not* appear to be robust is "tender offer." The inclusion of other variables (such as toeholds and hostility) tends to affect conclusions as to whether offer premiums are higher, the same, or lower in tender offers than in merger bids. Additional specification analysis is needed to fully sort out the endogenous from truly exogenous forces in the data.

[94] The coefficient on the runup variable is 0.80. This means that a dollar increase in the target runup causes the bidder to raise the offer price by 80 cents on average. Betton, Eckbo, and Thorburn (2008b) also show that offer markups (either $ln(p_{initial}/p_{-1})$ or $ln(p_{final}/p_{-1})$) are significantly decreasing in the runup. Thus, there is partial substitution between runups and markups.

[95] Since most of the hostile bids occurred in the 1980s, this is consistent with the finding that offer premiums in hostile bids are no lower than those for nonhostile offers.

[96] Betton and Eckbo (2000), Officer (2003) and Betton, Eckbo, and Thorburn (2007) use systems of equations and various corrections for self-selection. See Li and Prabhala (2007) for a survey of self-selection models in corporate finance.

## 4. Takeover gains

In this section, we present estimates of abnormal stock returns to bidders and targets around takeover contests, as well as in the post-merger period. Given the large number of papers providing abnormal returns estimates in takeovers, we limit the review to more recent studies with large samples of 1,000 or more bidder firms, such as those listed in Table 6. Studies are included in the table only if announcement-induced abnormal returns to bidders are in fact reported. This excludes large-sample studies such as Schwert (1996) and Bates and Lemmon (2003), where the main focus is on targets or some other deal aspect and where bidder returns may be estimated and used for purposes of cross-sectional regressions—but average announcement returns are not reported. It also excludes almost all studies before SDC became available as a convenient online data source.[97]

### *4.1. Econometric caveats*

Abnormal stock returns measure only the unanticipated component of the total economic effect of the event. Given the difficulty in predicting target firms, partial anticipation of the bid announcement does not pose much of an econometric problem for studies of target takeover gains. Most researchers agree that one captures most, if not all, of the total target gains by comparing the offer price to the pre-offer target share price within two months of the first bid. As illustrated in this section, the bulk of the target pre-offer runup typically actually occurs within 10 days of the bid.

It is also widely understood that partial anticipation can severely complicate estimation of gains from bidding. Any partial anticipation must somehow be accounted for to avoid underestimating the value implications. In simple environments with only a single possible event, the announcement effect equals the valuation effect times one minus the probability that the merger event will occur. It is thus attenuated toward zero, creating a bias against rejection of the null of zero gains from bidding. Malatesta and Thompson (1985) directly model the information arrival process and conclude that bidder stock returns include a component due to partial anticipation of future acquisition activity. Eckbo, Maksimovic, and Williams (1990) model the probability of the takeover event and conclude that this probability affects estimates of bidder takeover gains. The conclusion from these studies is that partial anticipation of bidding activity is an important empirical issue when the researcher fails to reject the hypothesis of zero abnormal stock returns to bidders.

Another approach to dealing with partial anticipation is through various sampling techniques. For example, Schipper and Thompson (1983) sample firms that announce entire acquisition programs. Since this announcement capitalizes a whole series of future expected acquisitions (rather than responding to a single-acquisition announcement), power to detect true acquisition gains is enhanced. Their evidence is consistent with

[97] Two exceptions in Table 6 are Loderer and Martin (1990) and Betton and Eckbo (2000), who use large hand-collected samples.

Table 6
Large-sample ($N > 1{,}000$) estimates of announcement-induced average cumulative abnormal stock returns (ACAR) to U.S. bidders.

| Study | Sample | Average announcement return: ACAR (day $\tau_1$, day $\tau_2$)($*$ =significant at 10% level) |
|---|---|---|
| Loderer and Martin (1990) | $N = 1{,}135$ completed mergers, 274 completed tender offers, and 3,296 "other" acquisitions (not classifiable as merger or tender offer) by public acquirers, where the offer is announced in the *Wall Street Journal*, 1996–1984. | ACAR($-5$, 0) is 1.7%* for 970 cases 1966–1968, 0.57%* for 3,401 cases 1960–1980, and $-0.1$% for 801 cases 1981–1984. Bidder announcement returns smaller for larger bidders and decreasing in the relative size of the target firm. |
| Betton and Eckbo (2000) | Initial and rival bidders in $N = 1{,}353$ tender offer contests for public targets, 1971–1990. | (1) Day 0 is the initial bid date: ACAR($-60$, 0) is 1.3% for initial bidders and 2.2% for rival bidders. (2) Day 0 is the second bid date: ACAR($-60$, 0) is 1.2% for initial bidders, and 6.1%* for rivals. |
| Fuller, Netter, and Stegemoller (2002) | $N = 3{,}135$ takeovers, 1990–2000, by 539 public acquirers with at least 5 successful control bids within three years. Minimum deal size is \$1 million. | ACAR($-2$, 2) is 1.8%* for total sample of bidders, $-1.0$%* when target is public, 2.1%* when target is private, and 2.8%* when target is a subsidiary. |
| Akbulut and Matsusaka (2003) | $N = 3{,}466$ successful mergers between public firms, 1950–2002. | ACAR($-2$, 1) is 1.2%* for "related" acquisitions (bidder and target have at least one 3-digit SIC code in common) and 1.1%* for unrelated acquisition. |
| Officer (2003, 2004) | $N = 2{,}511$ attempted mergers and tender offers between public firms, 1988–2000 (Officer, 2003). | ACAR($-3$, 3) is $-1.2$%* for the total sample. |
| Moeller, Schlingemann, and Stulz (2004, 2005) | $N = 12{,}023$ acquisitions, 1980–2001. Minimum deal value is \$1 million and 1% of the acquirer's assets. | ACAR($-1$, 1) is 1.1%* for total sample, 2.3%* for small acquirers, and 0.1% for large acquirers. Using dollar values, bidders loose a total of \$221 billion in market capitalization over day $-1$ to $+1$. This aggregate loss is driven by a small number of very large deals concentrated to the 1998–2001 period. |
| Bhagat, Dong, Hirshleifer, and Noah (2005) | $N = 1{,}018$ tender offers for public targets. | ACAR($-5$, 5) is 0.2% with a median dollar return of $-1.2$ million. |
| Song and Walkling (2005) | $N = 3{,}389$ acquisitions, 1985–2001. Minimum deal value is \$10 million. | ACAR($-1$, 0) for the first bidder after a 12-month dormant period in the industry is 0.7%*, and 0.04% for subsequent bidders. Consistent with an attenuation effect of partial anticipation of takeover activity. |

(*Continued*)

Table 6 (*Continued*)

| Study | Sample | Average announcement return: ACAR (day $\tau_1$, day $\tau_2$)($*$ =significant at 10% level) |
|---|---|---|
| Bradley and Sundaram (2006) | $N = 12{,}476$ completed acquisitions by 4,116 public companies, 1990–2000. | ACAR(−2, 2) is 1.4% for the total sample, −0.7% for public targets, and 1.9% when the target is private. Bidding firms experience a large stock price runup over the two-years period preceding the bid. This runup is greater for stock bids than for cash bids. |
| Savor (2006) | $N = 1{,}484$ (159 failed and 1,335 successful) merger bids, 1990–2000. The bid is nonhostile and all-cash (359 successful cases) or all-equity (976 successful cases). Minimum deal size is 5% of bidder market value. | ACAR(−1, 1) is −3.5%* for all-stock bidders and 1.0%* for all-cash bidders. Similar results for the full sample of failed acquirers. |
| Dong, Hirshleifer, Richardson, and Teoh (2006) | $N = 3{,}137$ merger bids and tender offers between public firms, 1978—2000. | ACAR(−1, 1) ranges from −0.2% (when target is ranked as relatively "undervalued") to −1.8 (when target is ranked as relatively "overvalued")% |
| Moeller, Schlingemann, and Stulz (2007) | $N = 4{,}322$ all-cash and all-stock bids, 1980–2002. Minimum deal value is \$1 million and 1% of the acquirer's assets. | ACAR(−1, 1) for the total sample is 0.8%. When target is public, ACAR(−1, 1) is −2.3% in all-stock deals and 0.7% in all-cash deals. When the target is private, ACAR(−1, 1) is 3.4% in all-stock deals. |
| Bargeron, Schlingemann, Stulz, and Zutter (2007) | $N = 1{,}292$ completed all-cash takeovers of US public targets by private and public bidders, 1990–2005. | Average target announcement $CAR_{(-1,1)}$ is 32% for public bidders and 22% for private bidders. |
| Betton, Eckbo, and Thorburn (2007, 2008b) | $N = 10{,}806$ initial control bids for public targets: 7,076 merger bids from 1980–2002 and 3,730 tender offers from 1973–2002. | ACAR(−1, 1) is −1.2%* for total sample and −0.15%* if the bidder has a toehold. In Betton, Eckbo, and Thorburn (2008b), ACAR(−1, 1) is −1.9%* for merger offers, and an in significant 0.3% for tender offers. |
| Betton, Eckbo, and Thorburn (2008c) | $N = 15{,}987$ initial control bids by public bidders for public or private targets, 1980–2005: 13,985 merger bids and 1,468 tender offers. | ACAR(−1, 1) is 0.69% with a significantly *negative* $z$-statistic of −3.9 for initial bidders in mergers, and 0.76 (insignificant) for initial bidders in tender offers. Large public bidders acquiring public targets and paying with allcash produces ACAR(−1, 1) of −2,2%*. Small public bidders acquiring private targets in allstock offers produces ACAR(−1, 1) of 6,5%*. Details are in Tables 7, 8 and 9 in this survey. |

(*Continued*)

Table 6 (*Continued*)

| Study | Sample | Average announcement return: ACAR (day $\tau_1$, day $\tau_2$)($*$ =significant at 10% level) |
|---|---|---|
| Hackbarth and Morellec (2008) | $N = 1{,}086$ completed takeovers between public firms, 1985–2002. Minimum transaction value of \$50 mill., and regulated and financial firms ar excluded. | ACAR($-1$, 1) is $-0.5\%^*$. Bidder risk changes around the acquisition events are found to be consistent with a neoclassical investment model. |

the hypothesis that future expected acquisitions have positive net present value as a group. Song and Walkling (2000, 2005) select takeover announcements that follow a dormant period—with no previous takeovers in the industry of the bidder for a minimum of 12 months. Presumably, these announcements come as a relative surprise to the market, adding power to reject the null of zero bidder abnormal returns. Perhaps as a direct result, the authors report significantly positive bidder announcement returns.

Takeover announcements may also reveal new information about the quality of the bidder's management team—regardless of the value of the proposed acquisition per se. This further confounds the interpretation of bidder announcement returns as gains from merger activity. One approach is to formally model the signaling problem and test for its existence using cross-sectional regressions with bidder announcement returns as dependent variable (Eckbo, Giammarino, and Heinkel, 1990). Fuller, Netter, and Stegemoller (2002) approach this issue by selecting a sample of frequent acquirers (firms that acquire five or more targets within a three-year period). This sampling strategy helps control for certain bidder characteristics in the cross section.

Finally, because bidder managers time takeovers based on private information, consistent estimation of parameters in cross-sectional models with bidder returns as the dependent variable requires a correction for self-selection (Eckbo, Maksimovic, and Williams, 1990). While such cross-sectional regressions are commonly presented in the literature, this (or other equivalent) correction is rarely implemented. However, the recent review of Li and Prabhala (2007) is likely to increase general awareness of the importance of providing unbiased estimates in these cross-sectional models.[98]

### *4.2. Runup- and announcement-period returns*

We estimate the average daily abnormal stock return for firm $j$ over event window $k$ as the event parameter $AR_{jk}$ in the value-weighted market model

$$r_{jt} = \alpha_j + \beta_j r_{mt} + \sum_{k=1}^{K} AR_{jk} d_{kt} + \epsilon_{jt}, \quad t = day\{-293, \ldots, end\} \tag{16}$$

[98] Note that self-selection poses an econometric issue in cross-sectional regressions with the target abnormal return as dependent variable only to the extent that the target self-selects the timing of the acquisition.

where $r_{jt}$ is the return (in logarithmic form) to firm $j$ over day $t$, $r_{mt}$ is the value-weighted market return, and $d_{kt}$ is a dummy variable that takes a value of one if day $t$ is in the $k$th event window and zero otherwise.[99] This conditional event parameter estimation yields identical abnormal return estimates as the more standard residual analysis technique, but is more efficient in terms of using the available return data. Moreover, the regression easily incorporates variable-length event windows across takeovers, and it produces estimates of standard errors of the abnormal returns directly.[100]

Day 0 is the day of the initial control bid, and the ending date is the earlier of the day of the control last bid in the contest plus 126 trading days and the effective date + 126. If the target delisting date is between the date of the last control bid and the effective date, then the contest end is set to the target delisting date. The runup and announcement abnormal returns are estimated using three event windows ($K = 3$). The three event windows are $[-41,\ -2]$ (the runup period), $[-1,\ 1]$ (the announcement period), and $[2,\ end]$. The estimation uses Ordinary Least Squares (OLS) with White's heteroscedastic-consistent covariance matrix and requires a minimum of 100 days of nonmissing returns during the estimation period.

The cumulative abnormal return (CAR) to firm $j$ over event period $k$ is

$$\mathrm{CAR}_{jk} = \omega_k AR_{jk} \tag{17}$$

where $\omega_k$ is the number of trading days in the event window. In a sample of N firms, the average cumulative abnormal return (ACAR) is

$$\mathrm{ACAR}_k = (1/N)\sum_j \mathrm{CAR}_{jk} \tag{18}$$

The $z$-values are determined as

$$z = (1/\sqrt{N})\sum_j AR_{jk}/\sigma_{AR_{jk}} \tag{19}$$

and $\sigma_{AR_{jk}}$ is the estimated standard error of $AR_{jk}$. Under the null of $\mathrm{ACAR} = 0$, $z \sim N(0, 1)$ for large N. The combined bidder and target abnormal returns are determined by weighting the bidder and target abnormal returns by the market capitalization on day $-42$.

The twin Tables 7 and 8 detail the average abnormal return estimates (CAR) for the runup period $(-42, -2)$, the announcement period $(-1, 1)$, classified by market capitalization (Panel B), the public status of the bidder and target firms (Panel C), merger

[99] The return analysis is limited to ordinary shares. Missing returns are dealt with as follows: A succession of less than six missing returns are backfilled by allocating the cumulative return equally over the missing days. For example, if there are three missing days and then a return of 10%, each missing day and the subsequent nonmissing day would be allocated a return on 2.5%.

[100] For reviews of event study econometrics, and the conditional event parameter approach used here, see Thompson (1985, 1995), MacKinlay (1997), and Kothari and Warner (2007).

Table 7

Cumulative abnormal stock returns (CAR) to targets and bidders (individually and combined) relative to the initial bid date. Sample of control contests for U.S. targets, 1980–2005.

See the text for the details of the abnormal return estimation. The average market capitalization on day −42 for the target ($V_T$) and bidder ($V_B$), and for the ratio $V_T/V_B$, are reported in brackets (in \$1,000). Day 0 is the day of the initial control bid. The combined bidder and target abnormal returns are determined by weighting the bidder and target abnormal returns by the market capitalization on day −42.

| | Target CAR | | | Initial Bidder CAR | | | Combined CAR | | |
|---|---|---|---|---|---|---|---|---|---|
| | N (Average $V_T$) | Runup (−41,−2) | Ann'ct (−1,1) | N (Average $V_B$) | Runup (−41,−2) | Ann'ct (−1,1) | N (Av. $V_T/V_B$) | Runup (−41,−2) | Ann'ct (−1,1) |
| **A: Entire sample** | | | | | | | | | |
| Mean | \$9,298 | 0.0680 | 0.1461 | 15,987 | 0.0049 | 0.0073 | 4,803 | 0.0072 | 0.0106 |
| Median | (\$641,951) | 0.0516 | 0.1234 | (3,752,262) | −0.0024 | −0.0005 | (0.4445) | 0.0070 | 0.006 |
| Z | | 25.2701 | 102.2990 | | −2.1463 | −2.5297 | | 4.2496 | 14.605 |
| % positive | | 0.6231 | 0.8271 | | 0.4924 | 0.4939 | | 0.5259 | 0.5640 |
| **B: Subsamples based on market capitalization on day −42** | | | | | | | | | |
| Lowest quartile | | | | | | | | | |
| Mean | 2,324 | 0.1019 | 0.1472 | 3,995 | 0.0492 | 0.0404 | 1,200 | 0.0009 | −0.000 |
| Median | (11,207) | 0.0708 | 0.1200 | (28,762) | 0.0154 | 0.0087 | (0.0195) | −0.0005 | −0.001 |
| Z | | 11.9283 | 46.8157 | | 5.5730 | 21.7874 | | 0.5042 | 0.1303 |
| % positive | | 0.6248 | 0.7900 | | 0.5357 | 0.5827 | | 0.4958 | 0.4858 |
| Highest quartile | | | | | | | | | |
| Mean | 2,323 | 0.0365 | 0.1327 | 10,480 | −0.0122 | −0.0049 | 1,201 | 0.0155 | 0.0279 |
| Median | (2,372,966) | 0.0385 | 0.1156 | (5,698,863) | −0.0073 | −0.0035 | (1.3891) | 0.0149 | 0.019 |
| Z | | 11.2752 | 51.6770 | | −6.0704 | −17.5109 | | 4.1736 | 15.032 |
| % positive | | 0.6117 | 0.8429 | | 0.4757 | 0.4599 | | 0.5612 | 0.6295 |

(*Continued*)

Table 7 *(Continued)*

| | Target CAR | | | Initial Bidder CAR | | | Combined CAR | | |
|---|---|---|---|---|---|---|---|---|---|
| | N | Runup (−41,−2) | Ann'ct (−1,1) | N | Runup (−41,−2) | Ann'ct (−1,1) | N | Runup (−41,−2) | Ann'ct (−1,1) |
| **C: Subsamples based on the public status of the bidder and target firms** | | | | | | | | | |
| **Public target** | | | | | | | | | |
| Mean | 9,298 | 0.0680 | 0.1461 | 6,301 | 0.0065 | −0.0087 | 4,803 | 0.0072 | 0.0106 |
| Median | | 0.0516 | 0.1234 | | 0.0012 | −0.0066 | | 0.0070 | 0.0067 |
| Z | | 25.2701 | 102.2990 | | 0.5387 | −19.0462 | | 4.2496 | 14.6057 |
| % positive | | 0.6231 | 0.8271 | | 0.5034 | 0.4269 | | 0.5259 | 0.5640 |
| **Private target** | | | | | | | | | |
| Mean | | | | 9,686 | 0.0040 | 0.0176 | | | |
| Median | | | | | −0.0051 | 0.0029 | | | |
| Z | | | | | −3.1918 | 12.1118 | | | |
| % positive | | | | | 0.4852 | 0.5375 | | | |

Table 8

Cumulative abnormal announcement returns in control contests (continued from Table 15.7 ).

| | Target CAR | | | Initial Bidder CAR | | | Combined CAR | | |
|---|---|---|---|---|---|---|---|---|---|
| | N | Runup (−41,−2) | Ann'ct (−1,1) | N | Runup (−41,−2) | Ann'ct (−1,1) | N | Runup (−41,−2) | Ann'ct (−1,1) |
| **D: Subsamples based on form of initial bid** | | | | | | | | | |
| **Merger** | | | | | | | | | |
| Mean | 6,836 | 0.0619 | 0.1338 | 13,995 | 0.0050 | 0.0069 | 3,939 | 0.0071 | 0.0060 |
| Median | | 0.0481 | 0.1134 | | −0.0024 | −0.0008 | | 0.0070 | 0.0037 |
| Z | | 20.7051 | 88.2153 | | −2.2479 | −3.8858 | | 3.5536 | 7.7429 |
| % positive | | 0.6181 | 0.8212 | | 0.4921 | 0.4920 | | 0.5268 | 0.5380 |
| **Tender offer** | | | | | | | | | |
| Mean | 2,320 | 0.0868 | 0.1881 | 1,468 | 0.0060 | 0.0076 | 837 | 0.0090 | 0.0335 |
| Median | | 0.0693 | 0.1707 | | 0.0006 | 0.0011 | | 0.0073 | 0.0232 |
| Z | | 14.9492 | 52.7321 | | 0.5420 | 0.9110 | | 2.6312 | 18.4987 |
| % positive | | 0.6427 | 0.8573 | | 0.5014 | 0.5123 | | 0.5245 | 0.6941 |
| **E: Subsamples based on method of payment of initial offer** | | | | | | | | | |
| **All-Cash** | | | | | | | | | |
| Mean | 2,846 | 0.0765 | 0.2023 | 1,857 | −0.0039 | 0.0081 | 996 | 0.0011 | 0.0285 |
| Median | | 0.0523 | 0.1797 | | −0.0027 | 0.0025 | | 0.0039 | 0.0170 |
| Z | | 15.0345 | 65.3668 | | −1.1140 | 7.7136 | | 1.4279 | 18.6150 |
| % positive | | 0.6283 | 0.8949 | | 0.4890 | 0.5315 | | 0.5120 | 0.6837 |

(*Continued*)

Table 8 *(Continued)*

| | Target CAR | | | Initial Bidder CAR | | | Combined CAR | | |
|---|---|---|---|---|---|---|---|---|---|
| | N | Runup (−41,−2) | Ann'ct (−1,1) | N | Runup (−41,−2) | Ann'ct (−1,1) | N | Runup (−41,−2) | Ann'ct (−1,1) |
| **E: Subsamples based on method of payment of initial offer** | | | | | | | | | |
| **All-Stock** | | | | | | | | | |
| Mean | 2,163 | 0.0680 | 0.1396 | 5,189 | 0.0096 | 0.0025 | 1,909 | 0.0092 | −0.0030 |
| Median | | 0.0533 | 0.1215 | | −0.0003 | −0.0044 | | 0.0083 | −0.0016 |
| Z | | 12.6045 | 46.7639 | | 0.8922 | −10.0304 | | 3.0591 | −2.5634 |
| % positive | | 0.6301 | 0.8174 | | 0.4993 | 0.4531 | | 0.5343 | 0.4819 |
| **F: Subsamples based on time period of initial offer** | | | | | | | | | |
| **1991–1995** | | | | | | | | | |
| Mean | 1,601 | 0.0608 | 0.1344 | 3,654 | 0.0102 | 0.0169 | 941 | 0.0012 | 0.0131 |
| Median | | 0.0485 | 0.1141 | | −0.0027 | 0.0017 | | 0.0019 | 0.0079 |
| Z | | 9.0822 | 43.0355 | | −0.9762 | 6.6646 | | 0.4272 | 6.0191 |
| % positive | | 0.6121 | 0.8189 | | 0.4910 | 0.5216 | | 0.5090 | 0.5770 |
| **1996–2000** | | | | | | | | | |
| Mean | 3,008 | 0.0818 | 0.1564 | 5,464 | 0.0039 | 0.0074 | 1,816 | 0.0099 | 0.0072 |
| Median | | 0.0674 | 0.1372 | | −0.0017 | −0.0008 | | 0.0108 | 0.0052 |
| Z | | 16.6761 | 56.6973 | | −0.3038 | −1.4063 | | 3.6605 | 5.9602 |
| % positive | | 0.6483 | 0.8328 | | 0.4941 | 0.4929 | | 0.5391 | 0.5435 |

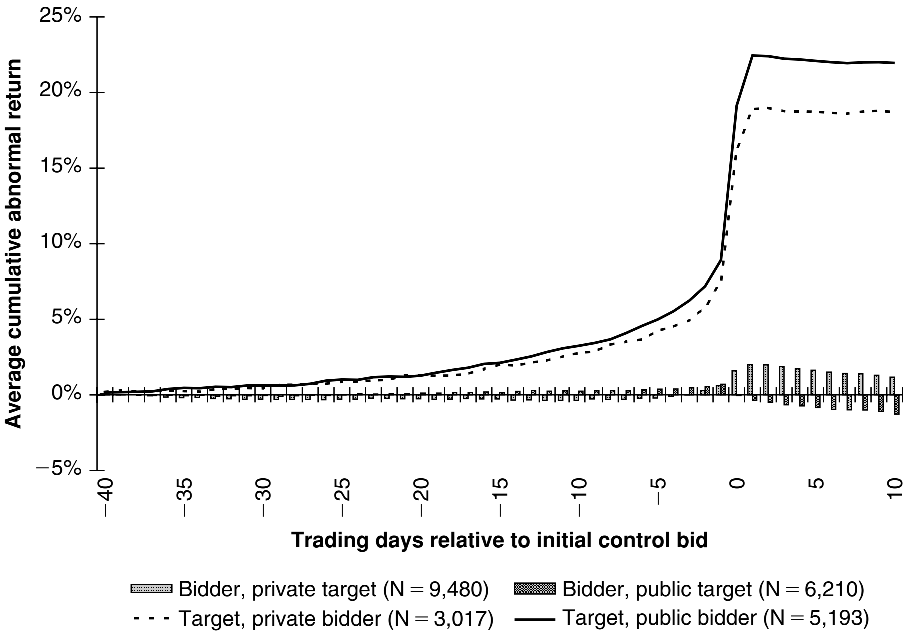



Fig. 10. Percent average cumulative abnormal stock returns to targets and initial bidders from day −40 through day 10 relative to the initial control bid. U.S. targets 1980–2005.

v. tender offer (Panel D), the payment method (Panel E), and, finally, the time period (Panel F). CAR is shown for the target, the initial bidder, and the value-weighted sum of the bidder and target CARs. For illustrative purposes, Figure 10 plots the daily cumulative abnormal returns from day −40 through day 10 relative to the initial offer announcement, classified by the public status of the bidder and target.[101] The cumulative abnormal returns to targets are somewhat greater when the bidder is public than when the bidder is private. Moreover, bidder returns are somewhat greater when the target is private than when the target is public.

Several overall conclusions emerge from the results in Tables 7 and 8 that are broadly consistent with the conclusions from the extant literature, including those listed in Table 6:

(1) *Target CARs*

(a) The average target CAR is positive and significant in all samples, over both the runup and the announcement period.

[101] The cumulative abnormal returns shown in the graph are estimated by including a dummy variable for each of the days in the (−42, +10) interval and adding the estimated dummy coefficients.

(b) The runup typically constitutes about one-third of the total runup plus announcement CAR. The largest target CAR occurs in all-cash offers (Panel E), where the sum of the runup and the announcement CAR is 28%.

(2) *Combined CARs (value-weighted)*

(a) The average combined CAR is positive and significant over the runup period for 9 of the 10 sample categories, and insignificant for the lowest size-quartile bidders (Panel B). The average combined runup-period CAR for the total sample of 4,803 cases is 0.7% with a $z$-value of 4.3.
(b) The average combined CAR is positive and significant for the announcement period for 8 of the 10 samples, insignificant in one (Panel E, for bidders in the lowest size quartile), and significantly negative in one (Panel E, when the payment method is all-stock). The average combined announcement-period CAR for the total sample of 4,803 cases is 1.06% with a $z$-value of 14.6.
(c) For the total sample (Panel A), the sum of the combined CAR for the runup and announcement periods is a significant 1.79%.

(3) *Bidder CARs*

(a) Announcement-period CAR is 0.73% for the total sample, but with a negative and significant $z$-statistic of $-2.53$.[102] The median CAR is $-0.05\%$, and the percentage of bidders with negative CAR is 49%.
(b) The average announcement-period bidder CAR is significantly positive for the lowest bidder size-quartile (Panel B), when the target is private (Panel C), in all-cash bids (Panel E), and in the period 1991–1995 (Panel F). It is significantly negative for bidders in the highest size-quartile (Panel F), when the target is public (Panel C), when the initial bid is a merger (Panel D), and when the payment method is all-stock (Panel E).
(c) The runup period bidder CAR is positive but largely insignificant, typically in the range 0.05% to 0.10%. Bidders in the lowest size quartile have a significantly positive average runup of 4.9%, and the average runup is a significant -1.2% for bidders in the highest quartile (Panel B). In these two subsamples, the runup is greater than the announcement return (and of the same sign).

This confirms several of the conclusions of the studies listed in Table 6, in particular Fuller, Netter, and Stegemoller (2002), Moeller, Schlingemann, and Stulz (2004, 2005), Bradley and Sundaram (2006), Savor (2006), Moeller, Schlingemann, and Stulz (2007), Bargeron, Schlingemann, Stulz, and Zutter (2007), and Betton, Eckbo, and Thorburn (2007, 2008b,c). Table 9 further highlights the impact of key offer characteristics on bidder announcement returns. The combination of large bidder (here in the upper-size quartile), payment in all-stock, and the target being a public company represents

[102] The average CAR and its $z$-statistic may differ in sign.

Table 9

Summary of initial bidder three-day announcement-period abnormal returns, 1980–2005.

Initial bidder cumulative abnormal returns for the window −1, 1 relative to the initial control bid. Large bidders are bidders in the upper quartile of market capitalization on day −42 (in constant 2000 dollars) and small bidders are bidders in the lower quartile of market capitalization on day −42. The cutoff values for the upper and lower quartiles are $134 million and $2.2 billion respectively. The method of payment is determined from the SDC 100% cash or 100% stock consideration field. The public status of the target is determined from SDC. ** represents significance at the 1% level (2 sided test).

| | Public targets | | Private targets | |
|---|---|---|---|---|
| **Sample** | N | CAR(−1, 1) | N | CAR(−1, 1) |
| **A: Large bidders** | | | | |
| All-stock | 769 | −0.0221** | 445 | 0.0010 |
| All-cash | 439 | −0.0030** | 88 | 0.0026** |
| **B: Small bidders** | | | | |
| All-stock | 495 | −0.0006 | 872 | 0.0646** |
| All-cash | 190 | 0.0306** | 184 | 0.0176** |

a worst-case scenario with average bidder announcement-period CAR of −2.21%. The best-case scenario is the combination of a small bidder (lower size-quartile), private target, and, again, *all-stock* as payment. This produces an average bidder announcement-period CAR of 6.46%. Thus, a major driver of negative bidder returns is not, as previously thought, the all-stock payment. Rather, the two key drivers appear to be the target's status as public or private and the bidder size. As shown next, bidder size was particularly large in 1999 and 2000, which suggests that the bidder size effect may also represent a unique time-period effect.

### *4.3. Dollar returns*

Figure 11 presents an annual scatter plot of the three-day announcement period bidder abnormal returns CAR(−1, 1) (Panel A) and the raw bidder dollar change from closing of day −2 to closing of day 1. As first noticed by Moeller, Schlingemann, and Stulz (2004, 2005), the distributions of the CAR(−1, 1) and the dollar change are dramatically different. Betton, Eckbo, and Thorburn (2008c) extend the sample period to 2005 and discover that the period 1998–2000 is unusual not only relative to the pre–1998 period, but also relative to the post–2000 years. Figure 12 further illuminates the role and effect of bidder size. Panel A plots bidder market values (in constant 2000 dollars) as of day −2. Clearly, bidders in the 1998–2000 period were unusually large.

Betton, Eckbo, and Thorburn (2008c) examine the distribution of dollar differences in Figure 11B. They identify 125 firms in the lower 1% and 129 firms in the upper 1%. In *both* groups, the dominant sector was manufacturing and the dominant firm was Cisco.

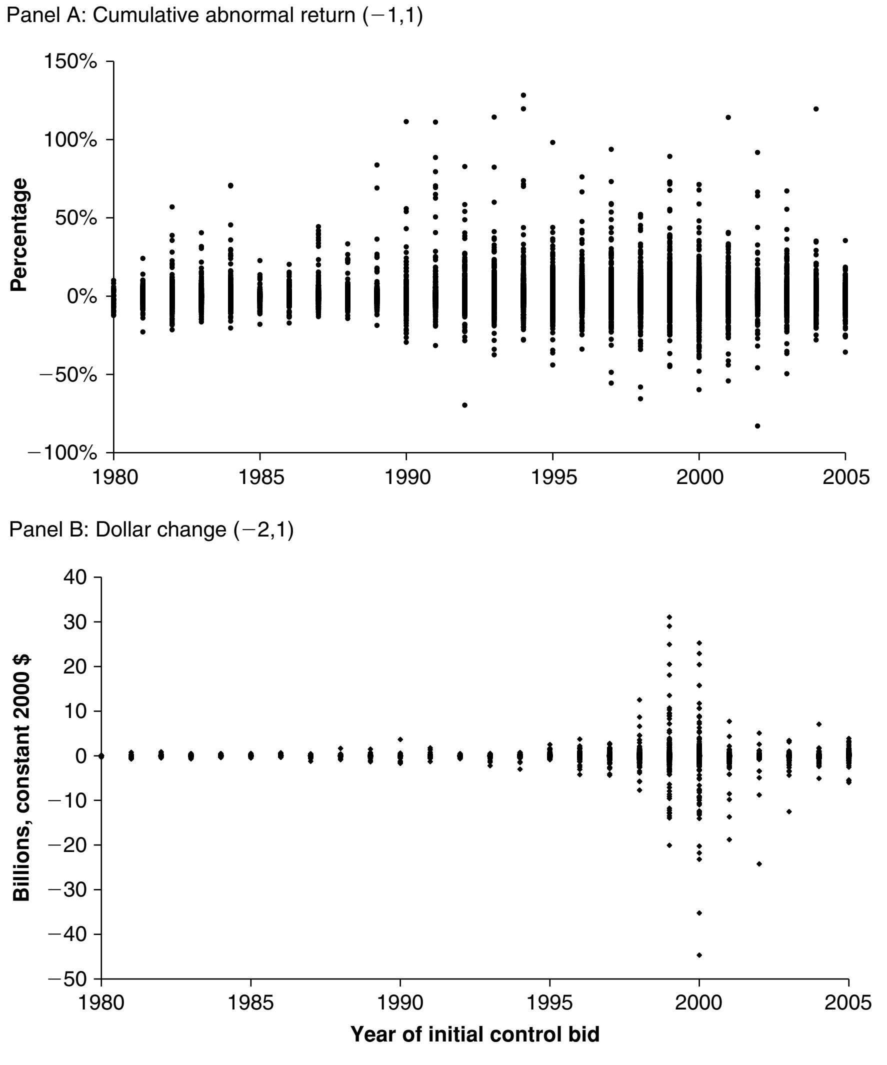
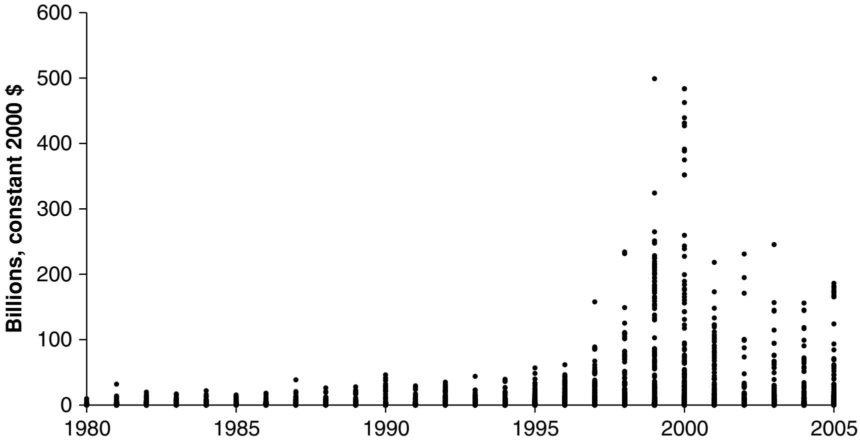



Fig. 11. Announcement-period abnormal returns and dollar-changes for 12,898 successful initial bidders, 1980–2005. Panel A is a scatter plot of the announcement period abnormal stock returns, CAR(−1, 1). Panel B is a scatter plot of the bidders' announcement-period dollar changes. Dollar changes are calculated as the change in market capitalization from day −2 to day 1 (relative to initial control bid) and converted to constant 2000 dollars using the CPI.

Panel C: Market value day $-2$




Panel D: Aggregate dollar abnormal returns

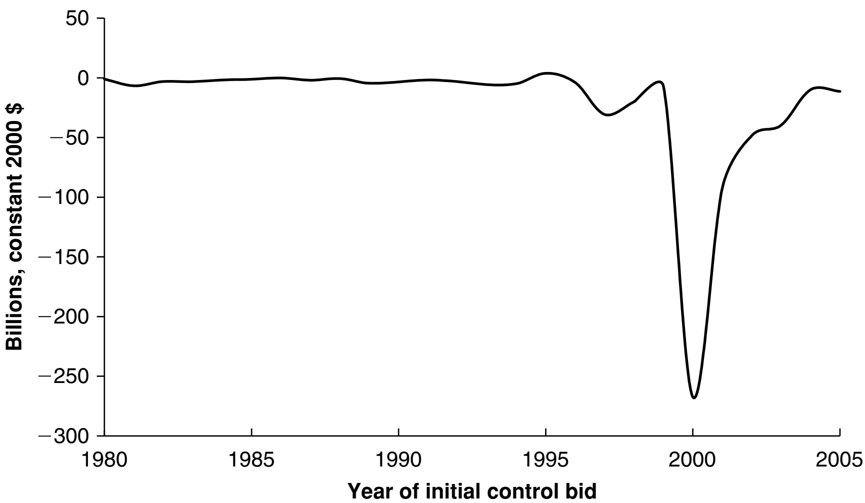



Fig. 12. The market values and announcement-period aggregate dollar abnormal return to 12,898 successful initial bidders, 1980–2005. Panel C is a scatter plot of the market value in constant 2000 dollars of successful initial bidders on day $-2$ relative to the initial control bid announcement. Panel D is a plot of the aggregate dollar abnormal returns earned by successful initial bidders over the window $(-2,1)$ Aggregate dollar abnormal returns are calculated by multiplying the bidder market capitalization on day $-2$ by the cumulative abnormal return and then summing over the year.

Cisco appears with 62 deals in the total sample, with an average (constant dollar) market capitalization of \$180 billion. Other frequent acquirers are Union Planters with 40 deals (market cap \$2.5 billion) and BancOne with 40 deals (market cap \$8.1 billion).[103] Of these 62 deals made by Cisco, 26 appear in the group with the highest 1% CAR(−1, 1), with 10 bids in 99 and 6 bids in 2000.[104] Furthermore, Cisco appears 17 times in the lower 1% group (distributed evenly over the three-year period 1999–2001).[105]

Panel D in Figure 12 plots the aggregate dollar CAR(−1, 1) for each sample year (combining Panel A of Figure 11 and Panel C, Figure 12). The large negative spike in the years 1999 and 2000 is what Moeller, Schlingemann, and Stulz (2005) characterize as a "wealth destruction on a massive scale." It is massive indeed; yet, it is important not to forget that it is caused by a few very large firms that decided to bid in this particular period and that, on average, made value-decreasing acquisitions. Note that Panel D of Figure 12 does not eliminate overlapping abnormal returns to frequent acquirers (which may be one reason why the spike is greater here than in Moeller, Schlingemann, and Stulz (2005)). Also, removing Cisco from the sample changes the minimum of the spike to −\$198 billion from −\$267 billion. The ultimately unanswered question is whether the spike is a bidder size effect or a year fixed effect (or a combination of the two). At this point, there appears to be no explanation for why the large firms decided to enter the market for corporate control in 1998–2001, and then only to leave again.

Finally, Figure 13 shows the frequency distribution for the dollar announcement abnormal return for the total sample of successful initial bidders, classified by the time period and the method of payment (all-stock or all-cash). Panel A covers the total sample, while Panel B is restricted to the 1995–2005 period. There is very little difference between the two panels (in both panels, all-stock offers are slightly skewed relative to all-cash bids). Thus, the distribution in Panel B is not noticeably affected by the extreme cases from the 1998–2000 period. Until we reach a better understanding of the unique 1998–2000 period, estimates of the expected gains from bidding are best obtained from overall distributions such as those in Figure 13.

### *4.4. Estimating expected bidder gains*

Referring back to Figure 2, let $\mathrm{CAR}_s$ and $\mathrm{CAR}_f$ denote average bidder gains conditional on the offer succeeding or failing, respectively. Moreover, let $\pi(x_j)$ denote the market's estimate of the probability that an offer by bidder $j$ will succeed conditional on the offer characteristics $x_j$. As discussed in Section 3, important offer characteristics include the offer premium, toehold, payment method, and hostility. The bid announcement causes

[103] The largest bidders are CitiGroup (market cap \$245 billion and two deals), Microsoft (market cap \$190 billion and seven deals).
[104] The next most common bidders in the upper 1% group are Johnson & Johnson with six cases and Tyco with five cases.
[105] The next most common bidder in the lower 1% group is Lucent with six bids.

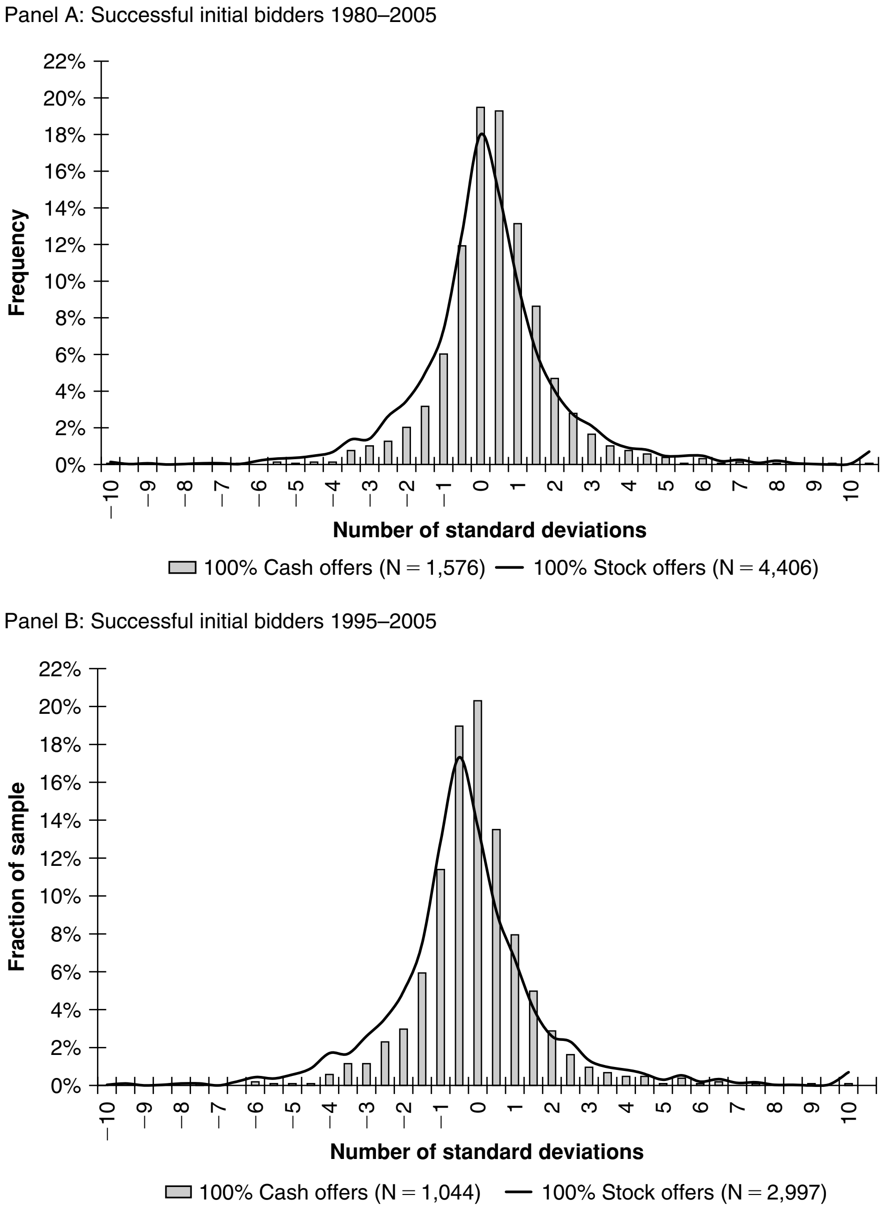



Fig. 13. Standardized dollar abnormal returns to successful initial bidders by method of payment, 1980–2005 (Panel A) and 1995–2005 (Panel B). Dollar abnormal returns are calculated as the change in market capitalization (in constant 2000 dollars) from day −2 to day 1 relative to initial control bid. Dollar abnormal returns are then standardized by the historical average and standard deviation of non-overlapping 3 day dollar market value changes measured over the period −293 to −42 relative to the initial control bid. Sample windows are (−45,−42), (−48,−45), etc.

rational investors to impound the *expected* bidder takeover gain into the bidder's share price, generating the following announcement return:

$$\mathrm{CAR}(-1,\ 1) = \mathrm{CAR}_s\pi(x) + \mathrm{CAR}_f(1 - \pi(x)) \tag{20}$$

The empirical objective is to estimate the bidder gain $\mathrm{CAR}_s$ from successful takeovers. The common procedure is to form the average cumulative return in the subsample of ex-post successful bids. This average is either the average $\mathrm{CAR}(-1, 1)$ for successful bids or the average abnormal bidder return cumulated all the way through the end of the contest (at which point $\pi = 1$). Note that, since these ex-post averages necessarily restrict the sample to successful bids, they ignore information in the abnormal returns to ultimately unsuccessful bids. Also, cumulation to the end of the contest (typically, six months for mergers) adds noise relative to that of the three-day estimate $\mathrm{CAR}(-1,\ 1)$.

Betton and Eckbo (2000) develop an alternative estimation procedure that exploits the information in *all* initial bids (also the ultimately unsuccessful ones) in order to extract an estimate of $\mathrm{CAR}_s$. The idea is to view Equation (20) as a cross-sectional regression where $\mathrm{CAR}_j(-1,\ 1)$ is the dependent variable, $\pi(x_j)$ is the regressor, and $\mathrm{CAR}_s$ and $\mathrm{CAR}_f$ are estimated directly as regression parameters. Using a sample of 1,353 initial tender offers (both successful and unsuccessful), Betton and Eckbo (2000) find that the parameter $\mathrm{CAR}_s$ for bidders is statistically insignificantly different from zero. Thus, the expected net bidder return from initiating tender offers is nonnegative. Moreover, they estimate $\mathrm{CAR}_f$ to be significantly positive, which they suggest in part reflects the expected gain to the unsuccessful bidder from selling its toehold in the target to the ultimately winning (rival) bidder.[106]

This alternative estimation procedure also allows one to test the effect on bidder expected returns of changing one or more of the offer parameters in the vector $x$. That is, when estimated, the right-hand-side of Equation (20) forms the predicted (conditional) value $E[\mathrm{CAR}(-1,\ 1)|\ x]$. As modeled in Equation (20), changes in $x$ affect bidder expected gains by changing $\pi(x)$. Tests of the bidder valuation impact of changing the offer parameters $x$ amount to testing whether the partial derivative of $E[\mathrm{CAR}(-1,\ 1)|\,x]$ with respect to $x$ is significantly different from zero. For example, both Betton and Eckbo (2000) and Betton, Eckbo, and Thorburn (2007) report that this partial derivative with respect to the bidder's toehold is positive and significant.

### 4.5. *Post-takeover (long-run) abnormal returns*

Several studies report evidence of post-merger underperformance, particularly when using the matched-firm buy-and-hold technique (implemented below). For example, Rau and Vermaelen (1998) find that merged firms with low book-to-market ratio tend

[106] In Betton and Eckbo (2000), 48% of all initial bidders have a positive toehold. Their sample period is 1971–1990, with the largest toehold frequency prior to the mid-1980s (consistent with Figure 8).

to underperform, while Loughran and Vijh (1997) report underperformance when the merger is paid in stock (while all-cash mergers overperform). Moeller, Schlingemann, and Stulz (2004) report insignificant long-run (36-month) post-merger performance. Harford (2005) finds some evidence suggestive of relatively poor post-merger performance for the largest bidders. Moeller, Schlingemann, and Stulz (2005) find significantly negative long-run buy-and-hold returns to portfolios of "large loss deal" bidders. Rosen (2006) reports evidence that mergers that take place during periods of high general merger activity tend to have high pre-merger share prices followed by low post-merger performance.

There are at least three possible explanations for the post-merger underperformance. First, under behavioral arguments, the market slowly corrects its overvaluation of the merged firms' shares (Shleifer and Vishny, 2003; Baker, Ruback, and Wurgler, 2007). Second, a neoclassical argument is that the merger is a response to a negative industry shock and that the merged firm performs better than it would have without the merger—which may still be worse than the pre-merger performance (Harford, 2005). Third, the apparent underperformance is an artifact of the econometric methodology itself. The rest of this section sheds light on the third hypothesis.

We begin the long-run abnormal return analysis with the matched firm technique, and then we show results when returns are risk-adjusted using factor regressions applied to portfolios of merged firms.[107] Our sample drops to 15,298 mergers after imposing the following additional restrictions: (1) The sample period is 1980–2003 to allow a minimum of three years of post-merger stock returns. (2) The merged firm is found on CRSP and is listed on NYSE/AMEX/Nasdaq for at least one year following the year of the effective date of the merger. (3) The merged firm must have Compustat information on equity book-to-market ratio (B/M) to allow selection of a matched firm based on size and B/M.[108]

### *4.5.1. Buy-and-hold returns*

The typical buy-and-hold experiment involves buying the merged firm's stock in the month following the merger completion month (effective merger date) and holding the stock for a period of three to five years or until delisting, whichever comes first. In a sample of $N$ issues, the average return over a holding period of $T$ months is computed

[107] We thank Øyvind Norli for his generous programming assistance. The econometric methodology implemented below is identical to the one used by Eckbo, Masulis and Norli (2007) when estimating the long-run performance following security offerings.

[108] Book value is defined as "the Compustat book value of stockholders equity, plus balance sheet deferred taxes and investment tax credits (if available), minus the book value of preferred stock. Depending on availability, we use the redemption, liquidation, or par value (in that order) to estimate the value of preferred stock" (Fama and French, 1993, p. 8). If available on Compustat, the book value of equity is also measured at the end of the year prior to the year of the acquisition. If this book value is not available, we use the first available book value on Compustat starting with the acquisition year and ending with the year following the acquisition year.

as the average cumulative (T-period) return, also referred to as $\overline{\text{BHR}}$ (for buy-and-hold return):

$$\overline{\text{BHR}} \equiv \sum_{i=1}^{N} \omega_i \left[ \prod_{t=\tau_i}^{T_i} (1 + R_{it}) - 1 \right] \tag{21}$$

where $R_{it}$ denotes the return to stock $i$ over month $t$ and $\omega_i$ is stock $i$'s weight in forming the average holding-period return ($\omega_i = 1/N$ when equal-weighting). The effective holding period for stock $i$ is $T_i$, where $T_i$ in the analysis below is either five years or the time until delisting or the occurrence of a new merger, whichever comes first).[109,110] The matched-firm technique equates the *expected* return to merged firms with the *realized* return to a nonmerging firm, usually matched on firm characteristics such as industry, size, and book-to-market ratio. The abnormal or unexpected return BHAR is then

$$\text{BHAR}_{\text{Issuer}} \equiv \text{BHR}_{\text{Issuer}} - \text{BHR}_{\text{Matched firm}} \tag{22}$$

Table 10 shows average five-year percent buy-and-hold returns for our sample and for firms matched on size and B/M. The matched firms are selected from all CRSP-listed companies at the end of the year prior to the year of the merger completion and companies that are not in our sample of mergers for a period of five years prior to the offer date. Moreover, the matching procedure is as follows: We first select the subset of firms that have equity market values within 30% of the equity market value of the

109 Kothari and Warner (1997), Barber and Lyon (1997) and Lyon, Barber, and Tsai (1999) provide simulation-based analysis of the statistical properties of test statistics based on long-run return metrics such as BHR. Kothari and Warner (2007) survey the main statistical conclusions from this analysis.

110 An alternative to $\overline{\text{BHR}}$ is to estimate the average monthly return to a strategy of investing in the stocks of merged firms and hold these for up to $T$ periods. The $T$-period return would then be formed as the *cumulative average* (portfolio) return, or

$$\overline{\text{CMR}} \equiv \prod_{t=\tau}^{T} \left[ 1 + \frac{1}{\omega_t} \sum_{i=1}^{N_t} R_{it} \right] - 1$$

As noted by Kothari and Warner (2007), depending on the return generating process, the statistical properties of $\overline{\text{BHR}}$ and $\overline{\text{CMR}}$ can be very different. Notice also that while $\overline{\text{CMR}}$ represents the return on a feasible investment strategy, $\overline{\text{BHR}}$ does not. You obtain $\overline{\text{CMR}}$ by investing one dollar in the first security issue at the beginning of the sample period, and then successively rebalancing this initial investment to include subsequent issues as they appear (and $N$ increases), all with a $T$-period holding period. In contrast, $\overline{\text{BHR}}$ is formed in event time—and thus presumes prior knowledge of the magnitude of $N$. Thus, estimates of $\overline{\text{CMR}}$ are better suited than estimates of $\overline{\text{BHR}}$ to address the question of whether investors have an incentive to take advantage of a potential market mispricing of merged firms' securities. Most of the empirical studies using the matched firm technique report results based on $\overline{\text{BHR}}$, which we follow here. In the subsequent section, we discuss portfolio benchmark returns based on asset pricing models, which use the return concept $\overline{\text{CMR}}$ on a monthly basis, that is, without the $T$-period cumulation.

Table 10

Percent average five-year buy-and-hold stock returns ($\overline{\text{BHR}}$) for merged firms, non-merging matched firms, and the difference between merged and matched firms, 1980–2006.

Buy-and-hold percent returns are defined as:

$$\overline{\text{BHR}} \equiv \sum_{i=1}^{N} \omega_i \left[ \prod_{t=\tau_i}^{T_i} (1 + R_{it}) - 1 \right] \times 100.$$

The sampling of merged firms starts in February 1980 and ends in December 2003, while the return holding period is allowed to continue to December 2006. The total sample of merged firms with information on matched firms is 15,298. The non-merging matched firms are firms that did not merge in the previous five-year period and have similar total equity size and book-to-market ratio. When equal-weighting, $\omega_i \equiv 1/N$, and when value-weighting, $\omega_i = MV_i/MV$, where $MV_i$ is the firm's common stock market value at the start of the holding period and $MV = \sum_i MV_i$. The abnormal buy-and-hold returns shown in the column marked "Diff" represent the difference between the $\overline{\text{BHR}}$ in the "Merged" and "Match" columns. "N" is the total number of issues. The $p$-values for equal-weighted abnormal returns are $p$-values of the $t$-statistic using a two-sided test of no difference in average five-year buy-and-hold returns for issuer and matching firms. The $p$-values for the value-weighted abnormal returns are computed using $U \equiv \omega' x/(\sigma\sqrt{\omega'\omega})$, where $\omega$ is a vector of value weights and $x$ is the corresponding vector of differences in buy-and-hold returns for issuer and match. Assuming that $x$ is distributed normal $N(\mu, \sigma^2)$ and that $\sigma^2$ can be consistently estimated using $\sum_i \omega_i (x_i - \bar{x})^2$, where $\bar{x} = \sum_i \omega_i x_i$, $U$ is distributed $N(0, 1)$.

| | | Equal-weighted $\overline{\text{BHR}}$ | | | | Value-weighted $\overline{\text{BHR}}$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| Merger sample period | N | Merged | Matched | Diff | $p(t)$ | Merged | Matched | Diff | $p(t)$ |
| 1980–2003 | 15,298 | 62.6% | 84.6 | −21.9 | 0.000 | 32.6 | 49.6 | −17.1 | 0.000 |
| 1980–1989 | 3,815 | 83.6 | 95.1 | −11.5 | 0.003 | 102.0 | 113.9 | −12.0 | 0.120 |
| 1990–2003 | 11,483 | 55.7 | 81.1 | −25.4 | 0.000 | 26.7 | 44.2 | −17.5 | 0.000 |

merged firm. This subset is then ranked according to book-to-market ratios. The size and book-to-market matched firm is the firm with the book-to-market ratio, measured at the end of the year prior to the merger year, that is closest to the merged firm's ratio. Matched firms are included for the full five-year holding period or until they are delisted, whichever occurs sooner. If a match delists, a new match is drawn from the *original* list of candidates described earlier.

Table 10 shows that, when using either the total sample period 1980–2003 or the subperiod 1990–2003, merged firms on average underperform their matched firms whether $\overline{\text{BHR}}$ is formed using equal weights or value weights. For the total sample period, the difference between the equal-weighted $\overline{\text{BHR}}$ for merged and matched firms is $-21.9\%$ and $-17.1\%$ with value-weighting, both with $p$-values of 0.00. About 20% of the sample mergers take place in the 1980s, and here the underperformance is evident only for equal-weighted $\overline{\text{BHR}}$. For the subperiod 1990–2003 the underperformance estimates are again highly significant and slightly greater than for the total period $-25.4\%$ using the equal-weighted estimate of $\overline{\text{BHR}}$.

### 4.5.2. Portfolio performance estimation

An alternative to the buy-and-hold matched firm technique is to form portfolios of event firms rolling forward in calendar time and to estimate portfolio performance. Monthly portfolio (excess) returns are regressed on a set of risk factors presumed to generate expected returns. The regression intercept—or alpha—is the measure of average monthly abnormal return. We estimate alphas in a model with the following five risk factors:

$$r_{pt} = \alpha_p + \beta_1 \mathrm{RM} + \beta_2 \mathrm{SMB}_t + \beta_3 \mathrm{HML}_t + \beta_4 \mathrm{UMD} + \beta_5 \mathrm{LMH} + e_t \tag{23}$$

where $r_{pt}$ is the excess return to an equal-weighted portfolio of issuers, RM is the excess return on the CRSP value-weighted market index. SMB and HML are the Fama and French (1993) size and book-to-market factors. UMD is a momentum factor inspired by Carhart (1997) and constructed as the return difference between the one-third highest and the one-third lowest CRSP performers over the past 12 months. LMH is the Eckbo and Norli (2005) turnover factor, defined as a portfolio long in low-turnover stocks and short in high-turnover stocks.

We report estimates for three different portfolios: (1) the merging firms, (2) the nonmerging matched firms, and (3) the zero-investment portfolio that is long in merged firms and short in matched firms. The zero-investment portfolio has the advantage that it controls for any omitted risk factor with identical factor betas across issuer and matched firm, effectively combining the matched-firm and asset pricing techniques. For example, suppose the true set of risk factors is given by the vector $F$ and that only a subset $F_1$ of this vector is included in the regression model, with the complement vector $F_2$ omitted. Let B denote merged firm and M matched firm. The merger-match zero-investment portfolio regression is then

$$r_B - r_M = (\alpha_B - \alpha_M) + (\beta_{1B} - \beta_{1M})F_1 + \epsilon \tag{24}$$

where $\epsilon = (\beta_{2B} - \beta_{2M})F_2 + u$, where $u$ is a white noise error term. The definition of a "good" match is that $\beta_B$ is close to $\beta_M$. Given a good match, the zero-investment portfolio will have both a small alpha and values of beta close to zero. Alternatively, if the matching technique fails to control for important risk factors, then the zero-investment portfolio will contain significant factor loadings.

Table 11 reports the alphas and factor loadings (betas) for our three portfolios and the five-factor model. Portfolio formation starts in 1980 and ends in 2003. The table shows estimates for both equal weighting and value weighting of the firms in the portfolios. Given the large portfolios, the $R^2$ are high, approximately 0.94. Notice also that the zero-investment portfolios receive $R^s$ of close to 0.20, with several significant factor loadings, indicating that the usual size and B/M matching procedure typically yields firms that have different expected returns than the event firms. This in turn means that the "abnormal" return reported earlier in Table 11 in part reflects differences in *expected* returns for merged and matched firms.

The key result in Table 11 is that the alphas for the portfolio of merged firms are small and statistically insignificant for both equal- and value-weighted portfolios. Thus, we

Table 11

Alphas and factor loadings for five-year rolling portfolios of merged firms, non-merging matched firms, and long-short merged-matched firms, 1980–2006.

The merged-matched portfolio is a zero-investment portfolio that is long in the merged firms and short in the non-merging matched firms. The portfolios are either equal-weighted ("EW") or value-weighted ("VW"). The non-merging matched firms are firms that did not merge in the previous five-year period and have similar total equity size and book-to-market ratio. The portfolios are formed starting in February 1980: a firm is added to the portfolio in the month following the month of the effective merger date and held for the minimum of five years and until delisting. The merger sampling stops in 12/2003, yielding a total of 15,298 successful mergers with data on size and book-to-market matched firms. The abnormal return estimation ends in December 2006. Abnormal returns are estimated using the following asset pricing model:

$$r_{pt} = \alpha_p + \beta_1 \mathrm{RM} + \beta_2 \mathrm{SMB}_t + \beta_3 \mathrm{HML}_t + \beta_4 \mathrm{UMD} + \beta_5 \mathrm{LMH} + e_t$$

where $r_{pt}$ is the portfolio excess return, RM is the excess return on the CRSP value weighted market index, SMB and HML are the Fama and French (1993) size and book-to-market factors, UMD is a momentum factor constructed as the returns difference between the one-third highest and the one-third lowest CRSP performers over the past 12 months, and LMH is the Eckbo and Norli (2005) turnover factor (a portfolio long in low-turnover stocks and short in high-turnover stocks). The coefficients are estimated using OLS. Standard errors are computed using the heteroskedasticity consistent estimator of White (1980). The numbers in parentheses are $p$-values. $R^2$ is the adjusted R-squared.

| Portfolio | $\alpha_p$ | RM | SMB | HML | UMD | LMH | $R^2$ |
|---|---|---|---|---|---|---|---|
| EW-merger | 0.08<br>(0.434) | 1.05<br>(0.000) | 0.62<br>(0.000) | 0.26<br>(0.000) | −0.28<br>(0.000) | −0.13<br>(0.070) | 0.943 |
| EW-match | 0.23<br>(0.003) | 0.97<br>(0.000) | 0.52<br>(0.000) | 0.24<br>(0.000) | −0.19<br>(0.000) | −0.14<br>(0.019) | 0.949 |
| EW-zero | −0.15<br>(0.050) | 0.09<br>(0.000) | 0.11<br>(0.000) | 0.02<br>(0.333) | −0.09<br>(0.000) | 0.01<br>(0.785) | 0.239 |
| VW-merger | 0.02<br>(0.838) | 1.07<br>(0.000) | −0.08<br>(0.029) | −0.10<br>(0.016) | −0.05<br>(0.028) | −0.01<br>(0.828) | 0.935 |
| VW-match | 0.11<br>(0.076) | 1.00<br>(0.000) | −0.14<br>(0.000) | −0.06<br>(0.016) | −0.02<br>(0.266) | 0.07<br>(0.100) | 0.949 |
| VW-zero | −0.09<br>(0.288) | 0.06<br>(0.017) | 0.06<br>(0.027) | −0.03<br>(0.386) | −0.03<br>(0.181) | −0.09<br>(0.071) | 0.170 |

cannot reject the hypothesis of zero abnormal post-merger performance. Four of the five factor-mimicking portfolios have significant factor loadings, with the turnover-based factor producing a factor loading that is significant at the 10% level for equal-weighted portfolio returns.

Table 11 also shows that the empirical factor model misprices the equal-weighted portfolio of matched firms. The alpha of this portfolio is 0.23 with a $p$-value of 0.003. As a result, the portfolio of merged firms underperforms the matched portfolio (the alpha for the zero-investment portfolio is −0.15 with a $p$-value of 0.05). When equal-weighting returns, the factor model's mispricing of the matched-firm portfolio is less significant, and now the alpha of the zero-investment portfolio is insignificantly different from zero.

In sum, when using the rolling portfolio technique, there is no evidence of abnormal stock returns following mergers. Moreover, our evidence that matched firms have significantly different factor loadings than merged firms undermines the notion that the underperformance reported in Table 10 represents truly negative abnormal stock returns.

## 5. Bondholders, executives, and arbitrageurs

### *5.1. Takeovers and bondholder wealth*

Corporate mergers affect the wealth of the target and the acquiring firms' senior claimholders for the same reasons that they affect stockholders. Merger-induced synergies add security to outstanding bonds and therefore increase bond values, while value-reducing mergers reduce bond value. In addition, bondholders benefit from any co-insurance effect from combining the less than perfectly correlated cash flows of the bidder and target firms.[111] The coinsurance effect means that a merger that generates no synergies, and where the bidder firm neither overpays for the target nor manages to sell overpriced bidder stock to the target, nevertheless causes a wealth transfer from stockholders to bondholders (Galai and Masulis, 1976). The magnitude of this wealth transfer depends on the sensitivity of the bond payments to changes in firm value (bond risk), with greater potential valuation impact on ex-ante riskier bonds.[112] The coinsurance effect also reduces the risk of firm-specific human capital. This argument has led to a concern that entrenched managers seek empire building through conglomerate merger activity primarily in order to hedge the risk of their firm-specific human capital.

A difficulty facing bond studies is the lack of access to high-frequency data on bond values, particularly prior to the 1980s. One of the primary data sources is the Lehman Brothers Fixed Income Database. Most bonds do not have published transaction prices, and many of the reported prices are matrix prices. Matrix prices are reported when the bond does not trade or a dealer quote is unavailable. The matrix consists of prices of similar bonds that did trade, based on characteristics such as bond rating and maturity. Obviously, the effect of a merger does not show up in matrix prices (for other bonds), reducing power to reject the null of no price impact of the merger.

Kim and McConnell (1977) examine 2,286 mergers but find price data for 44 bonds of 39 firms only. In their sample of 2,870 mergers, Asquith and Kim (1982) find prices for 62 bonds, while the sample in Eger (1983) includes 33 acquirer bonds and 6 target bonds. The Dennis and McConnell (1986) bond sample contains 67 bonds of 39 acquirers and 27 bonds for 21 targets. Maquieira, Megginson, and Nail (1998) identify a large sample of 504 acquirer bonds and 124 target bonds of firms involved in 260 stock-for-stock mergers

[111] Levy and Sarnat (1970), Lewellen (1971), Higgins and Schall (1975).

[112] There is also a maturity effect: When the bonds of the bidder and target firms have different maturities, the shorter maturity bonds effectively gain seniority after the merger. This seniority effect is valuable because of the larger merged firm's asset base.

(the bulk of which took place prior to 1980). More recently, Billett, King, and Mauer (2004) examine 940 mergers and acquisitions from the period 1979–1997, identifying 818 bonds of 265 target firms and 3,083 bonds of 831 acquiring firms. Moreover, Penas and Unal (2004) use a sample of 282 bonds in 66 mergers between commercial banks during the period 1991–1997.[113]

The early studies found mixed evidence for the wealth effects of mergers on bidder bonds: excess bond returns (typically computed as the difference between monthly total return and the return on a bond index matched on rating and remaining maturity) are significantly positive in Eger (1983) and Maquieira, Megginson, and Nail (1998); insignificant in Kim and McConnell (1977) and Asquith and Kim (1982); and negative (marginally significant) in Dennis and McConnell (1986). Billett, King, and Mauer (2004) find zero or negative bidder bond excess returns, while Penas and Unal (2004) document significantly positive bidder bond returns for their sample of commercial bank mergers.

Early studies of target bond returns report insignificant excess returns to *target* bonds. This finding is surprising, as one would expect target bondholders to benefit from the typically large asset-base increase that comes with a merger with a bidder that is often several times larger than the target. However, with improved data, both Billett, King, and Mauer (2004) and Penas and Unal (2004) report significantly positive excess returns to target bonds. This finding may also reflect the increased use of event risk covenants in bonds issued in the 1990s.[114] Penas and Unal (2004) conclude that the bond market views bank mergers as default-risk-reducing events. Billett, King, and Mauer (2004) conclude that there is no evidence of wealth transfers in the data, or that positive synergies expected from the corporate combinations tend to overshadow any wealth transfer that do exist.

### *5.2. Takeovers and executive compensation*

Does the structure of CEO compensation packages affect the quality of takeover decisions? Or, as Lehn and Zhao (2006) put it: "Are Bad Bidders Fired?" The literature on optimal compensation presumes that a strong pay-performance sensitivity helps promote better acquisition decisions.[115] There is evidence that target firms tend to underperform prior to becoming targets.[116] Moreover, Mitchell and Lehn (1990), Martin

[113] Warga and Welch (1993) also use this Lehman Brothers Fixed Income Database to study the bond wealth effects of leveraged buyouts, while Eberhart and Siddique (2002) use these bond data in their study of long-run bond returns following securities offerings. See Eckbo and Thorburn (2008a) and Eckbo, Masulis, and Norli (2007) for reviews.

[114] A typical event risk covenant for mergers requires the company to repurchase the outstanding bonds at the full principal amount plus accrued interest, effectively insuring the bond against potentially value-decreasing control events (Lehn and Poulsen, 1992; Nash, Netter, and Poulsen, 2003).

[115] See, for example, Murphy (1999) and Aggarwal (2008) for comprehensive reviews of the literature on executive compensation and pay-performance sensitivity.

[116] Asquith (1983), Malatesta (1983). There is also evidence of poor operating performance prior to divisional sales. See Eckbo and Thorburn (2008a) for a review.

and McConnell (1991), Agrawal and Walkling (1994), and Kini, Kracaw, and Mian (1995, 2004) document that targets of hostile bids tend to show a prior history of value-decreasing acquisitions and that CEO turnover increases after hostile bids. Offenberg (2008) find evidence that CEOs of larger firms are more likely to be replaced following a series of poor acquisitions than CEOs of smaller firms. This is consistent with a disciplinary role played by the market for corporate control.

With the spread of the poison pill defense and the subsequent decline of hostile takeovers after the 1980s, the market for corporate control may have become a court of last resort—with internal governance structures being the primary mechanism for disciplining poor managers.[117] Huson, Parrino, and Starks (2001) find that changes in the intensity of the takeover market over the period 1976–1994 are *not* associated with changes in the sensitivity of CEO turnover to firm performance. Their evidence suggests that changes in external and internal governance mechanisms have not significantly changed the likelihood that the CEO of a poorly performing firm will be replaced. They also suggest that the effectiveness of internal monitoring mechanisms is not dependent on the intensity of the takeover market. With data from 1979 through 1998, Kini, Kracaw, and Mian (2004) conclude that the corporate takeover market intercedes when internal control mechanisms are relatively weak or ineffective.

Lehn and Zhao (2006) show that managers who undertake value-reducing acquisitions in the period 1990–1998 face a significantly higher probability of being replaced than managers who make value-enhancing acquisitions, either by internal governance, takeovers, or bankruptcy. They also show that CEOs who cancel an acquisition after observing a reduction in their company's stock price face significantly lower replacement risk than their counterparts who proceed with value-reducing acquisitions. Among firms not subjected to takeover or bankruptcy, they find no association between a firm's governance characteristics and the probability that the CEOs who make value-reducing acquisitions are replaced.

Lehn and Zhao (2006) conclude that "corporate governance and the external market for corporate control generally work well in disciplining managers who pursue acquisitions to the detriment of their stockholders." Moreover, they interpret their evidence of a lack of association between the CEO replacement probability and specific governance characteristics following bad takeovers as an indication that governance structures are on average optimally chosen. While this is one possible interpretation, an alternative view (which they recognize) is that governance structure is irrelevant as to the firing decision for the sample firms. Given the endogeneity of the governance structure (where the CEO herself plays a role), additional research is necessary to discriminate between these two positions.[118]

Lehn and Zhao (2006) also present evidence of relevance for the "market-driven acquisition" hypothesis of Shleifer and Vishny (2003) discussed above. This hypothesis implies that acquisitions that are followed by poor long-run bidder stock returns may

[117] Mikkelson and Partch (1997), Holmstrom and Kaplan (2001).

[118] Himmelberg, Hubbard, and Palia (1999), Bhagat and Jefferis (2002), Hermalin and Weisbach (2003).

nevertheless be in the interest of bidder stockholders, provided the alternative of no merger would have been even worse. For example, it is tempting (with hindsight) to characterize AOL/Time Warner merger as a successful attempt by AOL's CEO Stephen Case to use overvalued stock as currency to acquire Time Warner's "hard" assets:

> From our perspective, the central feature of this acquisition is not technological synergies, but rather the attempt by the management of the overvalued AOL to buy the hard assets of Time Warner to avoid even worse returns in the long run. In this acquisition, as in other deals involving high-technology acquirers with overvalued stock prices, long-run acquirer returns appear to be poor. However, according to our model, these returns are not as negative as they would have been had the acquisitions not taken place. When future writers condemn the merger spree of the late 1990s as manifesting misguided policies on the part of acquirers, they should focus on the alternative of not making these acquisitions. (Shleifer and Vishny, 2003, p. 295)

The market-driven acquisition hypothesis implies that the bidder prefers cash as payment method when bidder stock is sufficiently undervalued. Cash acquisitions must generate value through synergies (as opposed to selling overvalued stock) for the bidder management to act in their shareholders' interest. Thus, while poor bidder performance following all-stock mergers is consistent with bidder value-maximizing behavior, poor performance following all-cash mergers is not.

Lehn and Zhao (2006) find a significant inverse relation between long-run returns after acquisitions and the probability that CEOs are replaced. More importantly, CEOs of acquiring firms with negative bidder returns are equally likely to be replaced, regardless of whether they used stock or cash as the method of payment in the acquisition. This finding challenges the prediction of Shleifer and Vishny (2003) and instead suggests that stock acquisitions (as well as cash acquisitions) associated with negative long-run bidder returns are destructive of value.

Several recent papers provide evidence on CEO compensation changes (other than turnover) following acquisition activity. Bliss and Rosen (2001) study bank mergers over the period 1985–1995, a period characterized by overcapacity and frequent mergers. Mergers are found to have a net positive effect on bidder firm CEO compensation, mainly via the effect of size on compensation. Compensation increases even if the merger causes the acquiring bank's stock price to decline (which is typical upon merger announcement). However, CEOs with more stock-based compensation are less likely to make an acquisition, suggesting that bank managers are motivated by their compensation contracts.

Datta, Iskandar-Datt, and Raman (2001) study 1,719 acquisitions over the period 1993–1998 and separate the acquirers into whether the equity-based compensation of their respective CEOs is above (high) or below (low) the median. While the market reaction to the merger announcements is insignificantly different from zero on average, it is significantly positive for bidder CEOs with high equity-based compensation and significantly negative when the equity-based compensation is low. Moreover, the compensation structure impacts the target selection: high equity-based managers tend to seek out targets with relatively high market-to-book ratio (growth targets), whereas CEOs in the low-incentive compensation group tend to acquire targets with low growth prospects. Thus, it appears that managers with high equity-based compensation are willing to take on riskier and more valuable acquisition projects than managers with low equity-based compensation.

Grinstein and Hribar (2004) examine M&A bonuses (typically all-cash) paid to CEOs of bidder firm after 327 large merger deals over the period 1993–1999. Bonuses are larger for larger deals. Other than size, CEO power is the single most powerful variable explaining the cross-sectional variation in M&A bonuses. Much as in Bebchuk and Fried (2003), CEO power is measured as the CEO's ability to influence directors (and thereby the compensation decision). A CEO gains influence as a chairman of the board, as a member of the nominating committee, as the proportion of insiders on the board increases, and as board size increases. The size and power variables explain much more of the variation in bonuses than variables capturing CEO skill, effort, and performance. Moreover, the deal announcement-induced abnormal stock return is significantly lower (more negative) in the sample of CEOs with high power than those with low power. Moeller (2005) also concludes that targets with powerful CEOs receive lower takeover premiums. However, Bauguess, Moeller, Schlingemenn, and Zutter (2007) present evidence that inside (managerial) ownership has a positive relation with target returns, whereas active-outside (nonmanaging director) ownership has a negative relation with target returns. They suggest that the latter effect reflects outsiders' willingness to share gains with the bidder.

Harford and Li (2007) also study how CEO pay and pay-performance sensitivity are affected by acquisitions. With a sample of 1,508 mergers completed over the period 1993–2000, they show that bidding firm CEOs receive substantial rewards in the form of new stock and options grants following acquisitions. While a poorly performing acquisition reduces the value of the CEO's portfolio of stocks and options obtained prior to the acquisition, the new post-acquisition grants more than compensate for this personal value reduction. As a result, "CEO's pay and wealth are completely insensitive to poor post-acquisition performance, but CEO's wealth remains sensitive to good post-acquisition performance" (p. 919). Interestingly, they show that bidding firms with stronger boards retain the sensitivity of their CEO's compensation to poor post-acquisition performance.

Harford and Li (2007) also document that compensation changes around major capital expenditures are much smaller and more sensitive to performance than those following acquisitions. That is, similar to conclusions made by Andrade and Stafford (2004), external and internal expansion decisions are treated fundamentally differently by the board. This difference may be rooted in the greater degree of uncertainty and information asymmetry surrounding acquisitions, which may allow the CEO to demand (and receive) some degree of protection for the downside risk to her personal wealth.

Walkling and Long (1984), Cotter and Zenner (1994), Wulf (2004) and Hartzell, Ofek, and Yermack (2004) all present evidence on acquisition-induced compensation of *target* firm CEOs. Hartzell, Ofek, and Yermack (2004) analyze a sample of 311 negotiated mergers between 1995 and 1997. A key question is what package of capital gains, cash, and subsequent employment do target CEOs accept in exchange for relinquishing control? Another important issue is whether target CEOs sacrifice premiums paid to their own outside shareholders in return for a favorable golden handshake.[119] Consistent with earlier studies, they conclude that "acquirers overtly pay certain CEOs to surrender managerial control over the firm's assets, or equivalently, that some CEOs "purchase" executive jobs in the buyer by foregoing cash payments that they might otherwise have obtained" (p. 39). Also, they present some evidence of an inverse association between selling shareholder premia and unusual bonuses received by the target CEO as a reward to "step aside." However, since their study uses a sample of completed mergers only, it does not provide information on the sort of packages that other target CEOs turn down in attempted mergers that were not completed. Thus, as the authors recognize, the study does not conclusively indicate that the large CEO packages come at the expense of target shareholders.

Finally, there is some evidence that board structure and director compensation affect the outcome of takeovers. Byrd and Hickman (1992) and Cotter, Shivdasani, and Zenner (1997) find that boards dominated by outsider directors increase value for their shareholders during an acquisition attempt. Harford (2003) documents the effect of a takeover bid on target directors, both financially and in terms of its effect on the number of future board seats held by those directors. He finds that directors are rarely retained following a completed offer and that target directors as a group hold fewer directorships after a takeover, suggesting that the target board seat is difficult to replace. Moreover, he shows that for outside directors, the direct financial impact of a completed merger is largely negative. In sum, failing as a monitor imposes a personal cost on outside directors.

### *5.3. Merger arbitrage*

#### *5.3.1. Arbitrage positions*

After the announcement of a takeover bid, the target stock price adjusts upward but typically still trades at a discount from the offer price. The difference between the offer price and the post-announcement market price is called the arbitrage spread. Merger arbitrage (or risk arbitrage) is a specialized investment strategy that tries to profit from this spread. Specifically, it is a bet on the likelihood that the proposed transaction closes. If the bid (or a rival bid) is successful and the target is acquired, the arbitrageur captures the price differential. If the takeover fails and the target remains independent, however, the target stock tends to fall back to pre-bid levels and the arbitrage position has to be closed at a loss. Since the position carries the transaction risk, it is not an arbitrage in the true (riskless) sense of the word. It is, however, designed to be neutral to subsequent market movements and to price fluctuations between the bidder and the target if the deal succeeds.

For a cash bid, a merger arbitrage position simply involves a long position in the target stock. When the acquisition is consummated, the target stock is exchanged for cash. With a positive arbitrage spread, the cash received at closing will exceed the initial investment in the target stock, hence generating a profit. In contrast, if the takeover fails

[119] Yermack (2006) presents evidence on severance packages more generally, in a sample of *Fortune 500* companies.

and the target stock price falls, the speculative position has to be sold at a loss equal to the price decline in the target stock.

The arbitrage position in a stock-for-stock transaction is more complex, since target shareholders are offered acquirer stock as payment. Here, the arbitrage position consists of a long target stock and a short acquirer stock in the same proportion as the exchange ratio. For example, with an offer of two acquirer shares for each target share, the arbitrage position is long one target share and short two acquirer shares. If the bid is subsequently revised, the arbitrage position must be adjusted to reflect the new exchange ratio. When the transaction closes, the arbitrageur receives in return for the target share the promised number of acquirer shares, which are used to cover the short position. The profit from a successful arbitrage position in a stock deal is the difference between the price of the short acquirer stock and the price of the target at the point in time when the position is established. If the bid fails, the arbitrageur will likely incur a loss from selling its target share holdings. The effect of closing out the short position in the acquirer is more uncertain: if the bidder stock falls, there may be an offsetting gain; and if the bidder stock appreciates, there may be additional losses.

Jindra and Walkling (2004) examine arbitrage spreads for 362 cash tender offers of publicly traded U.S. targets between 1981 and 1995. They document large cross-sectional variations in the initial arbitrage spread, with one-quarter of the targets exhibiting a negative spread (i.e., a trading price exceeding the offer price) and an average spread of 2% (median 2%). Arbitrage spreads are greater for lengthier contests and smaller for hostile targets, and they suggest that spreads reflect market anticipation of the duration and price resolution of the offer.

### *5.3.2. Arbitrage gains*

The magnitude of arbitrage returns depends on several factors, including the size of the arbitrage spread, the probability that the deal closes, the length of time that the arbitrageur must hold the position, and the target stock price development, if the deal fails. Several empirical studies document that merger arbitrage strategies tend to generate substantial excess returns. The largest abnormal returns have been documented for cash tender offers. For a sample of 295 cash tender offers from 1962 to 1980, Bhagat, Brickley and Loewenstein (1987) document an average target excess return of 2% from two days after the tender offer announcement to the day prior to the expiration of the offer (on average 29 days). Dukes, Frohlich, and Ma (1992) analyze 761 cash tender offers identified from 14D-1 filings in 1971–1985. They find average daily raw returns of 0.5%, or holding-period returns of 25%, for the average arbitrage-position holding period of 52 days. Jindra and Walkling (2004) report an abnormal monthly return of 2% for investments in the target stock from the day after the initial bid until bid resolution. Although continuous reinvestment at similar returns is unlikely, these studies indicate annualized excess returns ranging from 25% to over 100%.

Studies involving a mix of cash and stock, as well as tender and merger offers, also document positive, though smaller returns to merger arbitrage. Larcker and Lys (1987) examine a sample of 111 13-D filings in 1977–1983 that state arbitrage or participation in a takeover proposal as the purpose of the investment and are associated with an acquisition offer. They show that an arbitrage position held from the announcement date to the resolution of the offer (median of 31 days) generates a cumulative excess return of on average 5% (median 3%). Karolyi and Shannon (1999) study 37 takeover bids for Canadian publicly traded targets in 1997. They find an average abnormal return to a merger arbitrage strategy of 5% over a 57-day average investment period. Baker and Savasoglu (2002) report monthly abnormal returns of almost 1% from a merger arbitrage strategy for a portfolio of 1,901 U.S. takeover offers between 1981 and 1996.

The studies reviewed above collectively suggest that merger arbitrage strategies systematically generate excess risk-adjusted returns. The literature proposes various explanations for the existence of these returns. One is that risk arbitrageurs may be compensated for carrying the risk of deal failure. Jensen (1986) points to three important roles played by merger arbitrageurs for which they should be compensated: (1) they help value alternative offers; (2) they provide risk-bearing services for investors who do not want the uncertainty associated with the outcome of the takeover offer; and (3) they help resolve the free-rider problems of small, diffuse shareholders who cannot organize to negotiate directly with competing bidders for the target. Moreover, transactions costs and other practical constraints may limit the possibilities of successfully implementing an arbitrage strategy.

Larcker and Lys (1987) argue that the excess returns constitute compensation to arbitrageurs for assembling costly information related to the outcome of the bid. They show that the ex-post fraction of successful bids is significantly higher than the success probability implied by the arbitrage spread, suggesting that arbitrageurs have gathered private information about the deal outcome. In contrast, Cornelli and Li (2002) argue that the private information may be endogenous to the creation of the arbitrage position itself. The probability of offer success is positively related to the increased participation of arbitrageurs, since they are more likely to tender their target shares. The arbitrageur's investment in the target will therefore create an informational advantage, which can explain the profits earned by arbitrageurs. The model in Cornelli and Li (2002) predicts that the more liquid the stock, the easier it is to hide trades and the larger the arbitrage profits.[120]

Hsieh and Walking (2005) examine the importance of merger arbitrageurs for the market for corporate control using a sample of 680 all-cash and all-stock takeover offers during the period 1992–1999. They find that arbitrage holdings increase in offers that are likely to be successful, and suggest that this is evidence of the participation of passive arbitrageurs, whose accumulation of target stock does not affect the outcome of the deal. Hsieh and Walkling further find that these changes in arbitrage holdings are positively correlated to the probability of bid success, bid premia, and arbitrage returns. They interpret this as evidence of the involvement of active arbitrageurs, who influence the outcome and the terms of the deal.

[120] Gomes (2001) makes a similar argument where the entry of merger arbitrageurs creates large blocks of target shares that can hold out to a freezeout and hence forces the bidder to offer a higher preemptive bid.

### 5.3.3. Limits to arbitrage

The significance of transactions costs in limiting profits from merger arbitrage is investigated by Mitchell and Pulvino (2001). For a sample of 4,750 mergers from 1963 to 1998, they document annual excess returns to merger arbitrage of 10% when ignoring transactions costs. When accounting for transactions costs, such as brokerage commissions and the price impact of trading, the annual excess returns are reduced to 4%. Thus, transactions costs appear to limit but not entirely eliminate the excess profits generated by merger arbitrage strategies. Mitchell and Pulvino (2001) further show that merger arbitrage returns are correlated with market returns in a nonlinear way. Specifically, the returns are positively correlated in down markets but uncorrelated in flat and appreciating markets. They suggest that the excess returns are compensation to arbitrageurs for providing liquidity, especially during severe market downturns.

Arbitrage activity may be limited in practice because these investments are risky and require capital (Shleifer and Vishny, 1997). It is obvious that merger arbitrage in cash offers require capital, since the investor takes a long position in the target stock. Because the lender of a stock typically demands the short-sale proceeds as collateral, merger arbitrage positions require capital in stock-for-stock transactions too. Baker and Savasoglu (2002) propose that the capacity of arbitrageurs to carry risk is limited by the transaction risk and the size of their arbitrage position. They report that merger arbitrage returns increase with target size and a measure for the ex-ante deal completion risk. Moreover, there is some evidence that subsequent arbitrage profits are negatively related to changes in the supply of arbitrage capital.

Although merger arbitrage tends to be a profitable strategy, these trading strategies periodically generate large losses, primarily caused by unexpected deal failure (Baker and Savasoglu, 2002). Such liquidity events could affect the available supply of risk capital and hence the presence of arbitrageurs in subsequent deals. Officer (2007) examines the direct effect of two liquidity events—large arbitrage losses and the announcement of large deals-on arbitrage spreads. For a sample of 4,593 all-cash and all-stock offers in 1985–2004, he finds that risk returns are negatively related to big arbitrage losses, but this is attributable to the deal itself and has no contagion to other deals or spreads on pending deals. Overall, Officer (2007) finds little evidence indicating that large losses would cause withdrawal of arbitrage funds to the extent that it affects pricing in other merger and acquisition transactions.

Trading volumes typically increase in connection with the announcement of a takeover offer. Estimates of the target ownership by merger arbitrageurs following a takeover announcement ranges from 15% (Hsieh and Walking, 2005) to 35% (Officer, 2007). Yet, Geczy, Musto and Reed (2002) suggest that merger arbitrage strategies may be limited by short-selling constraints. They show that it is relatively expensive to borrow acquirer stock compared to other company stocks, in particular when the acquirer is small.

Mitchell, Pulvino, and Stafford (2004) study the effects of merger arbitrage for 2,130 mergers announced between 1994 and 2000. They document a substantial arbitrage activity after the announcement of a takeover offer. In the announcement month, the acquiring firm's short interest increases with a median of 40% in fixed-exchange-ratio stock mergers (i.e., where the exchange ratio is determined in the merger agreement). Interestingly, there is no corresponding change in the short interest for mergers where the arbitrage position does not involve a concurrent shorting of the acquirer stock, such as cash mergers or floating-exchange-ratio mergers (where the acquirer offers stock equivalent to a specific dollar value). The level of short interest falls dramatically when the merger closes. Also, the announcement effect of stock mergers is related to the change in short interest that occurs in the month of the announcement, suggesting a relationship between the arbitrage spread and the level of arbitrage activity.

To single out the effect of arbitrage trading activity, Mitchell, Pulvino, and Stafford (2004) further examine a subsample of 64 floating-exchange-ratio mergers. During the pricing period, which typically lasts 10 days and ends 5 days prior to merger closing, the corresponding number of acquirer stock is determined. In this type of stock mergers, the short selling of acquirer stock typically takes place during the pricing period. Since most of the deal uncertainty has already been resolved at this point, the effect of the short-selling pressure is no longer confounded with the revelation of new information about the merger. Importantly, the short interest increases significantly and there is a negative abnormal drift in the acquirer stock price of 3% during the pricing period for the floating-exchange-ratio mergers. Mitchell, Pulvino, and Stafford (2004) conclude that the short-selling by merger arbitrageur causes downward price pressure that accounts for almost half of the negative announcement return for acquirers in stock-financed mergers.

Overall, mergers and acquisitions of publicly traded firms attract substantial merger arbitrage activity. Such merger arbitrage strategies, betting on the closing of the transaction, seem to systematically generate positive excess returns. These returns reflect limits to arbitrage from transaction costs as well as compensation for carrying transaction risk.

## 6. Takeovers, competition and antitrust

In Section 4, we concluded that the typical merger produces significantly positive combined announcement-induced abnormal stock returns to bidders and targets. A standard interpretation is that the wealth effect is the present value of future expected increases in the merging firms' operating margins (the spread between future revenues and costs). In this section, we review studies that attempt to tease out whether the wealth effect predominantly originates in cost reductions (efficiency effects) or in revenue increases (market power effects).

### *6.1. Efficiency v. market power: predictions*

Eckbo (1983) and Stillman (1983) develop a test approach based on stock prices rather than product price data to infer the anticompetitive significance of horizontal mergers.[121] On the one hand, the (combined) abnormal stock returns to the bidder, and the target

[121] Examples of merger studies examining product price and output data are Barton and Sherman (1984) on microfilm; Borenstein (1990), Werden, Joskow, and Johnson (1991), Kim and Singal (1993) and Singal (1996)

cannot be used to discriminate between efficiency and market power hypotheses: these returns represent the *net* effect of expected cost reductions and revenue increases. On the other hand, merger-induced changes in expected future product and factor prices translate into abnormal stock returns to industry rivals (as well as upstream suppliers and downstream customers) of the merging firms. In particular, a collusive, anticompetitive merger raises the product price and thus benefits the nonmerging rivals as well. This means that evidence of a negative industry wealth effect of a merger announcement is inconsistent with the merger having collusive, anticompetitive effects on its industry.

A positive industry wealth effect is necessary but not sufficient to conclude in favor of the collusion hypothesis. The reason is that the industry wealth effect of an efficient merger may be either positive or negative. On the one hand, scale-increasing efficient mergers tend to have a negative impact on the industry's equilibrium product price, which harms rival firms and by itself causes a negative industry wealth effect.[122] On the other hand, news of the merger may reveal positive information about the value of the resources controlled by the rival firms. That is, the merger may reveal increased demand for resources owned by other firms, causing a positive revaluation of these rivals. For example, the increased demand may lead to expectations of future merger activity, resulting in a positive "in-play" effect on rival firms from the announcement of the initial merger. In sum, the efficiency hypothesis does not restrict the abnormal returns to industry rivals.

As summarized in Table 12, which is reproduced here from Eckbo and Wier (1985), these predictions can be refined further by distinguishing between public announcements that either increase or decrease the probability of a merger in the industry. The table adds predatory pricing as a variant of the market power hypothesis. The predation theory holds that the merger provides an incentive for the bidder firm to increase output and drive product prices down until rivals exit—at which point output is cut back to the monopoly level. Thus, both predation and productive efficiency arguments predict a lowering of the product price (albeit in the short run under the predation argument), which harms rivals.

An event decreasing the probability of the merger is the announcement of a decision by U.S. antitrust authorities (Department of Justice or Federal Trade Commission) to challenge the proposed merger with violation of antitrust laws (Section 7 of the 1914 Clayton Act).[123] As is seen in Table 12, the only pattern of abnormal stock returns to rival firms at once inconsistent with the market power hypothesis and consistent

on airlines; Akhavein, Berger, and Humphrey (1997), Prager and Hannan (1998) and Focarelli and Panetta (2003) on banking; and Dafny (2005) on hospital mergers.

[122] Rivals may minimize the negative product price impact by racing to adopt similar technological innovations as the merging firms—prompting industry merger waves.

[123] Section 7 of the Clayton Act replaced Section 2 of the 1890 Sherman Act as the principal federal antitrust law regulating corporate mergers and acquisitions. A *potential* threat to competition constitutes an offense under this law, and it is not necessary to prove a horizontal relationship between the combining firms. Furthermore, anticipated or demonstrated economic efficiencies are not a defense against the illegality of a merger that may lessen competition.

Table 12

Predicted abnormal returns to merging firms and their industry rivals under market power and productive efficiency hypotheses, classified by whether the event increases or decreases the probability of merger in the industry.

Examples of positive information effects on rival firms are the case where the merger announcement reveals possibilities for efficiency gains also available to non-merging rivals and the case where the merger signals an increase in demand for resources generally owned throughout the industry of the merging firms.

| Theory predicting the source of the merger gains | Abnormal returns to merging firms | Abnormal returns to industry rivals |
|---|---|---|
| **A. Events increasing the probability of merger (e.g. initial merger proposal announcement)** | | |
| **Market Power:** | | |
| (1) Collusion | *Positive* (monopoly rents) | *Positive* (monopoly rents) |
| (2) Predatory pricing | *Positive* (monopoly rents) | *Negative* (costs of price war) |
| **Productive Efficiency:** | | |
| (3) Productivity increase | *Positive* (cost savings) | *Negative* (competitive disadvantage) |
| (4) Information | *Positive* (undervalued resources) | *Zero or Positive* (undervalued resources and/or opportunities for productivity increases) |
| **B. Events decreasing the probability of merger (e.g. antitrust complaint blocks the merger)** | | |
| **Market Power:** | | |
| (1) Collusion | *Negative* (loss of monopoly rents) | *Negative* (loss of monopoly rents) |
| (2) Predatory pricing | *Negative* (loss of monopoly rents) | *Positive* (avoiding price war) |
| **Productive Efficiency:** | | |
| (3) Productivity increase | *Negative* (loss of cost savings) | *Positive* (avoiding competitive disadvantage) |
| (4) Information | *Zero* (information already out) | *Zero* (information already out) |

with the efficiency hypothesis is one where the rivals experience nonnegative returns in response to *both* probability-increasing and probability-decreasing events. Moreover, the collusion hypothesis is rejected unless one observes positive rival returns to the initial merger proposal followed by negative returns to news of the antitrust action. The predation theory is rejected unless a price pattern opposite to the pattern under the collusion theory is observed.

This shows that information on the abnormal return to rival firms in principle has the power to test market power hypotheses. This is true even if a given merger has a combination of productive efficiency and market power effects (so the rival firm performance reflects the net effect of the two). Tests of the predictions in Table 12 do, however, presume that collusion and predation are mutually exclusive market power effects (since you would otherwise be netting out positive and negative rival effects at both announcements). It is common in the theoretical literature, as well as in the practice of antitrust policy, to treat these two market power theories as separate.

Further refinements of the predictions in Table 12 are possible. Schumann (1993) suggests that market power theories may have different implications for rivals with small versus large market shares. Fee and Thomas (2004) and Shahrur (2005) follow the suggestion of Eckbo (1983) to examine the wealth effects of mergers also for customers and suppliers. The two papers develop similar predictions, reproduced here as Table 13 from Table 1 in Shahrur (2005). The major focus in these two studies is on the buyer power hypothesis (last column)—that is, the possibility that the merger increases the monopsony power of the combined firm over its input suppliers. In this case, the merger benefits the merging firms and (possibly) its industry rivals at the expense of upstream suppliers. Consumers benefit as well, provided some of the increased monopsony rents are passed on downstream from the merging firms' industry. Evidence on customer performance also helps resolve a possible ambiguity from looking at rival firm performance alone. For example, while evidence of a positive rival firm performance in response to the merger proposal announcement does not discriminate between collusion and efficiency, collusion is rejected if customers also benefit.

Tests of predictions such as those in Tables 12 and 13 are likely to pick up an in-play effect in the abnormal returns to rival firms in response to merger announcements. The in-play effect, which motivates the positive information effect predicted by hypothesis (4) of Table 12, occurs when the merger event increases the probability that the rivals may become targets. An in-play effect follows naturally from the fact that rival firms use similar production technologies and own some of the same (and possibly scarce) productive resources. A takeover may signal increased resource scarcity, causing a positive revaluation of every firm holding those resources. The findings of most of the the studies discussed below are consistent with such a positive industry information effect.[124]

Banerjee and Eckard (1998) and Fridolfson and Stennek (2005) suggest that since a successful merger bid eliminates rival firms as potential merger partners for the target, there could be a negative out-of-play effect for these rivals. Such a negative effect might attenuate a positive effect due to market power. In their sample of largely conglomerate takeovers (where there arguably are no market power effects), Betton and Eckbo (2000) document *positive* rival firm performance when the rival learns that it has lost the target. While the idea of an out-of-play effect is interesting and consistent with formal competitive takeover models such as Akdogu (2007b) and Molnar (2008), we are unaware of evidence favoring a significant out-of-play effect on rival firms.

[124] Exceptions are Eckbo (1992), Akdogu (2007a), and Becher, Mulherin, and Walkling (2008), who find a negative industry wealth effect of multi-industry horizontal merger announcements in Canada, and in single-industry studies of the U.S. telecommunications and utility firms, respectively.

Table 13
Predicted abnormal returns to merging firms, rivals, customers, and suppliers.
*Source*: Shahru (2005)

| | Productive efficiency | Collusion | Buyer power |
|---|---|---|---|
| **Merging firms:** | *Positive*<br>More-efficient production will result in higher infra-marginal rents to the merging firms | *Positive*<br>Higher likelihood of collusion will result in increased monopoly rents to the merging firms (Eckbo, 1983) | *Positive*[a]<br>Lower input prices due to intensified competition among suppliers (Snyder, 1996) |
| **Rivals:** | *Unrestricted*<br>Positive: information regarding industry-wide restructuring. Negative: more-intense competition in the industry due to a new, more-efficient combined firm (Eckbo, 1983) | *Positive*<br>Higher likelihood of collusion will result in increased monopoly rents to rival firms (Eckbo, 1983) | *Positive*[a]<br>Lower input prices due to more intense competition among suppliers (Snyder, 1996) |
| **Customers:** | *Unrestricted*[b]<br>Positive: scale-increasing mergers. Negative: scale-decreasing mergers | *Negative*<br>Restricted output in the takeover industry results in lower demand for suppliers' output | *Unrestricted*[c]<br>Positive: benefit from lower input costs for merging firms. Negative: supplier underinvestment |
| **Suppliers:** | *Unrestricted*[b]<br>Positive: scale-increasing mergers. Negative: scale-decreasing mergers and/or more-efficient combined firm | *Negative*<br>Restricted output in the takeover industry results in lower demand for suppliers' output | *Negative*[a]<br>The increased buyer power of the merging firms will intensify competition among suppliers (Snyder, 1996) |

[a]Efficient mergers can be of the scale-increasing or the scale-decreasing types (see, e.g. Eckbo, 1992; Andrade and Stafford, 2004). If the merger is expansionary in nature, it should benefit customers. Suppliers can benefit from a scale-increasing merger as long as the positive effect of expansion is not outweighed by the adverse effect of the increased efficiency of the combined firm. Finally, an efficient merger of the scale-decreasing type can hurt customers and suppliers.
[b]Snyder (1996) shows that by creating a larger buyer, a horizontal merger can result in more intense competition among suppliers, which will benefit the merging firms and their rivals at the expense of suppliers.
[c]Customers may benefit from the increased buyer power if some of the gains resulting from lower input prices are passed on to them because of competition in the takeover industry. Customers can also suffer if the increased buyer power induces suppliers to underinvest.

## *6.2. Effects of merger on rival firms*

Table 14 lists a number of empirical studies providing estimates of the industry wealth effects of horizontal mergers, beginning with Eckbo (1983) and Stillman (1983). Eckbo

Table 14
Stock market studies examining industry wealth effects of horizontal mergers

| Study | Merger Sample | Selection of Industry Rival, Customers and Suppliers | Key Findings |
|---|---|---|---|
| Stillman (1983) | 11 U.S. horizontal mergers challenged by DOJ or FTC with violating Section 7 of the Clayton Act, 1964–72 | Rival firms identified in antitrust litigation reports and antitrust enforcement agency fact memoranda | Zero average abnormal returns to rivals. Results inconsistent with market power effects of the sample mergers |
| Eckbo (1983) | 191 U.S. horizontal and 68 vertical mergers between mining or manufacturing firms operating nationally, 1963–78. 65 horizontal and 11 vertical mergers were challenged by DOJ or FTC under Section 7 | Rival firms selected from the major target industry using a *five*-digit SIC-based product classification procedure created by the author. For challenged mergers, the relevant industry is identified using court documents | Rival firms earn zero or positive abnormal returns in response to **both** the initial merger proposal announcement and the subsequent antitrust complaint. Results inconsistent with the market power (collusion) effects of the horizontal mergers |
| Eckbo and Wier (1985) | 82 U.S. challenged horizontal mergers, 1963–81, including 65 from Eckbo (1983). 17 cases occurred after the passage of the 1978 Hart-Scott-Rodino Antitrust Improvements Act | Two sets of rivals: one based on five-digit SIC codes as in Eckbo (1983), and another identified in antitrust litigation reports | Rival firm performance inconsistent with market power either before or after 1978. This conclusion holds for both sets of rivals |
| Eckbo (1985) | 266 U.S. merger proposals (196 horizontal, 98 challenged), including those in Eckbo (1983) and Eckbo and Wier (1985), 1963–81 | Rivals selected using the 5-digit SIC code procedure of Eckbo (1983) | Industry wealth effect *negatively* related to merger-induced increase in industry concentration. Inconsistent with predictions of the Market Concentration Doctrine |
| Eckbo (1992) | 471 merger proposals (312 horizontal), 266 between U.S. firms and 205 between Canadian firms, 1963–82. 80 of the U.S. horizontal mergers were challenged under Section 7, none of the Canadian mergers were challenged | Rivals for both U.S. and Canadian mergers identified using the 5-digit SIC code procedure of Eckbo (1983) | No evidence of market power despite no antitrust deterrence of anticompetitive mergers in Canada until 1985. Industry wealth effect *negatively* related to merger-induced increase in industry concentration. |
| Schumann (1993) | 37 acquisitions challenged by the FTC, 1981–87 | Rival firms identified using antitrust litigation reports | Positive rival firm performance at merger proposal and zero or positive at antitrust complaint. At the complaint, rival returns lower the greater the merger-induced change in concentration |

*(Continued)*

Table 14 *(Continued)*

| Study | Merger Sample | Selection of Industry Rivals, Customers and Suppliers | Key Findings |
|---|---|---|---|
| Fee and Thomas (2004) | 554 proposed (four-digit SIC) U.S. horizontal mergers between publicly traded firms from SDC, 39 challenged under Section 7, 391 deals completed, 1981–1997 | Identifies single- and multiple-segment rivals in same 4-digit horizontal industry using Compustat industry segment information. The segment information also helps identify customers and suppliers | Evidence inconsistent with increased monopolistic collusion, but consistent with improved efficiency and buying power of merged firm |
| Shahrur (2005) | 463 successful (four-digit SIC) U.S. horizontal mergers and tender offers from SDC, 1987–99 | Identifies single-segment rivals, and customer and supplier firms, using Compustat industry segment information | Evidence inconsistent with increased monopolistic collusion, but consistent with improved efficiency and buying power of merged firm |
| Aktas, deBodt, and Roll (2006) | 290 proposed horizontal mergers in the European Union, of which 55 were subjected to "in depth investigation" for potential antitrust violation. Bidder is a non-EU firm in 104 cases | Rival firms identified in same country and industrial sector as target using Hoover's Online Database, European Commission Web Site, and Datastream | Negative rival abnormal performance around merger proposal announcements. Suggests the sample mergers enhance industry competitiveness |
| Bhattahcaryya and Nain (2006) | 615 successful (four-digit SIC) U.S. horizontal mergers, 1989–2000, from SDC | Use Bureau of Economic Analysis benchmark Input-Output tables to identify the fraction of the supplier industry's output sold to the merging industry | Conclude with increased buying power based on post-merger decline in supplier product prices |
| Akdogu (2007a) | N = 275, of which 115 (four-digit SIC) U.S. horizontal takeover bids in the telecommunications industry, 1996–2005 | Rivals firms identified using SDC and CRSP, using SIC code 4813 | Evidence of negative industry wealth effect of the acquisition bids. Conclude that acquirers are on average expected to gain competitive advantage from the takeovers. |
| Becher, Mulherin, and Walkling (2008) | 384 successful mergers between electric utilities, 1980–2004 | Rival firms are all public utilities with assets > \$500 mill. | Evidence inconsistent with market power but consistent with efficiency (synergy) effects of the horizontal mergers |

*(Continued)*

Table 14 *(Continued)*

| Study | Merger Sample | Selection of Industry Rivals, Customers and Suppliers | Key Findings |
|---|---|---|---|
| Shenoy (2008) | 453 successful vertical mergers, 1981–2004 | Use Bureau of Economic Analysis benchmark Input-Output tables to identify the main customer industries and rivals. Customer portfolios based on single-segment firms | Evidence fail to support market power but consistent with vertical mergers having efficiency effects |

(1983) examines intra-industry wealth effects of 191 horizontal mergers in the United States between 1963 and 1978, 65 of which were challenged by either the Department of Justice or the Federal Trade Commission with violating Section 7 of the Clayton Act. A sample of 68 vertical mergers, of which 11 were challenged, is also examined. For each merger, a set of horizontal competitors of the merging firms that were listed on the NYSE or the American Stock Exchange (ASE) at the time of the merger proposal announcement is identified.

The rivals are defined based on overlapping five-digit Standard Industrial Classification (SIC) codes. For the challenged mergers, the relevant product market is the one identified in court records as being the market "threatened" by the "anticompetitive" merger. For unchallenged mergers, the relevant product market is the target's major product line, as defined in Standard & Poor's Registry of Corporations. As shown by Eckbo and Wier (1985), the empirical results based on the five-digit SIC rivals are robust: They duplicate the tests using rivals identified by the Department of Justice (DOJ) or the Federal Trade Commission (FTC) as being relevant competitors, and they draw the same inferences. To test the hypotheses in Table 12, the paper reports estimates of the abnormal stock returns to the merging firms and their horizontal rivals relative to (i) the merger proposal announcement and (ii) the subsequent announcement that the DOJ or the FTC has filed a Section 7 complaint against the horizontal merger.

Eckbo (1983) reports that the observed sequence of abnormal returns across the proposal and antitrust complaint announcements does not follow the pattern predicted by the collusion hypothesis. Rivals of the 65 horizontal challenged mergers earn small but significantly positive abnormal returns around the merger proposal announcement, followed by zero or positive abnormal returns in response to the antitrust complaint announcement.[125] The antitrust complaint causes a negative average abnormal return of −10% to the *merging* firms. This means that the antitrust complaint comes as a surprise

[125] The industry wealth effect of the merger is estimated by first forming an equal-weighted portfolio of the rivals of a given target and then estimating the portfolio abnormal return. Let $AR(t_1, t_2)$ denote the average of these portfolio estimates, where the period of cumulation is $t_2 - t_1 + 1$ days around the event-announcement day (day 0). Eckbo (1983) reports $AR(-1,1) = 0.10\%$ and $AR(-10,5) = 1.17\%$ relative to the merger proposal (significant at the 5% level), and an insignificant $AR(-1,1) = 0.17\%$ relative to the antitrust complaint

to the market, which in turn means that the complaint announcement has the requisite power to test the market power hypothesis using rival firm returns.

This pattern of abnormal return to rival firms is inconsistent with the predictions as summarized in Table 12 of the collusion hypothesis, but it is jointly consistent with the efficiency and information arguments. Stillman (1983) performs a similar set of tests on 11 horizontal, challenged mergers for the period 1964–1972 and finds zero average abnormal rival stock returns relative to both merger announcements and antitrust complaints. Thus, both Eckbo and Stillman conclude against the market power hypothesis. Eckbo (1983) also reports that the average intra-industry wealth effect of unchallenged horizontal mergers is indistinguishable from the average intra-industry wealth effect of unchallenged vertical mergers. Since vertical mergers are unlikely to have collusive effects, this further supports the view that the horizontal unchallenged mergers in the sample were not expected to be anticompetitive.

Schumann (1993) also examines the effect of horizontal merger proposals and antitrust complaints on rival firms. His sample consists of 37 cases from 1981–1987 that were challenged by the FTC, a period with less antitrust intervention than in the sample periods of the earlier studies. Rival firms are identified using antitrust litigation reports, much as in Eckbo and Wier (1985). The results for the total sample, which indicates significantly positive rival returns at the proposal announcement and zero at the time of the antitrust complaint, are "remarkably similar to those reported in Eckbo (1983) and Eckbo and Wier (1985)" (Schumann, 1993, p. 681). For a subsample of 97 rivals with available data on market shares, Schumann also reports that rivals in the smallest market-share quartile have the largest abnormal returns and that these are significantly positive at both the proposal and complaint events. Following the predictions in Table 12, these findings contradict the collusion (market power) hypothesis.

Several studies also document significantly negative abnormal returns to rival firm portfolios in response to the announcement of horizontal mergers. Eckbo (1992) finds a negative industry wealth effect of horizontal merger announcements in Canada. Aktas, deBodt, and Roll (2006) study horizontal mergers in the European Union, several of which were subjected to a preliminary antitrust review. They report significantly negative rival abnormal returns. Akdogu (2007a) finds negative rival abnormal returns in response to horizontal takeover bids in the U.S. telecommunications industry. Becher, Mulherin, and Walkling (2008) document a significantly negative industry wealth effect in a large sample of horizontal mergers between U.S. electric utilities. All of these studies reject the market power hypothesis and conclude that the typical sample merger would likely enhance efficiency.

### *6.3. Effects of merger on suppliers and customers*

Fee and Thomas (2004) and Shahrur (2005) estimate the effect of horizontal mergers on rivals and, in particular, on upstream suppliers and downstream customers, over the

announcement. Eckbo (1983) also reports that complaints by the FTC cause a significantly *positive* industry wealth effect of $AR(0) = 0.74\%$.

period 1981–1999. These two studies provide new tests of market power theories based directly on the wealth effects for suppliers and customers. Moreover, by revisiting abnormal returns to rivals during a time period with relatively lax U.S. antitrust enforcement in the merger area, they provide a perspective on the generality of the findings of earlier studies on industry wealth effects.

Starting with the evidence on rival firms, both studies report a statistically significant, positive industry wealth effect in response to the merger proposal announcement. Abnormal returns to portfolios of single-segment rival firms average 0.54% in Fee and Thomas (2004) and 0.39% in Shahrur (2005). These results are close to the early results in Eckbo (1983), and to those in Song and Walkling (2000). The evidence confirms that the composition of the rival firm portfolios in this literature yields sufficient power to register industry wealth effects of horizontal mergers.[126] Moreover, Eckbo (1983) as well as Fee and Thomas (2004) report significantly positive rival firm abnormal returns in response to news of antitrust action against the proposed merger.[127] Thus, consistent with the earlier literature, Fee and Thomas (2004) reject the collusion hypothesis for the horizontal-merger gains.

Fee and Thomas (2004) and Shahrur (2005) identify customer and supplier information using Compustat's industry segment files. These files record information mandated by Federal Accounting Standards Board (FASB) rule No. 14 during their sample period. Under this rule, firms are required to report financial information for any industry segment comprising more than 10% of consolidated yearly sales, assets, or profits.[128] This reporting requirement also discloses the identity of any customer representing more than 10% of the total sales, as well as the company segment that was primarily responsible for those sales. Both studies also use sales data to identify suppliers and customers that are particularly dependent on the industry of the merged firm.

Under the monopoly (collusion) hypothesis, the merging firms and their rivals gain at the expense of customers. Fee and Thomas (2004) and Shahrur (2005) reject this hypothesis because they find no systematic evidence of customer losses, even for customers that are particularly reliant on the merging firm's industry. There is also evidence that the mergers with the largest gains to the merging firms also produce gains to customers. As Fee and Thomas (2004) conclude, "Taken together, the customer and rival results are strongly inconsistent with the monopolistic collusion hypothesis" (p. 457). Shahrur (2005) states that "Our overall evidence suggests that the lenient antitrust policy in recent years does not appear to have resulted in predominantly anticompetitive takeovers" (p. 95). These results support the conclusion in Eckbo (1992) that, when it comes to the need to use antitrust policy to strongly deter potentially anticompetitive mergers, "Judging from the evidence, there simply isn't much to deter" (p. 1028).

126 See Eckbo (1983, 1989, 1992), McAfee and Williams (1988), and Schumann (1993) for discussions of power issues in the selection of rivals.

127 Shahrur (2005) does not study antitrust events.

128 After 1998, SFAS No. 131 governs required segment disclosures.

Fee and Thomas (2004) and Shahrur (2005) find some evidence of losses to upstream suppliers of the merging firms and conclude that horizontal merger tends to increase buying power. Increased buying power follows if the merger increases monopsony power or if it forces upstream suppliers to be more efficient. Fee and Thomas (2004) argue that if the source of buying power is upstream efficiency, then the losses to suppliers will be asymmetric—with losses only to those suppliers that are not retained post-merger. That is, those suppliers that lose a bidding competition post-merger would suffer. Fee and Thomas do in fact find that the wealth effect for suppliers depends significantly on the supplier's ability to retain its product-market relationship with the merged entity. Only the suppliers that are terminated experience negative abnormal returns around the merger announcement and significant negative cash flow changes post-merger. Suppliers that are retained experience increases in market share and do not show evidence of abnormal stock returns or changes in operating performance. The authors therefore conclude that the effect of the merger on suppliers reflects efficiency-increasing buying power. Shahrur (2005) reaches a similar conclusion: "Along with the evidence in Fee and Thomas (2004), our results suggest that industry consolidations can help increase the efficiency of upstream industries" (p. 96).

Bhattahcaryya and Nain (2006) and Shenoy (2008) also focus on vertical buying power. Bhattahcaryya and Nain (2006) sample 615 successful horizontal mergers and fail to find a significant announcement effect on the horizontal rivals. However, they find evidence of a reduction in the product price paid to upstream suppliers, which is consistent with increased buying power. Moreover, they find some evidence indicating that the upstream suppliers, perhaps feeling the squeeze from the increased buying power, restructure to counter the effect of the downstream horizontal merger. The authors suggest that the net effect of all this may have been to leave the market value of the horizontal rivals of the merging firms unchanged. Finally, Shenoy (2008) studies the industry wealth effects of 453 successful *vertical* mergers and concludes that these on average have efficiency effects. This evidence is also consistent with the effects of vertical mergers first reported by Eckbo (1983).

### *6.4. Some implications for antitrust policy*

#### *6.4.1. The market concentration doctrine*

The U.S. government selects Section 7, Clayton Act, cases against horizontal mergers largely on the basis of market share and industry concentration.[129] The government agencies' reliance on structural standards for selection of merger cases is rooted in one

[129] The Justice Department's Merger Guidelines of 1968 state market shares that were likely to trigger an antitrust complaint. The critical aggregate market shares varied according to the four-firm market concentration ratios. For example, a merger between two firms each having 4% of the sales in a market with a four-firm concentration ratio of 75% or more was likely to be challenged. The department's 1982 Merger Guidelines use the Herfindahl Index of concentration and are somewhat less restrictive than the old guidelines, but their focus is also on market structure.

of the oldest propositions in industrial economics: the market concentration doctrine. This doctrine holds that the level of industry concentration is a *reliable* index of the industry's market power. The empirical implication is that a relatively high level of industry concentration, which in the presence of entry barriers is believed to facilitate intra-industry collusion or dominant-firm pricing, should be associated with relatively large industrywide monopoly rents.

A horizontal merger produces a measurable change in the industry's level of concentration and a change in the risk-adjusted present value of industry rents that is directly associated with the concentration change. Under the market concentration doctrine, this change in industry rents is positively correlated with the change in concentration. This value-based test in the *changes* of the two variables (industry rents and concentration) allows more specific inferences than can be drawn from a correlation between the levels of (accounting) profits and concentration.

Eckbo (1985, 1992) provides empirical tests of the market concentration doctrine by estimating cross-sectional regressions of the following form:

$$AR_i = \alpha_0 + \alpha_1 CR_i + \alpha_2 dCR_i + \beta' \mathbf{Z_i} + e_i \tag{25}$$

where $CR_i$ is a measure of the pre-merger level of concentration in the industry where the horizontal merger is taking place, $dCR_i$ is the change in concentration caused by the merger, $\mathbf{Z_i}$ is a set of firm- and industry-specific control variables, and $AR_i$ is the abnormal return to an equal-weighted portfolio of the rivals of the merging firms around the merger proposal announcement. Under the market concentration doctrine, and assuming the sample includes some anticompetitive mergers, one should find that $\alpha_2 > 0$. This is because the $AR$ of rivals of an anticompetitive merger represents increased monopoly rents, and the market concentration doctrine holds that the increase in monopoly rents will be larger the larger the increase in concentration caused by the merger. Furthermore, under the stronger proposition embedded in antimerger policy, which holds that a merger is more likely to have anticompetitive effects the larger the pre-merger *level* of concentration, one should also find evidence of $\alpha_1 > 0$.

While the form of Equation (25) is similar in spirit to the regression models typically estimated in the "structure-conduct-performance" (industrial organization) literature, there are some notable qualitative differences: For example, while the dependent variable $AR$ measures directly the market value of the increase in industry profits expected to follow from the (merger-induced) increase in industry concentration, the tradition has been to regress an accounting measure of the level of industry profits on the level of concentration. The traditional approach has been criticized on the grounds that accounting profits are a poor proxy for economic profits and that any cross-sectional variation in the level of industry profits can simply reflect differences in risk. This criticism does not apply here, since $AR$ is measured using market values and represents a risk-adjusted change in the level of industry rents. Equally important is the fact that since Equation (25) is specified in the form of changes in the central variables, $\alpha_2$ can be meaningfully interpreted without specifying a complete structural model relating the level of industry profits to concentration.

Eckbo (1985, 1992) uses the four-firm concentration ratio ($CR_4$) of the major four-digit SIC industry of the target firm to represent $CR$, while the change in the industry's Herfindahl Index ($dH$) measures $dCR$.[130] While data on $CR_4$ is generally available, the market shares of the bidder and target firms, which yield $dH$, were collected from case-related court records and publications. In the sample of challenged mergers in Eckbo (1985), the average level of $CR_4$ is 58% (ranging from 6 to 94%), while the average value of $dH$ is 3.3% (ranging from 0.02 to 24.2%).

Both studies report a statistically significant negative coefficient on $dH$. In Eckbo (1985), increasing $dH$ by 1% implies a reduction of 0.42% in the abnormal returns to the average portfolio of rival firms. Similar results emerge when one uses the abnormal returns to the *merging* firms as dependent variable. The author notes that since the regressions of the type in Equation (25) are based on challenged mergers, the results are biased in favor of the market concentration doctrine. Despite this potential bias, there is no evidence supporting the doctrine.[131]

### *6.4.2. Did the 1978 antitrust improvements act improve antitrust?*

The 1976 Hart-Scott-Rodino (HSR) Antitrust Improvements Act significantly increased the legal powers of the law enforcement agencies to obtain private information needed for judging a merger's anticompetitive impact before filing a complaint. The HSR Act addressed two perceived handicaps borne by the agencies charged with enforcing Section 7 of the Clayton Act. First, under the 1962 Antitrust Civil Process Act, the DOJ could not require third parties, such as competitors and trade associations, to provide information about corporate acquisitions until after a Section 7 complaint had been filed. This frequently caused the DOJ to drop an investigation altogether for lack of information or to file a "skeleton" complaint based on scanty data. HSR established the right of the DOJ to issue Civil Investigative Demands to the merging firms and to other parties not directly involved in the merger prior to filing a complaint. Second, until the HSR Act, the government could not require postponement of proposed acquisitions pending investigation. HSR required firms planning mergers to prenotify the FTC and the DOJ of the transaction, providing the agencies with time and information to prepare a case before merger consummation.

According to the FTC,[132] the notification requirements and delay assure that "virtually all significant mergers or acquisitions occurring in the United States will be reviewed by the antitrust agencies prior to the consummation of the transaction." Moreover, the

[130] $CR_4 \equiv \sum_{i=1}^{4} s_i$, and $H \equiv \sum_{i=1}^{n} s_i^2$, where $s_i$ is the market share of firm $i$ (in $CR_4$ the sum is over the four firms with the largest market shares) and $n$ is the total number of firms in the industry. The change in the Herfindahl Index caused by the merger between firms $i$ and $j$ in the same industry is therefore given by $dH = 2s_i s_j$.

[131] "[T]he evidence systematically rejects the antitrust doctrine even for values of industry concentration and market shares which,over the past four decades, have been considered critical in determining the probability that a horizontal merger will have anticompetitive effects" (Eckbo, 1992 p. 1028).

[132] 6 FTC Ann. Rep. to Cong. concerning HSR ACT 11 (1983).

information provided by the parties "usually is sufficient for the enforcement agencies to make a prompt determination of the existence of any antitrust problems raised by the transaction." These assurances notwithstanding, Eckbo and Wier (1985) compare the anticompetitive significance of horizontal mergers challenged before and under HRS and find no significant difference in their effect on rival firms. Moreover, they conclude that the pattern of abnormal stock returns to the industry rivals is inconsistent with the mergers having collusive anticompetitive effects both before and under the HSR. Based on this, they argue that HSR is unlikely to have significantly increased the precision with which defendants are chosen by the antitrust agencies.

Why would the antitrust process, which many believe is designed to protect consumer interests, result in blocking efficient mergers? Bittlingmayer and Hazlett (2000) suggest bureaucratic self-interest, political extraction, and private benefits. In this context, it is important to keep in mind that, while preventing efficient mergers harms consumers, the *rivals* of the merging firms benefit as they avoid having to face competition from an increasingly efficient merged firm. The rivals can indeed form a politically strong interest group in situations where they perceive a significant threat to their existing industry equilibrium. This industry capture theory is highlighted by Posner (1969) who asserts that the FTC is significantly impaired in its task of promoting the public interest; Posner claims that its investigations are initiated "at the behest of corporations, trade associations, and trade unions whose motivation is at best to shift the costs of their private litigation to the taxpayer and at worst to harass competitors" (p. 88).

A case in point is Chrysler's vocal opposition to the joint venture between GM and Toyota in 1983. At the time the venture was announced, Chrysler demanded publicly that the FTC take action to stop the venture because it would "harm competition" in the automobile industry. An alternative interpretation of Chrysler's opposition is that it suspected the venture would make GM a tougher competitor, placing Chrysler at a competitive disadvantage. In fact, Eckbo (1990a) finds significant abnormal returns of −9% to Chrysler upon the announcement of the GM-Toyota joint venture. More recent cases in point include the airline industry, where Slovin, Sushka, and Hudson (1991) conclude that Civil Aeronautics Board interventions during 1965 to 1988 reduced competition and favored collusion among existing carriers. Bittlingmayer and Hazlett (2000) study the effect on the software and computer industry of 54 antitrust enforcement actions against Microsoft over the period 1991–1997, and strongly reject the thesis that these actions would enhance efficiency. Also, Aktas, deBodt, and Roll (2006), who study rival firm performance following antitrust interventions against mergers in the European Union, find evidence consistent with antitrust policy being used to protect EU firms from outside competition.

Since the anticompetitive significance of a horizontal merger does not represent a directly observable characteristic, policy makers are forced to rely on largely untested theories to justify their decisions. As noted by Stigler (1982), the economics profession has supplied "precious little" in the way of *tested* knowledge to support the market share and concentration criteria that (still) form the basis for U.S. antimerger policy. As long as those responsible for antimerger policy continue to insist on rigid structural standards for evaluating the competitive effects of mergers, it is reasonable, given the evidence, that special interest groups, including those representing relatively inefficient producers and/or a rigid workforce, will continue to exploit antitrust policy toward merger.

## 7. Summary and conclusions

Table 15 summarizes key findings across the various topics we have surveyed. Here, we draw broad inferences from these findings and point to interesting but unresolved issues.

### *7.1. Takeover activity*

While there are clear patterns of merger waves in the data, there is little agreement on the basic sources of the waves. Under neoclassical theories, basic sources include industry-specific technological and demand shocks, regulatory changes, and liquidity constraints. Under behavioral arguments, mergers are driven by attempts to sell overpriced assets and securities and herding behavior. There is evidence that regulatory changes and liquidity factors predict industry waves. There is also evidence of greater average market-to-book ratios during periods of merger waves, which may (but need not) indicate overvaluation. On the other hand, additional research is needed on the extent to which bidders select stock as payment in response to market-to-book ratios and on whether the presence of mixed cash-stock offers (which are typically as frequent as all-stock offers) are consistent with equilibria in which targets willingly accept overpriced bidder shares.

Perhaps the most straightforward way to advance our understanding of aggregate merger activity is to model the takeover process from basic, microeconomic principles. One does not get something from nothing—so this requires imposing various restrictive (but hopefully testable) assumptions on production technologies and market structures. The theoretical literature on the optionality of corporate investments is a promising avenue, as are models of industry competition in which industry shocks force rival firms to restructure. Empirical research tailored to such modeling efforts is only starting to emerge.

Important stylized facts from the aggregate takeover activity in the 1980–2005 period include (1) the stability of horizontal combinations at 30%–40% of the total takeover population, (2) negotiations (as opposed to open auction) as the preferred route to acquiring control (3) the sharp drop-off in hostile takeovers after 1988, (4) the large increase in volume and deal values involving public bidders toward the end of the 1990s, (5) the predominance of all-cash and mixed cash-stock offers in tender offers, (6) the rise of mixed stock-cash offers to become the most frequently used payment method in mergers by 2001, and (7) the dramatic fall in toehold bidding since the mid-1980s.

Additional research is needed to sort out the competing theories for the sharp drop in hostile takeovers and what this drop means for the market for corporate control to function effectively as the court of last resort. While takeover activity depends on market liquidity factors (and thus fell during the credit crunch of the late 1980s), it is also important to

Table 15

Summary of empirical results on corporate takeovers, classified by research topic

| Topic |
| --- |
| **A. Takeover Activity:** |
| (1) Merger waves (clustering of takeovers) tend to occur in periods of market booms. They occurred in the late 1800s and early 1900s (“the monopolization wave”), the late 1960s (“the conglomerate wave”), the mid 1980s (“the refocusing wave”), and the late 1990s (“the strategic/global wave”). |
| (2) There is substantial evidence of industry-clustering of mergers. Regulatory changes and macroeconomic liquidity variables are better predictors of industry merger waves than are market-to-book ratios. |
| (3) In the period 1996–2000, when market valuations were particularly high, the sum of all-cash and mixed cash-stock bids was equal to the number of all-stock bids. Also, in this period, the proportion all-stock offers was the same as during the previous five-year periods. |
| (4) Despite strong merger patterns, predicting target firms with any accuracy has proven difficult. |
| (5) Target firms increasingly initiate the takeover process by soliciting bid indications from a set of potential negotiating partners. The bidder that is selected is recorded as the first bidder in SEC registration documents and therefore by data bases such as SDC (Thomson Financial). |
| (6) When organizing all SDC control bids into contest for U.S. targets, there were a total of 35,727 control contests. Of these, the initial bidder proposed a merger in 28,994 cases and made a public tender offer in another 4,500 cases (the balance being 2,224 controlling-block trades). |
| (7) In constant 2000 dollars, the merger deal was valued at \$436 million on average (median (\$35 mill.), while the deal value of the average tender offer was \$480 (median \$79 mill.). |
| (8) SDC provides information on the payment method for about half of the cases. Of these, 26% were all-cash deals, 37% were all-stock deals, and 37% were mixed cash-stock deals. All-cash and mixed offers have similar deal sizes, slightly above all-stock deals. |
| (9) A total of 590 initial bids are classified as “hostile” and another 435 deals are “unsolicited”. Hostile bids have substantially higher than average deal values. |
| (10) In approximately thirty percent of all deals, the initial bidder and target operate in the same four-digit SIC industry (horizontal takeover). The two most active takeover sectors are Manufacturing, and Finance/Insurance/Real Estate. |
| (11) Two-thirds of the 35,727 initial bidders are public companies, while 37% of the targets are public. In 44% of the initial bids, a public bidder is pursuing a private target (the largest single group of takeovers), with an average deal value of \$114 mill. (median \$23 mill.). The total number of deals involving either a public bidder or target rose sharply in the 1990s. |
| (12) Of the 35,727 initial bidders, 11% were foreign companies (primarily Canada and the UK). Deals involving foreign bidders are relatively large. |
| (13) The time from the initial offer to the effective takeover date averages 108 trading days (median 96) when the initial bid is a tender offer, and 71 days (median 49) for merger bids. In cases where there are more than one control bid for the target, the time from the first to the second bid averages 40 trading days (median 19). |
| (14) The likelihood that the initial bidder wins the target is higher when the bidder has a toehold, when the payment method is all-cash, when the bid form is tender offer, and when the bidder is a public company. The probability of winning is lower for targets with poison pills, and when the target reaction is negative. All bids fail (no bidder wins) in 22% of the cases, with a greater failure probability for private bidders. |

*(Continued)*

Table 15 *(Continued)*

| Topic |
|---|
| **B. The Payment Method:** |
| (15) Bidders initiating takeover bids for U.S. targets over the period 1980–2005 offered all-cash as payment in 26% of the cases, all-stock in 37%, and a mix of stock of cash in 37%. |
| (16) The majority of tender offers are all-cash or a mix of cash and stock. while the majority of merger bids are in the form of all-stock (with the exception of the 1980–85 period where most merger bids offered a mixed cash-stock payment). |
| (17) In the two subperiods 1990–1995 and 1996–2000, the percentage all-stock offers in initial merger bids were approximately 55% in *both* period. This means that (1) nearly half of the initial merger bids in the 1990s use some cash as payment, and (2) the percentage all-stock merger bids remained unaffected by the significant runup in overall market valuations in the 1996–2000 period. |
| (18) The payment method choice is in part determined by tax considerations, the degree of information asymmetry between the bidder and the target, the degree of market mispricing of bidder stock, and by corporate control considerations. Stock offers are more likely the greater the bidder's asset size and market-to-book ratio. Stock offers are less likely the greater the bidder management's shareholdings and the greater the dispersion in analyst forecast of bidder earnings. |
| (19) Offer premiums are greater in all-cash offers than in all-stock offers. The probability that the initial bidder wins the target is lower for all-stock offers than for cash offers. |
| (20) When the target is public, bidder announcement returns are on average negative in all-stock offers and greater in all-cash and mixed cash-stock offers than in all-stock offers. Moreover, bidder announcement-induced stock returns are increasing in the cash-portion of the (mixed) offer. |
| (21) When the target is a private company, stock offers generate positive bidder announcement returns that are as high—if not higher—than for all-cash bids. |
| **C. Toehold Bidding:** |
| (22) The frequency of toehold bidding in friendly mergers and tender offers has fallen dramatically since the 1980s. Over the 1990–2002 period, 7% of bidders initiating a takeover had toeholds, and only 2% had toeholds acquired in the market shortly prior to launching the bid. |
| (23) Toehold bidding remains common in hostile bids (50% frequency). |
| (24) Toeholds are large when they exist: on average 20%. |
| (25) Toehold bidders tend to pay lower offer premiums and win the target more often than zero-toehold bidders. |
| (26) The presence of a bidder toehold attenuates the drop in the target share price when all bids fail. |
| (27) Since bidder toehold benefits mirror target toehold costs (lower offer price, greater probability of target management being replaced) toehold bidding may be viewed as aggressive by the target. Thus, approaching the target with a toehold may cause some otherwise friendly targets to refuse negotiations. Consistent with this, the data indicates a significantly negative association between the likelihood of the initial bidder approaching with a toehold and the expected value of resistance costs (including the opportunity loss of a termination agreement). |

*(Continued)*

Table 15 *(Continued)*

| Topic |
|---|
| **C. Bid Jumps and Markup Pricing:** |
| (28) The average offer premium in successful single-bid takeover contests is somewhat higher than the average initial offer premium in multi-bid contests. This is consistent with the greater premium preempting competition in ex-post successful single-bid cases. |
| (29) Bid revisions are substantial, with an average bid jump from the first to the second bid in the contest of 10% (a 31% change in the offer premium). |
| (30) A dollar increase in the pre-offer target share price runup causes the initial bidder to mark up the total offer premium by $0.80. |
| (31) Markup pricing notwithstanding, bidder takeover gains are increasing in the target runup. Thus, takeovers with greater target runups are more profitable for *both* bidder and target firms, which may also explain why bidders agree to (partial) markup pricing. |
| (32) Toehold acquisitions during the runup period bidder increases the target runup. When the toehold is acquired by the initial bidder, however, the markup is reduced. No such markup reduction is observed when the toehold is acquired by another investor. |
| **D. Takeover Defenses:** |
| (33) The presence of a majority of independent directors on the board of the target is viewed by the court as a strong indication of satisfaction of the fiduciary duty of loyalty. |
| (34) Delaware case law sanctions the right to "just say no" to an unsolicited takeover bid. That is, the board may determine in good faith that the continuing independence of the corporation is in the long-term best interest of the corporation and its stockholders. |
| (35) If the board's defensive response is not "draconian" (i.e., it is neither coercive nor preclusive) but "within the range of reasonableness" given the perceived threat, the board is protected by the business judgement rule. A defense that is deemed preclusive because it frustrates, impedes or disenfranchises a shareholder vote is unlikely to be upheld. |
| (36) The twin defense of staggered board election and a poison pill ("shareholder rights plan") is "draconian" in the eyes of many economists but not the court. However, "dead hand" pills (where only directors not up for election may vote to rescind the pill) have been struck down. |
| (37) The fraction of "hostile" (sum of unsolicited bids and bids where target is explicitly hostile) drops sharply after 1989, from more than 20% in the 1980s to less than 3% by the end of the 1990s. |
| (38) Offer premiums are no lower for targets with poison pills. |
| (39) There is a small but significantly negative market reaction to the adoption of strong antitakeover amendments such as poison pills and staggered board. The market reacts positively to antigreenmail amendments provided these occur when a takeover is rumored. |
| **E. Targets in Bankruptcy:** |
| (40) There is a trend towards market-based mechanisms for resolving Chapter 11 cases, including sale of the firm to a bidder. Target firms that are sold spend less time in Chapter 11, which lowers bankruptcy costs. Acquirers tend to be in the same industry, and premiums paid are on average lower than in takeovers of non-bankrupt firm in the same industry. |

*(Continued)*

Table 15 *(Continued)*

| Topic |
|---|
| (41) Premiums paid for targets sold in mandatory, open, first-price, all-cash bankruptcy auctions in Sweden suggest the possibility that the auction mechanism may work well for the typical Chapter 11 case as well (which is of a similar size as the Swedish sample firm). |
| (42) The average mandatory auction receives three bids and lasts two months; three-quarters of the auctioned firms are sold as going concern; the prices paid in these going-concern sales do not exhibit fire-sale discounts; and competition among bidders appear to force insiders to pay premiums comparable to those paid by outsiders. |
| (43) The bankrupt firm's major creditor (bank) often finances a bidder in the auction, which pushes the auction towards overbidding. Post-bankruptcy operating performance is found to be at par with non-bankrupt industry rivals, regardless of overbidding incentives, suggesting that the auction leads to a relatively efficient restructuring of the target firm. |
| **F. Offer Premiums:** |
| (44) Large-sample evidence on offer premiums are only starting to emerge. This evidence indicates that both the initial and final offer premiums are<br>• greater after the 1980s;<br>• greater for public bidders;<br>• greater in all-cash offers;<br>• lower for toehold bidders;<br>• increasing in the target runup (markup pricing);<br>• decreasing in target total equity capitalization and grater if the target's book-to-market ratio exceeds the industry median market-to-book ratio;<br>• greater in the presence of substantial dispersion in analysts' earnings forecasts;<br>• lower when the bidder CEO is female, and the higher the target board's proportion of female directors (provided that the female directors are independent appointees).<br>• unaffected by either the presence of a target poison pill or target hostility to the initial bid; |
| **G: Takeover Gains:** |
| (45) The average target cumulative average abnormal stock return (CAR) is positive and significant, both over the runup period and the announcement period. The runup constitutes about one-third of the total runup plus announcement CAR. The largest target CAR occurs in all-cash offers. |
| (46) The average, value-weighted combined CAR to bidders and targets is positive and significant over both the runup period and the announcement period. For the overall sample used here, the sum of the combined CAR for the runup- and announcement periods is a significant 1.79%. |
| (47) Bidder announcement period CARs average close to zero for the overall sample, with 49% of the bidders having negative CAR. The combination large bidder (here in the upper size quartile), payment in all-stock, and the target being a public company represents a "worst-case scenario" with average bidder announcement-period CAR of a significant −2.21%. The "best-case scenario" for the bidder is the combination of a small bidder (lower size-quartile), private target and *all-stock* as payment. This produces a significant average bidder announcement-period CAR of 6.46%. |
| (48) The major driver of negative bidder returns is not, as previously thought, the all-stock payment. Rather, the two key drivers are the target's status a public or private, and bidder size. |

*(Continued)*

Table 15 *(Continued)*

| Topic |
|---|
| (49) Bidder size was particularly large in 1999 and 2000. These years were unusual relative to years before *and* years after. Cisco, with a (constant 2000 dollar) market capitalization of $180 billion was the dominant bidder in *both* the upper 1% and lower 1% tails of the distribution of bidder abnormal announcement returns. Removing Cisco from the sample reduces the aggregate bidder dollar wealth loss in 1999–2000 period by almost $100 billion. |
| (50) Studies of long-run abnormal stock returns use either the matched-firm technique or Jensen's alpha (regression constant in an asset pricing model) to measure expected return to the merged firms in the sample. With 15,298 successful takeovers completed during the period 1980–2003, we show that long-run returns are significantly negative based on the matched-firm technique and insignificantly different from zero based Jensen's alpha. |
| (51) The standard matched-firm procedure identifies firms that have significantly different factor loadings than the event firms—which undermines their role as "matches". |
| (52) A zero-investment portfolio strategy which is long in the merged firms and short in the matched firms fail to produce long-run abnormal stock returns which are significantly different from zero, even for the sample of all-stock mergers. |
| **H. Bondholders, Management, and Arbitrageurs:** |
| (53) Studies of excess returns to bondholders of bidder and target firms find zero or negative gains to bidder bondholders and positive gains to target bondholders. There is no evidence of a wealth transfer from stockholders to bondholders due to a coinsurance effect of mergers. As of the 1990s, target bondholders are often fully protected via event risk covenants. |
| (54) Some target firms, particularly those receiving hostile bids, underperform prior to becoming targets. Moreover, CEO turnover increases after hostile bids. These findings indicate a disciplinary role played by the market for corporate control. There is, however, indications that this external control mechanism represents a "court of last resort". |
| (55) There is evidence that managers undertaking value-reducing acquisitions face a greater probability of being replaced than do managers undertaking value-increasing acquisitions. That is, bad bidders risk being fired. |
| (56) There is evidence that CEO compensation (other than turnover) changes following acquisition activity. The market reaction to merger announcements tends to be positive and greater for CEOs with above-average equity-based compensation, suggesting that compensation affects the quality of managerial investment decisions. |
| (57) CEOs with high equity-based compensation tend to seek out targets with relatively high market-to-book ratios (growth firms). This is consistent with high equity compensation inducing risk-taking behavior. |
| (58) Empirical measures of CEO "power" helps explain the cross-sectional variation in M&A bonuses. Deal announcement induced abnormal stock returns tend to be lower for CEOs with greater "power", suggesting that power may be misused. |
| (59) While a poorly performing acquisition reduces the value of the CEO's portfolio of stocks and options, there is evidence that the value of post-acquisition grants more than compensates for this value reduction. This indicates that CEOs face combination of low downside risk and high upside potential from making good acquisition decisions. |

*(Continued)*

Table 15 *(Continued)*

| Topic |
|---|
| (60) There is evidence that some target firm CEOs may be sacrificing takeover premium in return for a "golden handshake" from the bidder (to step aside and relinquish control). |
| (61) There is evidence that boards dominated by outside directors tend to increase value for their shareholders during an acquisition attempt. Target directors are rarely retained after a completed takeover, and their number of board seats and income levels tend to drop. This indicates that failing as a monitor imposes a personal cost on directors. |
| (62) There is substantial evidence of increased trading activity in the bidder and and target shares following merger announcements. In all-cash offers, merger (risk) arbitrageurs purchase target shares without shorting the bidder shares. In all-stock deals, arbitrageurs short the bidder stock using the exchange rate. If the exchange ratio is floating, the short sales are postponed until the final pricing has been set and the floating ratio has been fixed. |
| (63) There are substantial (risk-adjusted) returns to merger arbitrage strategies. Moreover, the short selling activity appears to put downward pressure on the acquirer stock price that may account for almost half of the negative announcement return for acquirers of stock-financed mergers. |
| **I. Mergers, Competition and Antitrust:** |
| (64) Merger-induced changes in product and factor prices directly translate into abnormal stock returns to the merging firms' industry rivals, upstream suppliers and downstream customers. Market power theories (collusion, predation, buying power) and productive neoclassical efficiency theories make empirically testable predictions for these abnormal stock returns. Such tests complement and extend traditional product price analysis seen in industrial economics. |
| (65) The power of tests based on stock returns depend on the accurate identification of related firms (rivals, customers, suppliers). Since much of the available evidence indicates significant contagion effects of horizontal merger announcements on these related firms, the tests appear to have sufficient power. Related firms are identified using four-digit SIC codes, Compustat industry segment information, and the Bureau of Economic Analysis Input Output tables. |
| (66) The tests utilize two sets of samples: Mergers that have been challenged with violation of antitrust laws (or, in the European Union, reviewed for such violation), and non-challenged mergers. For challenged mergers, the tests exploit *two* events with (typically) opposing implications for the industry wealth effects, thus increasing power to reject. |
| (67) The empirical studies typically conclude against horizontal market power effects of horizontal mergers, whether or not these were challenged. That is, the observed wealth effects on horizontal rivals and downstream (corporate) customers do not support increased market power. Some studies find traces of monopsony (buying power) effects vis-a-vis upstream suppliers. |
| (68) A horizontal merger *causes* a measurable increase in industry concentration (equal to twice the product of the market shares of the bidder and target when using the Herfindahl measure of concentration). The classical market concentration doctrine holds that increases in concentration reliably increases the industry's market power and thus industry monopoly rents. Since the abnormal returns to industry rivals directly measures changes in industry rents, regressing the merger-induced rival abnormal returns on the change in industry concentration provides a powerful test of the market concentration doctrine. Empirical tests reject the doctrine. |

establish how draconian antitakeover devices such as the staggered board and poison pill defense contributed to the fall.

The choice between merger and tender offer is interesting but has received little attention. There is some evidence that this decision is impacted by industry competition. This is hardly surprising as the likelihood of attracting rival competition in an auction setting depends on industry structure as well as asset characteristics. This is a fertile area for future research, both empirical and theoretical, and it ultimately links back to our understanding of takeover waves. Moreover, there are some indications that the target (and not the bidder) is increasingly initiating takeovers and thus determines the acquisition form. The economics of the selection process behind target-initiated deals is an exciting area for future research.

### *7.2. Bidding strategies and offer premiums*

Bidders initiating takeover bids for U.S. targets over the period 1980–2005 offered all-cash as payment in 26% of the cases, all-stock in 37%, and a mix of stock of cash in 37%. The majority of tender offers are all-cash or a mix of cash and stock. While the majority of merger bids are in the form of all-stock (with the exception of the 1980–1985 period where most merger bids offered a mixed cash-stock payment). As pointed out earlier, all-cash and mixed cash-stock offers are predominant in tender offers. Moreover, mixed stock-cash offers rose to become the most frequently used payment method in mergers by 2001. In the two subperiods 1990–1995 and 1996–2000, the percentage of all-stock offers in initial merger bids was approximately 55% in *both* period. This means that (1) nearly half of the initial merger bids in the 1990s use some cash as payment, and (2) the fraction of merger bids where the payment is all-stock remained unaffected by the significant runup in overall market valuations in the 1996–2000 period.

The choice of payment method is strategic for several reasons, including tax effects, its impact on the conditional expected value of the bid to asymmetrically informed bidders and targets, and corporate control considerations. The evidence indicates that stock offers are more likely the greater the bidder's asset size and market-to-book ratio. Stock offers are less likely the greater the bidder management's shareholdings and the greater the dispersion in analyst forecast of bidder earnings. Moreover, offer premiums are greater in all-cash offers than in all-stock offers, and the probability that the initial bidder wins the target is lower for all-stock offers than for cash offers.

The pervasive negative market reaction to all-stock merger bids by public bidders is typically compared to the average negative market reaction to seasoned equity offers. The comparison is appealing since the timing of the equity issue is determined endogenously by the issuer in both events, and thus involves some degree of adverse selection. On the other hand, in terms of the *issue method*, stock swaps in mergers are closer to private placements than they are to an underwritten seasoned equity offering—and there is substantial evidence that the market reaction to private placement is positive on average (Eckbo, Masulis, and Norli, 2007). Moreover, the market reaction to merger stock swaps is positive when the target is private. Also, formal tests of signaling theories for the payment method choice have received mixed success. Additional research is needed to establish the empirical relevance of asymmetric information arguments for the strategic payment choice.

The dramatic fall in toehold bidding coincides with the rise of structural takeover defenses beginning in the 1980s. In theory, toehold bidding conveys the substantial strategic advantage of rival bidders, particularly in a common-value setting. Since many of these advantages come at the expense of the *target*, some targets may be reluctant to negotiate if the bidder has a toehold. If so, acquiring a toehold prior to attempting friendly merger negotiations may backfire: if the target refuses, the bidder foregoes not only things like a termination agreement but also the opportunity to examine the target books—which is crucial for pricing the merger.

Another way to put this is that a toehold must be large to be worth it—larger than 10% by some (conservative) estimates. This argument may go a long way in explaining the dual observation that toeholds are large (on average 20%) when they exist and that they occur mostly in hostile bids. An interesting and hitherto unexplored empirical issue is whether toeholds are important in other jurisdictions, in particular those with highly concentrated shareownership and a set of takeover regulations and corporate governance practices that differ from those in the U.S.

The average offer premium in successful single-bid takeover contests is somewhat higher than the average initial offer premium in multi-bid contests. While this is consistent with the greater premium preempting competition in ex-post successful single-bid cases, systematic empirical tests of preemption are almost nonexistent. Bid revisions are substantial when the initial bid attracts competition and/or is revised by the initial bidder. The average bidjump from the first to the second bid in the contest is 10 percentage points, a 31% change in the offer premium.

Another interesting jump is the markup of the offer price above the target stock price on the day before the offer is announced. There is substantial evidence that a dollar increase in the pre-offer target share price runup causes the initial bidder to mark up the total offer premium by almost a dollar ($0.80). Interestingly, *bidder* takeover gains are also found to be increasing in the target runup, which raises issues concerning the true nature of the markup pricing phenomenon itself. It appears that takeovers with greater pre-bid target runups are more profitable for *both* bidder and target firms, which may explain why bidders agree to the (partial) markup.

A useful approach to investigating the markup pricing phenomenon further is to document in much greater detail the bidder's pricing process during merger negotiations. An analogy here may be the structure of the pricing process in seasoned equity offerings and in initial public offerings. Which parties are involved? What role do fairness reports play for the pricing process? If bidders, in fact, react by revising the offer price in response to the target runup, how is the runup analyzed? Is the reverse causality at play, that is, is the offer price set high ex ante in profitable takeovers, which when rumored drives the runup in the target price ex-post?

Delaware case law sanctions the right to "just say no" to an unsolicited takeover bid. Moreover, if the board's defensive response is not draconian (i.e., it is neither coercive nor preclusive), the board is protected by the business judgment rule. The twin defense of staggered board election and a poison pill (shareholder rights plan) is draconian in the eyes of many economists but not the court. However, "dead-hand" pills (where only incumbent directors may vote to rescind the pill) have been struck down.

There is a small but significantly negative market reaction to the adoption of strong antitakeover amendments such as poison pills and staggered board. The market reacts positively to antigreenmail amendments provided these occur when a takeover is rumored. Offer premiums appear to be as high (if not higher) for targets with poison pills than targets with no pill in place. Since pills can be adopted any time, and in particular in response to a bid ("morning after pill"), the power of tests that compare offer premiums in pill-targets with no-pill-targets is questionable and should be examined further. Understanding the true economic effects of defenses such as staggered boards and poison pills is important, not the least for the ongoing public policy debate over antitakeover measures.

There is a trend toward market-based mechanisms for resolving Chapter 11 cases, including sale of the firm to a bidder. Target firms that are sold spend less time in Chapter 11, which lowers bankruptcy costs. Acquirers tend to be in the same industry, and premiums paid are on average lower than in takeovers of nonbankrupt firms in the same industry. Premiums paid for targets sold in mandatory, open, first-price, all-cash bankruptcy auctions in Sweden suggest the possibility that the auction mechanism may also work well for the typical Chapter 11 case (which is of a similar size as the Swedish sample firms). Importantly, the Swedish auction prices do not exhibit fire-sale discounts, contradicting a central presumption behind the creation of Chapter 11 back in 1978. The growing use of auction-related mechanisms in the United States is likely to have lowered bankruptcy costs. By how much remains an important question for future research.

Large-sample evidence on offer premiums are only starting to emerge. This evidence indicates that both the initial and final offer premiums were greater after the 1980s; greater for public bidders; greater in all-cash offers; lower for toehold bidders; increasing in the target runup (markup pricing); decreasing in target total equity capitalization and greater if the target's book-to-market ratio exceeds the industry median market-to-book ratio; greater in the presence of substantial dispersion in analysts' earnings forecasts; lower when the bidder CEO is female, and the higher the target board's proportion of female directors (provided that the female directors are independent appointees); and unaffected by either the presence of a target poison pill or target hostility to the initial bid.

Several variables used by researchers to explain the offer premium are themselves endogenous choice variables (payment method, toehold, hostility, termination agreements, bidder's public status). Some of the effects stated earlier appear robust to corrections for endogeneity (including systems of equations and Heckman procedures). One variable that does *not* appear to be robust, however, is tender offer. The inclusion of other variables (such as toeholds and hostility) appears to affect conclusions as to whether offer premiums are higher, the same, or lower in tender offers than in merger bids. Additional work is needed to sort this issue out—and may also affect the conclusion so far that poison pills have a neutral effect on offer premiums.

### *7.3. Takeover gains*

Becoming a target is a significant surprise event; thus target total gains are measured relatively precisely by the offer premium (typically, relative to the target market price two months prior to the first offer announcement) or, alternatively, by target cumulative abnormal returns over the same period. Consistent with the evidence on offer premiums (above), the average target cumulative average abnormal stock return (CAR) is positive and significant, over both the runup and the announcement period. The target runup constitutes about one-third of the total runup plus announcement CAR. The largest target CAR occurs in all-cash offers.

The average, value-weighted combined CAR to bidders and targets is positive and significant over both the runup period and the announcement period. For the overall sample used here, the sum of the combined CAR for the runup and announcement periods is a significant 1.79%. Bidder announcement-period CARs average close to zero for the overall sample, with 49% of the bidders having negative CAR. The combination large bidder (here in the upper size quartile), payment in all-stock, and the target being a public company represents a worst-case scenario, with average bidder announcement-period CAR of a significant −2.21%. The best-case scenario for the bidder is the combination of a small bidder (lower size-quartile), private target, and *all-stock* as payment. This produces a significant average bidder announcement-period CAR of 6.46%.

The major driver of negative bidder returns is not, as previously thought, the all-stock payment. Rather, the two key drivers are the target's status as public or private and bidder size. Bidder size was particularly large in 1999 and 2000. These years were unusual relative to years before *and* years after. Cisco, with a market capitalization of $180 billion (constant 2000 dollars) was the dominant bidder in *both* the upper 1% and lower 1% tails of the distribution of bidder abnormal announcement returns. Removing Cisco from the sample reduces the aggregate bidder dollar wealth loss in the 1999–2000 period by almost $100 billion. An important but unanswered question is whether the negative spike is truly a bidder size effect or a year effect (or a combination of two). At this point, there appears to be no explanation for why the large firms decided to enter the market for corporate control in 1998–2001, only then to leave again.

Studies of long-run abnormal stock returns use either the matched-firm technique or Jensen's alpha (regression constant in an asset pricing model) to measure expected return to the merged firms in the sample. With 15,298 successful takeovers completed during 1980–2003, we show that long-run returns are significantly negative based on the matched-firm technique but insignificantly different from zero-based Jensen's alpha. Of the two methods, only the latter can actually be replicated using a portfolio investment strategy. We also show that the standard matched-firm procedure identifies firms that have significantly different factor loadings than the event firms—which undermines their role as "matches." A zero-investment portfolio strategy that is long in the merged firms and short in the matched firms fails to produce long-run abnormal stock returns that are significantly different from zero, even for the sample of all-stock mergers. Overall, the long-run performance evidence presented here does not support the hypothesis that merged firms underperform.

### *7.4. Bondholders, executives and arbitrage*

Studies of bondholder returns have suffered from limited access to data on bond market values. However, bond data are improving. Recent studies of excess returns to bondholders of bidder and target firms find zero or negative gains to bidder bondholders and positive gains to target bondholders. There is no evidence of a wealth transfer from stockholders to bondholders due to a coinsurance effect of mergers. As of the 1990s, target bondholders are often fully protected via event risk covenants. Bondholder wealth effects of a variety of corporate control decisions seem a fertile area for future research.

There is evidence that managers undertaking value-reducing acquisitions face a greater probability of being replaced than do managers undertaking value-increasing acquisitions. That is, bad bidders risk being fired. Some target firms, particularly those receiving hostile bids, underperform prior to becoming targets. However, CEO turnover increases after hostile bids, indicating a disciplinary role played by the market for corporate control. There is also evidence that CEO compensation (other than turnover) changes following acquisition activity. The market reaction to merger announcements tends to be positive and greater for CEOs with above-average equity-based compensation, suggesting that compensation affects the quality of managerial investment decisions.

CEOs with high equity-based compensation tend to seek out targets with relatively high market-to-book ratios (growth firms). This is consistent with high equity compensation inducing risk-taking behavior. Moreover, while a poorly performing acquisition reduces the value of the CEO's portfolio of stocks and options, there is evidence that the value of post-acquisition grants more than compensates for this value reduction. This indicates that CEOs face the combination of low downside risk and high upside potential from making good acquisition decisions.

There is also some evidence that target firm CEOs may be sacrificing takeover premium in return for a golden handshake from the bidder (to step aside and relinquish control). Empirical measures of CEO power helps explain the cross-sectional variation in M&A bonuses. Moreover, deal announcement-induced abnormal stock returns tend to be lower for CEOs with greater power, suggesting that power may be misused. This raises the question of what role boards play in monitoring takeover activity. There is evidence that boards dominated by outside directors tend to increase value for their shareholders during an acquisition attempt. Target directors are rarely retained after a completed takeover, and their number of board seats and income levels tend to drop. This suggests that failing as a monitor imposes a personal cost on directors, which helps align the interest of directors and shareholders.

Merger arbitrage (or risk arbitrage) is a specialized investment strategy that tries to profit from the spread between the offer price and the target stock market price conditional on the offer having been made. It is essentially a (risky) bet on the likelihood that the proposed transaction will go through. Arbitrage gains depend on several factors, including the size of the arbitrage spread, the probability that the deal closes, the length of time that the arbitrageur must hold the position, and the target stock price development if the deal fails. Average gains are significantly positive, with the largest abnormal returns reported for cash tender offers. In addition to bearing deal failure risk, merger arbitrageurs provide a service in terms of providing deal-related information, liquidity, and helping resolve the free rider problems among small, diffuse target shareholders. Transaction costs, such as brokerage commissions and price impact of trading, limit arbitrage returns. There is evidence that short-selling by merger arbitrageur causes downward price pressure that accounts for almost half of the negative announcement return for acquirers in stock-financed mergers.

### *7.5. Competition and antitrust*

Merger-induced changes in product and factor prices directly translate into abnormal stock returns to the merging firms' industry rivals, upstream suppliers, and downstream customers. Market power theories (collusion, predation, buying power) and productive neoclassical efficiency theories make empirically testable predictions for these abnormal stock returns. Such tests complement and extend the traditional product price analysis seen in industrial economics. The empirical studies typically conclude against the horizontal market power effects of horizontal mergers. That is, the observed wealth effects on horizontal rivals and downstream (corporate) customers do not support increased market power. Some studies find traces of monopsony (buying power) effects vis-à-vis upstream suppliers.

A horizontal merger causes a measurable increase in industry concentration (equal to twice the product of the market shares of the bidder and target when using the Herfindahl measure of concentration). The classical market concentration doctrine holds that an increase in concentration reliably increases the industry's market power and thus industry monopoly rents. Since the abnormal returns to industry rivals directly measure changes in industry rents, regressing the merger-induced rival abnormal returns on the change in industry concentration provides a powerful test of the market concentration doctrine. Empirical tests reject the doctrine.

The power of tests based on stock returns depends on the accurate identification of related firms (rivals, customers, suppliers). Since much of the available evidence indicates significant contagion effects of horizontal merger announcements on these related firms, the tests appear to have sufficient power. Related firms are identified using four-digit SIC codes, Compustat industry segment information, and the Bureau of Economic Analysis Input Output tables. The tests utilize two sets of samples: mergers that have been challenged with violation of antitrust laws (or, in the European Union, reviewed for such violation) and nonchallenged mergers. For challenged mergers, the tests exploit two events with (typically) opposing implications for the industry wealth effects, thus increasing power to reject.

In the future the interaction of industrial and financial economics, where econometric methods traditionally used in corporate finance are applied to interesting phenomena in industrial economics, is likely to increase in importance. While most of the attention thus far has centered on testing theories of monopoly, the econometric method applies equally well to an examination of alternative efficiency theories of corporate investment. For example, an industry-based theory of merger waves may be couched in terms of the valuation effects for related firms and may be tested using the event study methodology. Similarly, behavioral arguments for things like clustering of merger activity and post-merger underperformance have hitherto untested implications for the event-induced valuation effect across industry rivals.

## References

Adams, R. and D. Ferreira, 2007, "One Share, One Vote: The Empirical Evidence," Working Paper, University of Queensland and London School of Economics & Political Science.

Aggarwal, R., 2008, "Executive Compensation," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 2, Chapter 17 (Elsevier/North-Holland, Handbooks in Finance Series).

Aghion, P., O. Hart and J. Moore, 1992, "The Economics of Bancruptcy Reform," *Journal of Law, Economics and Organization*, 8, 523–546.

Agrawal, A. and R. A. Walkling, 1994, "Executive Careers and Compensation Surrounding Takeover Bids," *Journal of Finance*, 49, 985–1014.

Akbulut, M. E. and J. G. Matsusaka, 2003, "Fifty Years of Diversification Announcements," Working Paper, Marshall School of Business, University of Southern California.

Akdogu, E., 2007a, "Gaining a Competitive Edge Through Acquisitions: Evidence from the Telecommunications Industry," *Journal of Corporate Finance*, forthcoming.

Akdogu, E., 2007b, "Value-maximizing Managers, Value-increasing Mergers and Overbidding," Working Paper, Southern Methodist University.

Akhavein, J. D., A. N. Berger and D. B. Humphrey, 1997, "The Effects of Megamergers on Efficiency and Prices: Evidence from a Bank Profit Function," *Review of Industrial Organization*, 12, 95–139.

Akhavein, J. D., A. N. Berger and D. B. Humphrey, 2007, "Negotiation under the Threat of an Auction: Friendly Deals, ex ante Competition and Bidder Returns," Working Paper, University of California at Los Angeles.

Aktas, N., E. deBodt and R. Roll, 2006, "Was European M&A Regulation Protectionist?" *Economic Journal*, forthcoming.

Ambrose, K. W. and W. L. Megginson, 1992, "The Role of Asset Structure, Ownership Structure and Takeover Defenses in Determining Acquisition Likelihood," *Journal of Financial and Quantitative Analysis*, 27, 575–589.

Amihud, Y., B. Lev and N. Travlos, 1990, "Corporate Control and the Choice of Investment Financing: The Case of Corporate Acquisitions," *Journal of Finance*, 45, 603–616.

Andrade, G., M. Mitchell and E. Stafford, 2001, "New Evidence and Perspectives on Mergers," *Journal of Economic Perspectives*, 15, 103–120.

Andrade, G. and E. Stafford, 2004, "Investigating the Economic Role of Mergers," *Journal of Corporate Finance*, 10, 1–36.

Ang, J. S. and Y. Cheng, 2006, "Direct Evidence on the Market-driven Acquisition Theory," *Journal of Financial Research*, 29, 199–216.

Asquith, P., 1983, "Merger Bids, Uncertainty and Stockholder Returns," *Journal of Financial Economics*, 11, 51–83.

Asquith, P., R. Bruner and D. Mullins, 1987, "Merger Returns and the Form of Financing," *Proceedings of the Seminar on the Analysis of Securities Prices*, 34, 115–146.

Asquith, P., R. Bruner and D. Mullins, 1983, "The Gains to Bidding Firms from Merger," *Journal of Financial Economics*, 11, 121–139.

Asquith, P. and E. H. Kim, 1982, "The Impact of Merger Bids on the Participating Firms' Security Holders," *Journal of Finance*, 37, 1209–1228.

Austen-Smith, D. and P. O'Brien, 1986, "Takeover Defenses and Shareholder Voting," Working Paper, Sloan School of Management, Massachusetts Institute of Technology.

Bagnoli, M. and B. L. Lipman, 1996, "Stock Price Manipulation Through Takeover Bids," *RAND Journal of Economics*, 27, 124–147.

Bagnoli, M. and B. L. Lipman, 1988, "Successful Takeover Without Exclusion," *Review of Financial Studies*, 1, 89–110.

Baird, D. G., and R. K. Rasmussen, 2003, "Chapter 11 at Twilight," *Stetson Law Review*, 56, 673–699.

Baird, D. G., 1993, "Revisiting Auctions in Chapter 11," *Journal of Law and Economics*, 36, 633–653.

Baird, D. G., 1986, "The Uneasy Case for Corporate Reorganizations," *Journal of Legal Studies*, 15, 127–147.

Baker, M. and S. Savasoglu, 2002, "Limited Arbitrage in Mergers and Acquisitions," *Journal of Financial Economics*, 64, 91–15.

Baker, M., R. S. Ruback and J. Wurgler, 2007, "Behavioral Corporate Finance," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 1, Chapter 4 (Elsevier/North-Holland, Handbooks in Finance Series).

Banerjee, A. and W. E. Eckard, 1998, "Are Mega-Mergers Anticompetitive? Evidence from the First Great Merger Wave," *RAND Journal of Economics*, 29, 803–827.

Barber, B. M. and J. D. Lyon, 1997, "Detecting Long-run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics," *Journal of Financial Economics*, 43, 341–372.

Bargeron, L., 2005, "Do Shareholder Tender Agreements Inform or Expropriate Shareholders?" Working Paper, University of Pittsburgh.

Bargeron, L., F. P. Schlingemann, R. M. Stulz and C. Zutter, 2007, "Why Do Private Acquirers Pay So Little Compared to Public Acquirers?" *Journal of Financial Economics*, forthcoming.

Barton, D. M. and R. Sherman, 1984, "The Price and Profit Effect of Horizontal Merger: A Case Study," *Journal of Industrial Economics*, 33, 165–177.

Bates, T. W., D. A. Becher and M. L. Lemmon, 2008, "Board Classification and Managerial Entrenchment: Evidence from the Market for Corporate Control," *Journal of Financial Economics*, 87, 656–677.

Bates, T. H. and M. L. Lemmon, 2003, "Breaking Up Is Hard to do? An Analysis of Termination Fee Provisions and Merger Outcomes," *Journal of Financial Economics*, 69, 460–504.

Bauguess, S. W., S. B. Moeller, F. P. Schlingemenn and C. J. Zutter, 2007, "Ownership Structure and Target Returns," *Journal of Corporate Finance*, forthcoming.

Bebchuk, L. A., 1985, "Toward Undistorted Choice and Equal Treatment in Corporate Takeovers," *Harvard Law Review*, 98, 1695–1808.

Bebchuk, L. A., 1988, "A New Approach to Corporate Reorganizations," *Harvard Law Review*, 101, 775–804.

Bebchuk, L. A. and H. F. Chang, 1992, "Bargaining and The Division of Value in Corporate Reorganization," *Journal of Law, Economics and Organization*, 8, 253–279.

Bebchuk, L. A., J. C. Coates and G. Subramanian, 2002, "The Powerful Anti-takeover Force of Staggered Boards: Theory, Evidence and Policy," *Stanford Law Review*, 54, 887–951.

Bebchuk, L. A. and A. Cohen, 2005, "The Costs of Entrenched Boards," *Journal of Financial Economics*, 78, 409–433.

Bebchuk, L. A. and A. Ferrell, 2004, "What Matters in Corporate Governance?" Working Paper, Harvard University.

Bebchuk, L. A. and J. M. Fried, 2003, "Executive Compensation as an Agency Problem," *Journal of Economic Perspectives*, 17, 71–92.

Becher, D. A., J. H. Mulherin and R. A. Walkling, 2008, "Industry Shocks and Merger Activity: An Analysis of U.S. Public Utilities," Working Paper, Drexel University.

Becht, M., P. Bolton and A. Roll, 2003, "Corporate Governance and Control," in George Constantinides, M. Harris and R. Stulz (Ed.), *Handbook of the Economics of Finance*, vol. 1, Chapter 1, 1–109 (North-Holland).

Berglof, E. and M. Burkart, 2003, "European Takeover Regulation," *Economic Policy*, 36, 171–213.

Berkovitch, E. and M. Narayanan, 1990, "Competition and the Medium of Exchange in Takeovers," *Review of Financial Studies*, 3(2), 153–174.

Berkovitch, E. and N. Khanna, 1991, "A Theory of Acquisition Markets: Mergers versus Tender Offers, and Golden Parachutes," *Review of Financial Studies*, 4, 149–174.

Berle, A. A. Jr. and G. C. Means, 1932, "*The Modern Corporation and Private Property*," Macmillan, New York.

Betker, B. L., 1995. "An Empirical Examination of Prepackaged Bankruptcy," *Financial Management*, 24, 3–18.

Betton, S. and B. E. Eckbo, 2000, "Toeholds, Bid Jumps, and Expected Payoff in Takeovers," *Review of Financial Studies*, 13, 841–882.

Betton, S., B. E. Eckbo and K. S. Thorburn, 2007, "Merger Negotiations and the Toehold Puzzle," *Journal of Financial Economics*, forthcoming.

Betton, S., B. E. Eckbo and K. S. Thorburn, 2008a, "Does Industry Competition Affect Takeover Bids?" Working Paper, Tuck School of Business at Dartmouth.

Betton, S., B. E. Eckbo and K. S. Thorburn, 2008b, "Markup Pricing Revisited," Working Paper, Tuck School of Business at Dartmouth.

Betton, S., B. E. Eckbo and K. S. Thorburn, 2008c, "Massive Wealth Destruction? Bidder Gains Revisited," Working Paper, Tuck School of Business at Dartmouth.

Bhagat, S., M. Dong, D. Hirshleifer and R. Noah, 2005, "Do Tender Offers Create Value? New Methods and Evidence," *Journal of Financial Economics*, 76, 3–60.

Bhagat, S. and R. H. Jefferis, 2002, *The Econometrics of Corporate Governance Studies* (MIT Press, Cambridge, MA).

Bhagat, S. and R. H. Jefferis, 1991, "Voting Power in the Proxy Process: The Case of Antitakeover Charter Amendments," *Journal of Financial Economics*, 30, 193–226.

Bhagat, S. J., A. Brickley and U. Loewenstein, 1987, "The Pricing Effects of Interfirm Cash Tender Offers," *Journal of Finance*, 42:4, 965–986.

Bhattahcaryya, S. and A. Nain, 2006, "Horizontal Acquisitions and Buying Power: A Product Market Approach," Working Paper, Ross School of Business, University of Michigan.

Billett, M. T., T.-H. D. King and F. C. Mauer, 2004, "Bondholder Wealth Effects in Mergers and Acquisitions: New Evidence from the 1980s and 1990s," *Journal of Finance*, 59, 107–135.

Bittlingmayer, G. and T. W. Hazlett, 2000, "DOS *kapital*: Has Antitrust Action Againt Microsoft Created Value in the Computer Industry?" *Journal of Financial Economics*, 55, 329–359.

Bliss, R. T. and R. J. Rosen, 2001, "CEO Compensation and Bank Mergers," *Journal of Financial Economics*, 61, 107–138.

Boone, A. L. and J. H. Mulherin, 2007a, "Do Termination Provisions Truncate the Takeover Bidding Process?" *Review of Financial Studies*, 20, 461–489.

Boone, A. L. and J. H. Mulherin, 2007b, "How Are Firms Sold?" *Journal of Finance*, 62, 847–875.

Borenstein, S., 1990, "Airline Mergers, Airport Dominance, and Market Power," *American Economic Review*, 80, 400–404.

Bradley, M., 1980, "Interfirm Tender Offers and the Market for Corporate Control," *Journal of Business*, 53, 345–376.

Bradley, M., and M. Rosenzweig, 1992, "The Untenable Case for Chapter 11," *Yale Law Journal*, 101, 1043–1095.

Bradley, M. and A. Sundaram, 2006, "Acquisitions and Performance: A Re-assesment of the Evidence," Working Paper, Duke University.

Bradley, M. and L. M. Wakeman, 1983, "The Wealth Effects of Targeted Share Repurchases," *Journal of Financial Economics*, 11, 301–328.

Brickley, J., R. Lease and C. Smith, 1988, "Ownership Structure and Voting on Takeover Amendments," *Journal of Financial Economics*, 20, 267–293.

Bris, A., I. Welch and N. Zhu, 2006, "The Costs of Bankruptcy: Chapter 7 Liquidation versus Chapter 11 Reorganization," *Journal of Finance*, 61, 1253–1303.

Brown, D. T. and M. D. Ryngaert, 1991, "The Mode of Acquisition in Takeovers: Taxes and Asymmetric Information," *Journal of Finance*, 46, 653–669.

Bruner, R. F., 2004, *Applied Mergers & Acquisitions* (John Wiley & Sons, Hoboken, NJ).
Bulow, J., M. Huang and P. Klemperer, 1999, "Toeholds and Takeovers," *Journal of Political Economy*, 107, 427–454.
Burch, T., 2001, "Locking Out Rival Bidders: The Use of Lockup Options in Corporate Mergers," *Journal of Financial Economics*, 60, 103–141.
Burkart, M., 1995, "Initial Shareholdings and Overbidding in Takeover Contests," *Journal of Finance*, 50, 1491–1515.
Burkart, M., D. Gromb and F. Panunzi, 1998, "Why Takeover Premia Protect Minority Shareholders," *Journal of Political Economy*, 106, 647–677.
Burkart, M. and F. Panunzi, 2003, "Mandatory Bids, Squeeze-out, Sell-out and the Dynamics of the Tender Offer Process," Working Paper, European Corporate Governance Institute.
Burkart, M. and F. Panunzi, 2006, "Takeovers," in Ph. Hartmann X. Freicas and C. Mayer (Ed.), *Financial Markets and Institutions: An European Perspective*, Forthcoming (Oxford University Press).
Byrd, J. and K. Hickman, 1992, "Do Outside Directors Monitor Managers? Evidence from Tender Offer Bids," *Journal of Financial Economics*, 32, 195–222.
Cain, M. D., D. Denis and D. K. Denis, 2005, "Earnouts: A Study of Contracting in Acquisition Agreements," Working Paper, Purdue University.
Carhart, M. M., 1997, "On the Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.
Carleton, W., D. Guilkey, R. Harris and J. Stewart, 1983, "An Empirical Analysis of the Role of the Medium of Exchange in Mergers," *Journal of Finance*, 38, 57–82.
Chang, S., 1998, "Takeovers of Privately Held Targets, Methods of Payment, and Bidder Returns," *Journal of Finance*, 53, 773–784.
Chang, T. and A. Schoar, 2007, "The Effect of Judicial Bias in Chapter 11 Reorganization," Working Paper, Massachusetts Institute of Technology.
Chatterjee, S., K. John and A. Yan, 2008, "Takeover Premium and Divergence of Opinions," Working Paper, School of Business, Fordham University.
Chemmanur, T. J. and I. Paeglis, 2003, "The Choice of the Medium of Exchange in Acquisitions: A Direct Test of the Double-sided Asymmetric Information Hypothesis," Working Paper, Boston College.
Chen, L. H. and H. Sami, 2006, "Does the Use of Fariness Opinions Impair the Acquirers' Abnormal return? The Litigation Risk Effect," Working Paper, Arizona State University.
Claessens, S., S. Djankov and L. H. P. Lang, 2000, "The Separation of Ownership and Control in East Asian Corporations," *Journal of Financial Economics*, 58, 81–112.
Coase, R., 1937, "The Nature of the Firm," *Economica*, 4, 386–405.
Coates, J. C., 2000, "Takeover Defenses in the Shadow of the Pill: A Critique of the Scientific Evidence," *Texas Law Review*, 79, 271–382.
Comment, R. and G. W. Schwert, 1995, "Poison or Placebo? Evidence on the Deterrent and Wealth Effects of Modern Antitakeover Measures," *Journal of Financial Economics*, 39, 3–43.
Comment, R. and G. Jarrell, 1987, "Two-Tier and Negotiated Tender Offers: The Imprisonment of the Free-riding Shareholder," *Journal of Financial Economics*, 19, 283–310.
Cornelli, F. and D. D. Li, 2002, "Risk Arbitrage in Takeovers," *Review of Financial Studies*, 15:3, 837–868.
Cotter, J., A. Shivdasani and M. Zenner, 1997, "Do Independent Directors Enhance Target Shareholder Wealth During Tender Offers?" *Journal of Financial Economics*, 43, 195–218.
Cotter, J. and M. Zenner, 1994, "How Managerial Wealth Affects the Tender Offer Process," *Journal of Financial Economics*, 35, 63–97.
Cremers, K. J. Martijn, V. B. Nair and K. John, 2008, "Takeovers and the Cross-section of Returns," *Review of Financial Studies*, forthcoming.
Dafny, L., 2005, "Estimation and Identification of Merger Effects: An Application to Hospital Mergers," NBER Working Paper No. W11673.
Daniel, K. and D. Hirshleifer, 2008, "A Theory of Costly Sequential Bidding," Working Paper, Anderson Graduate School of Management, UCLA.
Danielsen, M. G. and J. M. Karpoff, 1998, "On the Uses of Corporate Governance Provisions," *Journal of Corporate Finance*, 4, 347–371.

Dann, L. Y. and H. DeAngelo, 1983, "Standstill Agreements, Privately Negotiated Stock Repurchases, and the Market for Corporate Control," *Journal of Financial Economics*, 11(April), 275–300.

Dasgupta, S. and R. G. Hansen, 2007, "Auctions in Corporate Finance," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 1, Chapter 3, 87–143 (Elsevier/North-Holland, Handbooks in Finance Series).

Dasgupta, S. and K. Tsui, 2004, "Auctions with Cross-Shareholdings," *Economic Theory*, 24, 163–194.

Dasgupta, S. and K. Tsui, 2003, "A 'Matching Auction' for Targets with Heterogeneous Bidders," *Journal of Financial Intermediation*, 12, 331–364.

Datta, S., M. Iskandar-Datt and K. Raman, 2001, "Executive Compensation and Corporate Acquisition Decisions," *Journal of Finance*, 56, 2299–2336.

DeAngelo, H. and E. M. Rice, 1983, "Antitakeover Charter Amendments and Stockholder Wealth," *Journal of Financial Economics*, 11, 329–359.

DeMarzo, P. M., I. Kremer and A. Skrzypacz, 2005, "Bidding with Securities: Auctions and Security Design," *American Economic Review*, 95, 936–959.

Dennis, D. and J. McConnell, 1986, "Corporate Mergers and Security Returns," *Journal of Financial Economics*, 16(2, June), 143–187.

Djankov, S., R. La Porta, F. Lopez-de-Silanes and A. Shleifer, 2007, "The Law and Economics of Self-Dealing," *Journal of Financial Economics*, forthcoming.

Dlugosz, J., R. Fahlenbrach, P. A. Gompers and A. Metrick, 2006. "Large Blocks of Stock: Prevalence, Size, and Measurement," *Journal of Corporate Finance*, 12, 594–618.

Dong, M., D. Hirshleifer, S. Richardson and S. H. Teoh, 2006, "Does Investor Misvaluation Drive the Takeover Market?" *Journal of Finance*, 61, 725–762.

Dukes, W. P., C. J. Frohlich and C. K. Ma, 1992, "Risk Arbitrage in Tender Offers," *Journal of Portfolio Management*, 18:4, 47–55.

Dyck, A. and L. Zingales, 2004, "Private Benefits of Control: An International Comparison," *Journal of Finance*, 59, 533–596.

Eberhart, A. C. and A. Siddique, 2002, "The Long–Term Performance of Corporate Bonds (and Stocks) Following Seasoned Equity Offerings," *Review of Financial Studies*, 15, 1385–1406.

Eckbo B. E., 2008, "Bid Strategies and Takeover Premiums: A Review," *Journal of Corporate Finance*, forthcoming.

Eckbo B. E., 1992, "Mergers and the Value of Antitrust Deterrence," *Journal of Finance*, 47, 1005–1029.

Eckbo B. E., 1990a, "Competition and Wealth Effects of Horizontal Mergers," in M. Mathewson, M. Trebilcock and M. Walker (Ed.), *The Law and Economics of Competition Policy*, Chapter 9, 297–332 (The Fraser Institute).

Eckbo B. E., 1990b, "Valuation Effects of Greenmail Prohibitions," *Journal of Financial and Quantitative Analysis*, 25, 491–505.

Eckbo B. E., 1989, "The Role of Stock Market Studies in Formulating Antitrust Policy Towards Horizontal Mergers," *Quarterly Journal of Business and Economics*, 28, 22–38.

Eckbo B. E., 1988, "The Market for Corporate Control: Policy Issues and Capital Market Evidence," in R. S. Khemani, D. Shapiro and W. T. Stanbury (Ed.), *Mergers, Corporate Concentration and Corporate Power in Canada*, Chapter 7, 143–225. (Canadian Institute for Research on Public Policy, Montreal).

Eckbo B. E., 1985, "Mergers and the Market Concentration Doctrine: Evidence from the Capital Market," *Journal of Business*, 58, 325–349.

Eckbo B. E., 1983, "Horizontal Mergers, Collusion, and Stockholder Wealth," *Journal of Financial Economics*, 11, 241–272.

Eckbo B. E., Ronald M. G. and R. L. Heinkel, 1990, "Asymmetric Information and the Medium of Exchange in Takeovers: Theory and Tests," *Review of Financial Studies*, 3, 651–675.

Eckbo, B. E. and H. Langohr, 1989, "Information Disclosure, Method of Payment, and Takeover Premiums: Public and Private Tender Offers in France," *Journal of Financial Economics*, 24, 363–403.

Eckbo, B. E., V. Maksimovic and J. Williams, 1990, "Consistent Estimation of Cross-sectional Models in Event Studies," *Review of Financial Studies*, 3, 343–365.

Eckbo, B. E., R. W. Masulis and Ø. Norli, 2007, “Security Offerings,” in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 1, Chapter 6 (Elsevier/North-Holland, Handbooks in Finance Series).

Eckbo, B. E. and Ø. Norli, 2005, “Liquidity Risk, Leverage and Long-Run IPO Returns,” *Journal of Corporate Finance*, 11, 1–35.

Eckbo, B. E. and K. S. Thorburn, 2008a, “Corporate Restructuring: Breakups and LBOs,” in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 2, Chapter 16 (Elsevier/North-Holland, Handbooks in Finance Series).

Eckbo, B. E. and K. S. Thorburn, 2008b, “Creditor Financing and Overbidding in Bankruptcy Auctions,” *Journal of Corporate Finance*, forthcoming.

Eckbo, B. E. and K. S. Thorburn, 2007, “Automatic Bankruptcy Auctions and Fire-Sales,” *Journal of Financial Economics*, forthcoming.

Eckbo, B. E. and K. S. Thorburn, 2003, “Control Benefits and CEO Discipline in Automatic Bankruptcy Auctions,” *Journal of Financial Economics*, 69, 227–258.

Eckbo, B. E. and K. S. Thorburn, 2000, “Gains to Bidder Firms Revisited: Domestic and Foreign Acquisitions in Canada,” *Journal of Financial and Quantitative Analysis*, 35, 1–25.

Eckbo, B. E. and S. Verma, 1994, “Managerial Shareownership, Voting Power, and Cash Dividend Policy,” *Journal of Corporate Finance*, 1, 33–62.

Eckbo, B. E. and P. Wier, 1985, “Antimerger Policy under the Hart-Scott-Rodino Act: A Reexamination of the Market Power Hypothesis,” *Journal of Law and Economics*, 28 (1, April), 119–150.

Eger, Carol E., 1983, “An Empirical Test of the Redistribution Effect in Pure Exchange Mergers,” *Journal of Financial and Quantitative Analysis*, 4, 547–572.

Emery, G. and J. Switzer, 1999, “Expected Market Reaction and the Choice of Method of Payment for Acquisitions,” *Financial Management*, 28(4), 73–86.

Faccio, M. and L. H. P. Lang, 2002, “The Ultimate Ownership of Western European Corporations,” *Journal of Financial Economics*, 65, 365–395.

Faccio, M. and R. W. Masulis, 2005, “The Choice of Payment Method in European Mergers and Acquisitions,” *Journal of Finance*, 60, 1345–1388.

Faccio, M., J. J. McConnell and D. Stolin, 2006, “Returns to Acquirers of Listed and Unlisted Targets,” *Journal of Financial and Quantitative Analysis*, 41, 197–220.

Fama, E. F. and K. R. French, 1993, “Common Risk Factors in the Returns on Stocks and Bonds,” *Journal of Financial Economics*, 43, 3–56.

Fee, C. E. and S. Thomas, 2004, “Sources of Gains in Horizontal Mergers: Evidence from Customers, Supplier, and Rival Firms,” *Journal of Financial Economics*, 74, 423–460.

Field, L. C. and J. M. Karpoff, 2002, “Takeover Defenses at IPO Firms,” *Journal of Finance*, 57, 1857–1889.

Fishman M. J., 1988, “A Theory of Preemptive Takeover Bidding,” *RAND Journal of Economics*, 19, 88–101.

Fishman M. J., 1989, “Preemptive Bidding and the Role of the Medium of Exchange in Acquisitions,” *Journal of Finance*, 44, 41–57.

Focarelli, D. and F. Panetta, 2003, “Are Mergers Beneficial to Consumers? Evidence from the Market for Bank Deposits,” *American Economic Review*, 93, 1151–1171.

Franks, J., C. P. Mayer and S. Rossi, 2005, “Ownership: Evolution and Regulation,” Working Paper, London Business School.

Franks, J. R., R. S. Harris and C. Mayer, 1988, “Means of Payment in Takeovers: Results for the U.K. and the U.S.,” in A. Auerbach (Ed.), *Corporate Takeovers*. (NBER, University of Chicago Press).

Fridolfson, S.-O. and J. Stennek, 2005, “Why Mergers Reduce Profits, and Raise Share Prices—a Theory of Preemptive Mergers,” *Journal of the European Economic Association*, 3, 1083–1104.

Fuller, K., J. Netter and M. Stegemoller, 2002, “What do Returns to Acquiring Firms Tell Us? Evidence from Firms that Make Many Acquisitions,” *Journal of Finance*, 57, 1763–1793.

Galai, D. and R. W. Masulis, 1976, “The Option Pricing Model and the Risk Factor of Stock,” *Journal of Financial Economics*, 3, 53–81.

Gaughan, P. A., 2007, *Mergers, Acquisitions, and Corporate Restructurings*. 4th edition (John Wiley & Sons, Hoboken, NJ).

Geczy, C. C., D. K. Musto and A. V. Reed, 2002, "Stocks are Special Too: An Analysis of the Equity Lending Market," *Journal of Financial Economics*, 66, 241–269.

Ghosh, A. and W. Ruland, 1998, "Managerial Ownership and the Method of Payment for Acquisitions, and Executive Job Retention," *Journal of Finance*, 53(2), 785–797.

Gilson, R. J., M. S. Scholes and M. A. Wolfson, 1988, "Taxation and the Dynamics of Corporate Control: The Uncertain Case for Tax-Motivated Acquisitions," in J. C. Coffe, L. Lowenstein and S. Rose-Ackerman (Ed.), *Knights, Raiders and Targets: The Impact of the Hostile Takeover*. (Oxford University Press, New York).

Goldman, E. and J. Qian, 2005, "Optimal Toeholds in Takeover Contests," *Journal of Financial Economics*, 77, 321–346.

Gomes, A., 2001. "Takeovers, Freezeouts and Risk Arbitrage," Working Paper, University of Pennsylvania.

Gompers, P., J. Ishi and A. Metrick, 2003, "Corporate Governance and Equity Prices," *Quarterly Journal of Economics*, 118, 107–155.

Gordon, L. A. and J. Pound, 1993, "Information, Ownership Structure, and Shareholder Voting: Evidence from Shareholder–Sponsored Corporate Governance Proposals," *Journal of Finance*, 68, 697–718.

Gort, M., 1969, "An Economic Disturbance Theory of Mergers," *Quarterly Journal of Economics*, 83, 624–642.

Grinstein, Y. and P. Hribar, 2004, "CEO Compensation and Incentives—Evidence from M&A Bonuses," *Journal of Financial Economics*, 73, 119–143.

Grossman, S. and O. Hart, 1980a, "Disclosure Law and Takeover Bids," *Journal of Finance*, 35, 323–334.

Grossman, S. J. and O. D. Hart, 1980b, "Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation," *Bell Journal of Economics*, 11, 42–64.

Hackbarth, D. and E. Morellec, 2008, "Stock Returns in Mergers and Acquisitions," *Journal of Finance*, forthcoming.

Hansen, R. G., 1985, "Auctions with Contingent Payments," *American Economic Review*, 75, 862–865.

Hansen, R. G., 1987, "A Theory for the Choice of Exchange Medium in the Market for Corporate Control," *Journal of Business*, 60, 75–95.

Harford, J., 2005, "What Drives Merger Waves?" *Journal of Financial Economics*, 77, 529–560.

Harford, J., 2003, "Takeover Bids and Target Directors' Incentives: The Impact of a Bid on Directors' Wealth and Board Seats," *Journal of Financial Economics*, 69, 51–83.

Harford, J., 1999, "Corporate Cash Reserves and Acquistions," *Journal of Finance*, 54(6, December), 1969–1997.

Harford, J., and K. Li, 2007, "Decoupling CEO Wealth and Firm Performance: The Case of Acquiring CEOs," *Journal of Finance*, 62, 917–949.

Harris, M. and A. Raviv, 1988, "Corporate Control Contests and Capital Structure," *Journal of Financial Economics*, 20, 55–86.

Hartzell, J. C., E. Ofek and D. Yermack, 2004, "What's in It for Me? CEOs Whose Firms are Acquired," *Review of Financial Studies*, 17, 37–61.

Hasbrouck, J., 1985, "The Characteristics of Takeover Targets," *Journal of Banking and Finance*, 9, 351–362.

Hayn, C. 1989, "Tax Attributes as Determinants of Shareholder Gains in Corporate Acquisitions," *Journal of Financial Economics*, 23, 121–153.

Helwege, J., C. Pirinsky and R. M. Stulz, 2007, "Why Do Firms Become Widely Held? An Analysis of the Dynamics of Corporate Ownership?" *Journal of Finance*, forthcoming.

Hermalin, B. E. and M. S. Weisbach, 2003, "Boards of Directors as an Endogenously Determined Institution: A Survey of the Economic Literature," *Federal Reserve Bank of New York Economic Policy Review*, 9, 7–26.

Heron, R. A. and E. Lie, 2004, "A Comparison of the Motivations for and the Information Content of Different Types of Equity Offerings," *Journal of Business*, 77, 605–632.

Heron, R. A. 2006, "On the Use of Poison Pill and Defensive Payouts by Takeover Targets," *Journal of Business*, 79, 1783–1807.

Higgins, R. C. and L. D. Schall, 1975, "Corporate Bankruptcy and Conglomerate Merger," *Journal of Finance*, 30, 93–111.

Himmelberg, C., R. G. Hubbard and D. Palia, 1999, "Understanding the Determinants of Manager Ownership and the Link between Ownership and Firm Performance," *Journal of Financial Economics*, 53, 353–384.

Hirshleifer, D., 1995, "Mergers and Acquisitions: Strategic and Informational Issues," in R. A. Jarrow and V. M. W. T. Ziemba (Ed.), *Finance*, vol. 9 of *Handbooks in Operation Research and Management Science*, Chapter 26, 839–885. (North-Holland).

Hirshleifer, D. and I. P. L. Ping, 1990, "Facilitation of Competing Bids and the Price of a Takeover Target," *Review of Financial Studies*, 2, 587–606.

Hirshleifer, D. and S. Titman, 1990, "Share Tendering Strategies and the Success of Hostile Takeover Bids," *Journal of Political Economy*, 98, 295–324.

Holderness, C. G., 2006, "The Myth of Diffuse Ownership in the United States," *Review of Financial Studies*, forthcoming.

Holderness, C. G., R. S. Kroszner and D. P. Sheehan, 1999, "Were the Good Old Days that Good? Changes in Management Stock Ownership Since the Great Depression," *Journal of Finance*, 54, 435–469.

Holmstrom, B. and S. Kaplan, 2001, "Corporate Governance and Merger Activity in the U.S.: Making Sense of the 1980's and 1990's," Working Paper, NBER no W8220.

Holmstrom, B. and B. Nalebuff, 1992, "To the Raider Goes the Surplus? A Reexamination of the Free Rider Problem," *Journal of Economics and Management Strategy*, 1, 37–62.

Hotchkiss, E. S., 1995, "Post-bankruptcy Performance and Management Turnover," *Journal of Finance*, 50, 3–21.

Hotchkiss, E. S., K. John, R. Mooradian and K. S. Thorburn, 2008, "Bankruptcy and the Resolution of Financial Distress," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 2, Chapter 14 (Elsevier/North-Holland, Handbooks in Finance Series).

Hotchkiss, E. S. and R. M. Mooradian, 1997, "Vulture Investors and the Market for Control of Distressed Firms," *Journal of Financial Economics*, 43, 401–432.

Hotchkiss, E. S., 1998, "Acquisitions as a Means of Restructuring Firms in Chapter 11," *Journal of Financial Intermediation*, 7, 240–262.

Hotchkiss, E. S., 2003, "Auctions in Bankruptcy," *Journal of Corporate Finance*, 9, 555–574.

Hsieh, J. and R. A. Walking, 2005, "Determinants and Implications of Arbitrage Holdings in Acquisitions," *Journal of Financial Economics*, 77, 605–648.

Huang, Y.-S. and R. A. Walkling, 1987, "Abnormal Returns Associated with Acquisition Announcements: Payment Method, Acquisition Form, and Managerial Resistance," *Journal of Financial Economics*, 19, 329–349.

Hubbard, G. and D. Palia, 1999, "A Re-examination of the Conglomerate Merger Wave in the 1960's: An Internal Capital Markets View," *Journal of Finance*, 54, 1131–1152.

Huson, M. R., R. Parrino and L. Starks, 2001, "Internal Monitoring Mechanism and CEO Turnover: A Long-term Perspective," *Journal of Finance*, 56, 2265–2297.

Jarrell, G. A. and M. Bradley, 1980, "The Economic Effects of Federal and State Regulations of Cash Tender Offers," *Journal of Law and Economics*, 23(2, October), 371–407.

Jarrell, G. A., J. A. Brickley and J. M. Netter, 1988, "The Market for Corporate Control: The Empirical Evidence Since 1980," *Journal of Economic Perspectives*, 2, 49–68.

Jarrell, G. A. and A. B. Poulsen, 1987, "Shark Repellents and Stock Prices: The Effects of Antitakeover Amendments since 1980," *Journal of Financial Economics*, 19, 127–168.

Jarrell, G. A. and A. B. Poulsen, 1989, "Stock Trading before the Announcement of Tender Offers: Insider Trading or Market Anticipation?" *Journal of Law, Economics and Organization*, 5, 225–248.

Jennings, R. H. and M. A. Mazzeo, 1993, "Competing Bids, Target Management Resistance, and the Structure of Takeover Bids," *Review of Financial Studies*, 6, 883–910.

Jensen, M., 1993, "The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems," *Journal of Finance*, 48 (July), 831–880.

Jensen, M. C., 1989, "Eclipse of the Public Corporation," *Harvard Business Review*, September–October, 61–74.

Jensen, M. C., 1986, “Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers,” *American Economic Review*, 76, 323–329.

Jensen, M. C., and W. Meckling, 1976, “Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure,” *Journal of Financial Economics*, 3, 305–360.

Jensen, M. C. and K. J. Murphy, 1990, “Performance Pay and Top-Management Incentives,” *Journal of Political Economy*, 98, 225–264.

Jensen, M. C. and R. S. Ruback, 1983, “The Market for Corporate Control,” *Journal of Financial Economics*, 11, 5–50.

Jindra, J. and R. A. Walkling, 2004, “Speculation Spreads and the Market Pricing of Proposed Acquisitions,” *Journal of Corporate Finance*, 10, 495–526.

Jovanovic, B. and P. L. Rousseau, 2004, “Mergers as Reallocation,” Working Paper, New York University.

Jovanovic, B. and P. L. Rousseau, 2002, “The Q-Theory of Mergers,” *American Economic Review*, 92, 198–204.

Karolyi, G. A. and J. Shannon, 1999, “Where’s the Risk in Risk Arbitrage?” *Canadian Investment Review*, 12, 11–18.

Kim, E. H. and V. Singal, 1993, “Mergers and Market Power: Evidence from the Airline Industry,” *American Economic Review*, 83, 549–569.

Kim, E. H. and J. J. McConnell, 1977, “Corporate Mergers and the Co-insurance of Corporate Debt,” *Journal of Finance*, 32, 349–365.

King, M. R. and M. Padalko, 2005, “Pre-bid Run-ups Ahead of Canadian Takeovers: How Big is the Problem?” Working Paper 2005-3, Bank of Canada.

Kini, O., W. Kracaw and S. Mian, 2004, “The Nature and Discipline by Corporate Takeovers,” *Journal of Finance*, 59, 1511–1552.

Kini, O., W. Kracaw and S. Mian, 1995. “Corporate Takeovers, Firm Performance, and Board Composition,” *Journal of Corporate Finance*, 1, 383–412.

Kisgen, D. J., J. Qian and W. Song, 2006, “Are Fairness Opinions Fair? The Case of Mergers and Acquisitions,” Working Paper, Boston College.

Kohers, N. and J. Ang, 2000, “Earnouts in Mergers: Agreeing to Disagree and Agreeing to Stay,” *Journal of Business*, 73, 445–476.

Kohers, N., G. Kohers and T. Kohers, 2007, “Glamour, Value, and the Form of Takeover,” *Journal of Economics and Business*, 59, 74–87.

Kothari, S. P. and J. B. Warner, 2007, “Econometrics of Event Studies,” in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 1, Chapter 1 (Elsevier/North-Holland, Handbooks in Finance Series).

Kothari, S. P. and J. B. Warner, 1997, “Measuring Long-horizon Security Price Performance,” *Journal of Financial Economics*, pp. 301–339.

La, P. R., F. Lopez-de-Silanes and A. Shleifer, 1999, “Corporate Ownership Around the World,” *Journal of Finance*, 54, 471–518.

Larcker, D. F. and T. Lys, 1987, “An Empirical Analysis of the Incentives to Engage in Costly Information Acquisition: The Case of Risk Arbitrage,” *Journal of Financial Economics*, 18, 111–126.

Lease, R., J. McConnell and E. Tashjian, 1996, “Prepacks: An Empirical Analysis of Prepackaged Bankruptcies,” *Journal of Financial Economics*, 40, 135–162.

Lehn, K. and A. B. Poulsen, 1992, “Contractual Resolution of Bondholder-Stockholder Conflicts in Leveraged Buyouts,” *Journal of Law and Economics*, 34, 645–673.

Lehn, K. and M. Zhao, 2006, “CEO Turnover After Acquisitions: Are Bad Bidders Fired?” *Journal of Finance*, 61, 1759–1811.

Levi, M., K. Li and F. Zhang, 2008, “Mergers and Acquisitions: The Role of Gender,” Working Paper, Sauder School of Business, University of British Columbia.

Levy, H. and H. Sarnat, 1970, “Diversification, Portfolio Analysis, and the Uneasy Case for Conglomerate Mergers,” *Journal of Finance*, 25, 795–802.

Lewellen, W., 1971, “A Pure Financial Rationale for the Conglomerate Merger,” *Journal of Finance*, May.

Li, K. and N. R. Prabhala, 2007, "Self-Selection Models in Corporate Finance," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 1, Chapter 2, 37–86 (Elsevier/North-Holland, Handbooks in Finance Series).

Linn, S. and J. McConnell, 1983, "An Empirical Investigation of the Impact of Anti-Takeover Amendments on Common Stock Prices," *Journal of Financial Economics*, 11, 361–400.

Lipton, M. and E. H. Steinberger, 2004, "*Takeovers & Freezeouts*," vol. 1. (Law Journal Press, New York).

Loderer, C. and K. Martin, 1990, "Corporate Acquisitions by Listed Firms: The Experience of Comprehensive Sample," *Financial Management*, 19, 17–33.

Loughran, T. and A. M. Vijh, 1997, "Do Long-term Shareholders Benefit from Corporate Acquisitions?" *Journal of Finance*, 52, 1765–1790.

Lyon, J. D., B. M. Barber and C.-L. Tsai, 1999, "Improved Methods for Tests of Long-run Abnormal Stock Returns," *Journal of Finance*, 54, 165–201.

MacKinlay, A. C., 1997, "Event Studies in Economics and Finance," *Journal of Economic Literature*, 35, 13–39.

Mahoney, J. H. and J. T. Mahoney, 1993, "An Empirical Investigation of the Effect of Corporate Charter Amendments on Stockholder Wealth," *Strategic Management Journal*, 14, 17–31.

Makhija, A. K. and R. P. Narayanan, 2007, "Fairness Opinions in Mergers and Acquisitions," Working Paper, Ohio State University.

Maksimovic, V. and G. Phillips, 2007, "Conglomerate Firms and Internal Capital Markets," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 1, Chapter 8, 423–479 (Elsevier/North-Holland Handbooks in Finance Series).

Maksimovic, V., G. Phillips and N. R. Prabhala, 2008, "Post-Merger Restructuring and the Boundaries of the Firm," Working Paper, University of Maryland.

Maksimovic, V. and G. Phillips, 2002, "Do Conglomerate Firms Allocate Resources Inefficiently across Industries?" *Journal of Finance*, 57, 721–768.

Maksimovic, V. and G. Phillips, 2001, "The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and are There Efficiency Gains?" *Journal of Finance*, 56, 2019–2065.

Maksimovic, V. and G. Phillips, 1998, "Asset Efficiency and Reallocation Decisions of Bankrupt Firms," *Journal of Finance*, 53, 1495–1532.

Malatesta, P., 1983, "The Wealth Effect of Merger Activity and the Objective Functions of Merging Firms," *jfe*, 11(1–4 April), 155–181.

Malatesta P. and R. Thompson, 1985, "Partially Anticipated Events: A Model of Stock Price Reaction with an Application to Corporate Acquisitions," *Journal of Financial Economics*, 14, 237–250.

Malatesta, P. H. and R. A. Walkling, 1988, "Poison Pill Securities: Stockholder Wealth, Profitability, and Ownership Structure," *Journal of Financial Economics*, 20, 347–376.

Malekzadeh, A. R, V. B. McWilliams and N. Sen, 1998, "Implications of CEO Structural Ownership Powers, Board Ownership and Composition on the Market's Reaction to Antitakeover Charter Amendments," *Journal of Applied Business Research*, 14, 54–62.

Manne, H. G., 1965, "Mergers and the Market for Corporate Control," *Journal of Political Economy*, 73, 110–120.

Maquieira, C. P., W. L. Megginson and L. A. Nail, 1998, "Wealth Creation Versus Wealth Redistribution in Pure Stock-for-stock Mergers," *Journal of Financial Economics*, 48, 3–33.

Martin, K., 1996, "The Method of Payment in Corporate Acquisitions, Investment Opportunities, and Management Ownership," *Journal of Finance*, 51, 1227–1246.

Martin K. J. and J. McConnell, 1991, "Corporate Performance, Corporate Takeovers, and Management Turnover," *Journal of Finance*, June, 671–687.

Martynova, M. and L. Renneboog, 2007, "Sources of Transaction Financing in Corporate Takeovers," Working Paper, Tilburg University.

Martynova, M. and L. Renneboog, 2006, "Mergers and Acquisitions in Europe," in Luc Renneboog (Ed.), *Advances in Corporate Finance and Asset Pricing*, Chapter 2, 13–75. (Elsevier).

Martynova, M. and L. Renneboog, 2005, "Takeover Waves: Triggers, Performance and Motives," Working Paper, Tilburg University.

Masulis, R. W., C. Wand and F. Xie, 2007, "Corporate Governance and Acquirer Returns," *Journal of Finance*, 62, 1851–1889.

Matsusaka, J. G., 1996, "Did Tough Antitrust Enforcement Cause the Diversification of American Corporations?" *Journal of Financial and Quantitative Analysis*, 31, 283–294.

Matsusaka, J. G., 1993, "Takeover Motives During the Conglomerate Merger Wave," *RAND Journal of Economics*, 24, 357–379.

McAfee, R. and M. Williams, 1988, "Can Event Studies Detect Anticompetitive Mergers?" *Economics Letters*, 28, 199–203.

McWilliams, V. B. and N. Sen, 1997, "Board Monitoring and Antitakeover Amendments," *Journal of Financial and Quantitative Analysis*, 32, 491–505.

Meulbroek, L., 1992, "An Empirical Analysis of Illegal Insider Trading," *Journal of Finance*, 47, 1661–1699.

Mikkelson, W. and M. Partch, 1997, "The Decline of Takeovers and Disciplinary Managerial Turnover," *Journal of Financial Economics*, 44, 205–228.

Mikkelson, W. H. and M. M. Partch, 1989, "Managers' Voting Rights and Corporate Control," *jfe*, 25, 263–290.

Mikkelson, W. H. and R. S. Ruback, 1991, "Targeted Repurchases and Common Stock Returns," *rand*, 22(4), 544–561.

Mikkelson, W. H, and R. S. Ruback, 1985, "An Empirical Analysis of the Interfirm Equity Investment Process," *Journal of Financial Economics*, 14, 523–553.

Mitchell, M. and K. Lehn, 1990, "Do Bad Bidders Become Good Targets?" *Journal of Political Economy*, 98, 372–398.

Mitchell, M. and J. H. Mulherin, 1996, "The Impact of Industry Shocks on Takeover and Restructuring Activity," *Journal of Financial Economics*, 41, 193–229.

Mitchell, M. and T. Pulvino, 2001, "Characteristics of Risk and Return in Risk Arbitrage," *Journal of Finance*, 56 (6), 2135–2175.

Mitchell, M. and E. Stafford, 2004, "Price Pressure Around Mergers," *Journal of Finance*, 59 (1), 31–63.

Moeller, S. B., F. P. Schlingemann and R. M. Stulz, 2007, "How do Diversity of Opinion and Information Asymmetry Affect Acquirer Returns?" *Review of Financial Studies*, 20, 2047–2078.

Moeller, S. B., F. P. Schlingemann and R. M. Stulz, 2005, "Wealth Destruction on a Massive Scale? A Study of Acquiring Firm Returns in the Recent Merger Wave," *Journal of Finance*, 60, 757–782.

Moeller, S. B., F. P. Schlingemann and R. M. Stulz, 2004, "Firm Size and the Gains from Acquisitions," *Journal of Financial Economics*, 73, 201–228.

Moeller, T., 2005, "Let's Make a Deal! How Shareholders Control Impacts Merger Payoff," *Journal of Financial Economics*, 76, 167–190.

Molnar, J., 2008, "Preemptive Horizontal Mergers: Theory and Evidence," Working Paper, Bank of Finland.

Morellec, E. and A. Zhdanov, 2005, "The Dynamics of Mergers and Aquisitions," *Journal of Financial Economics*, 77, 649–672.

Mork, R., A. Shleifer and R. W. Vishny, 1988, "Characteristics of Targets of Hostile and Friendly Takeovers," in Alan J. Auerbach (Ed.), *Corporate Takeovers: Causes and Consequences*. National Bureau of Economic Research, Chicago, IL.

Moussawi, R., 2004, "Shareholders and Antitakeover Provisions," Working Paper, University of Texas at Dallas.

Mulherin, J. H. and A. L. Boone, 2000, "Comparing Acquisitions and Divestitures," *Journal of Corporate Finance*, 6, 117–139.

Munro, J. R., 1989, "Takeovers: The Myths Behind the Mystique," *Vital Speeches*, May 15.

Murphy, K. J., 1999, "Executive Compensation," in O. Ashenfelter and D. Card (Ed.), *Handbook of Labor Economics*, vol. 3b of *Handbook of Labor Economics*, Chapter 38, 2485–2563 (Elsevier/North Holland).

Myers, S. C., 1984, "The Capital Structure Puzzle," *Journal of Finance*, 39, 575–592.

Myers, S. C., and N. S. Majluf, 1984, "Corporate Financing and Investment Decisions when Firms Have Information that Investors do not Have," *Journal of Financial Economics*, 13, 187–221.

Nash, R., J. Netter and A. B. Poulsen, 2003, "Determinants of Contractual Relations between Shareholders and Bondholders: Investment Opportunities and Restrictive Covenants," *Journal of Corporate Finance*, 9, 201–232.

Nathan, K. S. and T. B. O'Keefe, 1989, "The Rise in Takeover Premiums: An Exploratory Study," *Journal of Financial Economics*, 23, 101–119.

O'Brien, A. P., 1988, "Factory Size, Economies of Scale, and the Great Merger Wave of 1898–1902," *Journal of Economic History*, 48, 639–649.

Offenberg, D., 2008, "Firm Size and the Effectiveness of the Market for Corporate Control," *Journal of Corporate Finance*, forthcoming.

Officer, M. S., 2007, "The Price of Corporate Liquidity: Acquisition Discounts for Unlisted Targets," *Journal of Financial Economics*, 83, 571–598.

Officer, M. S., 2006, "The Market Pricing of Implicit Options in Merger Collars," *Journal of Business*, 79, 115–136.

Officer, M. S., 2004, "Collars and Renegotiations in Mergers and Acquisitions," *Journal of Finance*, 59, 2719–2743.

Officer, M. S., 2003, "Termination Fees in Mergers and Acquisitions," *Journal of Financial Economics*, 69, 431–467.

Officer, M. S., A. B. Poulsen and M. Stegemoller, 2007, "Information Asymmetry and Acquirer Returns," *Review of Financial Studies*, forthcoming.

Palepu, K. G., 1986, "Predicting Takeover Targets: A Methodological and Empirical Analysis," *Journal of Accounting and Economics*, 8, 3–35.

Penas, M. F. and H. Unal, 2004, "Gain from Bank Mergers: Evidence from the Bond Markets," *Journal of Financial Economics*, 74, 149–179.

Posner, R., 1969, "The Federal Trade Commission," *University of Chicago Law Review*, 37, 47–89.

Prager, R. A. and T. H. Hannan, 1998, "Do Substantial Horizontal Mergers Generate Significant Price Effects? Evidence from the Banking Industry," *Journal of Industrial Economics*, 46, 433–452.

Rau, P. R. and T. Vermaelen, 1998, "Glamour, Value and the Post-Acquisition Performance of Acquiring Firms," *Journal of Financial Economics*, 49, 223–253.

Rhodes-Kropf, M. and D. T. Robinson, 2008, "The Market for Mergers and the Boundaries of the Firm", *Journal of Finance* 63, 1169–1211.

Rhodes-Kropf, M., D. T. Robinson and S. Viswanathan, 2005, "Valuation Waves and Merger Activity: The Empirical Evidence," *Journal of Financial Economics*, 77, 561–603.

Rhodes-Kropf, M. and S. Viswanathan, 2004, "Market Valuation and Merger Waves," *Journal of Finance*, 59, 2685–2718.

Ritter, J. R., 2003, "Investment Banking and Security Issuance," in G. Constantinides, M. Harris, and R. Stulz (Ed.), *Handbook of the Economics of Finance: Corporate Finance*, Chapter 5, 254–304. (Elsevier/North-Holland, Handbooks in Economics Series).

Roll, R., 1986, "The Hubris Hypothesis of Corporate Takeovers," *Journal of Business*, 59, 437–467.

Rose, M. J., 2008, "Heterogeneous Impacts of Staggered Boards by Ownership Concentration," *Journal of Corporate Finance*, forthcoming.

Rosen, R. J., 2006, "Merger Momentum and Investor Sentiment: The Stock Market Reaction to Merger Announcements," *Journal of Business*, 79, 987–1017.

Rose, S., 1977, "The Determination of Financial Structure: The Incentive Signalling Approach," *Bell Journal of Economics*, 8, 23–40.

Ryngaert, M., 1988, "The Effect of Poison Pill Securities on Stockholder Wealth," *Journal of Financial Economics*, 20, 377–417.

Savor, P., 2006, "Do Stock Mergers Create Value for Acquirers?" Working Paper, Wharton School of Business.

Schipper, K. and R. Thompson, 1983, "Evidence on the Capitalized value of Merger Activity for Acquiring Firms," *Journal of Financial Economics*, 11, 85–119.

Schlingemann, F. P., 2004, "Financing Decisions and Bidder Gains," *Journal of Corporate Finance*, 10, 683–701.

Schumann, L., 1993, "Patterns of Abnormal Returns and the Competitive Effects of Horizontal Mergers," *Review of Industrial Organization*, 8, 679–696.

Schwert, G. W., 1996, "Markup Pricing in Mergers and Acquisitions," *Journal of Financial Economics*, 41, 153–192.

Schwert, G. W., 2000, "Hostility in Takeovers: In the Eyes of the Beholder?" *Journal of Finance*, 55, 2599–2640.

Servaes, H. and A. Tamayo, 2007, "Waking up the Neighbors: How Industry Peers Respond to Control Threats?" Working Paper, London Business School.

Servaes, H., 1991, "Tobin's q and the Gains from Takeovers," *Journal of Finance*, 46, 409–419.

Shahrur, H., 2005, "Industry Structure and Horizontal Takeovers: Analysis of Wealth Effects on Rivals, Suppliers, and Corporate Customers," *Journal of Financial Economics*, 76, 61–98.

Shapley, L., 1953, "A Value for N-Person Games," in H. Kuhn and A. Tucker (Ed.), *Contributions to the Theory of Games, Volume II, Annals of Mathematics Studies*, 28, 307–317. (Princeton, Princeton University Press).

Shenoy, J., 2008, "An Examination of the Efficiency, Foreclosure, and Collusion Rationales for Vertical Takeovers," Working Paper, Georgia State University.

Shivdasani, A., 1993, "Board Composition, Ownership Structure, and Hostile Takeovers," *Journal of Accounting and Economics*, 16, 167–198.

Shleifer, A. and R. Vishny, 2003, "Stock Market Driven Acquisitions," *Journal of Financial Economics*, 70, 295–311.

Shleifer, A. and R. W. Vishny, 1997, "The Limits of Arbitrage," *Journal of Finance*, 52 (1), 35–55.

Shleifer, A. and R. W. Vishny, 1992, "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," *Journal of Finance*, 47, 1343–1366.

Shleifer, A. and R. W. Vishny, 1991, "Takeovers in the '60s and the '80s: Evidence and Implications," *Strategic Management Journal*, 12, 51–59.

Shleifer, A. and R. W. Vishny, 1989, "Management Entrenchment: The Case of Manager-Specific Investments," *Journal of Financial Economics*, 25, 123–139.

Shleifer, A. and R. W. Vishny, 1986b, "Large Shareholders and Corporate Control," *Journal of Political Economy*, 94, 461–488.

Shleifer, A. and R. W. Vishny, 1986a, "Greenmail, White Knights, and Shareholders' Interest," *RAND Journal of Economics*, 17, 293–309.

Singal, V., 1996, "Airline Mergers and Competition: An Integration of Stock and Product Market Effects," *Journal of Business*, 69, 233–268.

Slovin, M. B., M. E. Sushka and C. D. Hudson, 1991, "Deregulation, Contestability, and Airline Acquisitions," *Journal of Financial Economics*, 30, 231–251.

Smith, R. L. and J.-H. Kim, 1994, "The Combined Effect of Free Cash Flow and Financial Slack on Bidder and Target Stock Return," *Journal of Business*, 67, 281–310.

Snyder, C. M., 1996, "A Dynamic Theory of Countervailing Power," *RAND Journal of Economics*, 27, 747–769.

Song, M. H. and R. A. Walkling, 2005, "Anticipation, Acquisitions and Bidder Returns," LeBow College of Business Working Paper.

Song, M. H. and R. A. Walkling, 2000, "Abnormal Returns to Rivals of Acquisition Targets: A Test of the Acquisition Probability Hypothesis," *Journal of Financial Economics*, 55, 143–172.

Spatt, C., 1989, "Discussion: Strategic Analyses of Takeover Bids," in Sudipto Bhattacharya and George Constantinides (Ed.), *Financial Markets and Incomplete Information: Frontiers of Modern Financial Theory*, vol. 2, 106–121. (North-Holland).

Spiller, P. T., 1983, "The Differential Effect of Airline Regulation on Individual Firms and Markets: An Empirical Analysis," *Journal of Law and Economics*, 26, 655–689.

Stigler, G. J., 1950, "Monopoly and Oligopoly Power by Merger," *American Economic Review*, 40, 23–34.

Stigler, G. J., 1982, "The Economist and the Problem of Monopoly," *American Economic Review*, 72, 1–11.

Stillman, R., 1983, "Examining Antitrust Policy Toward Horizontal Mergers," *Journal of Financial Economics*, 11, 225–240.

Strömberg, P., 2000, "Conflicts of Interests and Market Illiquidity in Bankruptcy Auctions: Theory and Tests," *Journal of Finance*, 55, 2641–2692.

Stulz, R., 1988, "Managerial Control of Voting Rights: Financing Policies and the Market for Corporate Control," *Journal of Financial Economics*, 20, 25–54.

Sundaramurthy C., J. M. Mahoney and J. T. Mahoney, 1997, "Board Structure, Antitakeover Provisions, and Stockholder Wealth," *Strategic Management Journal*, 18, 231–245.

Thompson, R., 1995, "Empirical Methods in Event Studies in Corporate Finance," in Robert A. Jarrow and Vojislav Maksimovic William T. Ziemba (Ed.), *Finance*, vol. 9 of *Handbooks in Operation Research and Management Science*, Chapter 29, 963–992. (North-Holland).

Thompson, R., 1985, "Conditioning the Return-Generating Process on Firm-Specific Events: A Discussion of Event Study Methods," *Journal of Financial and Quantitative Analysis*, 20, 151–168.

Thorburn, K. S., 2000, "Bankruptcy Auctions: Costs, Debt Recovery, and Firm Survival," *Journal of Financial Economics*, 58, 337–368.

Toffanin, M., 2005, "Examining the Implications of Financing Choice for Cash Acquisitions," Unpublished Master of Science Dissertation, Concordia University.

Travlos, N. G., 1987, "Corporate Takeover Bids, Method of Payment, and Bidding Firms' Stock Returns," *Journal of Finance*, 42, 943–963.

Walkling, R., 1985, "Predicting Tender Offer Success: A Logistic Analysis," *Journal of Financial and Quantitative Analysis*, 20, 461–478.

Walkling, R. A. and M. S. Long, 1984, "Agency Theory, Managerial Welfare, and Takeover Bid Resistance," *RAND Journal of Economics*, 15, 54–68.

Warga, A. and I. Welch, 1993, "Bondholder Losses in Leveraged Buyouts," *Review of Financial Studies*, 6, 959–982.

Wasserstein, B., 2000, *Big Deal: Mergers and Acquisitions in the Digital Age*. (Warner Books, New York).

Werden, G. J., A. S. Joskow and R. Johnson, 1991, "The Effects of Mergers on Price and Output: Two Case Studies from the Airline Industry," *Managerial and Decision Economics*, 12, 341–352.

White, H., 1980, "A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," *Econometrica*, 48, 817–838.

Wulf, J., 2004, "Do CEOs in Mergers Trade Power for Premium? Evidence from Mergers of Equals," *Journal of Law, Economics and Organization*, 20, 60–101.

Yermack, D., 2006, "Golden Handshakes: Separation Pay for Retired and Dismissed CEOs," *Journal of Accounting and Economics*, 41, 237–256.

Yook K. C., 2003, "Larger Returns to Cash Acquisitions: Signaling Effect or Leverage Effect?" *Journal of Business*, 76, 477–498.

Zingales, L., 1994, "The Value of the Voting Right: Study of the Milan Stock Exchange Experience," *Review of Financial Studies*, 7, 125–148.

*Chapter 16*

# CORPORATE RESTRUCTURING: BREAKUPS AND LBOs

B. ESPEN ECKBO

*Tuck School of Business at Dartmouth*

KARIN S. THORBURN

*Tuck School of Business at Dartmouth*

## Contents



*Handbook of Empirical Corporate Finance, Volume 2*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

## Abstract

This chapter surveys the empirical literature on corporate breakup transactions (divestitures, spinoffs, equity carveouts, tracking stocks), leveraged recapitalizations, and leveraged buyouts (LBOs). Many breakup transactions are a response to excessive conglomeration and reverse costly diversification discounts. The empirical evidence shows that the typical restructuring creates substantial value for shareholders. The value-drivers include elimination of costly cross-subsidizations characterizing internal capital markets, reduction in financing costs for subsidiaries through asset securitization and increased divisional transparency, improved (and more focused) investment programs, reduction in agency costs of free cash flow, implementation of executive compensation schemes with greater pay-performance sensitivity, and increased monitoring by lenders and LBO sponsors. Buyouts after the turn of the century created value similar to LBOs of the 1980s. Recent developments include club deals (consortiums of LBO sponsors bidding together), fund-to-fund exits (LBO funds selling the portfolio firm to another LBO fund), a highly liquid (until mid-2007) leveraged loan market, and evidence of persistence in fund returns (perhaps because brand-sponsors borrow at better rates). Perhaps the greatest challenge to the restructuring literature is to achieve a modicum of integration of the analysis across transaction types. Another challenge is to produce precise estimates of the expected return from buyout investments in the presence of limited data on those portfolio companies that do not return to public status.

**Keywords**

restructuring, breakup, divestiture, spinoff, equity carveout, tracking stock, leveraged recapitalization, leveraged buyout, highly leveraged transaction

## 1. Introduction

Shocks to the corporate economic environment may give rise to severe organizational inefficiencies. For example, a vertically integrated firm may find that long-term contracts and/or spot market purchases of a key input have become more efficient. Or increased general capital market liquidity may have rendered internal capital markets a relatively costly divisional funding mechanism for conglomerates. High leverage may be optimal as financial innovations and expertise make it less expensive to manage financial distress. Financial innovations and general market liquidity may also render it optimal to securitize an entire division. The result is increased divisional managerial focus. In this chapter, we collectively refer to the transactions that implement these and other changes in asset composition, financial contracting, and ownership structure as "corporate restructurings".

We focus the survey on garden-variety restructuring procedures used to securitize and sell off part of the firm. Takeovers—the perhaps ultimate form of corporate restructuring—are reviewed in Chapter 15 of this Handbook (Betton, Eckbo, and Thorburn, 2008). However, we include leveraged buyouts (LBOs) in which the entire firm is acquired by a financial buyer such as a buyout fund. We also deal with issues of financial distress only tangentially, as the evidence surrounding restructurings in bankruptcy are covered in Chapter 14 (Hotchkiss, John, Mooradian, and Thorburn, 2008).

It is useful to classify corporate restructurings into two broad groups: breakups and highly leveraged transactions. Breakup transactions focus primarily on the separation of company assets and therefore include divestitures, spinoffs, equity carveouts, and tracking stock. Highly leveraged transactions involve a significant increase of debt in the firm's capital structure, either through a debt-financed special dividend in a leveraged recapitalization or in an LBO.[1]

Corporate restructurings may be initiated by the firm's top-level management, by divisional managers, or by outside sponsors like buyout funds. Occasionally, the restructuring is defensive, arising in response to a control threat from the market for corporate control. Regardless of who initiates the transaction, the parties are seeking to improve operating efficiency, increase cash flow, and, ultimately, enhance firm profitability. In breakup transactions, assets are transferred to higher-value users, while highly leveraged transactions involve optimizing capital structure, improving managerial incentives and achieving tax efficiency.

The survey is organized as follows. We begin in Section 2 with a brief discussion of the so-called diversification discount and potential costs of diversification, which motivate many breakup transactions. Sections 3 through Section 6 then detail the structure and consequences of various types of breakup transactions, including divestitures (Section 3), spin-offs (Section 4), equity carveouts (Section 5), and tracking stock (Section 6).

[1] We do not survey recapitalizations that do not involve extensive use of leverage. Examples include state privatizations (Megginson and Netter, 2001), conversions from mutual to stock companies (Masulis, 1987), and stock repurchases. Stock repurchases are reviewed in Chapter 10 of this Handbook (Kalay and Lemmon, 2008).

Next we turn to highly leveraged transactions, including leveraged recapitalizations and leveraged buyouts (LBOs). Section 9 concludes the survey.

## 2. Restructurings and the boundaries of the firm

### 2.1. Breakup transactions

The economic boundary of the firm may be defined as the point where within-firm transactions start to become more costly than arms-length (across market) transactions. There are numerous theories for why within-firm transactions may economically dominate market transactions, ranging from transactions costs (Coase, 1937) to agency costs and costs of imperfect contracting and moral hazard (Jensen and Meckling, 1976, Klein, Crawford and Alchian, 1978, Williamson, 1985, Grossman and Hart, 1985, Jensen, 1986, Hart and Moore, 1990). Alternatives outright ownership of resources include renting (long- or short-term contracts) and "spot" market transactions to ensure continued operations of the firm. These organizational alternatives have different implications for corporate taxes, firm-specific resource specialization and development of appropriable quasi-rents (which in turn lead to bargaining issues and potential for opportunistic behavior), investment decisions, risk-sharing and financing costs.

An asset such as an operating plant may have greater value as a division of a conglomerate than as a stand-alone "pure play" entity. The degree to which conglomerates generate value depends on the managerial skills and the nature of the resources required to operate efficiently within an industry. The value of using shared resources, such as managerial time and internal capital, differs across firms and industries as well as through time. As the boundaries of the firm change over time, some firms respond by undertaking expansions (mergers and acquisitions), breakups (divestitures, spinoffs) and recapitalizations (leveraged recaps and buyouts). Breakup transactions create value when synergies from conglomeration become negative, that is, when the costs of keeping the company's assets together exceed the benefits from doing so.

As emphasized by Maksimovic and Phillips (2007) (Chapter 8 of this Handbook), the corporate finance literature on conglomeration took off with the discovery of the "conglomerate discount" by Lang and Stulz (1994) and Berger and Ofek (1995). The discount is measured as the difference between the market value of the diversified firm and the sum of the estimated values of the (nontraded) divisions. The latter are estimated using multiples from single-segment (pure play) competitors. Berger and Ofek (1995) report a diversification discount of 13 to 15% in the 1986–1991 period. Subsequent empirical work has extended and reinterpreted the early results. Maksimovic and Phillips (2007) conclude that diversified firms *predominantly* behave like value maximizers given their productivity and that internal capital markets tend to facilitate the efficient transfer of resources. However, they also point to ambiguities reflecting econometric issues of endogeneity and self-selection, as well as choice of data, at various steps of the overall test strategy.

The typical breakup transaction reviewed below is shown to generate substantial shareholder value. This evidence is consistent with both the empire-building hypothesis and the value-maximizing self-selection hypothesis for the average observed diversification discount. Whether managers of firms breaking up are value-maximizers or empire-builders, the breakup may be an optimal response to exogenous changes in the economic boundaries of the firm. Reversing costs of excessive conglomeration may be a by-product of downsizing. Diversified firms undertaking breakup transactions are, however, more likely to be facing significant diversification costs than a random sample of conglomerates. Consequently, firms busting up are prime candidates for examining the potential nature of diversification costs.

The literature provides several examples of diversification costs and how they may distort investment. Scharfstein and Stein (2000) describe conditions under which top management inefficiently allocates too much funds to divisions with poor investment opportunities (cross-subsidization). Rajan, Servaes and Zingales (2000) argue that investment choices may be distorted because top management cannot commit to future distribution of funds until a surplus has been realized. Goldman (2004) models the resource allocation inside a multidivision firm of a manager with stock-based compensation and shows that the investment incentives improve after a spinoff of a division.

Another potential cost of diversification is related to executive compensation: since the division is a private entity, it is difficult to tie divisional manager compensation directly to the underlying value of the operations under their control. Stock-based compensation policies may be critical to induce optimal investment decisions and to retain managerial talent in a competitive labor market. A separate listing of subsidiary stock resolves such compensation issues, lowering agency costs and increasing market value.[2]

Breakup transactions may also result because conglomeration accentuates costly information asymmetries between investors and corporate insiders. Nanda and Narayanan (1999) model a diversified firm's decision to divest a division that is undervalued by the market. Outside investors observe the aggregated (conglomerate) cash flow only, while management also observes the divisional cash flows. Without detailed divisional information, the market rationally assigns an average performance to each division. This pooling results in undervaluation of the well-performing division and overvaluation of the poorly performing division. In this situation, it may he optimal to divest the overvalued (underperforming) division in order to lower the cost of capital for the undervalued division.

A related information-based argument is that conglomerates operating in a wide range of industries are more difficult for analysts to value correctly. This is true both because analysts tend to specialize in certain industries and because divisions may be relatively opaque in terms of financial information. A breakup may lead to increased coverage by financial analysts and improved quality of the information available to investors. Liu (2005) further maintains that a breakup allows outsiders more generally to discover firm

[2] See Aron (1991) for a model of this effect in the context of spinoffs.

value at a lower cost. As a result, high-value firms may undertake breakups in order to separate themselves from low-value firms.

Are there too few breakups? Boot (1992) argues that self-interested managers are reluctant to sell assets because a divestiture may signal poor managerial quality. Lang, Poulsen, and Stulz (1995) also point out that managers who value control may be reluctant to sell assets in order to promote operating efficiency alone. In this situation, an active market for corporate control may be required to force more divestitures. Financial distress is another scenario which may force even non-value-maximizing managers of financially constrained firms to divest assets in order to raise capital (see also Hotchkiss, John, Mooradian and Thorburn, 2008, Chapter 14 of this Handbook).

The above arguments emphasize how breakups create value by reversing negative synergies. A divisional or asset sale may also be the result of the demand side: the assets may simply be worth more under the buyer's control. That is, the buyer may be a higher-quality manager, and the divisional resources may offer a greater potential for synergies when merged with the acquiring firm. Selling the asset at a premium may serve the interest of all parties involved. Finally, corporate breakups may be forced by direct legal actions under antitrust or bankruptcy court, or by regulatory changes changing the economic boundary of the firm.

### *2.2. Highly leveraged transactions*

In a highly leveraged transaction, the focus of the restructuring is on the economic effects of the leverage increase. Whether undertaking a debt-financed dividend (leveraged recap), or a leveraged purchase of a division or the entire firm (LBO, where the firm goes private), it is the leverage increase rather than any concomitant asset restructuring that provides the main economic motivation for the transaction. As a result, LBOs tend to involve financial (as opposed to strategic) buyers, such as buyout funds.

The literature points to several possible sources of gains in leverage-increasing transactions. Under the classical trade-off theory of debt (see Frank and Goyal, 2008, Chapter 12 of this Handbook), firms move to a higher level of debt in order to capitalize on the corporate debt tax shield provided by the (U.S.) tax law. In addition to the potential for corporate tax benefits, the literature emphasizes beneficial managerial incentive and monitoring effects of higher leverage. Some highly leveraged firms may also gain a strategic advantage in product markets. On the other hand, high leverage is not for everyone: under conditions of financial distress, a debt overhang tends to prevent efficient investments (Myers, 1977).

In terms of managerial incentives, Ross (1977) presents a signaling model in which managers who face personal bankruptcy costs signal their private information about higher future expected cash flows by committing to a greater corporate debt level. In the vernacular of Jensen (1986), entrenched managers prefer to overinvest rather than pay out the firm's "free cash flow" as dividends (where free cash flow is defined as corporate liquid funds in excess of what is required to fund all positive net present value projects). A leveraged recapitalization, where the firm increases its debt without retaining the proceeds (thus increasing leverage ratios), reduces Jensen's overinvestment problem by precommitting to disgorge future cash flows in the form of interest payment. Jensen (1986) further argues that the greater risk of financial distress associated with higher leverage also helps discipline managerial investment policies. Stulz (1990) formalizes this intuition and shows that high leverage is particularly valuable when investment opportunities are poor, even if the free cash flow is negative.

Increasing leverage also allows wealth constrained managers to hold a greater percentage of total equity after the transaction is completed. For example, in a leveraged recapitalization, the debt may be paid out as cash dividend to non-managerial stockholders and as a stock dividend (or a cash dividend that is immediately reinvested in the firm) to managers. In an LBO, the managers may roll over their equity investment, while other equity-holders are paid out, again increasing managers' fractional equity ownership. The incentive effect of such greater managerial equity ownership helps reduce manager-shareholder conflicts of interest. Garvey (1992) explores the conditions under which leverage and management equity ownership are complementary in reducing the overinvestment problem of free cash flow.[3]

Highly leveraged transactions may also lead to improved monitoring by banks, and by the LBO sponsor who has its own money at risk in the transaction. Jensen (1989) argues that active governance by buyout sponsors and high-powered managerial incentives, combined with the pressure from high leverage, provides an incentive structure that is superior to that of public firms with dispersed ownership and weak governance. He even suggests that the LBO organizational form may "eclipse" the traditional corporate form, a prediction that has yet to be proven (we present evidence on the frequency of LBO transactions in Section 8 below).

Moreover, highly leveraged transactions may cause wealth transfers across the firm's various constituencies. For example, bonds that lack protective covenants may become more junior in the capital structure, resulting in a bondholder loss (benefiting shareholders). It is also possible that incumbent managers participating in a leveraged buyout have inside information about the firm's future prospects, expropriating selling shareholders. Muller and Panunzi (2004) argue that the LBO sponsor may be in a position to expropriate minority shareholders by merging the firm with the raider's leveraged acquisition subsidiary. Perotti and Spier (1993) present a model in which the firm gains bargaining power in contracting renegotiations by temporarily increasing leverage. Specifically, after retiring equity through a junior debt issue, shareholders threaten to underinvest in valuable new projects unless employees concede to wage reductions. Finally, there is a growing literature linking leverage to the firm's strategic position in product markets. See the reviews of Maksimovic (1995) and Parsons and Titman (2008), Chapter 13 of this Handbook, for reviews of this literature.

[3] See also Garvey (1995) for an analysis of managerial incentive effects of leverage.

We now turn to a detailed description of the empirical evidence on breakups and highly leveraged transactions. In the course of discussing the evidence, we return to several of the hypotheses outlined above.

## 3. Divestitures

A divestiture is the sale of a portion of the firm's assets to a third party—typically another company or a private equity fund—in a private transaction. The assets sold may be a division, segment, subsidiary, or product line. In return, the seller typically receives cash, but sometimes also securities or a combination of both. The proceeds from the sale are reinvested in the remaining business or distributed to the firm's claim holders. While eliminating some assets, the selling firm continues to exist in essentially the same form as before. Divestitures may trigger a substantial tax liability: the difference between the proceeds from the sale and the firm's tax basis in the assets is a capital gain or capital loss, which is taxed at the corporate tax rate.

### *3.1. Transaction volume*

In 2006, U.S. corporations announced 3,375 divestitures with a total deal value of $342 billion (*source*: *Mergerstat Review*). The line in Panel A of Figure 1 shows the annual number, and the bars show the total dollar volume of U.S. divestitures over the period 1980–2006. The number of transactions was relatively stable between 1980 and 1995. Since the mid-1990s, however, the divestiture activity has tripled and reached record high levels in 2005–2006.

The most aggressive divester in 2006 was UTEK (26 divestitures), followed by General Electric (17), Clear Channel Communications (11), El Paso (10), and Federated Department Stores (10). Two of the sellers, General Electric and El Paso, were also among the most aggressive divesters in the previous year. In addition, General Electric was listed as the most aggressive U.S. acquirer in 2006 and 2005, with 30 and 28 acquisition announcements, respectively.

The total divestiture activity tracks closely the merger and acquisition (M&A) activity in the economy. Panel B of Figure 1 shows the annual number of U.S. divestitures as a percentage of all U.S. takeovers from 1970 and forward. While the number of divestitures increased sharply in the second half of the 1990s, it fell behind the even greater increase in M&A volume over the same period. This trend was reversed once the takeover activity slowed after the turn of the century. In 2006, divestitures made up 32% of all M&A transactions, somewhat below the annual average of 38% over the whole 1970–2006 period.

### *3.2. Valuation effects*

Panel A of Table 1 shows the stock price reaction of the divesting firm for 18 selected studies with announcement dates in years 1963 through 1999. The studies typically

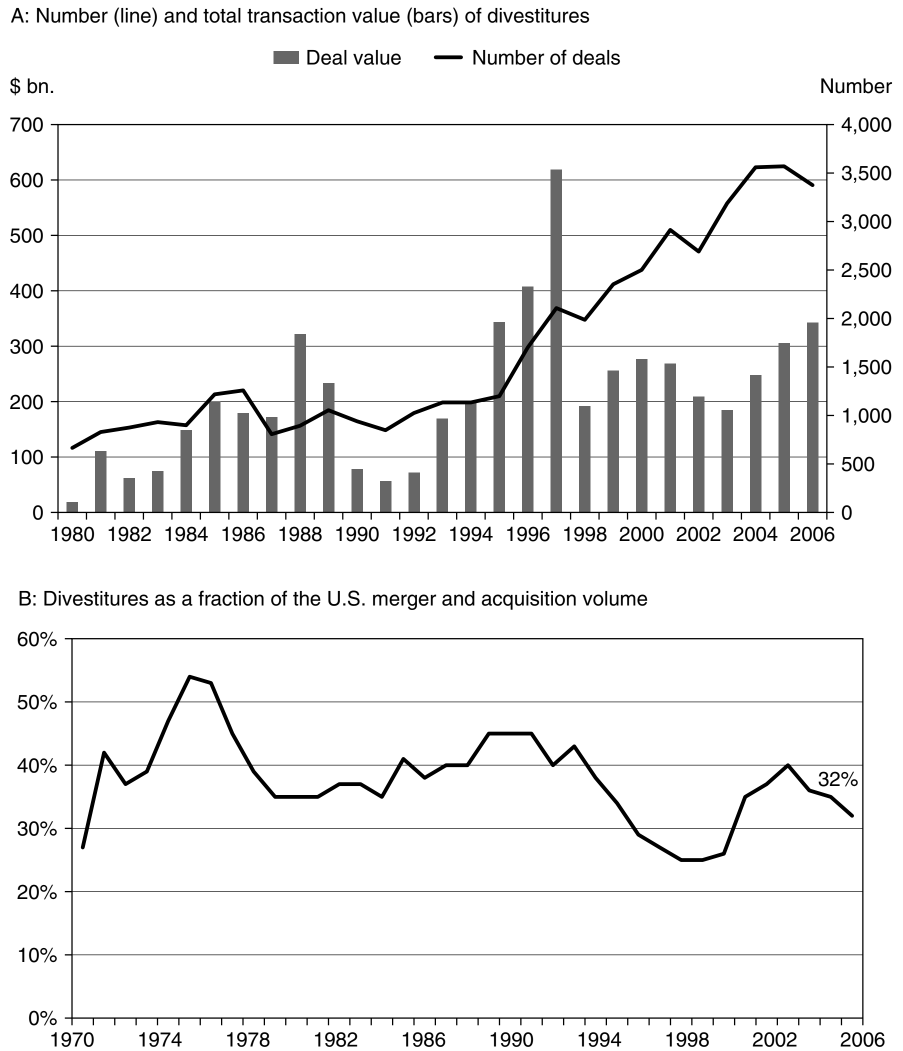



Fig. 1. Annual volume of U.S. divestitures, 1970–2006.
*Source*: Mergerstat Review.

report the cumulative abnormal stock return (CAR) over the two-day interval (−1, 0) where day 0 is the announcement day.[4] The average CAR for the announcements are

[4] A typical approach is to estimate the parameters using a single-factor market model over approximately a year prior to the event: $R_{jt} = \alpha_j + \beta_j R_{mt} + \epsilon_{jt}$, where $R_{jt}$ is the stock return of firm $j$ and $R_{mt}$ is the market return on day $t$. The abnormal return $AR_{j\tau}$ over event day $\tau$ is computed as $AR_{j\tau} = R_{j\tau} - (\hat{\alpha}_j + \hat{\beta}_j R_{m\tau})$, where $\hat{\alpha}_j$ and $\hat{\beta}_j$ are the coefficient estimates from the time-series regression. The cumulative abnormal return is $CAR(\tau_1, \tau_2) = \Sigma_{\tau=\tau_1}^{\tau_2} AR_{j\tau}$, where $\tau_1$ and $\tau_2$ define the event window relative to the announcement day 0.

Table 1

Cumulative abnormal returns (CAR) for divestiture announcement of 3,700 sellers and 1,243 buyers in 19 selected studies, 1963–1999

Relative size is the average ratio of the sales price of the divested assets to the pre-deal total assets (TA) and market value of equity (MVE) of the seller and buyer, respectively.

| Study | CAR | | Relative size | | Sample size | Time period | Event window |
|---|---|---|---|---|---|---|---|
| | Mean | Median | TA | MVE | | | |
| **Panel A: Seller returns:** | | | | | | | |
| Alexander, Benson, and Kampmeyer (1984) | 0.3% | | | | 53 | 1964–1973 | [−1,0] |
| Linn and Rozeff (1984) | 1.6 % | | | | 77 | | [−1,0] |
| Rosenfeld (1984) | 2.3% | | | | 62 | 1969–1981 | [−1,0] |
| Jain (1985) | 0.5% | | | | 1,062 | 1976–1978 | [−1,0] |
| Klein (1986) | 1.1% | | | | 202 | 1970–1979 | [−2,0] |
| Hite, Owers, and Rogers (1987) | 1.5% | | 16% | | 114 | 1963–1981 | [−1,0] |
| Hirschey, and Zaima (1989) | 1.6% | | | | 170 | 1975–1982 | [−1,0] |
| Hirschey, Slovin, and Zaima (1990) | 1.5% | | | 38% | 75 | 1975–1982 | [−1,0] |
| Afshar, Taffler, and Sudarsanam (1992) | 0.7% | | | 10% | 178 | 1985–1986 | [−1,0] |
| Sicherman and Pettway (1992) | 0.9% | | | 30% | 278 | 1980–1987 | [−1,0] |
| John and Ofek (1995) | 1.5% | 0.8% | | 39% | 258 | 1986–1988 | [−2,0] |
| Lang, Poulsen, and Stulz (1995) | 1.4% | 0.7% | 11% | 69% | 93 | 1984–1989 | [−1,0] |
| Loh, Bezjak, and Toms (1995) | 1.5% | | | | 59 | 1980–1987 | [−1,0] |
| Slovin, Sushka, and Ferraro (1995) | 1.7% | 0.7% | 33% | 17% | 179 | 1980–1991 | [0,1] |
| Hanson and Song (2000) | 0.6% | 0.3% | | 27% | 326 | 1981–1995 | [−1,1] |
| Mulherin and Boone (2000) | 2.6% | 1.6% | | 18% | 139 | 1990–1999 | [−1,1] |
| clubb and Stouraitis (2002) | 1.1% | 0.5% | | 14% | 187 | 1984–1994 | [−1,0] |
| Dittmar and Shivdasani (2003) | 3.4% | | 31% | | 188 | 1983–1994 | [−1,1] |
| Sample size weighted seller average | 1.2% | | 25% | 27% | 3,700 | 1963–1999 | |
| **Panel B: Buyer returns:** | | | | | | | |
| Jain (1985) | 0.5% | | | | 304 | 1976–1978 | [−1,0] |
| Hite, Owers, and Rogers (1987) | 0.6% | | 19% | | 105 | 1963–1981 | [−1,0] |
| Sicherman and Pettway (1992) | 0.5% | | | | 278 | 1980–1987 | [−1,0] |
| Datta and Iskandar-Datta (1995) | 0.0% | | | 13% | 63 | 1982–1990 | [−1,0] |
| John and Ofek (1995) | 0.4% | −0.5% | | 72% | 167 | 1986–1988 | [−2,0] |
| Hanson and Song (2000) | 0.5% | 0.2% | | | 326 | 1981–1995 | [−1,1] |
| Sample size weighted buyer average | 0.5% | | 19% | 25% | 1,243 | 1963–1995 | |

positive—ranging from 0.3% to 3.4% across the different samples—and almost all of the estimates are statistically significant at the 1%-level (two-sided $t$-test against zero). The sample-size-weighted average CAR for the combined sample of 3,700 divestitures is 1.2%. In sum, the evidence indicates that the average divestiture increases the value of the selling firm.

As further shown in the table, firms sell one-quarter or less of their total assets in the average transaction. Several studies find that the seller firm announcement returns are increasing in the relative size of the divested assets (Zaima and Hearth, 1985; Klein, 1986; Mulherin and Boone, 2000). It is possible that the returns on asset sales are independent of the size of the assets, so that relatively larger assets have a greater impact on the parent firm's return. This is similar to the effect of the relative size of the target on bidder returns documented in the takeover literature and reviewed in Betton, Eckbo, and Thorburn (2008). (See Chapter 15 of this Handbook.)

Klein (1986) reports that the disclosure of the sales price is central to the market's assessment of the transaction. She finds a positive seller stock price reaction only when the price is disclosed at the initial divestiture announcement. Firms that fail to announce the transaction price have CARs close to zero. The significance of price disclosure is confirmed by Afshar, Taffler, and Sudarsanam (1992) and Sicherman and Pettway (1992). Clubb and Stouraitis (2002) find that the announcement returns tend to increase with the difference between the sales price and an estimated value of the assets in their current use. Overall, this suggests that the market's valuation of the transaction depends on the sales price relative to the value of the assets when operated by the firm.

The abnormal returns on divestiture announcements are positive also for buyers, although they are of a smaller magnitude than for sellers. For six selected studies with data for the period 1963–1995 and listed in Panel B, the average buyer announcement CAR ranges from 0.0% to 0.6%. The sample-size-weighted buyer ACAR is 0.5% for the combined sample of 1,243 divestiture announcements. Sicherman and Pettway (1992) document a size effect in the buyer's stock price reaction similar to that of sellers; that is, buyer returns tend to increase with the relative size of the acquired assets.

While both sellers and buyers appear to gain from a divestiture, most of the gains tend to accrue to the selling (divesting) firm. In each individual transaction, however, the division of the total gains depends on the relative bargaining strength of the two parties. Sicherman and Pettway (1992) use a debt downgrade prior to the asset sale as an indication of a weaker bargaining position vis-à-vis the buyer. As expected, they find significantly lower CARs for sellers whose debt was downgraded prior to the transaction.

Moreover, the value creation is conditional on the successful completion of the divestiture. Hite, Owers, and Rogers (1987) show that the seller stock price drops back to its initial level if a previously announced divestiture is canceled. In addition, announcement returns are positive for buyers completing the transaction, but insignificant for buyers in transactions that subsequently fail.

### *3.3. Drivers of value creation in divestitures*

The positive announcement returns for sellers and buyers indicate that divestitures generally create value. We now turn to the evidence on the potential reasons for this value creation.

### 3.3.1. Increase in corporate focus

The typical divestiture involves sales of assets that are outside of the diversified firm's core business, and it results in an increased focus of the remaining operations. John and Ofek (1995) show that three-quarters of divested segments are unrelated to the seller's core business, defined as its primary four-digit Standard Industry Classification (SIC) code. Moreover, using various measures for firm focus, they find that sellers become more focused after the divestiture. Their focus measures include a sales-based Herfindahl Index across the firm's business segments, the total number of business lines reported by the firm, and whether the divested division is outside the firm's core business.

Schlingemann, Stulz, and Walkling (2002) find that firms tend to divest noncore segments that are relatively small. Maksimovic and Phillips (2001) and Kaplan and Weisback (1992) show that firms are more likely to sell peripheral assets. Kaiser and Stouraitis (2001) describe how Thorn EMI successfully raise cash by selling unrelated assets, reinvesting the proceeds in the company's core business. In sum, divested assets are typically outside the firm's core business, and the asset sales result in an increased focus of the firm's remaining operations.

John and Ofek (1995) find that the divestment announcement returns are positively related to measures capturing the increase in focus. Moreover, the operating profitability of the remaining assets increases after a divestiture, but only for the firms that become more focused. Denis and Shome (2005) show that large firms downsizing their assets become more focused and increase their operating performance. Berger and Ofek (1999) document average CARs of 7% for focusing-related announcements by diversified firms. Overall, there is substantial evidence that the value creation from divestitures is related to the resulting increase in the selling firm's focus.

### 3.3.2. Elimination of negative synergies

If the divested segment has negative synergies with other divisions of the diversified firms, the divestiture will create value simply by eliminating these negative synergies. Dittmar and Shivdasani (2003) examine the investment efficiency of divesting firms and find that segment sales are associated with a reduction of the diversification discount. Moreover, they document significant improvements in the investment decisions of the firm's remaining segments after the divestiture. Specifically, the investment level increases for segments that underinvest relative to single-segment firms and decreases for segments that overinvest relative to their peers. They also find that the announcement returns are higher the greater the subsequent reduction in the diversification discount and the greater the improvement in segment investments. Overall, their evidence suggests that divestitures create value by reducing costly cross-subsidization of inefficient investments in the diversified firm.

Colak and Whited (2007) reach a very different conclusion, addressing the endogeneity of breakup decisions. They confirm that firms selecting a divestiture or spin-off are different from their peers: the firms that restructure are typically larger and more diversified, and are in relatively fast-growing industries. Controlling for these differences, they show that although spin-offs and divestitures are associated with improved investment efficiency, these improvements are not directly caused by the restructuring itself.

Kaplan and Weisback (1992) examine whether divestitures are evidence of failed acquisitions. Studying a sample of 271 large firms acquired between 1971 and 1982, they find that 44% of the targets were sold by the end of 1989. Only one-third of the divested segments are classified as failed acquisitions, however, based on accounting profitability and comments by managers and the business press. Kaplan and Weisback (1992) conclude that acquirers sell businesses that they have improved or that they once had synergies with but no longer do. See also Fluck and Lynch (1999) for a model where diversifying acquisitions are made to help finance marginally profitable projects, to subsequently be divested once the projects are profitable and can generate the necessary funds internally.

#### *3.3.3. Better fit with the buyer*

As discussed above, a divestiture will create value if the assets are worth more to the buyer than the value in their current use. A buyer could, for example, have substantial synergies or superior management skills. John and Ofek (1995) find that seller announcement returns are higher when the buyer has some comparative advantage in managing the assets, such as a buyer operating in the same industry as the divested division or a leveraged buyout group.

Using U.S. Bureau of Census data, Maksimovic and Phillips (2001) examine the effect of asset sales on the productivity at the plant level. They show that divestitures are more likely in business cycle upturns, when the assets are less productive than industry benchmarks, when the selling division is less efficient than the buyer, and when the firm has more efficient divisions in other industries. They conclude that most divestitures result in productivity gains by redeploying assets from relatively low-productivity sellers to higher-ability buyers.

Datta, Iskandar-Datta, and Raman (2003) also study the efficiency of the reallocation of assets in divestitures. They use Tobin's $q$, defined as the ratio between the market value and the replacement cost (here the book value) of the assets, as a proxy for management's capability to manage the assets. They find that the announcement returns are highest for transactions where the buyer has a relatively high $q$ and the seller has a relatively low $q$, possibly because the assets are transferred to a better managed firm. Overall, the evidence suggests that divestitures create value by transferring assets to higher-valuation buyers.

### *3.4. Corporate governance*

#### *3.4.1. Agency issues*

Although divestitures may be required to maximize shareholder wealth, some incumbent managements resist such actions. Berger and Ofek (1999) find that announcements of focus-increasing transactions often are preceded by corporate control and incentive-altering events, including management turnover, outside shareholder pressure, changes in management compensation, and unsuccessful takeover attempts. Gillan, Kensinger, and Martin (2000) describe how Sears announced the divestiture of financial services and refocused on retail first after a long period of poor performance and coincident with substantial pressure from institutional investor activists. This suggests that the restructuring may have been postponed until it could no longer wait.

Consistent with a reluctance to sell assets, the monitoring of and incentives provided to top management are critical to the value created by a divestiture. Tehranian, Travlos, and Waegelein (1987) document significantly higher announcement returns for divesting firms that provide long-term performance plans to their top executives. Hirschey and Zaima (1989) find higher announcement returns for divestitures by companies with concentrated ownership than sales by widely held firms. Also, the returns are higher for firms where insiders are net-buyers of the firm's stock over the preceding six-month period. Hanson and Song (2000) further show that divestiture gains are increasing in the fraction of outside directors on the board and the percentage equity ownership of the management team. Pointing to the importance of banks as monitors, Hirschey, Slovin, and Zaima (1990) find some evidence of higher announcement returns for firms with bank debt. Overall, firms with better monitoring and more managerial share ownership seem to make divestitures that create more value.

The proceeds received by the divesting firm may be reinvested in the firm's remaining operations, used to retire debt, or distributed to shareholders. Lang, Poulsen, and Stulz (1995) and Kaiser and Stouraitis (2001) show that the announcement returns are positive when the proceeds are used to pay back debt, but insignificant for firms that reinvest the proceeds. Slovin, Sushka, and Ferraro (1995) also find higher announcement returns when the proceeds are paid out. This suggests that management may employ the funds inefficiently if retained by the firm.

Bates (2005) examines the payout and retention decision for 400 large asset sales between 1990 and 1998. He finds that the probability of retaining the cash proceeds increases in the divesting firm's growth opportunities, measured by its market-to-book ratio. However, firms retaining the proceeds consistently overinvest (have higher capital expenditure) relative to their industry peers. Also, the higher the equity ownership of officers and directors, the more likely it is that the sale proceeds are paid out. The evidence is again consistent with investment inefficiencies associated with retention of proceeds from asset sales.

### *3.4.2. Financial distress*

Several studies indicate that asset sales are used as a way of generating cash when the firm is financially constrained. Divestiture announcements are typically preceded by a period of negative stock returns (Alexander, Benson, and Kampmeyer, 1984; Jain, 1985; Hanson and Song, 2003) and poor operating performance (Lang, Poulsen, and Stulz, 1995; Schlingemann, Stulz, and Walkling, 2002; Brown, James, and Mooradian, 1994).

Moreover, firms with high leverage are more likely to sell assets (Ofek, 1993; Kruse, 2002). Officer (2007) shows that selling firms have lower cash balances, cash flow, and bond ratings than size- and industry-matched control firms, all of which suggests that the sellers are liquidity constrained. Also, Nixon, Roenfeldt, and Sicherman (2000) find that financially distressed firms prefer a divestiture to a spin-off, which does not generate cash. In addition, Asquith, Gertner, and Scharfstein (1992), Ofek (1993) and others show that firms in financial distress frequently sell assets as part of the restructuring process.

The optimal use of proceeds from asset sales changes when the firm is in financial distress. The firm's ability to pay dividends to shareholders is typically limited by debt covenants at this point, and the choice stands between reinvestment in the business or repayment of debt. For a sample of distressed firms, Brown, James and Mooradian (1994) show that shareholder announcement returns are significantly higher when the proceeds are retained by the firm rather than used to repay debt. Also as expected, bondholder announcement returns are higher when the proceeds are used to pay off debt. They suggest that creditor influence over distressed firms may force asset sales that benefits the firm's creditors at the detriment of shareholders. Datta and Iskander-Datta (1996) find that divestitures by financially distressed firms generate positive announcement returns for bondholders but not for shareholders.

Schleifer and Vishny (1992) argue that financially distressed firms sell assets at depressed prices to lower-valuation industry outsiders because higher-valuation industry insiders are liquidity constrained. Pulvino (1998) finds that financially constrained airlines sell aircraft at lower prices than their unconstrained competitors. Moreover, Officer (2007) shows that acquisition multiples are lower when the parent firm has experienced negative abnormal stock returns over the year leading up to the sale and when the corporate loan spread above treasury rates are high. Examining firms auctioned in Swedish bankruptcy, however, Eckbo and Thorburn (2007) reject the fire-sale hypothesis: they find little evidence of fire-sale discounts when assets are sold as going-concerns.[5]

Liquidity may be a factor in the decision to sell assets. Kim (1998) documents that managers sell their most liquid assets first, before selling more illiquid assets. Moreover, Mulherin and Boone (2000) and Schlingemann, Stulz, and Walking (2002) show that breakup transactions tend to cluster in industries where the aggregate corporate transaction volume is large, that is, in industries with relatively liquid markets for corporate assets.

## 4. Spinoffs

In a spinoff, a public company distributes its equity ownership in a subsidiary to its shareholders. The distribution is a pro-rata dividend, and parent shareholders receive

[5] See Chapter 14 (Hotchkiss, John, Mooradian, and Thorburn, 2008) for a more detailed review of asset restructurings by financially distressed firms.

subsidiary stock in proportion to their ownership in the parent firm. The spinoff involves a complete separation of the two firms. After the spinoff, the subsidiary becomes a publicly traded company with a unique ticker symbol and an independent board of directors. In contrast to a divestiture, a spinoff does not generate any cash proceeds for the parent company. Also, since the spinoff involves a public listing of shares, it has higher transaction costs and takes longer time than a divestiture.

A spinoff may be structured as a tax-free transaction if it qualifies under Section 355 of the Internal Revenue Code. Among the most important requirements under Section 355 are (i) the parent must have control of the subsidiary (own at least 80% of the voting rights) prior to the distribution; (ii) the parent must distribute control (at least 80% of the votes) to shareholders and retain no practical control of the subsidiary; (iii) the spinoff must have a valid business purpose; and (iv) the parent or the subsidiary cannot be acquired within two years after the spinoff. If the spinoff qualifies under Section 355, there is no tax on the distribution of stock, at neither the parent nor the shareholder level. Most spinoffs in the United States are structured as tax-free transactions.

If a spinoff does not qualify under Section 355, however, the distribution is taxed as a property dividend. The parent recognizes a gain equal to the difference between the fair market value of the subsidiary and the parent's tax basis in the subsidiary, similar to a capital gain. This gain is taxed at the corporate tax rate. Moreover, shareholders pay a dividend tax on the fair market value of the subsidiary (the distributed subsidiary stock).

The condition under Section 355 requiring that the subsidiary is not acquired within two years of the spinoff is outside the parent company's control. Yet, a potential acquisition of the subsidiary after a tax-free spinoff would trigger an often substantial tax liability at the parent company level. To transfer the cost of this potential liability to the subsidiary and thus ultimately the acquirer, it is common practice that the subsidiary contractually commits to pay any such future tax liability of the parent, if the subsidiary is acquired within two years of the spinoff.

Maydew, Schipper, and Vincent (1999) compare 52 tax-free spinoffs with 218 divestitures in the period 1987–1995. They find that tax costs average 8% of the divested assets. They suggest that managers prefer a taxable assets sale when the sales price is high enough to offset the associated tax cost.

### *4.1. Transaction volume*

Using data from Thompson SDC Platinum (SDC), Figure 2 plots the annual number (line) and total deal value (bars) of spinoffs announced between 1985 and 2007. As shown in Panel A, the number of U.S. spinoffs soared in the second half of the 1990s and reached a peak in year 2000 with over 90 transactions. The aggregate spinoff value peaked in 1999 with a total market capitalization of $144 billion. While the interest for spinoffs plummeted with the burst of the internet bubble, the deal activity has recently recovered. In 2007, a total of 90 spinoffs were announced in the United States to a total value of almost $80 billion. The largest U.S. spinoffs announced in 2006–2007

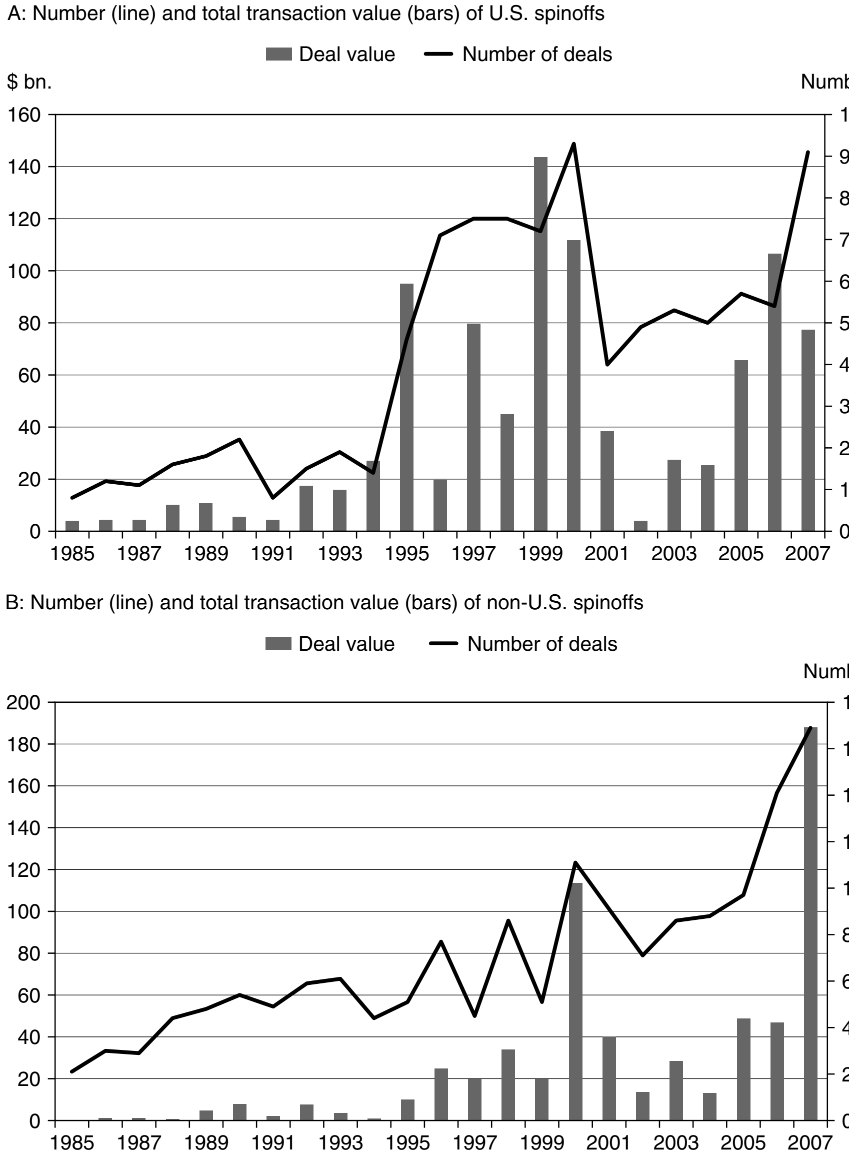



Fig. 2. Annual volume of spinoff announcements, 1985–2007.

*Source*: SDC

were Kraft Foods (market capitalization of $51 billion), Tyco Healthcare Group ($23 billion), Tyco Electronics ($19 billion), Duke Energy Corp-Natural Gas ($15 billion), and Discover Financial Services ($15 billion).

Panel B shows the annual number and total deal value of non-U.S. spinoff transactions. The international volume of spinoffs has been growing relatively steadily since the mid-1990s, reaching an all-time high in 2007, with a total market value of $188 billion across 169 transactions. Some of the largest spinoffs in 2006–2007 outside the United States include Philip Morris International, Switzerland (market value of $108 billion); SK Corp-Petrochemical Business, South Korea ($17 billion); HydroOGK, Russian Federation ($12 billion); and Experian Ltd., the United Kingdom ($11 billion).

### 4.2. *Valuation effects*

The results from 19 selected studies estimating shareholder gains from spinoff announcements are listed in Table 2. The samples contain a total of 2052 spinoffs announced between 1962 and 2000. Shareholder average cumulative abnormal returns are significantly positive and range from 1.7 to 5.6% across the various studies. The lowest average

Table 2

Cumulative abnormal returns (CAR) for 2,052 spinoffs in 19 selected studies, 1962–2000

Relative size is the ratio of the market value of equity of the spunoff subsidiary and the parent company prior to the spinoff.

| Study | CAR | | Relative size | | Sample size | Time period | Event window |
|---|---|---|---|---|---|---|---|
| | Mean | Median | Mean | Median | | | |
| Miles and Rosenfeld (1983) | 3.3% | | | 10% | 55 | 1963–1980 | [0,1] |
| Hite and Owers (1983) | 3.3% | | | 7% | 123 | 1963–1981 | [−1,0] |
| Schipper and Smith (1983) | 2.8% | | | 20% | 93 | 1963–1981 | [−1,0] |
| Rosenfeld (1984) | 5.6% | | | | 35 | 1969–1981 | [−1,0] |
| Vijh (1994) | 2.9% | 2.1% | 29% | 18% | 113 | 1964–1990 | [−1,0] |
| Allen, Lummer, McConnell, and Reed (1995) | 2.1% | | | | 94 | 1962–1991 | [−1,0] |
| Slovin, Sushka, and Ferraro (1995) | 1.3% | 1.6% | 33% | 24% | 37 | 1980–1991 | [0,1] |
| Daley, Mehrotra, and Sivakumar (1997) | 3.4% | 1.4% | | | 85 | 1975–1991 | [−1,0] |
| Best, Best, and Agapos (1998) | 3.4% | | | | 72 | 1979–1993 | [−1,0] |
| Desai and Jain (1999) | 3.8% | | 29% | 18% | 144 | 1975–1991 | [−1,1] |
| Krishnaswami and Subramaniam (1999) | 3.1% | 1.9% | 31% | 14% | 118 | 1979–1993 | [−1,0] |
| Mulherin and Boone (2000) | 4.5% | 3.6% | 22% | 14% | 106 | 1990–1999 | [−1,1] |
| Gertner, Powers, and Scharstein (2002) | 3.9% | 2.2% | 24% | 19% | 160 | 1982–1996 | [−1,0] |
| Wruck and Wruck (2002) | 3.6% | | | | 172 | 1985–1995 | [−1,0] |
| Burch and Nanda (2003) | 3.7% | 3.2% | 24% | 20% | 106 | 1979–1996 | [−2,1] |
| Maxwell and Rao (2003) | 3.6% | 2.6% | 25% | 19% | 80 | 1976–1997 | [−1,0] |
| Seoungpil and Denis (2004) | 4.0% | 3.1% | 25% | 17% | 150 | 1981–1988 | [−1,1] |
| Veld and Veld-Merkoulova (2004) | 1.7% | 0.6% | | | 156 | 1987–2000 | [−1,0] |
| McNeil and Moore (2005) | 3.5% | | 25% | 23% | 153 | 1980–1996 | [−1,1] |
| Sample size weighted average | 3.3% | | 26% | 18% | 2,052 | 1962–2000 | |

CAR of 1.7% is for a sample of 156 European spinoffs announced in 1987–2000 and examined by Veld and Veld-Merkoulova (2004). Combining the 19 studies, the sample-size-weighted abnormal announcement return is 3.3%.

The average CAR of 3.3% in spinoffs is higher than the 1.2% average CAR for divestitures reported above. Recall, however, that buyers also tend to experience positive announcement returns in divestitures (average CAR of 0.5%). In contrast, the total gains from a spinoff is reflected in the parent company stock. Thus, some of the difference in announcement returns between spinoffs and divestitures could be explained by buyers sharing in the value creation from the latter transaction.

Table 2 further shows that the market value of the subsidiary is about one-quarter that of its parent in the average spinoff. As for divestitures, the announcement returns for spinoffs are increasing in the relative size of the subsidiary. Miles and Rosenfeld (1983) show that shareholder CARs are on average greater in spinoffs of subsidiaries with a market value exceeding 10% of the parent company's market value compared to spinoffs of relatively small subsidiaries. In addition, Alli, Ramirez, and Yung (2001) find insignificant announcement returns for 47 spinoffs that are subsequently withdrawn, as if the market anticipates the withdrawal at the time of the announcement.

The evidence of positive announcement returns for spinoffs is compelling. Some studies also report long-term returns following spinoffs. Cusatis, Miles, and Woolridge (1983) estimate the buy-and-hold stock returns for parents and subsidiaries spun off in the 1965–1988 period. They find positive average returns for holding periods of 24 and 36 months compared with portfolios of industry and size-matched stocks. McConnell, Ozbilgin, and Wahal (2001) investigate portfolios of parents and subsidiaries in 89 spinoffs between 1989 and 1995. In contrast to the earlier work, they find little evidence of higher average buy-and-hold returns compared to portfolios matched on size and book-to-market. Also, using the Fama and French (1993) three-factor model as a benchmark, they reject the hypothesis that the portfolios of spinoff companies exhibit abnormal returns.

### *4.3. Drivers of value creation in spinoffs*

#### *4.3.1. Increased corporate focus*

As with divestitures, a potential source of value creation in spinoffs is an increase in corporate focus resulting from the elimination of unrelated divisions. Daley, Mehrotra, and Sivakumar (1997) report that the positive announcement returns are limited to spinoffs that increase corporate focus, defined as the parent and subsidiary having different two-digit SIC industry codes. They document substantial improvements in the return on assets for parents in focus-increasing spinoffs, but not for parents where the spunoff subsidiary is in a related industry. Moreover, Desai and Jain (1999) find that focus-increasing spinoffs have significantly higher announcement returns, long-run abnormal stock returns, and improvements in operating performance than do non-focus increasing spinoffs.

Burch and Nanda (2003) estimate the change in the parent firm's diversification discount from the year prior to the year after the spinoff. They find that the diversification discount is reduced when the spinoff increases corporate focus, but not otherwise. Overall, the evidence suggests that shareholder gains in spinoffs are associated with a subsequent increase in firm focus.

#### *4.3.2. Elimination of negative synergies*

The separation of an unrelated business segment may further reduce any negative synergies that exist between the subsidiary and the rest of the firm. Gertner, Powers, and Scharstein (2002) examine whether spinoffs help eliminate value-reducing cross-subsidization in diversified firms. They show that the subsidiary's investment decisions become much more sensitive to the firm's investment opportunities after the spinoff. Specifically, the total capital expenditure decreases for firms in low Tobin's $q$ industries and increases for firms in high $q$ industries. These changes take place primarily for subsidiaries whose operations are unrelated to the parent's core business and in spinoffs generating higher announcement returns.

Seoungpil and Denis (2004) further find that, prior to the spinoff, parent firms trade at a discount to and invest less in their high-growth (high $q$) divisions than do their stand-alone peers. Following the spinoff, however, the diversification discount is eliminated and investments have increased for the high-growth segments. Also, McNeil and Moore (2005) show that subsidiary capital expenditures move toward industry levels after the spinoff, for both previously rationed and subsidized divisions. Announcement returns are greater when parent firms allocate capital in a seemingly inefficient way, defined as rationing high $q$ and subsidizing low $q$ spunoff divisions, as is the reduction in the diversification discount. Overall, the evidence indicates that spinoffs create value by improving the investment decisions in diversified firms.

Allen, Lummer, McConnell, and Reed (1995) propose that spinoffs provide a way to unwind unsuccessful prior acquisitions. They examine a sample of 94 spinoffs in which the spunoff entity previously had been acquired by the parent firm. Their evidence suggests that the original acquisition was value destroying: the average acquisition announcement return is negative both for the acquirer and for the target and bidder combined. Moreover, the spinoff announcement return is positive and negatively correlated to the acquisitions return; that is, the greater the anticipated loss from the acquisition, the larger the expected gain from the spinoff. While not identifying a unique source for the value creation in spinoffs, these results are consistent with the elimination of negative synergies between parent and subsidiary.

#### *4.3.3. Wealth transfer from bondholders*

A spinoff may increase shareholder value at the expense of the parent firm's creditors by reducing the total assets of the firm. Also, if the spinoff increases the volatility of the cash flows of the two separate firms the expected payoff to debtholders will decrease, with a corresponding gain to shareholders (Galai and Masulis, 1976). MacMinn and Brockett (1995) further argue that a spinoff could transfer wealth from liability claimants by removing corporate assets from their reach. Nevertheless, the impact of a spinoff on debtholders is limited by the existence of restrictive debt covenants. Hite and Owers (1983) find insignificant bondholder abnormal returns for a sample of 31 spinoff announcements in 1963–1981, as do Schipper and Smith (1983).

In a case study of Marriott, however, Parrino (1997) documents a significant drop in the value of Marriott's bonds following its spinoff announcement. At the same time, shareholder announcement returns were positive, suggesting a wealth transfer from bondholders. Maxwell and Rao (2003) examine monthly bond return data for a sample of 80 spinoffs between 1976 and 1997. They find that parent bondholders tend to experience a price decline after the spinoff announcement. The average abnormal bond return (adjusted for the Treasury rate) in the month of the spinoff is $-0.9\%$, and decreasing in the relative size of the spunoff subsidiary. Consistent with a bondholder loss, credit ratings are more likely to be downgraded than upgraded subsequent to the spinoff. They find, however, that the combined value of the publicly traded debt and equity increases, suggesting that a wealth transfer from bondholders could only explain part of the shareholder gains.

#### *4.3.4. Information asymmetries*

The aggregation of financial data across divisions may exacerbate informational asymmetries between outside investors and insiders for diversified firms. Krishnaswami and Subramaniam (1999) examine whether spinoffs reduce such information gaps, using the dispersion in analysts' forecasts and analysts' forecast error as a measure for the information asymmetry. They find that spinoffs are more common for firms with relatively high levels of information asymmetry compared to their industry rivals. The announcement returns are higher for firms with a greater degree of information asymmetry, and the information gap tends to decrease after the spinoff. Best, Best, and Agapos (1998) also find that spinoff announcement returns are increasing in financial analysts' earnings forecast errors. Overall, this suggests that one source of value creation in spinoffs is the mitigation of information asymmetries.

Analysts play an important role in producing and disseminating information about the firm. Gilson, Healy, Noe, and Palepu (2001) study changes in the coverage by financial analysts for a sample of 103 focus-increasing spinoffs and equity carveouts over the period 1990–1995. They document a 45% increase in analyst coverage in the three years following a breakup. The new analysts tend to be specialists in the subsidiary's industry. Moreover, the accuracy of the earnings forecast improves by 30 to 50%, and in particular for the industry specialists. In sum, increases in corporate focus seem to improve the information provided by analysts, in both quality and quantity.

Huson and MacKinnon (2003) further show that analysts tend to revise upwards their short-term earnings forecast in response to a spinoff. Also, idiosyncratic stock return volatility increases following a spinoff, and more so when the spunoff subsidiary is unrelated to the parent firm. They conclude that the stock price becomes more sensitive to firm-specific information, which benefits informed traders relative to uninformed traders.

### *4.3.5. Clientele effects*

Previously combined into a single security, the spinoff creates the opportunity to hold the subsidiary stock separately. This expansion of investors' opportunity set increases liquidity and opportunities for investor diversification. In a sample of 113 spinoffs during 1964 to 1990, Vijh (1994) finds abnormal stock returns of 3.0% on the spinoff ex date, that is, the day that the subsidiary starts trading separately, accompanied by an increased trading volume. He attributes the positive returns to higher demand for the parent and subsidiary stocks once they have been separated.

Abarbanell, Bushee, and Ready (2003) show that institutional investors rebalance their portfolio holdings in parents and their spunoff subsidiaries dependent on the fund's investment style and fiduciary restrictions. However, they find little evidence that such rebalancing trades lead to abnormal price pressures for parents or subsidiaries around the spinoff. Chemmanur and He (2007) examine the trading of institutional investors in 66 spinoffs between 1999 and 2004. They find large imbalances in the post-spinoff trading of parent and subsidiary stock: 46% of the trades are in the opposite direction, and trades in the same direction are heavily concentrated in one of the firms. This imbalance increases in the measure of information asymmetry and the difference in beta risk and growth rates between the parent and subsidiary. Overall, spinoffs seem to relax a trading constraint that existed prior to distribution of the subsidiary stock.

### *4.3.6. Increased probability of a takeover*

The fact that it is possible after the spinoff to acquire control of the division through a stock purchase increases the probability that the division will become a future takeover target. The spinoff may also increase the probability that the parent will become a target as the parent is now a smaller and more focused firm. Cusatis, Miles, and Woolridge (1993) examine 146 tax-free spinoffs over the period 1965–1988 and show that both the parent and the spunoff subsidiary are indeed more likely to become takeover targets, compared to a set of control firms matched on size and industry. They suggest that two pure plays created by a spinoff are more attractive as targets than the combined company. Most of the takeovers occur two to three years after the spinoff, possibly to protect the tax-free status of the spinoff. Given the large premiums typically paid in control transactions, they attribute the positive abnormal stock returns at the time of the spinoff to the increased probability of being acquired.

### 4.4. Corporate governance

Self-interested managers may be reluctant to downsize assets under their control. Ahn and Walker (2007) study the importance of effective corporate governance for firms' decision to spin off a subsidiary. Their sample is 102 spinoffs between 1981 and 1997. They find that firms conducting a spinoff have greater stock ownership by outside board members, and smaller and more heterogeneous boards relative to their peers. Following the spinoff, parent firms increase their market-to-book ratios and reduce the diversification discount. They conclude that effective governance increases the likelihood of a spinoff, which is a value-increasing strategy.

Wruck and Wruck (2002) examine the management team of the spunoff subsidiary. They show that 21% of spinoff top managers are outsiders, while 48% of the insiders are parent company top managers rather than division heads. They argue that subsidiary managers lack the corporate governance expertise required when the former division becomes publicly traded. Announcement returns are highest for spunoff subsidiaries led by a parent firm's top manager and a division head, combining corporate governance and operating expertise.

In a spinoff, the parent management can design the governance structure of the subsidiary without seeking approval from shareholders. Daines and Kausner (2004) find that the charters of spunoff subsidiaries include substantially more takeover defenses than do the charters of a sample of size- and industry-matched IPO firms, where shareholders have a say on the corporate charter. Moreover, the spunoff firms tend to have more takeover protection than do their parents. Thus, it appears that managers prefer more takeover defenses than shareholders do.

Pyo (2007) find that pay-performance sensitivity increases for subsidiary CEOs after a spinoff. The higher the pay-performance sensitivity, the greater the improvements in operating performance post-spinoff. Seward and Walsh (1995) propose that the likelihood of becoming a takeover target should be higher for spunoff firms with little CEO equity incentives. They find that the takeover probability—hostile as well as friendly—increases with the CEO's stock and option ownership in the spunoff subsidiary. While not discussed by Seward and Walsh (1995), it is possible that CEOs with relatively low pay-performance sensitivity also adopt more takeover defenses in the spunoff firm.

Allen (2001) examines the post-spinoff trades of senior managers, directors, and blockholders in 193 public subsidiaries and their parents over the period 1978–1991. He finds that insiders who trade during the first year following the spinoff earn excess returns of 36% over the subsequent 12-month period. He suggests that insiders take advantage of the spinoff as an opportunity to use private information on the relative prospects of the parent and the subsidiary.

### 4.5. Splitoffs

A splitoff is similar to a spinoff in that the subsidiary becomes an independent company with a separate stock listing. The splitoff, however, involves an exchange offer, where shareholders are offered to exchange parent company stock for subsidiary stock. Thus, the splitoff effectively resembles a stock repurchase, where the parent company buys back its own shares using subsidiary stock as consideration. As a result of the exchange offer, the ownership structure in the parent and the subsidiary are different post-splitoff (depending on the extent to which parent shareholders participate in the exchange offer). Similar to a spinoff, a splitoff does not generate any new cash to the parent company. The tax treatment is also the same as for a spinoff.

Splitoffs are rare, partly because the valuation of the subsidiary stock is critical for the exchange offer. A splitoff is therefore always preceded by an equity carveout, which helps establish the market value of the subsidiary stock. Recent transactions include McDonald's splitoff of 51% of its interest in Chipotle Mexican Grill, announced in April 2006 and valued at $660 million; Viacom's splitoff of Blockbuster in 2004; and General Motors splitoff of Hugh Electronics in 2003.

We are unaware of any systematic empirical evidence on splitoffs—reflecting the limited number of transactions.[6] Given the similarity with spinoffs, the research on spinoffs is likely relevant for splitoffs as well. In addition, some value may be created in splitoffs from the repurchase of parent stock, for example, by signaling that the parent stock is undervalued.

## 5. Equity carveouts

An equity carveout is a partial initial public offering (IPO) of the stock in a subsidiary. The subsidiary gets its own management team and a separate board of directors. It becomes subject to all financial and other reporting requirements of public companies, such as 10-K reports and proxy statements filed with the Securities and Exchange Commission (SEC).[7]

The parent company often retains a controlling interest, creating a public minority interest in the subsidiary. There are several reasons for the retention of a majority ownership of the voting rights: Retention of at least 80% allows consolidation for tax purposes and the opportunity to subsequently undertake a tax-free spinoff, while retention of 50% or more permits consolidation for accounting purposes. Allen and McConnell (1998) show that parent firms on average retain 69% (median 80%) of the subsidiary's shares, while Vijh (2002) reports a median parent ownership of 72%. Of course, since the subsidiary becomes a publicly traded company of its own, the carveout does reduce the parent's control over its former wholly owned subsidiary.

The shares offered in the IPO may be sold either by the subsidiary itself (a primary issue) or by the parent company (a secondary issue). A primary issue has no tax consequence, while a secondary issue is taxable to the parent as a capital gain. Because

[6] For a case study, see E.I. du Pont de Nemours and Company: the Conoco splitoff (A), HBS 9-202-005.

[7] See Hand and Skantz (1998) for an analysis of the accounting choice for equity carveouts under SAB 51. Allen (1998) describes the equity carveout strategy of Thermo Electron, which carved out 11 subsidiaries during 1983–1995.

of this difference in tax treatment, the majority of equity carveouts are primary issues. The parent company may require the proceeds or leave the proceeds in the subsidiary. The proceeds are streamed back to to the parent using the following procedure: (i) prior to the carveout, the subsidiary issues a tax-free dividend to the parent in the form of a note (debt obligation); (ii) after the carveout, the proceeds from the IPO are used to repay the note.

### *5.1. Transaction volume*

Figure 3 shows the annual distribution of equity carveouts worldwide from 1985 to 2007, using data from SDC. Most of the carveout transactions are outside the United States. The worldwide carveout volume peaked in the first half of the 1990s, in both numbers and dollar values. The total market value of subsidiary IPOs reached $80 billion in 1993, and there were over 500 equity carveout transactions in 1994. The late 1990s saw a second surge in the dollar volume of carveouts ($70 billion), however, without a corresponding increase in the number of transactions. In recent years, only a handful of equity carveout transactions have taken place each year.

Most large carveouts in 2006/2007 took place outside the United States. The way SDC classifies carveouts, this transaction category also contains subsidiaries carved out by the government (state privatizations). The largest equity carveouts in 2007 include France

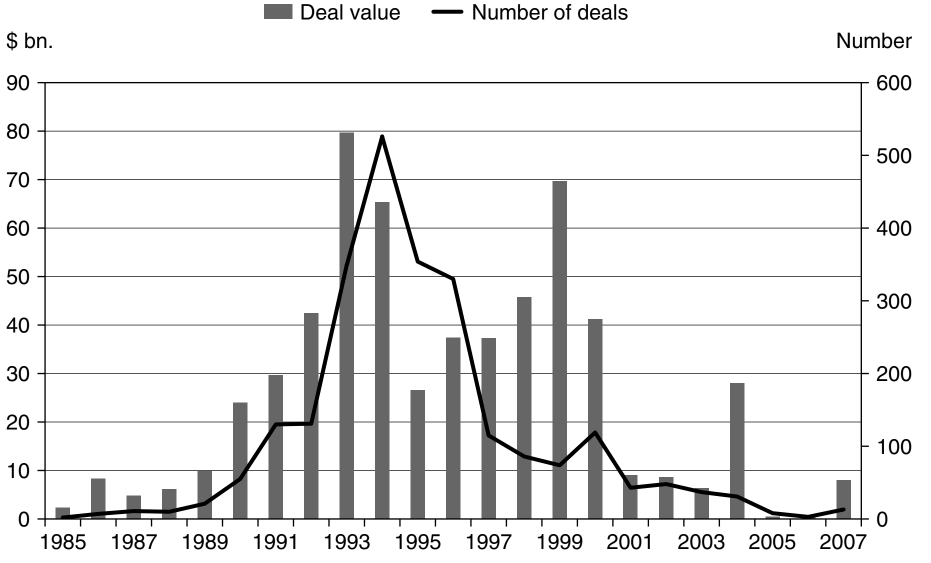



Fig. 3. Annual worldwide volume of equity carveouts, 1985–2007.

*Source*: SDC

Table 3

Cumulative abnormal returns for 1,050 equity carveout announcements in 8 selected studies, 1965–2002

CAR is the parent cumulative abnormal stock return over the event window relative to the announcement of the equity carveout. Relative size is the ratio of the market value of equity of the carved-out subsidiary and its parent company.

| Study | CAR | | Relative size | | Sample size | Time period | Event window |
|---|---|---|---|---|---|---|---|
| | Mean | Median | Mean | Median | | | |
| Schipper and Smith (1986) | 1.8% | | | 8% | 76 | 1965–1983 | [−4,0] |
| Klein, Rosenfeld, and Beranek (1991) | 2.7% | | | | 52 | 1966–1983 | [−4,0] |
| Slovin, Sushka, and Ferraro (1995) | 1.2% | 1.5% | 45% | 31% | 32 | 1980–1991 | [0,1] |
| Allen and McConnell (1998) | 2.1% | | 20% | 14% | 186 | 1978–1993 | [−1,1] |
| Vijh (1999, 2002) | 1.9% | | | 18% | 336 | 1980–1997 | [−1,1] |
| Mulherin and Boone (2000) | 2.3% | 0.8% | 37% | 17% | 125 | 1990–1999 | [−1,1] |
| Hulburt (2003) | 1.6% | 1.1% | | 30% | 172 | 1981–1994 | [−1,0] |
| Wagner (2004) | 1.7% | | 32% | 22% | 71 | 1984–2002 | [−1,1] |
| Sample size weighted average | 1.9% | | 33% | | 1,050 | 1965–2002 | |

Telecom SA (IPO proceeds of $3.6 billion); China Agri-Inds Holding Ltd., Hong Kong ($2.5 billion); Bank of Beijing, China ($2.0 billion); Qatar Airways, Qatar; and Kiora Holding Pty Ltd., Australia.

### 5.2. *Valuation effects*

Equity carveouts are viewed favorably by the market. Table 3 shows the parent cumulative abnormal announcement stock return for eight selected studies of equity carveouts over the period 1965–2002. The average announcement return is positive and significant across all samples, ranging from 1.2 to 2.7%. The sample-size-weighted average is 1.9% for the total of 1050 cases. The announcement returns for a sample of German firms average 1.7%, which is similar to the returns for U.S. firms (Wagner, 2004). Interestingly, the positive returns found for equity carveouts are in stark contrast to announcements of seasoned equity offerings, upon which the parent stock price typically falls.[8]

The average carved-out subsidiary across the studies in Table 3 has a market value of about one-third that of its parent. As for other breakup transactions, the announcement returns are found to be increasing in the relative size of the carved-out subsidiary (Allen and McConnell, 1998; Vijh, 2002). Vijh (1999) estimates long-term (three-year) abnormal stock returns for both parent companies and the carved-out subsidiaries, and finds that these are insignificantly different from zero using a variety of benchmarks.

[8] See Eckbo, Masulis, and Norli (2007) (Chapter 6 of this Handbook) for a review of security offerings.

### *5.3. Drivers of value creation in equity carveouts*

Equity carveouts separate the subsidiary from its parent. After the carveout, transactions between the two companies must take place at arms length. As a result, many of the sources of value creation in spinoffs may also create value in carveouts.

#### *5.3.1. Increased focus*

Vijh (2002) examines a sample of 336 equity carveouts between 1980 and 1997. A majority of the motives offered for the carveout by the parent company involve lack of fit and focus, and a desire to restructure the operations. He shows that parents and subsidiaries in carveouts are typically in different industries, and documents that announcement returns are higher on average for carveouts of nonrelated subsidiaries.

The evidence on improvements in operating performance following carveouts is mixed. Hulburt, Miles, and Woolridge (2002) find that both parents and subsidiaries improve their operating performance relative to their industry peers in the year after the carveout. In contrast, Powers (2003) and Boone, Haushalter, and Mikkelson (2003) show that the subsidiary operating performance declines after the carveout. Boone, Haushalter, and Mikkelson (2003) find that the operating performance of the parent company improves only when it has completely divested its ownership in the subsidiary after four years.

#### *5.3.2. Financing subsidiary growth*

Information asymmetries between the firm and outside investors tend to increase the cost of capital (Myers and Majluf, 1984). Prior to the carveout, outside investors have access to the parent company's financial information, with information at the divisional level being less accessible. This opaqueness may increase the cost of funding divisional-level capital expenditures. Because a public listing of the subsidiary increases the quality of the financial information available to investors, Schipper and Smith (1986) suggest that equity carveouts help finance high-growth subsidiaries. Their data bears this out: in their sample, a frequently stated motive for the carveout is to enable the subsidiary to finance future growth. They also show that carved-out subsidiaries typically have higher price-earnings ratios than their parents, indicating higher growth rates.

Chen and Guo (2005) also report that parent firms prefer equity carveouts and divestitures to spinoffs when revenue growth and book-to-market ratios are high. Vijh (2002) further finds that, over a subsequent three-year period, both parents and their carved-out subsidiaries do a greater number of seasoned equity offerings than control firms matched by industry and size. In addition, the capital expenditures of the subsidiaries exceed those of their control firms. Overall, it appears that equity carveouts are used to increase financing opportunities and reduce financing costs for high-growth subsidiaries.

Michaely and Shaw (1995) document investment banking fees of 7% for carveouts and 2% for spinoffs in a sample of 61 carveouts and 30 spinoffs between 1981 and 1988.

They attribute the higher costs of carveouts to the greater scrutiny and more stringent disclosure standard associated with the continued control by the parent company. They also suggest that, because of the higher costs, carveouts are more attractive to firms with relatively low leverage that hold high-quality assets. Consistent with this, they find that larger less-leveraged parents with relatively large and low-risk subsidiaries tend to prefer a carveout to a spinoff.

### *5.3.3. Wealth transfers and information asymmetries*

Carveouts have the potential for transferring wealth to shareholders from other claimholders. For example, the separation of assets from the parent possibly reduces the cash flow and collateral available to bondholders. Allen and McConnell (1998) find, however, positive excess bond returns when firms announce a carveout, thus rejecting the bondholder wealth transfer hypothesis.

Nanda (1991) models an equity carveout using the adverse selection framework of Myers and Majluf (1984). In equilibrium, only undervalued parents with overvalued subsidiaries perform carveouts. Thus, carveouts cause a positive announcement effect on average (and there are no wealth transfers).[9] Slovin, Sushka, and Ferraro (1995) examine industry rivals of equity carveout firms. They postulate that the market's misvaluation may apply to all firms in the industry. For a sample of 32 carveouts between 1980 and 1991, they show that industry rivals of the carved-out subsidiaries experience negative announcement returns, consistent with the overvaluation argument. They also report insignificant abnormal returns to parent-company rivals. However, Hulburt, Miles, and Woolridge (2002) find negative returns for parent-company rivals as well, using a sample of 185 equity carveout announcements over the years 1981–1994. They argue this is evidence against the proposition that parents of carveouts tend to be undervalued.

Vijh (2006) examines the announcement returns to the seasoned equity offering (SEO) of 90 subsidiaries and 37 parents following equity carveouts. He documents negative returns to the issuer, but insignificant returns to the nonissuer, whether parent or subsidiary. Using a sample of equity carveouts from 1995–2002, Baltin and Brettel (2007) detect traces of market timing for the 1998–2000 "hot-market" period. Overall, the proposition that equity carveouts are designed to sell overvalued equity in the subsidiary receives mixed support.

### *5.3.4. Follow-on events*

Equity carveouts appear to be a temporary organizational form. A majority of equity carveouts are followed by a subsequent event. In Schipper and Smith (1986), two-thirds of 76 carved-out subsidiaries were later reacquired by the parent (23), divested

[9] By assuming the carveout's assets in place are sufficiently small relative to those of the parent, Nanda (1991) rules out the possibility that the parent of the carveout is also overvalued (which would result in a negative announcement effect of the carveout). Overvalued parents always prefer to issue their own shares.

entirely (17), spunoff (4), or liquidated (4). Moreover, Klein, Rosenfeld, and Beranek (1991) find that 44 of 52 carveouts (85%) are followed by a second event: 25 reacquisitions, 17 selloffs, and two spinoffs. Divestitures take place sooner than reacquisitions: three-quarters of the divestitures occur within three years of the carveout, compared to one-third of the reacquisitions. Also, the probability of a reacquisition is greater when the parent retains 80% or more of the subsidiary shares.

Klein, Rosenfeld, and Beranek (1991) argue that an equity carveout may be the first stage in a divestiture of a subsidiary. As noted above, the listing of the subsidiary's shares reduces informational asymmetries and exposes the subsidiary to the market for corporate control. Perotti and Rossetto (2007) model equity carveouts as a way for the parent to obtain information from the market on the value of the subsidiary as an independent entity. Though costly, the listing generates information about the optimal allocation of ownership of the subsidiary. Thus, the carveout improves the decision to exercise the option to sell or reacquire control, explaining the temporary nature of carveouts.

Gleason, Madura, and Pennathur (2006) document insignificant announcement returns for carveouts that are later reacquired. However, Klein, Rosenfeld, and Beranek (1991) show that parents experience significantly positive announcement returns when the follow-on event is a selloff, both at the initial equity carveout and at the subsequent divestiture. Moreover, the probability of becoming a target is higher for carved-out subsidiaries than for a sample of matched firms (Hulburt, 2003). This evidence is consistent with equity carveouts creating value by facilitating future corporate control events.

### *5.4. Agency issues*

Allen and McConnell (1998) argue that some managers avoid selling off assets because their compensation (both tangible and intangible) is tied to the size of the assets they manage. When the financing of the investments require an asset sale, the preference is to sell a minority stake in a subsidiary, maintaining assets under control. For a sample of 188 equity carveouts, they find that parent firms perform relatively poorly prior to initiating a carveout: parents have lower interest coverage ratios, higher leverage, lower operating profitability, and lower return on assets than their industry rivals. In sum, the sample parents of the carveouts were poor performers and cash constrained.

Allen and McConnell (1998) also find that the stock market's reaction to the carveout announcement is determined by the use of the proceeds. Firms announcing that the proceeds will be reinvested in the firm experience insignificant announcement returns, while the average CAR is a significant 7% for firms that will use the proceeds for debt repayment or a dividend. This suggests that the stock market may be concerned with inefficient investment decisions if the firm retains the proceeds. Schipper and Smith (1986) provide further evidence on managers' reluctance to relinquish control of the subsidiary. They document that, in a majority of cases, the president or CEO of the carved-out subsidiary is also a parent company manager.

Powers (2003) suggests that managers use their inside information about the subsidiary prospects in determining what fraction of subsidiary shares to sell to the public. He shows that the subsequent improvement in subsidiary operating performance tends to increase in the size of the ownership stake retained by the parent. Similarly, Atanasov, Boone, and Haushalter (2005) show that carved-out subsidiaries tend to have lower operating performance than their peers only when parents retain less than 50% ownership. Their interpretation is very different, however. They suggest that parent managers either self-select the carveout to avoid consolidating the subsidiary's financial results, or transfer wealth from the minority shareholders in nonconsolidated subsidiaries through inter-corporate transactions ex post.

## 6. Tracking stocks

Tracking stock—also called targeted stock or letter stock—is a separate class of parent company common stock whose dividends track the performance of a given division. That is, the holders of the tracking stock are entitled to the cash flow generated by this division, hence determining the value of the stock. The diversified company retains its legal form as one consolidated entity, however, with one and the same board of directors and top management team. There is no legal separation or transfer of assets, and the parent retains control of the division. As a result, the voting rights of the tracking stock is in the parent firm and not in the tracked division. These voting rights typically vary in proportion to the market value of the underlying division, but could also be fixed at the issue of the tracking stock.

There are several ways to distribute tracking stock. It can be issued to current shareholders as a dividend or used as payment in an acquisition. The most common way, however, is to sell the tracking stock in a public offering, raising cash for the parent firm. Once the tracking stock is listed, the underlying division files separate financial statements with the SEC. Thus, tracking stock creates a type of quasi-pure play, where the tracked division files its own financial statements and has its own stock, while still being part of the diversified firm. Since tracking stock is an issue of the company's own stock, it has no tax implications.

### *6.1. Transaction volume*

The first tracking stock was issued by General Motors (GM) in 1984 as part of the payment for Electronic Data Systems (EDS). The new stock, GM-E, allowed the selling shareholders—most notably Ross Perot, who continued in a management position—to participate in the upside of EDS, despite being part of a much larger company going forward. GM issued its second class of tracking stock, GM-H, in 1985 when acquiring Hughes Aircraft. The next company to issue tracking stock was USX, separating its steel division from its oil division (Marathon) in 1991.

In total, 32 U.S. companies have issued some 50 different tracking stocks to date, most of them in the 1990s. The market seems to have lost its appetite for tracking stock since the turn of the century. The most recent issues of tracking stock include Sprint PCS and CarMax Group in 2001, and AT&T Wireless and Disney's Go.Com in 2000. Carolina Group announced an issue in 2002 that was subsequently withdrawn. Internationally, there has been only a handful tracking stock issues, including Sony Communication Network in 2001 (Japan) and Alcatel Optronics (France) in 2000.

### *6.2. Valuation effects*

Announcements of tracking stock are received positively by the market. D'Souza and Jacob (2000) document an average abnormal two-day announcement return of 3.7% for 37 tracking stocks issued by 14 U.S. companies between 1984 and 1999. Billett and Mauer (2000), Elder and Westra (2000), Chemmanur and Paeglis (2001), and Harper and Madura (2002) also report positive tracking stock announcement ACARs of 2 to 3%. Notice, however, that, given the limited number of tracking stock issues, these studies use largely the same data.

The evidence on the long-run performance of tracking stock is inconclusive. Examining 19 firms issuing tracking stock, Chemmanur and Paeglis (2001) find that the stock of parent firms underperform industry indexes over a subsequent three-year period, while the average subsidiary outperforms its industry index. In contrast, Billett and Vijh (2004) document negative buy-and-hold returns for subsidiaries, but insignificant long-term excess returns for parents. Clayton and Qian (2004) further report insignificant long-run stock performance for tracking stock issuers. As discussed below, however, the strongest testament to a poor performance of tracking stock is the fact that they have almost entirely disappeared from the marketplace.

### *6.3. Drivers of value creation in tracking stock*

A tracking stock is akin to a "quasi-pure play." On the one hand, tracking stock allows the firm to retain its internal capital market, file a joint tax return, and share certain fixed costs and resources (Billett and Mauer, 2000; Danielova, 2008). On the other hand, the requirement to file separate financial statements with the SEC provides some degree of separation between a division and its parent. Also, the tracking stock makes it possible to give stock-based compensation to subsidiary managers.

Clayton and Qian (2004) examine whether the separate listings increase the demand for the parent and subsidiary stocks. They document an ex-date abnormal return of 3% for the parent company, suggesting that the quasi pure-play created by the tracking stock increases investor interest in the firm. However, Elder, Jain, and Kim (2005) fail to find any increase in the liquidity of the parent firm after the tracking stock issue. Instead, firms issuing tracking stock have relatively low stock-market liquidity and greater bid-ask spreads than comparable control firms. Overall, the evidence is inconclusive as to whether tracking stock increases investor demand to hold the diversified firm.

Logue, Seward, and Walsh (1996) argue that tracking stock is most useful for firms where the benefits of consolidation and integration outweigh the benefits from a complete separation. However, it is questionable whether tracking stock separates the divisions sufficiently to successfully create a pure-play stock. Not surprisingly, D'Souza and Jacob (2000) show that the returns of tracking stocks are more highly correlated with other common stocks of the same company than with other firms in the same four-digit SIC industry as the tracked division. We now turn to a discussion of the major failure of tracking stock.

### 6.4. *Agency issues*

Under U.S. corporate law, the board of directors has full discretion to transfer assets between wholly owned divisions (within contractual boundaries set by debt covenants). The assets underlying a tracking stock therefore lack legal protection from expropriation by the parent company.[10] Toward the end of the 1990s, firms issuing tracking stock started to explicitly warn investors of the risk of expropriation. For example, in its 1999 prospectus for tracking stock in its online broker, Donaldson, Lufkin, & Jenrette (DLJ) warned of a conflict of interest: "The board of directors may make decisions that favor DLJ at the expense of DLJ direct."

There are several examples of expropriation taking place. When GM in August 1995 announced its plan to spin off its tracking stock in EDS (GM-E), it first required EDS to make a one-time contribution of \$500 million to the parent (GM). EDS shareholders challenged this payment in Delaware court—and lost: the court's decision was that the board of directors has full discretion to transfer money within the corporation—tracking stock or not. Similarly, before U.S. Steel spun off the tracking stock in its oil division Marathon in 2001, it first transferred \$900 million of debt to Marathon. Not surprisingly, the stock of the steel division soared 19% on the day of this announcement.

The poor legal protection of the assets underlying a tracking stock is likely the major reason for the near-disappearance of this security. In fact, most of the tracking stocks have been reversed over the last decade. In a press release issued on December 16, 1999, Kerry Hoggard, chairman of Fletcher Challenge Ltd., said: "It is clear the the Group's capital structure is seen as complex by investors, is perceived to raise governance issues, and has resulted in a significant structural discount being applied to all our stocks. We cannot allow this to continue, and will move as quickly as possible to a full dismantling of the target share structure."

Billett and Vijh (2004) examine 11 announcements to remove the tracking stock structure. They find significant and positive excess stock returns of 14% to the dismantling announcement. Tracking stock in its current form may very well be a phenomenon of the past.

[10] Hass (1996) provides an in-depth discussion of the fiduciary duties of the company's directors as they relate to tracking stock.

## 7. Leveraged recapitalizations

A leveraged recapitalization (henceforth "recap") is a significant payout to shareholders financed by new debt borrowed against the firm's future cash flow. The company remains publicly traded, but with a substantially higher debt level. For a sample of 27 firms completing leveraged recaps over the period 1984–1988, Gupta and Rosenthal (1991) find a threefold increase in the average debt-to-total-capital ratio, from 22 of to 67%. Denis and Denis (1983) document that the median ratio of total debt to total assets increases from 45% to 86% for a sample of 39 recaps in 1984–1988. Moreover, studying 42 leveraged recaps between 1985 and 1989, Handa and Radhakrishnan (1991) report that the proposed payout averages 60% of the pre-recap market value of equity.

The cash distribution to shareholders is typically structured as a large, special, one-time dividend. Alternatively, the distribution could be in the form of a share repurchase or exchange offer. Management often forfeist the cash distribution on their shareholdings and instead takes additional stock. Consequently, leveraged recaps typically result in a substantial increase in managerial equity ownership. Handa and Radhakrishnan (1991) document that insider equity ownership increases by three times, while Gupta and Rosenthal (1991) report a doubling of the insider ownership (from 3.8 to 8.4%). In Denis and Denis (1993), the median ownership of officers, directors, and employees soars from 6 to 15%.

Prior to the widespread use of poison pills, leveraged recaps were sometimes used as a defense against a hostile takeover threat. See Denis (1990) for an analysis of leveraged recapitalizations as a takeover defense.

A leveraged recapitalization triggers a tax liability at the investor level. The tax depends on how the payout to shareholders is structured. For a special dividend, the amount distributed from the firm's retained earnings is taxed as a dividend. If the special dividend exceeds the retained earnings on the firm's balance sheet, the remaining cash distribution is a return of capital, treated as a capital gain. If the recap is structured as a share repurchase, the entire distribution is taxed as a capital gain.

The financial accounting for leveraged recapitalizations does not require any step-up of the company's assets. As a result, if the new debt exceeds the book value of the firm's equity, the company's book equity becomes negative following the recap. What appears like a leveraged buyout by a private equity sponsor is sometimes structured as a recap. Recap accounting can be used if the buyer acquires less than 94.9% of the firm's stock, and the owners of the minority interest, which must be widely held, are independent from the buyer.

### *7.1. Transaction volume*

Figure 4 shows the annual volume of leveraged recapitalization transactions announcements from 1985 through 2007, using data from SDC. The recap volume has largely followed the ups and downs of the high-yield debt markets. As shown in Panel A, in the United States, leveraged recaps were particularly popular in the late 1980s, with a

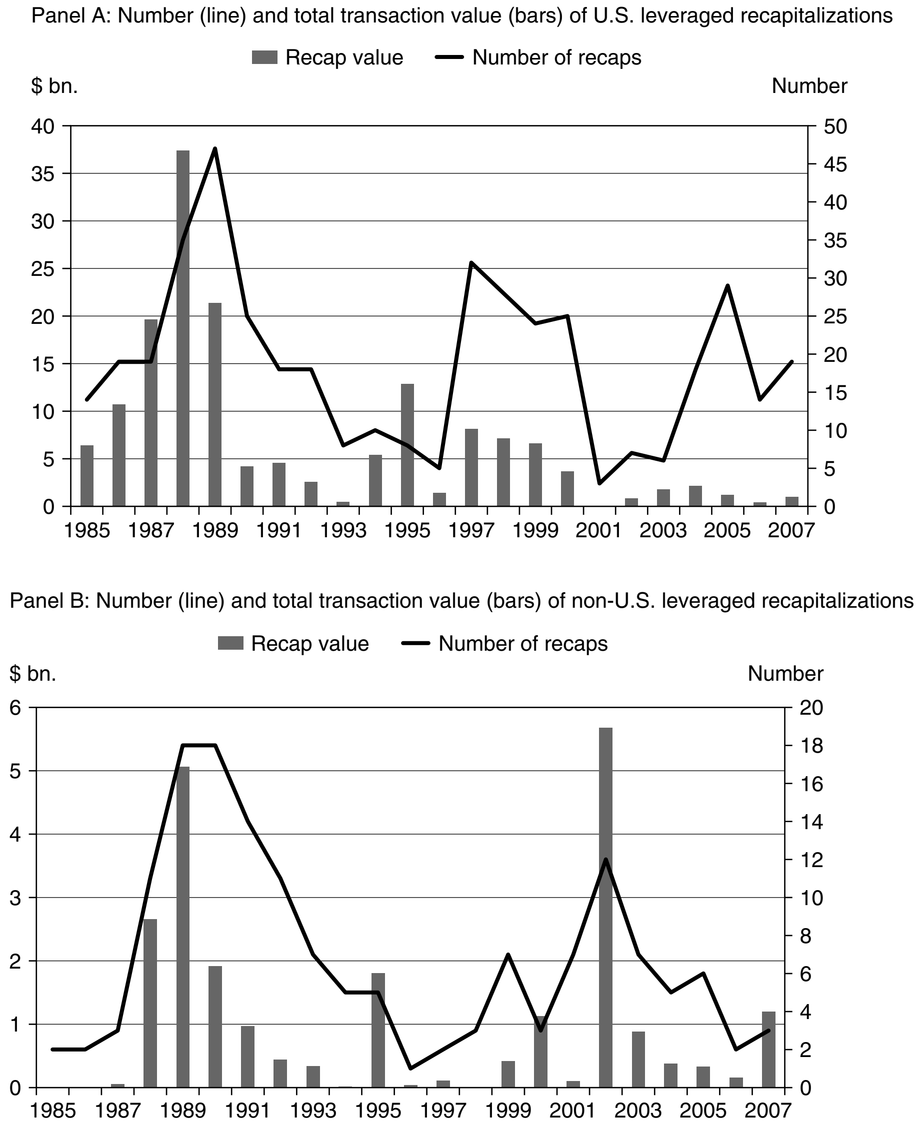



Fig. 4. Annual volume of leveraged recapitalizations, 1985–2007.

*Source*: SDC

peak in combined transaction value (bars) of $37 billion in 1988 and 47 recaps (line) in 1989. There was a smaller surge in recapitalization transactions in the period 1997–2000, and then again in 2005, however, without a corresponding increase in transaction size. Panel B shows the non-U.S. volume of leveraged recapitalizations. The international recap activity is generally lower and involves smaller amounts. Companies announcing leveraged recapitalizations in 2006–2007 include Charter Communications Inc, Palm Inc., Foster Wheeler Ltd., and Acadia Realty Trust.

### *7.2. Valuation effects*

The wealth effects of leveraged recapitalizations are substantial. For a sample of 44 recaps over 1985–1990, Bae and Simet (1998) find a two-day shareholder ACAR of 5.7%. Moreover, Handa and Radhakrishnan (1991) report an average two-day abnormal return of 5.5%, and Gupta and Rosenthal (1991) find an average announcement CAR of 5.9%. Moreover, Balachandran, Faff, and Nguyen (2004) document a three-day average CAR of 4.4% for a sample of 167 leveraged recapitalizations in Australia between 1989 and 2002.

Since the leveraged recapitalization may be a response to a corporate control threat, several studies measure the returns over a longer event window. Denis and Denis (1993) use a window starting 40 days prior to initiation, defined as the first indication of a takeover or the announcement of the recap, through completion of the recap. They estimate an average abnormal return of 32% (median 26%). Kaplan and Stein (1990) compute the cumulative abnormal stock return starting 40 days prior to the recap announcement, or the day of a hostile bid if there is one, through the recap completion. They find an average CAR of 45% (median 47%) for 12 leveraged recapitalizations between 1985 and 1988.

Kaplan and Stein (1990) further estimate the change in systematic risk of the firm's securities after the leveraged recap. The increase in the equity risk is relatively modest. Using daily returns and market-model estimates, the average equity beta increases by 37% from 1.01 to 1.38 after the recapitalization. They then make two different assumptions about the change in total asset risk from the transaction. Assuming that the systematic risk of the assets (asset beta) is constant, the implied debt beta averages 0.65. However, when they assume that the entire market-adjusted premium represents a reduction in fixed costs, the implied debt beta averages 0.40. Overall, leveraged recapitalizations generate substantial shareholder wealth and appears to be associated with a surprisingly small increase in equity systematic risk.

### *7.3. Drivers of value creation in leveraged recapitalizations*

As discussed earlier, the high debt in leveraged recapitalizations reduces the firm's free cash flow and hence managerial discretion over the investment decisions (Jensen, 1986). Denis and Denis (1993) examine the change in operating performance and investments for 29 completed recapitalizations between 1984 and 1988. They document large decreases in the undistributed cash flow (median −31%) and capital expenditures (median −35%), despite improvements in operating performance (median 21%) from the year prior to the year after the recap. Also, the post-recap cash flow covers only two-thirds of the pre-recap capital expenditures, forcing a reduction in the level of investments.

They further examine the market reaction for capital expenditure announcements and find a significantly negative ACAR over the five-year period prior to the recapitalization, suggesting a past pattern of overinvestment. Following the recap, the average number of announced investments drops from 1.2 to 0.3 per firm and year, with an average stock market reaction that is insignificantly different from zero. They conclude that the increased debt plays a central role in disciplining managers' investment decisions.

Consistent with these results, Wruck (1994) documents organizational and compensation changes in Sealed Air following its leveraged recapitalization in 1989. She suggests that the financial leverage was used as a tool to improve the internal control systems, which together with the high debt service created an environment that led to enormous performance improvements and value creation.

Peyer and Shivdasani (2001) study the efficiency of the internal allocation of investments after leveraged recapitalizations in 22 multidivisional firms between 1982 and 1994. Prior to the recap, companies allocate investments to high $q$ divisions. Following the recap, however, investments become less sensitive to division $q$ and more sensitive to division cash flow. While this may indicate that the internal allocation of capital becomes less efficient, the total level of capital expenditure declines, as do the firm's diversification discount. Peyer and Shivdasani conclude that the costs of distorted divisional investments are outweighed by the benefits of lower firm-level investments. Overall, leveraged recapitalizations appear to create value by curbing managerial overinvestment and improving operating performance.

Walker (1998) suggests that the benefits from leveraged recapitalizations are transitory, examining 39 recaps between 1985 and 1989. He finds that the recap firms have higher free cash flow prior to the recap than matching firms. However, the pre-recap level of capital expenditures is not significantly different from that of its peers. Moreover, operating performance increases from year $-1$ to $+1$ relative to the special dividend, but reverts in the subsequent years.

A leveraged recapitalization could be used to signal management's private information about the future cash flow of the firm. Healy and Palepu (1995) describe how managers at CUC International successfully undertake a leveraged recap in 1989 to communicate their optimistic beliefs about the firm's future cash flows to investors. Balachandran, Faff, and Nguyen (2004) examine if the positive information conveyed by a recap extends to other firms in the industry. They find insignificant stock returns for competitors of firms announcing a leveraged recapitalization, suggesting that the content of any new information is unique to the recap firm.

A large fraction of the leveraged recapitalizations in the late 1980s subsequently failed. Denis and Denis (1995) report that 9 (one-third) of 27 firms completing a leveraged recap between 1985 and 1988 became financially distressed. They find that the poor operating performance of the nine distressed firms is in line with that of their industry peers. Moreover, the stock market reacts negatively to announcements of asset sales as well as to economic and regulatory events associated with the demise of the high-yield market. They conclude that the incidence of distress is not related to poorly structured transactions, but rather to unexpected macroeconomic and regulatory developments.

## 8. Leveraged buyouts (LBO)

A leveraged buyout is the acquisition and delisting of an entire company or a division, financed primarily with debt. The buyer is typically a private equity fund managed by an LBO sponsor—or recently sometimes a consortium of funds. The sponsor raises debt to finance the majority of the purchase price and contributes an equity investment from the fund. The equity is injected into a shell company, which simultaneously borrows the debt and acquires the target.

The sponsor relies on the company's cash flow, often supplemented by assets sales, to service the debt. The objective is to improve operating efficiency and grow revenue for a 3–5 year period before divesting the firm. Debt is paid down over time and all excess returns accrue to the equity holders. The exit may be in the form of an IPO, a sale to a strategic buyer, or a sale to another LBO fund. While an IPO typically generates a higher valuation, it could take several years for the LBO fund to entirely unwind its holdings through the public markets.

Because of the heavy debt load, a target firm is traditionally characterized by a strong predictable cash flow, supported by a history of profitability. In addition, it is often in a mature industry, with low growth and limited need for additional capital expenditures. The industry scope of leveraged buyouts has increased over time, as has the importance of international deals. Also, while the conventional LBO involves a publicly traded target company, a majority of the leverage buyouts are of private firms.

A management buyout (MBO) is a leveraged buyout of a segment, a division or a subsidiary of a large corporation in which key corporate executives play a critical role. MBOs are generally smaller than traditional LBOs and, depending on the size of the transaction, a sponsor need not be involved. In the following, MBOs are singled out only if this term is explicitly used to characterize a sample.

### *8.1. Transaction volume*

The leverage buyout activity varies considerably over time. Figure 5 shows the annual number (line) and total deal value (bars) of LBOs announced between 1985 and 2007, using data from SDC. As shown in Panel A, a first surge in U.S. LBO activity occurred in the late 1980s. This is when landmark transactions such as KKR's buyouts of RJR Nabisco (worth \$25 billion) and Safeway (\$4 billion) took place. The economic recession in 1990–1991, combined with regulatory restrictions on investments in high-yield instruments, the bankruptcy of Drexel Burnham Lambert, and a reduction in new lending by commercial banks, put an abrupt end to this first wave of highly leveraged transactions.

Most of the transactions in the 1990s were LBOs of private companies and divisions. As the availability of debt financing soared in the mid-2000s, the public-to-private transaction reappeared in a second buyout boom. The total value of U.S. LBO transactions announced in 2006 and 2007 amounts to \$450 and \$410 billion, respectively. Recent large U.S. buyouts include Equity Office Properties (\$41 billion), HCA (\$33 billion),

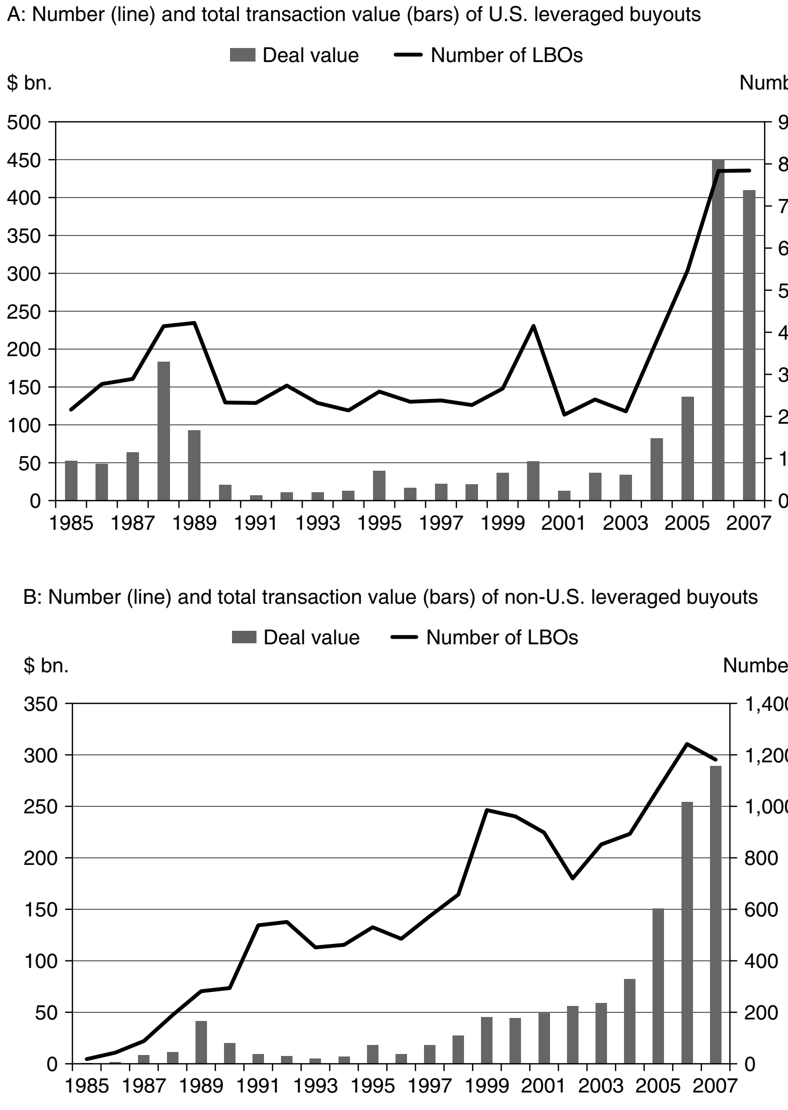



Fig. 5. Annual volume of leveraged buyouts, 1985–2007.

*Source*: SDC

TXU ($32 billion), Harrah's Entertainment ($28 billion), Clear Channel Communications ($27 billion), First Data ($26 billion), SLM ($26 billion), Kinder Morgan ($22 billion), and Hilton Hotels ($20 billion), to mention a few.

Panel B shows the corresponding LBO volume outside the United States. The number of non-U.S. buyouts has grown steadily since the mid-1980s, with a short dip in transaction volume in 2002 after the burst of the Internet bubble. The international LBO volume reached a record high in 2007 with a total deal value of $289 billion across almost 1,200 transactions. Large buyouts outside the United States announced in 2006–2007 include BCE, Canada ($51 billion); Alliance Boots, United Kingdom ($22 billion); BAA, Spain ($22 billion); Altadis, Spain ($18 billion); Thames Water, United Kingdom ($15 billion); and Vodafone KK, Japan ($14 billion).

Stromberg (2007) estimates the value of firms acquired in leveraged buyouts between 1970 and 2007 as a total of $3.6 trillion, three-quarters of which represent LBOs undertaken after 2000. This second wave of large LBOs has spurred a renewed interest in leveraged buyouts in academic research. Since the financing market turmoil began in mid-2007, however, only a limited number of large buyouts have been announced in the United States and internationally.

### *8.2. The LBO capital structure*

An LBO is financed with a mix of bank loans, high-yield debt, mezzanine debt, and private equity. The bank debt, which is often syndicated in the leveraged loan market, is secured and most senior in the capital structure. The interest rate is floating, generally quoted as a spread above the London Interbank Offering Rate (LIBOR). While the maturity varies with the firm's credit profile, it is commonly in the range of 5-8 years and always shorter than that of junior debt. The bank debt has to be amortized before any other claimholders are paid off. At times (but not in 2006/2007), cash sweeps are common, requiring the firm to use any excess cash flow for accelerated amortization of the bank loans.

The bank debt is typically structured as several tranches of term loans (A, B, C, and D), where the holder of Tranche A also provides a revolving credit facility. Term A, the pro-rata facility, is sold to traditional banks and is senior to the other tranches. In the second LBO wave, branches B, C, and D had minimal front-end amortization and were primarily sold to institutions and funds. The proportion leveraged bank loans in the capital structure varies, but was around 40% for U.S. buyouts closed in 2006–2007.

The remaining debt is raised from the subordinated debt markets. High-yield debt (junk bonds) is generally subordinated and/or unsecured. Interest is fixed, based on a spread to treasury bonds that varies with credit quality, and expressed as a coupon. This debt has a bullet maturity in 10 years and is as a rule callable at a premium. The high-yield bonds are typically sold to the public in a 144A offering, which requires a road show and hence takes time to close. It is therefore common practice to finance the high-yield portion through a bridge loan at deal closing, repaid within a year with the proceeds from the subsequent bond issue.

As an alternative to high-yield debt, which is publicly traded, the market for second lien loans took off in 2003. These loans are privately placed with hedge funds and Collateralized Loan Obligation (CLO) investors, and are secured in the firm's assets but subordinated to the bank loans. CLOs combine a large number of leveraged loans (first and second lien) into a pool, which itself is sliced in tranches sold to institutional

The debt multiple is the average ratio of the pro-forma total debt to adjusted EBITDA. The source is Standard & Poor's LCD.

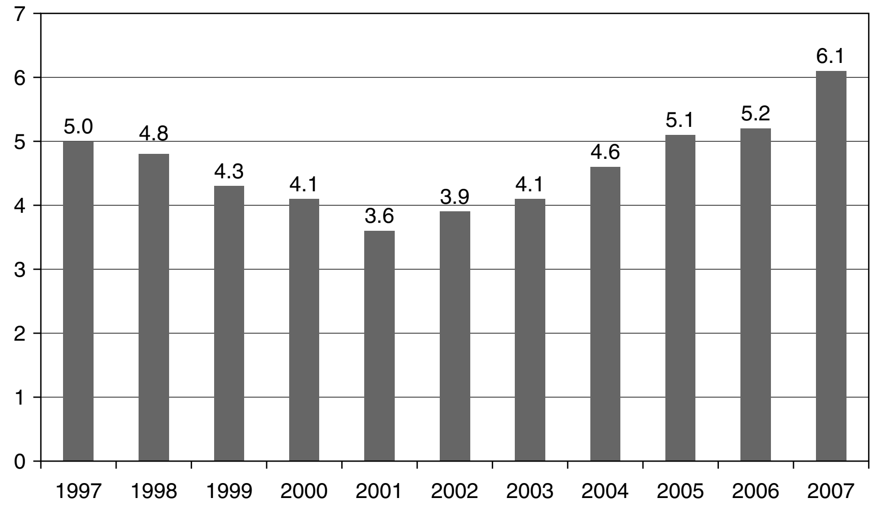



Fig. 6. Annual pro-forma debt multiples in LBOs, 1997–2007.

investors. In 2007, the total volume of second lien loans reached $30 billion (*Source*: Standard & Poor's LCD).

Figure 6 shows annual debt multiples, defined as the pro-forma ratio of total debt to adjusted EBITDA, in LBO transactions between 1997 and 2007. Debt multiples reached a low in 2001, when the average transaction raised 3.6 times EBITDA in the debt markets. The expansion of the debt markets and aggressive lending practices in 2007 are reflected in a much higher average debt multiple of 6.1 times EBITDA. After the financial market turmoil in mid-2007, however, credit markets are constrained, and debt multiples are considerably lower again.

In periods when access to high-yield debt and bank loans is limited, sponsors resort to mezzanine financing, which replaces or is subordinated to the high-yield bonds. It is sold in a private placement to funds and institutions, thus avoiding any public filing requirements. The mezzanine is a committed financing with individually negotiated terms. It is structured as a debt contract or preferred equity, with warrants and other "equity kickers" attached to increase its total returns. All or part of the interest expense or dividend is often in the form of additional securities rather than cash, so-called pay-in-kind (PIK). The use of mezzanine financing is more widespread in Europe, where the leveraged loan markets and high-yield bond markets lag those of the United States.

Private equity is the most junior in the capital structure. It typically has voting rights but no dividends. This equity is raised from pension funds, endowments, insurance companies, and wealthy individuals into a fund managed by a private equity partnership (the sponsor). Prominent LBO sponsors include Blackstone, Carlyle, and KKR. Most sponsors are paid a management fee of 2% on the fund's capital and receive a carried interest of 20% of the profits realized by the fund. In addition, some sponsors charge deal fees and monitoring fees to their portfolio companies. See Metrick and Yasuda (2007) for a detailed description and analysis of the fee structure in LBO funds. The capital raised for private equity is setting new record levels. In 2006, private equity funds had an inflow of $225 billion in new capital.

Panel A of Figure 7 shows the average equity contribution in LBOs from 1987 through 2007. The source is Portfolio Management Data. The deals in the end of the 1980s were extremely highly leveraged, with an average equity portion of 8–13% of the total capital. Over the last decade, most LBO transactions have had a substantially higher fraction of equity financing, with equity constituting on average one-third of the capital structure in recent years. Managers are generally required to co-invest in the buyout equity along with the LBO fund. If a manager has been involved in a prior buyout, she is asked to roll over a portion of her equity in the target firm. If it is a first-time LBO, managers may be offered to buy equity at a discount, or receive additional stock and options conditional on certain performance goals.

Panel B of Figure 7 shows the average price multiple in LBOs, defined as the ratio of the purchase price to the adjusted EBITDA, for the period 1997–2007. Average prices have risen from a low average multiple of 6.4 in 2001 to a high of 9.8 in 2007. The total funds raised in the buyout transaction are used for consideration to the seller as well as underwriter fees for the LBO debt (usually 1.5 to 2.5% of the principal amount) and call premiums on existing bonds.

Axelson, Jenkinson, Stromberg, and Weisbach (2007) document the financial structure of 153 large U.S. and European buyouts between 1985 and 2006. They find that the leverage of LBO firms is unrelated to debt levels of size- and industry-matched public firms. Instead, the leverage decreases in the interest rates prevailing at the time of the buyout. Prices also decline in interest rates, but are positively related to price multiples in public markets. They conclude that LBO capital structures are largely driven by the economywide cost of borrowing rather than firm-specific factors. See also Roden and Lewellen (1995) for an analysis of the structure of the LBO financing package.

### *8.3. Value creation in LBOs*

The total value created in a leveraged buyout is divided between the selling shareholders and the LBO investors. Table 4 shows the premiums paid in 1058 leveraged buyout transactions between 1973 and 2006 as reported by seven selected studies. The premium is defined as the final offer price in excess of the target stock price 20 to 60 days prior to the announcement of the bid. As shown in the table, the average premium ranges from 27 to 59% across the seven studies, with a sample-size-weighted average of 37%. The median premium ranges from 27 to 42%, with an average of 32%. It appears that premiums are generally somewhat lower in the 2000s compared to the 1980s. The exception is the

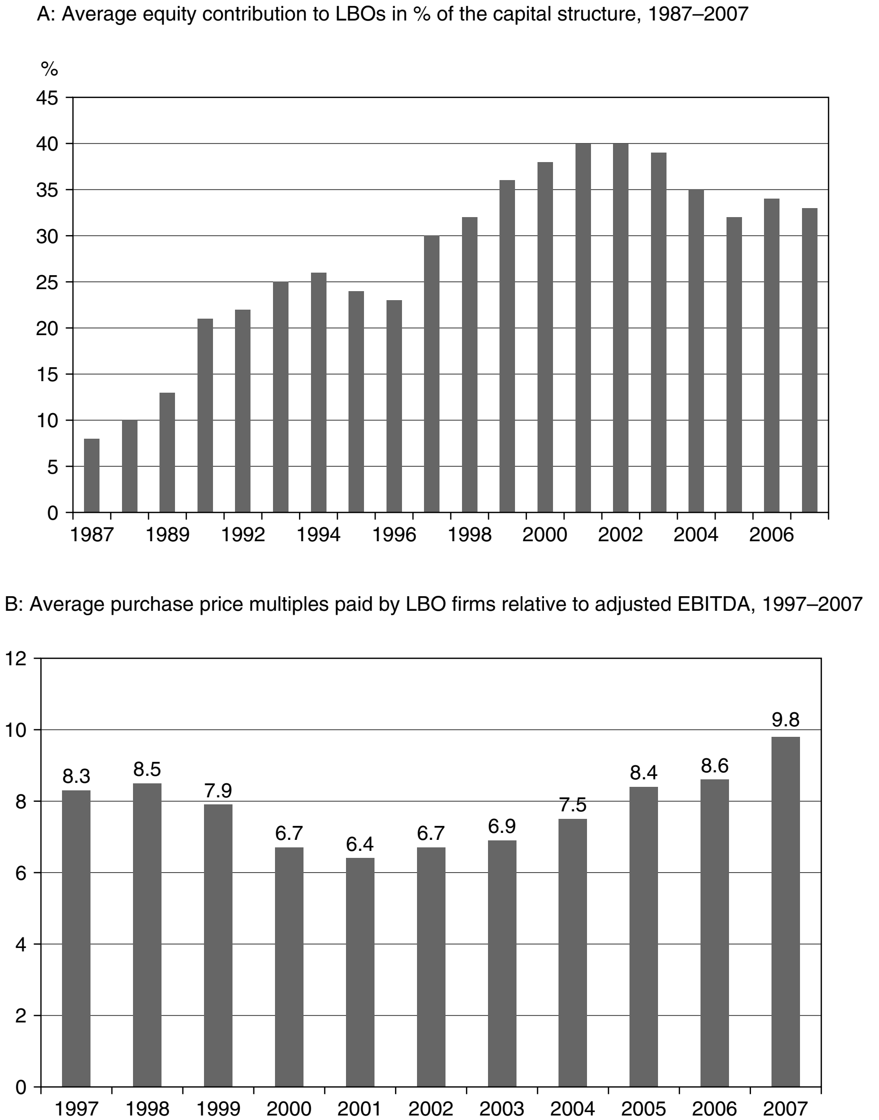



Fig. 7. Average annual % equity contribution and purchase multiples in LBOs.
*Source*: Portfolio Management Data.

study by Renneboog, Simons, and Wright (2007) of 177 buyouts in the United Kingdom between 1997 and 2003. They document an average premium of 40% (median 38%), which is higher than the contemporaneous LBO premiums of 27 to 29% in the United States (Billet, Jiang, and Lie, 2008; Guo, Hotchkiss, and Song, 2008).

Table 4

Premiums paid in 1,058 leverage buyouts for 7 selected studies, 1973–2006

The premium is the ratio between the final offer price and the pre-buyout stock price less one.

| Study | Premium | | Type of deal | Sample size | Time period | Day of pre-buyout stock price |
|---|---|---|---|---|---|---|
| | Mean | Median | | | | |
| DeAngelo, DeAngelo, and Rice (1984) | 59% | | LBO | 23 | 1973–1980 | −40 |
| Lehn and Poulsen (1989) | 36% | | LBO | 257 | 1980–1987 | −20 |
| Kaplan (1989b) | 46% | 42% | MBO | 76 | 1980–1985 | −40 |
| Harlow and Howe (1993) | 45% | | LBO | 121 | 1980–1989 | −20 |
| Renneboog, Simons, and Wright (2007) | 40% | 38% | LBO | 177 | 1997–2003 | −40 |
| Billet, Jiang, and Lie (2008) | 27% | 27% | LBO | 212 | 1990–2006 | −60 |
| Guo, Hotchkiss, and Song (2008) | | 29% | LBO | 192 | 1996–2006 | −20 |
| Sample size weighted average | 37% | 32% | | 1,058 | 1973–2006 | |

Several studies find two-day average CARs of 16 to 17% for LBO announcements in the 1980s (DeAngelo, DeAngelo, and Rice, 1984; Lehn and Poulsen, 1989; Slovin, Sushka, and Bendeck, 1991; Van de Gucht and Moore, 1998). For a sample of 641 LBOs in 1980–2001, Brown, Fee, and Thomas (2007) estimate an average announcement CAR of 19%. The announcement return reflects a combination of the market's estimate of the target gains from a deal and the likelihood that the deal succeeds. Overall, the target shareholders tend to make substantial gains in leveraged buyouts.

The second part of the equation is the returns realized by the LBO investors. These returns have been difficult to estimate since the buyout targets are taken private and often do not return to public ownership. Kaplan (1989a) estimates a median market-adjusted return of 28% (mean 42%) for investors in 25 MBOs that went public after on average 2.7 years. Muscarella and Versuypens (1990) examine the equity returns for 58 LBO firms that returned to public status after on average 2.9 years. Comparing the IPO price with the LBO price, they estimate an average annualized rate of raw return of 268%. This return is, however, not significantly different from the return of a hypothetical levered portfolio of S&P500 firms.

More recently, LBO fund return data has been available through self-reporting to Venture Economics. Jones and Rhodes-Kropf (2004) examine quarterly returns of 379 LBO funds formed between 1980 and 1999, using sponsor estimates of value changes. Estimating fund performance measured against the Fama and French (1993) three-factor model, they find fund excess returns (alphas) that are insignificantly different from zero.

Ljunqvist and Richarson (2003) use proprietary data from a large institutional investor. They study the returns to investments for 54 U.S. LBO funds raised between 1981 and 1993, observing the actual cash inflows and outflows of the funds. They find that the LBO funds outperform the stock market and have positive alphas. On a risk-adjusted basis, the excess return of the typical LBO fund is 5% annually. Groh and Gottschalg (2008) also find positive excess returns in a sample of 133 U.S. buyouts over 1984–2004.

Their benchmark is a portfolio of public market equivalents matched by systematic risk and timing, and they correct for self-selection.

Kaplan and Schoar (2005) study the returns net of costs for 169 LBO funds raised between 1980 and 2001. They estimate that the median fund underperforms the stock market index, generating 80% (mean 97%) of the return on the S&P500. However, for the subset of sponsors that have been around for at least five years, the median performance exceeds the S&P500 by 50% (mean 80%). They show that this performance is persistent, and they suggest that LBO sponsors may have different skills in managing portfolio companies. Phallipou and Gottschalg (2007) examine 739 LBO funds raised in 1980–1993, assuming that unrealized assets have zero value. They find that the funds net of cost underperform the S&P500 by 3% on average and confirm the persistence in LBO fund returns.

For a sample of 50 large U.K. buyouts over the 1997–2004 period, Acharya and Kehoe (2008) report positive industry-adjusted returns over the life of the deal. The returns are positively correlated to improvements in operating margins and asset growth. In contrast, Nikoskelainen and Wright (2007) find a median return to enterprise value of −5% (average 22%), adjusted for the return of the FTSE 100 index on the London Stock Exchange. They examine a sample of 321 exited buyouts in the United Kingdom between 1995 and 2004.

Overall, the total gains from LBOs are large, manifested in the substantial premiums paid to target shareholders. However, the evidence is inconclusive as to whether selling shareholders largely capture all the gains in leveraged buyouts. Depending on the sample, the benchmark portfolio, and assumptions about the value of assets that are not liquidated, the estimates of LBO fund abnormal returns range from positive to negative.

### *8.4. Drivers of value creation in LBOs*

#### *8.4.1. Operating efficiency*

As argued by Jensen (1986), the high leverage in buyouts may result in improved managerial investment decisions for firms with high cash flow and few growth opportunities. Lehn and Poulsen (1989) examine 263 LBOs in the 1980s. They find some evidence that firms with high levels of free cash flow are more likely to go private and that acquisition premiums increase with the target firm's cash flow. They conclude that the mitigation of agency problems associated with free cash flow are a major source of buyout gains.

Opler and Titman (1983) find that LBO targets have a combination of high cash flow and unfavorable investment opportunities (low $q$), and are more diversified than firms that don't become targets. In addition, buyouts are less likely for firms with high expenditures for research and development (R&D). Similarly, Long and Ravenscraft (1993) show that LBOs typically target firms with R&D expenditures below the industry average. Also, Bae and Simet (1998) find that LBO announcement returns are increasing in the free cash flow of the target firm. In contrast, Servaes (1994) finds no significant difference in the capital expenditure level between target firms in 99 going private transactions and their industry peers. Overall, however, the evidence suggests that the potential for incentive realignment in firms with high levels of free cash flow represents an important factor in the leveraged buyout decision.

If leverage successfully curbs overinvestment, this should show in the post-buyout operating performance. Kaplan (1989a) examines the performance of 48 large management buyouts between 1980 and 1986. He shows that the firms experience substantial increases in operating income (+42%), reductions in capital expenditure, and improvements of the net cash flow (+96%) over a three-year period following the buyout. Smith (1990) also reports significant performance improvements for 58 management buyouts in 1977–1986. She finds that operating returns, measured as operating cash flow per employee and per dollar of operating assets, increase significantly from the year prior to the year after the buyout. She examines changes in accounting line items and finds no evidence that repair and maintenance expenditures are postponed or that the R&D expenditures are reduced. Instead, the higher margins are a result of adjustments in the management of working capital.

Several other studies document improved operating efficiency after buyouts. Lichtenberg and Siegel (1990) examine data from the Longitudinal Business Database (LBD) of the U.S. Bureau of the Census for 131 LBOs in the period 1981–1986, with a total of 1132 plants. They show that plant total factor productivity (TFP) increases more than the industry average in the years following a leveraged buyout. Consistent with this finding, Harris, Siegel, and Wright (2005) find an above-industry increase in TFP for U.K. MBO plants in the 1990s. Moreover, Muscarella and Versuypens (1990) examine the performance of 72 LBO firms that went public again. They show that LBO firms reduce operating costs and experience significant improvements in their operating margins. Also, while there is a dramatic increase in leverage upon completion of the LBO, the debt ratios are gradually reduced before returning to public ownership.

The evidence of improvements in operating performance is weaker for more recent transactions. Guo, Hotchkiss, and Song (2008) examine 94 U.S. public-to-private LBOs between 1990 and 2005. They find that post-buyout gains in operating performance are comparable to or slightly exceed benchmark firms matched on industry and pre-buyout characteristics. The cash flow improvements are greater for firms with higher increases in leverage and when the CEO is replaced in the buyout transaction. Moreover, the median returns to LBO investors are 25% (average 57%) adjusted for Fama-French industry portfolio returns. Interestingly, the cash flow improvements and returns to capital are strongly related. However, due to the small magnitude of the cash flow gains, they suggest that recent transactions are not largely motivated by improving the operating efficiency of underperforming firms.

There is a concern that the trimmed organization and reduced capital expenditure may hurt the long-term prospects of LBO firms. Lerner, Sorensen, and Stromberg (2008) study a sample of 495 LBO firms that filed at least one successful patent application in the period 1986–2005. They show that firms continue to pursue high-impact patents after going private, concentrating their innovations in areas of historical core strengths. They conclude that leveraged buyouts promote a beneficial refocusing of the firm's patent portfolios.

Overall, the results suggest that leveraged buyouts target firms with free cash flow, where the leverage could help improve investment decisions by reducing managers' discretionary funds. There is convincing evidence of post-buyout improvements in operating performance and plant productivity. Also, while total capital expenditures decline, critical investments in R&D seem to continue.

### *8.4.2. Employment*

It appears that improvements in operating efficiency are associated with employee layoffs. Kaplan (1989a) finds that the median firm reduces its employee count by 12% relative to the industry from the year prior to the year after the buyout. Muscarella and Versuypens (1990) show that the average employment declines by 0.6% for LBO firms that subsequently went public. This job creation is in the bottom 10% of COMPUSTAT firms. Lichtenberg and Siegel (1990) report that white-collar compensation and employment decline in the years following the buyout. Moreover, for a sample of 33 LBOs in 1980–1984, Liebeskind, Wiersema, and Hansen (1992) report that LBO firms downsize the operations more than comparable firms in terms of number of employees, plants, and total revenues. In addition, there is some evidence that buyouts in the United Kingdom lead to modest declines in employment (Wright, Thompson, and Robbie, 1992; Amess and Wright, (2007).[11]

More recent evidence, however, suggests that the decline in LBO employment in existing facilities is outweighed by additional employment in new establishments, defined as new plants, offices, and retail outlets. Davis, Haltiwanger, Jarmin, Lerner, and Miranda (2008) examine LBD data for 5000 U.S. targets acquired in private equity transactions between 1980 and 1995. Consistent with previous work, they find that employment drops more in target establishments than in control firms following the buyout. However, the LBO firms create substantially more jobs in new establishments than their peers. They conclude that the private equity sponsors push the target firm to expand in new, higher-value directions. Overall, while LBO firms appear to trim their workforce to improve efficiency in existing production facilities, they also create additional job opportunities through new establishments.

### *8.4.3. Corporate governance*

Highly leveraged transactions lead to increased monitoring by banks and the LBO sponsor (who has its own money at stake). Jensen (1989) argues that the combination of active governance by buyout sponsors, high-powered managerial incentives, and

[11] Perotti and Spier (1993) present a strategic model of temporarily high leverage. They show how shareholders, by retiring equity through a junior debt issue, can credibly threaten to underinvest in valuable new projects unless employees concede to wage reductions.

pressures from high leverage, provide a corporate governance system and incentive structure that is superior to that of widely held public firms. He predicts that the LBO organization eventually will eclipse the traditional, widely held public companies to become the dominant organizational form. While this has not yet happened, there is little doubt that the LBO organization carries with it a relatively efficient governance structure.

A central governance characteristic of leveraged buyouts is a meaningful management equity participation. Kaplan (1989a) shows that the median equity ownership of the top management team increases from 6 to 23% for 76 MBOs in the 1980s. Moreover, Muscarella and Versuypens (1990) report that the most highly paid officer owns 18% of the LBO firm's equity prior to an IPO exit.

The equity ownership of the top management team is also substantial in more recent samples. Kaplan and Stromberg (2008) study 45 LBOs from 1996 to 2004. They find a median equity ownership of 6% for the CEO and 16% for the management team. Nikoskelainen and Wright (2007) report an average equity ownership of 37% (median 35%) for 321 U.K. buyouts over the 1995–2004 period. Acharya and Kehoe (2008) examine a sample of 59 large buyouts in the United Kingdom between 1997 and 2004. They document an equity ownership including options of 3% for the CEO and 13% for the top management team as a whole. In sum, leveraged buyouts provide significant equity-based incentives to top management that help align managerial incentives with shareholders' interests.

Furthermore, the concentration of ownership provides LBO sponsors with a strong incentive to monitor the firm closely. Baker and Wruck (1989) describe the organizational changes at O.M. Scott after its leveraged buyout in 1986. The board had five members, of which one was a manager and three represented the buyout sponsor. All board members owned stock. The board met quarterly, and an executive committee monthly. More importantly, one of the private equity partners served as a liaison between the LBO sponsor and the firm's managers. The operating partner, which functioned as an advisor and a consultant, spent several weeks at O.M. Scott after the buyout closed and was thereafter in telephone contact with the CEO daily. Baker and Wruck (1989) conclude that the close monitoring by the LBO sponsor, combined with the restrictions imposed by the high leverage and significant managerial shareholdings and bonus plans, led to a substantial improvement in O.M. Scott's operating performance and investment policies.[12]

The evidence suggests that LBO sponsors are also active monitors in more recent transactions. Cornelli and Karakas (2008) examine the board structure for 88 U.K. buyouts sponsored by a private equity firm over the 1998–2003 period. They find that, on average, the board size decreases by 15%, from 6.5 to 5.5 directors after the buyout. Moreover, outside directors are replaced by individuals representing the LBO sponsor, who controls on average 40% of the board seats. Also, the CEO is replaced in half of the buyout transactions.

[12] See also Denis (1994) for an analysis of the organizational changes at Safeway after its leveraged buyout in 1986.

Acharya and Kehoe (2008) show that LBO sponsors on average own 77% of the equity in their portfolio companies. The average sponsor holds 45% of the seats on a board with eight members that meet monthly. The sponsor engages through weekly, often informal, meetings with management over the due diligence phase and the first three months after closing. They also report that two-thirds of the LBO firm's top management is replaced within 100 days of the deal. In sum, buyout sponsors play an important role through active monitoring of the LBO firm.

Cressy, Munari, and Malipiero (2007) compare the operating performance of private equity-backed LBOs with that of comparable nonbuyout private firms matched on industry and size. Their sample is 122 U.K. buyouts in 1995–2002. They find a higher post-buyout operating profitability for the LBO firms, and particularly when the sponsor specializes in the target firm industry.

While the monitoring by LBO sponsors is an important governance mechanism in leveraged buyouts, managers sometime undertake MBOs without the involvement of a private equity sponsor. Fidrmuc, Roosenboom, and van Dijk (2008) examine the choice between an MBO and a sponsor-backed buyout across 129 U.K. leveraged buyouts in 1997–2003 and where management stayed in control. They find that MBO targets have lower market-to-book ratios, more cash on hand, and greater managerial ownership. They suggest that managers invite LBO sponsors when they need help to complete a deal, and they conclude that MBOs and sponsor-backed LBOs are complementary transactions.

Cotter and Peck (2001) analyze how the equity ownership of the LBO firm interacts with the structure of the buyout debt. Their sample is 64 LBO firms in 1984–1989, of which a buyout specialist owns majority control in 40 firms (63%). They find that firms controlled by an LBO sponsor use less short-term and/or senior bank debt to finance the transaction. Moreover, the LBO firm's operating performance increases with the use of senior debt only in deals where no buyout specialist is involved. They suggest that bank debt, having more restrictive covenants, and debt with shorter maturity, and thus higher debt service, both help motivate and monitor management in the absence of an active buyout specialist. See also Grinstein (2006) for an analysis of how the debt structure is used to commit investors to disciplinary actions against management.

In sum, leveraged buyouts are characterized by powerful corporate governance structures. First, management owns a substantial portion of the equity. Second, the ownership is concentrated with an LBO sponsor who actively monitors management. Third, the high leverage puts additional pressure to generate cash flow. Together, these mechanisms provide compelling incentives for managers to improve the efficiency of the LBO firm.

### *8.4.4. Wealth transfers from debtholders*

If the pre-buyout bonds lack protective covenants, the LBO firm may issue more senior debt. Bonds that lack protective covenants become more junior in the capital structure, resulting in a reduction in the value of those bonds. Thus, it is possible that some of the buyout gains represent wealth transfers from target firm debtholders. Marais, Schipper, and Smith (1989) examine a sample of leveraged buyouts between 1974 and 1985. They find positive average CARs for convertible securities and preferred stock, most of which are redeemed as part of the buyout. A majority of the nonconvertible debt claims remain outstanding without renegotiation after the buyout. This debt typically lacks covenants restricting additional borrowing with higher seniority, and there are pervasive downgradings of public debt following successful buyout proposals, suggesting bondholder losses.

Asquith and Wizman (1990) investigate the one-month return for 199 bonds of LBO targets in the 1980s. They find an average abnormal return of $-1\%$ across all bonds. However, these losses are concentrated to bonds with no covenant protection (mean return of $-3\%$). Bonds with strong covenant protection have insignificant returns. Overall, the losses to bondholders are small compared to the total gains accruing to shareholders in the same LBO. Warga and Welch (1993) document an average risk-adjusted LBO announcement return of $-7\%$ for 36 bonds. The bondholder losses, however, constitute at most 6% of the shareholder gains. They too conclude that bondholder expropriation is a minor source of gains in leveraged buyouts. See also Billet, Jiang and Lie (2008) for an examination of bond returns in leveraged buyouts. They suggest that bondholder wealth expropriation has declined with the increased use of change-in-control covenants.

Ippolito and James (1992) propose that LBOs could extract wealth from other stakeholders as well. They examine the termination of pension plans in 169 buyouts in the 1980s. They find that the incidence of pension terminations doubles following LBO announcements. However, many of these terminations are affiliated with plant closings or an adaption to terms offered by the competitors of the LBO firm. Brown, Fee, and Thomas (2007) examine the effect of leveraged buyouts on the firms' suppliers, using a sample of 157 suppliers of firms undertaking LBOs in 1990–2001. They document an average announcement CAR of $-1.3\%$ for the suppliers. Moreover, the negative returns are concentrated to suppliers with substantial relation-specific investments. Thus, some of the LBO gains may come from the financial leverage as a commitment device in negotiations with suppliers and other stakeholders.

Another group of stake holders in the buyout transaction are the LBO bank lenders. Kracaw and Zenner (1996) examine the wealth effects of highly leveraged transactions on the stock prices of lead banks of the leveraged-loan syndicate. They find significantly positive average CARs of 0.5% when the transaction is announced and another 0.4% when the bank financing is agreed upon. Moreover, the bank stock returns are increasing in the size of the highly leveraged transaction. In all, bank lenders are expected to make profits on financing highly leveraged transactions and not the opposite.

Demiroglu and James (2007) investigate whether brand-sponsors borrow at better terms. Examining a sample of 181 LBOs completed between 1997 and 2007, they find that buyouts sponsored by high-reputation partnerships pay narrower loan spreads, have fewer and less restrictive loan covenants, and borrow more at a lower cost from institutional loan markets. In addition, sponsor reputation is positively related to the amount of leverage used to finance the buyout. Moreover, Ivashina and Kovner (2008) study 1582 leveraged loans financing private equity sponsored LBOs between 1993 and 2005.

They show that transaction loan spreads decline in the sponsor's relationship (past business) with the bank and the potential for future bank business. In sum, larger LBO sponsors can borrow at better terms. It is possible that this competitive advantage could help explain the persistence in returns across LBO sponsors documented by Kaplan and Schoar (2005).

### *8.4.5. Wealth transfers from target shareholders*

While managers have a fiduciary duty to negotiate fair value in a buyout transaction, as acquirers of shares, they stand to gain from a low transaction value. By understating the true value of the target shares, they expropriate wealth from outside shareholders in the buyout. DeAngelo (1986) examines the accounting choices of 64 NYSE firms proposing an MBO during 1973–1982. Using a variety of tests, she fails to find any evidence that managers systematically understate earnings in the period leading up to the buyout. Perry and Williams (1994) employ a different methodology and a larger sample of 175 MBOs. In contrast, they find evidence of manipulation of the discretionary accruals that lowers the earnings in the year preceding the buyout announcement.

Kaplan (1989b) compares the financial forecasts that firms present at the time of a management buyout to subsequent performance. He finds that the actual post-buyout performance generally lags the forecast, rejecting the notion that managers capitalize on inside information in the MBO. Lee (1992) studies a sample of withdrawn MBO proposals to determine whether managers' proposals reveal information beyond the gains from the completed transaction. He finds that stock prices drop back to their pre-bid level after the withdrawal of the MBO proposal unless another bidder appears. He suggests that the wealth creation in LBOs primarily results from efficiency gains associated with the completed transaction rather than wealth transfers from pre-buyout shareholders. Moreover, Ofek (1994) finds that stock prices drop back to their pre-buyout level after MBO offers are canceled or rejected by the target boards. Also, there is no subsequent improvement in the operating performance of these firms. Overall, the evidence at large suggests that buyout gains come from other sources than expropriation of selling shareholders.

A relatively new practice is the so called club deals, where two or more private equity firms jointly sponsor an LBO. The equity portion in recent mega-deals may be too large for a single fund to finance on its own. However, a concern with these deals is that LBO sponsors may collude to limit competition, hence reducing the price paid to target shareholders. Indeed, the U.S. Department of Justice launched an inquiry in late 2006 into the effect of such private equity consortiums on takeover competition.

Officer, Ozbas, and Sensoy (2008) examine the collusion argument for a sample of 53 club deals and 133 single-sponsor LBOs completed between 1984 and 2007. Using target abnormal return estimates, they find that club deals are associated with significantly lower premiums than single-sponsor deals. Guo, Hotchkies, and Song (2008) show that club deals are associated with higher returns on the capital invested in the LBO. However, they also find higher returns for target shareholders, rejecting the proposal of lower prices. Boone and Mulherin (2008) examine 70 club deals and 94 single-sponsor deals over the 2003–2007 period. Based on SEC filings, they show that the level of takeover competition is significantly higher for both types of LBO bidders compared to a control sample of takeovers. Moreover, target abnormal returns are largely the same across the different bidder categories. In sum, there is little evidence that club deals limit bidder competition in LBOs.

Outside investors may play an active role in the buyout, protecting target shareholder interests. Peck (1996) examines block trades in 111 MBO bids between 1984 and 1987. She finds that acquisitions of equity blocks increase around MBO offers, peaking three months prior to the offer. The participation of these blockholders increases the probability that the MBO proposal fails and a rival bidder acquires the firm. For a sample of 196 LBOs in 1990–2006, Huang (2008) finds significant increases in hedge fund holdings prior to the bid. He shows that the initial buyout premium increases with the level of hedge fund ownership in the target. Thus, outside investors seem to play an important role in increasing target returns.

### *8.4.6. Taxes*

Interest expenses are deductible and therefore reduce the firm's cost of capital. In the 1980s, management could also choose to step up the value of the assets after the buyout, increasing depreciation deductions. Kaplan (1989b) estimates the value of potential tax benefits created in MBOs using a range of assumptions about the marginal tax advantage to debt and the debt retirement schedule. Depending on the assumptions, the median value of the tax benefits from interest deductions range from 13 to 130% of the premium paid to pre-buyout shareholders, or 5 to 53% of the market value of equity two months prior to the buyout. He finds a strong positive correlation between the total tax deductions and the premium, and suggests that taxes are an important source of gains in leveraged buyouts.

See also Schipper and Smith (1991) and Newbould, Chatfield, and Anderson (1992) for further analysis of tax deductions in leveraged buyouts. Jensen, Kaplan, and Stiglin (1989) estimate that leveraged buyouts have a positive overall effect on the tax revenue of the U.S. Treasury. Simulations of the net effect of leveraged buyout activity for the U.S. Treasury are found in Chatfield and Newbould (1996).

## *8.5. Industry effects*

Slovin, Sushka, and Bendeck (1991) propose that leveraged buyout announcements convey private information about the future prospects of the industry. Examining the stock price reaction of 940 industry rivals of 128 buyouts in the 1980s, they find a significant and positive rival average announcement CAR of 1.3%. The rival returns are greater for rivals that are smaller in size than the target firm. Phallipou and Gottschalg (2008) argue that LBO announcements signal the existence of an industrywide agency problem, encouraging industry rivals to improve their governance structure too. They document an increase in rival firm options awards, director share ownership, and CEO turnover following LBO activity. It is not clear, however, whether their results are specific to rivals in industries with LBO activity or reflect a general trend in corporate governance.

One of the potential costs with high leverage is that it reduces financial flexibility and makes the LBO firm vulnerable to price competition by rival firms. Chevalier (1995b) examines how a leveraged buyout affects the pricing behavior of the LBO firm and its rivals in a local market, using data from the supermarket industry. She shows that prices rise when rival firms are also highly leveraged and LBO firms have higher prices than their competitors. However, prices fall when rival firms have relatively low debt levels and a single competitor controls a large market share. She finds that these low prices increase the probability that the LBO firm will exit, and suggests that rivals attempt to prey on LBO chains.

Phillips (1995) examines how financial leverage interacts with product market decisions for four different industries where a major player initiated a leveraged buyout. In three of the industries, characterized by difficult entry and high leverage of rival firms, prices increase and industry output declines with the average industry debt ratio. In the fourth industry, characterized by low leverage of rivals and low barriers to entry, prices fall and industry output increases with the industry debt ratio.

Overall, the evidence indicates that firms' leverage decisions affect industry pricing and output. See also Dasgupta and Titman (1998) for an equilibrium model explaining the interaction between capital structure and product markets, Fulghieri and Nagarajan (1996) for a model on the strategic role of high leverage for deterring entry in monopolistic markets, and Chevalier (1995a) for further evidence. Also, Parsons and Titman (2007) (Chapter 13 of this Handbook) discuss empirical studies on the interactions between leverage and corporate strategy.

### *8.6. Organizational longevity and exit*

Are leveraged buyouts a transitory structure or a sustainable corporate form that lasts over a longer period of time? Jensen (1989) argues that the organizational form of a leveraged buyout is superior to public ownership for firms in low-growth industries, predicting long-lived LBO companies. In contrast, Rappaport (1990) claims that the lack of financial flexibility will ultimately harm the buyout firm and foresees a prompt return to the public equity markets. Kaplan (1991) examines 183 large leveraged buyouts completed between 1979 and 1986. He finds that the median LBO target remains in private ownership for seven years. Moreover, 45% of the LBO firms return to public ownership at some point. In a sample of 72 reversed IPOs, that is, LBOs that subsequently went public, Muscarella and Versuypene (1990) report that the average firm remains private for three years.

Stromberg (2007) studies holding periods and exits for 21,000 buyout transactions in 1970–2007. Of these buyouts 17,000 (80%) were backed by a financial sponsor. Given the large number of transactions in the 2000s, only 40% of the firms in his sample have exited. He finds that 39% of the exits are in the form of a sale to a strategic buyer. One quarter of the exits are a secondary buyout, that is, a sale to another LBO fund—an exit form that has increased in importance over the last decade. IPOs account for 13% of the exits. Moreover, despite the significant leverage used in buyouts, only 6% of exiting firms file for bankruptcy or initiates a financial restructuring. Stromberg (2007) further shows that the median firm stays in LBO ownership for nine years, and only 8% of the firms are sold within two years of the buyout. Overall, the evidence suggests that leveraged buyouts are a long-term organizational form for many firms.

Van de Gucht and Moore (1998) use a hazard model to estimate the probability that an LBO firm returns to public ownership for a sample of 343 LBOs over 1980–1992. They show that 27% of the firms reverse through an IPO after 3.5 years on average. Another 9% of the firms are sold to a publicly held company. Almost half of the firms remain private, and 12% file for bankruptcy. Moreover, the likelihood for an IPO is higher when the industry average market-to-book ratio is rising.

Degeorge and Zeckhauser (1993) study the decision to exit a buyout through a public offering for 62 reverse LBOs in the 1980s. They find that the IPO coincides with a peak in the buyout firm's operating performance. The stock of the reverse LBOs outperform comparison firms, however, suggesting that the market anticipates the subsequent decline in operating profitability. They conclude that LBO firms choose to go public when their performance is strong. Holthausen and Larcker (1996) further show that the accounting performance of LBO firms exceeds that of its industry rivals at the time of the IPO and for the following four years. See Liu (2006), Cao and Lerner (2006), and Cao (2008) for additional evidence on reverse LBOs.

Halpern, Kieschnick, and Rotenberg (1999) conjecture that there are two types of targets in leveraged buyouts. One is the classical public target with little managerial equity and high free cash flow. The other is a target that performs poorly because the manager has too much of her wealth invested in the firm and hence is suboptimally risk-averse. Examining 126 LBOs in 1981–1986, they find that their sample clusters into two groups. The first group has low prior managerial equity and takeover premiums that decrease in managerial equity. Moreover, the buyout is led by an outside sponsor, and the LBO firm is typically sold in an IPO or to a strategic buyer. The second group has high managerial equity and takeover premiums that increase in managerial equity. These buyouts are led by managers, and the LBO firm tends to remain private. In addition, managers in this group typically increase their ownership fraction but decrease the dollar investment in the LBO firm. The authors suggest that a partition into these two different types of target firms better describes the LBO population.

Why did so many deals fail in the early 1990s? Bruner and Eades (1992) examine the failure of Revco in 1988, only 19 months after its leveraged buyout. They simulate the ex-ante probability of survival, based on historical and predicted cash flows at the time of the deal. They conclude that the company was overleveraged from the closing of the deal, with little probability of successfully servicing its debt. Kaplan and Stein (1993) contend that the buyout market overheated toward the end of the 1980s, resulting in many poorly structured transactions. They find higher price multiples and leverage ratios, increased use of junk bonds with few restrictive covenants, and more money paid up-front to managers and investment banks.

## 9. Conclusions

In this chapter, we review the extant literature on corporate breakup transactions and highly leveraged transactions. For each individual transaction, we survey techniques, transaction volume, valuation effects and potential sources of restructuring gains. Corporate breakup transactions are optimal when the separation of the diversified firm's divisions increases firm value. The breakup transactions range from divestitures and spinoffs, which entirely separates a subsidiary from its parent, to equity carveouts and tracking stock, which preserves some parent control. The highly leveraged transactions result in the firm taking on substantial additional debt in its capital structure. This happens in leveraged recapitalizations and in leveraged buyouts.

A divestiture is a sale of a division or subsidiary in a private transaction. Asset sales generate cash to the parent firm on the one hand, but trigger a capital gains tax on the other. The average parent firm experiences an abnormal stock return of 1% and the average buyer a CAR of 0.5% when a divestiture is announced. These valuation effects have several explanations: (i) most divestitures involve divisions that are unrelated to the parent firm, increasing the corporate focus of the diversified firm; (ii) the parent firm's investment decisions tend to improve after the divestiture; and (iii) assets are often transferred to a higher valuation buyer. Moreover, it appears that managers are reluctant to sell assets, managers in firms with better corporate governance make better divestment decisions, and the retention of proceeds is associated with inefficient investments.

A spinoff is the separation of a subsidiary through a distribution of the stock to parent shareholders. Spinoffs can be completed without any tax implications, but also do not generate any cash to the parent. The parent stock price increases by 3% on average at the announcement of a spinoff. The value creation comes from: (i) increased corporate focus; (ii) elimination of cross-subsidization leading to improved investment decisions; (iii) reduced information asymmetries; and (iv) a higher probability of becoming a target. Investors rebalance their portfolios when the parent and subsidiary stocks start trading separately. Moreover, parent managers design the subsidiary corporate charter to include more takeover defenses compared to the parent firm itself as well as other IPO firms.

An equity carveout is a partial IPO of the subsidiary, where the parent typically retains a controlling stake. It generates cash (the IPO proceeds) but no tax. The average parent firm experiences an abnormal stock return of 2% at the announcement of an equity carveout. The gains in equity carveouts are attributed to: (i) an increase in corporate focus; and (ii) a reduction of the financing costs for high-growth subsidiaries. Equity carveouts are a temporary organizational form, and most carveouts are subsequently reacquired or sold off. It is possible that the carveout generates information about the value of the subsidiary as an independent company, improving the decision to exercise the option to sell out or buy back the subsidiary.

Tracking stock is a separate class of common stock in the parent company, tracking the performance of a given division. The tracking stock generates cash if it is offered to the public and has no tax implication. The average parent CAR is 3% on the announcement of a tracking stock issue. These announcement returns are, however, difficult to explain beyond an initial market infatuation with yet another breakup transaction. The tracking stock is a "quasi-pure" play in that it requires separate divisional SEC filings, but has voting rights in the parent. In fact, tracking stock trades like its corporate sibling divisions rather than its industry. It lends itself for expropriation since the corporate board, without legal remedy, can transfer funds from the tracked division to the rest of the company. As a result of such expropriation, most tracking stock issues have been dissolved.

A leveraged recapitalization is a large special dividend financed by debt, substantially increasing the firm's leverage. The average abnormal stock return is 5% on the announcement of a leveraged recapitalization and 20 to 30% through closing of the transaction. The gains in leveraged recapitalizations are attributed primarily to the incentive effects of debt: recap firms substantially cut their capital expenditures and increase operating profitability.

A leveraged buyout is an acquisition by private investors financed primarily by debt. Premiums paid to target shareholders in LBOs average 37%, and announcement CARs average 16–17%. The LBO gains are attributed to several sources: (i) improved investment and operating efficiencies; (ii) increased equity-based incentives to management; and (iii) strong monitoring by the LBO sponsor. Buyouts of the 2000s seem to have somewhat less improvements in operating efficiency, but in general create value similar to LBOs of the 1980s. Recent developments include club deals (consortiums of LBO sponsors bidding together), fund-to-fund exits (LBO funds selling the portfolio firm to another LBO fund), a highly liquid (until mid-2007) leveraged loan market, and evidence of persistence in fund returns (perhaps because brand-sponsors borrow at better rates).

In this survey, we have focused on the individual transactions and their associated empirical evidence. This is also how most of the literature progresses. A major drawback of this approach is the resulting lack of analysis of alternatives. That is, when a company self-selects a divestiture, what were reasonable alternative strategies? In what sense was divestiture superior to, say, a spinoff or an equity carveout? In what sense was going private via an LBO superior to a leveraged recapitalization? Are there systematic differences between public to private LBO transactions and private-to-private restructurings? Ideally, one would use a theoretical model to structure the answers to these types of questions. Perhaps the greatest challenge to the restructuring literature is to achieve a modicum of integration of the analysis across transaction types. Also, it is difficult to evaluate the expected return from buyout investments with only limited data on portfolio companies that do not return to public status within the sample period. We expect these issues to be resolved as both theories and data become more readily available in the future.

## References

Abarbanell, J. S., B. J. Bushee and J. S. Ready, 2003, "Institutional Investor Preferences and Price Pressure: The Case of Corporate Spin–offs," *Journal of Business*, 76, 233–261.

Acharya, V. and C. Kehoe, 2008, "Corporate Governance and Value Creation: Evidence from the Private Equity Market," Working Paper, London Business School.

Afshar, K. A., R. J. Taffler and P. S. Sudarsanam, 1992, "The Effect of Corporate Divestments on Shareholder Wealth: The UK Experience," *Journal of Banking and Finance*, 16, 115–135.

Ahn, S. and M. D. Walker, 2007, "Corporate Governance and the Spinoff Decision," *Journal of Corporate Finance*, 13, 76–93.

Alexander, G., P. Benson and J. Kampmeyer, 1984, "Investigating the Valuation Effects of Announcements of Voluntary Corporate Sell–offs," *Journal of Finance*, 39, 503–517.

Allen, J., 1998, "Capital Markets and Corporate Structure: The Equity Carve–outs of Thermo Electron." *Journal of Financial Economics*, 48, 99–124.

Allen, J. W., 2001, "Private Information and Spin-off Performance," *Journal of Business*, 74, 281–306.

Allen, J. W., S. L. Lummer, J. J. McConnell and D. K. Reed, 1995, "Can Takeover Losses Explain Spin–off Gains?" *Journal of Financial and Quantitative Analysis*, 30, 465–485.

Allen, J. W. and J. J. McConnell, 1998, "Equity Carve-outs and Managerial Discretion," *Journal of Finance*, 53, 163–186.

Alli, K., G. B. Ramirez and K. K. Yung, 2001, "Withdrawn Spin-offs: An Empirical Analysis," *Journal of Financial Research*, 24, 603–616.

Amess, K. and M. Wright, 2007, "The Wage and Employment Effects of Leveraged Buyouts in the UK," *International Journal of the Business of Economics*, 14, 179–195.

Aron, D. J., 1991, "Using the Capital Market as a Monitor: Corporate Spinoffs in an Agency Framework," *RAND Journal of Economics*, 22, 505–518.

Asquith, P., R. Gertner and D. Scharfstein, 1992, "Anatomy of Financial Distress: An Examination of Junk-bond Issuers," *Quarterly Journal of Economics*, 109, 625–658.

Asquith, P. and T. A. Wizman, 1990, "Event Risk, Covenants, and Bondholder Returns in Leveraged Buyouts," *Journal of Financial Economics*, 27, 195–213.

Atanasov, V., A. Boone and D. Haushalter, 2005, "Minority Shareholder Expropriation in U.S. Publicly-Traded Subsidiaries," Working Paper, Babson College.

Axelson, U., T. Jenkinson, P. Stromberg and M. S. Weisbach, 2007, "Leverage and Pricing in Buyouts: An Empirical Analysis," Working Paper, Stockholm Institute of Financial Research.

Bae, S. C. and D. P. Simet, 1998, "A Comparative Analysis of Leveraged Recapitalization versus Leveraged Buyout as a Takeover Defense," *Review of Financial Economics*, 7, 157–172.

Baker, G. and K. H. Wruck, 1989, "The Case of the O M Scott & Sons Company," *Journal of Financial Economics*, 25 (December), 163–190.

Balachandran, B., R. Faff and T. A. Nguyen, 2004, "The Intra-industry Impact of Special Dividend Announcements: Contagion versus Competition," *Journal of Multinational Financial Management*, 14, 369–385.

Baltin, M. and M. Brettel, 2007, "Equity Carve-outs—A Matter of Time?" Working Paper, Aachen University.

Bates, T. W., 2005, "Asset Sales, Investment Opportunities, and the Use of Proceeds," *Journal of Finance*, 60, 105–135.

Berger, P. G. and E. Ofek, 1995, "Diversifications Effect on Firm Value," *Journal of Financial Economics*, 37(1, January), 39–65.

Berger, P. G. and E. Ofek, 1999, "Causes and Effects of Corporate Refocusing Programs," *Review of Financial Studies*, 12, 311–345.

Best, R. W., R. J. Best and A. M. Agapos, 1998, "Earnings Forecasts and the Information Contained in Spinoff Announcements," *Financial Review*, 33, 53–67.

Betton, S., B. E. Eckbo and K. S. Thorburn, 2008, "Corporate Takeovers," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 2, Chapter 15, p. in press (Elsevier/North-Holland, Handbooks in Finance Series).

Billet, M. T., Z. Jiang and E. Lie, 2008, "The Role of Bondholder Weatlh Expropriation in lbo Transactions," Working Paper, University of Iowa.

Billett, M. T. and D. C. Mauer, 2000, "Diversification and the Value of Internal Capital Markets. *Journal of Banking and Finance*, 24, 1457–1490.

Billett, M. T. and A. M. Vijh, 2004, "The Wealth Effects of Tracking Stock Restructurings," *Journal of Financial Research*, 27, 559–583.

Boone, A., D. Haushalter and W. Mikkelson, 2003, "An Investigation of the Gains from Specialized Equity Claims," *Financial Management*, 32, 67–83.

Boone, A. L. and J. H. Mulherin, 2008, "Do Private Equity Consortiums Impede Takeover Competition?" Working Paper, University of Kansas.

Brown, D. T., C. E. Fee and S. E. Thomas, 2007, "Financial Leverage and Bargaining Power with Suppliers: Evidence from Leveraged Buyouts," Working Paper, University of Pittsburgh.

Brown, D. T., C. M. James and R. M. Mooradian, 1994, "Asset Sales by Financially Distressed Firms," *Journal of Corporate Finance*, 1, 233–257.

Bruner, R. and K. Eades, 1992, "The Crash of the Revco lbo: The Hypothesis of Inadequate Capital," *Financial Management*, 21 (1, Spring), 35–49.

Burch, T. R. and V. Nanda, 2003, "Divisional Diversity and the Conglomerate Discount: Evidence from Spinoffs," *Journal of Financial Economics*, 70, 69–98.

Cao, J., 2008, "What Role Does Private Equity Play When Leveraged Buyouts go Public?" Working Paper, Boston College.

Cao, J. and J. Lerner, 2006, "The Performance of Reverse Leveraged Buyouts," NBER Working Paper 12626.

Chatfield, R. E. and G. D. Newbould, 1996, "Leveraged Buyouts: Implications for U.S. Treasury Tax Receipts." *Quarterly Journal of Business and Economics*, 35, 51–65.

Chemmanur, T. and I. Paeglis, 2001, "Why Issue Tracking Stock," Working Paper, bcedu.

Chemmanur, T. J. and S. He, 2007, "Institutional Trading, Information Production, and Corporate Spin–offs," Working Paper, Boston College.

Chen, H.-L. and R.-J. Guo, 2005, "On Corporate Divestiture," *Review of Quantitative Finance and Accounting*, 24, 399–421.

Chevalier, J. A., 1995a, "Capital Structure and Product-Market Competition: Empirical Evidence from the Supermarket Industry," *American Economic Review*, 85, 415–435.

Chevalier, J. A., 1995b, "Do LBO Supermarkets Charge More? An Empirical Analysis of the Effects of LBOs on Supermarket Pricing," *Journal of Finance*, 50, 1095–1112.

Chou, D.-W., M. Gombola and F.-Y. Liu, 2006, "Earnings Management and Stock Performance of Reverse Leveraged Buyouts," *Journal of Financial and Quantitative Analysis*, 41, 407–438.

Clayton, M. J. and Y. Qian, 2004, "Wealth Gains from Tracking Stocks: Long-run Performance and Ex-date Returns," *Financial Management*, pp. 83–106.

Clubb, C. and A. Stouraitis, 2002, "The Significance of Sell-off Profitability in Explaining the Market Reaction to Divestiture Announcements," *Journal of Banking and Finance*, 26, 671–688.

Coase, R. 1937. "The Nature of the Firm," *Economica* 4, 386–405.

Colak, G. and T. M. Whited, 2007, "Spin–offs, Divestitures and Conglomerate Investment," *Review of Financial Studies*, 20, 557–595.

Cornelli, F. and O. Karakas, 2008, "Private Equity and Corporate Governance: Do lbos Have More Effective Boards?" Working Paper, London Business School.

Cotter, J. and S. Peck, 2001, "The Structure of Debt and Active Equity Investors: The Case of the Buyout Specialist. *Journal of Financial Economics*, 59, 101–147.

Cressy, R., F. Munari and A. Malipiero, 2007, "Playing to Their Strenghts? Evidence that Specialization in the Private Equity Industry Confers Competitive Advantage," *Journal of Corporate Finance*, 13, 647–669.

Cusatis, P. J., J. A. Miles and J. R. Woolridge, 1993, "Restructuring through Spinnoffs: The Stock Market Evidence," *Journal of Financial Economics*, 33, 293–311.

Daines, R. and M. Kausner, 2004, "Agents Protecting Agents: An Empirical Study of Takeover Defenses in Spinoffs," Working Paper, Stanford University.

Daley, L., V. Mehrotra and R. Sivakumar, 1997, "Corporate Focus and Value Creation: Evidence of Spinoffs," *Journal of Financial Economics*, 45, 257–281.

Danielova, A. N., 2008, "Tracking Stock or Spin–off? Determinants of Choice," *Financial Management*, 37, 125–139.

Dasgupta, S. and S. Titman, 1998, "Pricing Strategy and Financial Policy," *Review of Financial Studies*, 11, 705–737.

Datta, S., M. Iskandar-Datta and K. Raman, 2003, "Value Creation in Corporate Asset Sales: The Role of Managerial Performance and Lender Monitoring," *Journal of Banking and Finance*, 27, 351–375.

Datta, S. and M. Iskander-Datta, 1996, "Takeover Defenses and Wealth Effects on Securityholders: The Case of Poison Pill Adoptions," *Journal of Banking and Finance*, 20(7), 1231–1250.

Datta, S. and M. Iskander-Datta, 1995, "Corporate Partial Acquisitions, Total Firm Valuation and the Effect of Financing Method," *Journal of Banking and Finance*, 19(1), 97–115.

Davis, S. J., J. Haltiwanger, R. Jarmin, J. Lerner and J. Miranda, 2008, "Private Equity and Employment," Center for Economic Studies, CES-0807.

DeAngelo, H., L. DeAngelo and E. Rice, 1984, "Going Private: Minority Freezeouts and Stockholder Wealth," *Journal of Law and Economics* (October), 367–402.

DeAngelo, L. E., 1986, "Accounting Numbers as Market Valuation Substitutes: A Study of Management Buyouts of Public Stockholders," *Accounting Review*, 61, 400–420.

Degeorge, F. and R. Zeckhauser, 1993, "The Reverse LBO Decision and Firm Performance: Theory and Evidence," *Journal of Finance*, 48, 1323–1348.

Demiroglu, C. and C. James, 2007, "Lender Control and the Role of Private Equity Group Reputation in Buyout Financing," Working Paper, University of Florida.

Denis, D. J. and D. K. Denis, 1995, "Causes of Financial Distress Following Leveraged Recapitalizations," *Journal of Financial Economics*, 37, 129–157.

Denis, D. J. and D. K. Denis, 1993, "Managerial Discretion, Organizational Structure, and Corporate Performance: A Study in Corporate Performance," *Journal of Accounting and Economics*, 16, 209–237.

Denis, D. J., 1994, "Investment Opportunities and the Market Reaction to Equity Offerings," *Journal of Financial and Quantitative Analysis*, 29, 159–177.

Denis, D. J., 1990, "Defensive Changes in Corporate Payout Policy: Share Repuchases and Special Dividends," *Journal of Finance*, 46, 1433–1456.

Denis, D. K. and D. K. Shome, 2005, "An Empirical Investigation of Corporate Asset Downsizing," *Journal of Corporate Finance*, 11, 427–448.

Desai, H. and P. C. Jain, 1999, "Firm Performance and Focus: Long-run Stock Market Performance Following Spinoffs," *Journal of Financial Economics*, 54, 75–101.

Dittmar, A. and A. Shivdasani, 2003, "Divestitures and Divisional Investment Policies," *Journal of Finance*, 58, 2711–2743.

D'Souza, J. and J. Jacob, 2000, "Why Firms Issue Targeted Stock," *Journal of Financial Economics*, 56, 459–483.

Eckbo, B. E., R. W. Masulis and Ø. Norli, 2007, "Security Offerings," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 1, Chapter 6 (Elsevier/North-Holland, Handbooks in Finance Series).

Eckbo, B. E. and K. S. Thorburn, 2007, "Automatic Bankruptcy Auctions and Fire-sales," *Journal of Financial Economics*, forthcoming.

Elder, J., P. K. Jain and J.-C. Kim, 2005, "Do Tracking Stock Reduce Information Asymmetries? An Analysis of Liquididity and Adverse Selection," *Journal of Financial Research*, 28, 197–213.

Elder, J. and P. Westra, 2000, "The Reaction of Security Prices to Tracking Stock Announcements. *Journal of Economics and Finance*, 24, 36–55.

Fama, E. F. and K. R. French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 43, 3–56.

Fidrmuc, J. P., P. Roosenboom and D. van Dijk, 2008, "Do Private Equity Investors Crowd Out Management Buyouts," Working Paper, Erasmus University, Rotterdam.

Fluck, Z. and A. W. Lynch, 1999, "Why Do Firms Merger and Then Divest? A Theory of Financial Synergy," *Journal of Business*, 72, 319–346.

Frank, M. Z. and V. K. Goyal, 2008, "Tradeoff and Pecking Order Theories of Debt", in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 2, Chapter 12 (Elsevier/North-Holland, Handbooks in Finance Series).

Fulghieri, P. and S. Nagarajan, 1996, "On the Strategic Role of High Leverage in Entry Deterrence," *Journal of Banking and Finance*, 20, 1–23.

Galai, D. and R. W. Masulis, 1976, "The Option Pricing Model and the Risk Factor of Stock," *Journal of Financial Economics*, 3, 53–81.

Gertner, R., E. Powers and D. Scharstein, 2002, "Learning About Internal Capital Markets and Corporate Spin–offs," *Journal of Finance*, 57, 2479–2506.

Gillan, S., J. W. Kensinger and J. D. Martin, 2000, "Value Creation and Corporate Diversification: The Case of Sears Roebuck & Co.," *Journal of Financial Economics*, 55, 103–137.

Gilson, S. C., P. M. Healy, C. F. Noe and K. G. Palepu, 2001, "Analyst Specialization and Conglomerate Stock Breakups," *Journal of Accounting Research*, 39, 565–582.

Gleason, K., J. Madura and A. K. Pennathur, 2006, "Valuation and Performance of Reacquisitions Following Equity Carve-outs," *Financial Review*, 41, 229–246.

Goldman, E., 2004, "The Impact of Stock Market Information Production on internal resource Allocation," *Journal of Financial Economics*, 71, 143–167.

Grinstein, Y., 2006, "The Disciplinary Role of Debt and Equity Contracts: Theory and Tests," *Journal of Financial Intermediation*, 15, 419–443.

Groh, A. P. and O. Gottschalg, 2008, "Measuring the Risk-Adjusted Performance of U.S. Buyouts," Working Paper, University of Navarra.

Grossman, S. J. and O. D. Hart, 1986, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691–719.

Guo, S., E. S. Hotchkiss and W. Song, 2008, "Do Buyouts (still) Create Value?" Working Paper, Boston College.

Gupta, A. and L. Rosenthal, 1991, "Ownership Structure, Leverage, and Firm Value: The Case of Leveraged Recapitalizations," *Financial Management*, 20, 69–83.

Halpern, P., R. Kieschnick and W. Rotenberg, 1999, "On the Heterogeneity of Leveraged Going Private Transactions," *Review of Financial Studies*, 12, 281–309.

Hand, J. R. M. and T. R. Skantz, 1998, "The Economic Determinants of Accounting Choices: The Unique Case of Equity Carve-outs under SAB 51," *Journal of Accounting and Economics*, 24, 175–203.

Handa, P. and A. R. Radhakrishnan, 1991, "An Empirical Investigation of Leveraged Recapitalizations with Cash Payout as Takeover Defense," *Journal of Finance*, 2 (June), 731–745.

Hanson, R. C. and M. H. Song, 2003 "Long-term Performance of Divesting Firms and the Effect of Managerial Ownership," *jeaf*, 27, 321–336.

Hanson, R. C. and M. H. Song, 2000, "Managerial Ownership, Board Structure, and the Division of Gains," *Journal of Corporate Finance*, 6, 55–70.

Harper, J. T. and J. Madura. 2002, "Sources of Hidden Value and Risk within Tracking Stock," *Financial Management*, 31, 91–109.

Harris, R., D. S. Siegel and M. Wright, 2005 "Assessing the Impact of Management Buyouts on Economic Efficiency: Plant-level Evidence Form the United Kingdom," *Review of Economic Studies* 87, 148–153.

Hart, O. D. and J. Moore, 1990, "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98, 1119–1158.

Hass, J. J., 1996 "Directorial Fiduciary Duties in a Tracking Stock Equity Structure: The Need for a Duty of Fairness," *Michigan Law Review*, 94, 2089–2177.

Healy, P. M. and K. G. Palepu, 1995, "The Challenges of Investor Communication: The Case of Cuc International Inc.," *Journal of Financial Economics*, 38, 111–140.

Hirschey, M., M. B. Slovin and J. K. Zaima, 1990 "Bank Debt, Insider Trading, and the Return to Corporate Selloffs," *Journal of Banking and Finance*, 14, 85–98.

Hirschey, M. and J. K. Zaima, 1989 "Insider Trading, Ownership Structure, and the Market Assessment of Corporate Sell–offs," *Journal of Finance*, 44, 971–980.

Hite, G., J. Owers and R. Rogers, 1987 "The Market for Inter–Firm Asset Sales: Partial Sell–offs and Total Liquidations," *Journal of Financial Economics*, 18, 229–252.

Hite, G. L. and J. E. Owers, 1983, Security Price Reactions around Corporate Spin-off Announcements," *Journal of Financial Economics*, 12, 409–436.

Holthausen, R. W. and D. F. Larcker, 1996, "The Financial Performance of Reverse Leveraged Buyouts," *Journal of Financial Economics*, 42, 293–332.

Hotchkiss, E. S., K. John, R. Mooradian and K. S. Thorburn, 2008, "Bankruptcy and the Resolution of Financial Distress," In B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 2, Chapter 14. (Elsevier/North-Holland, Handbooks in Finance Series).

Huang, J., 2008, "When Bad Stocks Make Good Investments: The Role of Hedge Funds in Leveraged Buyouts," Working Paper, Boston College.

Hulburt, H. M., 2003, "Equity Carve–outs and Changes in Corporate Control," *Journal of Applied Business Research*, 19, 29–40.

Hulburt, H. M. J. A. Miles and J. R. Woolridge, 2002, "Value Creation From Equity Carve-outs," *Financial Management*, 31, 83–100.

Huson, M. R. and G. MacKinnon, 2003, "Corporate Spinoffs and Information Asymmetry between Investors," *Journal of Economics and Management Strategy*, 9, 481–503.

Ippolito, R. A. and W. H. James, 1992, "LBOs, Reversions and Implicit Contracts," *Journal of Finance*, 47, 139–167.

Ivashina, V. and A. Kovner, 2008, "The Private Equity Advantage: Leveraged Buyout Firms and Relationship Banking," Working Paper, Boston College.

Jain, P., 1985, "The Effect of Voluntary Sell–off Announcements on Shareholder Wealth," *Journal of Finance*, 40, 209–224.

Jensen, M. C., S. Kaplan and L. Stiglin, 1989, "Effects of LBOs on Tax Revenues of the U.S. Treasury," *Tax Notes*, 42, 727–733.

Jensen, M. C. and W. Meckling, 1976, "Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure," *Journal of Financial Economics*, 3, 305–360.

Jensen, M. C., 1989, "Eclipse of the Public Corporation," *Harvard Business Review*, September–October, 61–74.

Jensen, M. C., 1986, "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review*, 76, 323–329.

John, K. and E. Ofek, 1995, "Asset Sales and Increase in Focus," *Journal of Financial Economics*, 37, 105–126.

Jones, C. and M. Rhodes-Kropf, 2004, "The Price of Diversifyable Risk in Venture Capital and Private Equity," Working Paper, Columbia University.

Kaiser, K. M. J. and A. Stouraitis, 2001, "Reversing Corporate Diversification and the Use of the Proceeds from Asset Sales: The Case of Thorn Emi," *Financial Management*, 30, 63–102.

Kalay, A. and M. Lemmon, 2008, "Payout policy," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 2, Chapter 10 (Elsevier/North-Holland, Handbooks in Finance Series).

Kaplan, S., 1989a, "The Effects of Management Buyouts on Operating Performance and Value," *Journal of Financial Economics*, 24 (2, October), 217–254.

Kaplan, S., 1989b, "Management Buyouts: Evidence on Taxes as a Source of Value," *Journal of Finance*, 44, 611–633.

Kaplan, S. N., 1991, "The Staying Power of Leveraged Buyouts," *Journal of Financial Economics*, 29, 287–313.

Kaplan, S. N. and A. Schoar, 2005, "Private Equity Performance: Returns, Persistence, and Capital Flows," *Journal of Finance*, 60, 1791–1823.

Kaplan, S. N. and J. C. Stein, 1993, "The Evolution of Buyout Pricing and Financial Structure in the 1980s," *Quarterly Journal of Economics*, 108, 313–357.

Kaplan, S. N. and P. Stromberg, 2008, "Leverage Buyouts and Private Equity," Working Paper, University of Chicago.

Kaplan, S. N. and M. S. Weisback, 1992, "The Success of Acquisitions: Evidence from Divestitures," *Journal of Finance*, 47, 107–138.

Kim, C. E., 1998, "The Effects of Asset Liquidity: Evidence from the Contract Drilling Industry," *Journal of Financial Intermediation*, 7, 151–176.

Klein, A., 1986, "The Timing and Substance of Divestiture Announcements: Individual, Simultaneous and Cumulative Effects," *Journal of Finance*, 41, 685–696.

Klein, A., J. Rosenfeld and W. Beranek, 1991, "The Two Stages of an Equity Carve-out and the Price Response of Parent and Subsidiary Stock," *Managerial and Decision Economics*, 12, 449–460.

Klein, B., R. G. Crawford and A. A. Alchian, 1978, "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law and Economics*, 21, 297–326.

Kracaw, W. A. and M. Zenner, 1996, "The Wealth Effects of Bank Financing Announcements in Highly Leveraged Transactions," *Journal of Finance*, 51, 1931–1946.

Krishnaswami, S. and V. Subramaniam, 1999, "Information Asymmetry, Valuation, and the Corporate Spin–off," *Journal of Financial Economics*, 53, 73–112.

Kruse, T. A., 2002, "Asset Liquidity and the Determinants of Asset Sales by Poorly Performing Firms," *Financial Management*, 31, 107–129.

Lang, L., A. Poulsen and R. Stulz, 1995, "Asset Sales, Firm Performance, and the Agency Costs of Managerial Discretion," *Journal of Financial Economics*, 37, 3–37.

Lang, L. H. P. and R. M. Stulz, 1994, "Tobin's q, Corporate Diversification, and Firm Performance," *Journal of Political Economy*, 102, 1248–1280.

Lee, S., 1992, "Management Buyout Proposals and Inside Information," *Journal of Finance*, 47, 1061–1079.

Lehn, K. and A. Poulsen, 1989, "Free Cash Flow and Stockholder Gains in Going Private Transactions," *Journal of Finance*, 44, 771–787.

Lerner, J., M. Sorensen and P. Stromberg, 2008, "Private Equity and Long-run Investment: The Case of Innovation," Working Paper, Harvard University.

Lichtenberg, F. and D. Siegel, 1990, "The Effects of Leveraged Buyouts on Productivity and Related Aspects of Firm Behavior," *jfe*, 27, 165–194.

Liebeskind, J., M. Wiersema and G. Hansen, 1992, "LBOs, Corporate Restructuring, and the Incentive-Intensity Hypothesis," *Financial Management*, p. 73–88.

Linn, S. C. and M. S. Rozeff, 1984, "The Corporate Sell-off, *Midland Corporate Finance Journal*, 2, 17–26.

Liu, M. H., 2005, "Information Content of Conglomerate Stock Breakups," Working Paper, University of Kentucky.

Ljunqvist, A. and M. Richarson, 2003, "The Cash Flow, Return and Risk Characteristics of Private Equity," Working Paper, New York University.

Logue, D. E., J. K. Seward and J. P. Walsh, 1996, "Rearranging Residual Claims: A Case for Targeted Stock," *Financial Management*, 25, 43–61.

Loh, Charmen, R. B. Jennifer and H. Toms, 1995, "Voluntary Corporate Divestitures as Antitakeover Mechanisms," *Financial Review*, 30, 41–60.

Long, W. and D. Ravenscraft, 1993, "LBOs, Debt and R&D Intensity," *smj*, 14, 119–135.

MacMinn, R. D. and P. L. Brockett, 1995, "Corporate Spin-offs as a Value Enhancing Technique When Faced with Legal Liability," *Insurance: Mathematics and Economics*, 16, 63–68.

Maksimovic, V., 1995, "Financial Structure and Product Market Competition," in R. Jarrow, V. Maksimovic and B. Ziemba (Eds.), *Finance* (North-Holland, Series of Handbooks in Operations Research and Management Science), Chapter 27, 887–920.

Maksimovic, V. and G. Phillips, 2001, "The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains?" *Journal of Finance*, 56, 2019–2065.

Maksimovic, 2007, "Conglomerate Firms and Internal Capital Markets," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, Vol. 1, Chapter 9. (Elsevier/North-Holland, Handbooks in Finance Series).

Marais, L., K. Schipper and A. Smith, 1989, "Wealth Effects of Going Private for Senior Securities," *jfe*, 23, 155–191.

Masulis, R., 1987, "Changes in Ownership Structure: Conversions of Mutual Savings and Loans to Stock Charter," *Journal of Financial Economics*, 18, 29–60.

Maxwell, W. F. and R. P. Rao, 2003, "Do Spin–offs Expropriate Wealth from Bondholders?" *Journal of Finance*, 58, 2087–2108.

McConnell, J. J., M. Ozbilgin and S. Wahal, 2001, "Spin-offs, Ex Ante," *Journal of Business*, 74, 245–280.

McNeil, C. R. and W. T. Moore, 2005, "Dismantling Internal Capital Markets vis spinoff: Effects on Capital Allocation Efficiency and Firm Valuation," *Journal of Corporate Finance*, 11, 253–275.

Megginson, W. and J. Netter, 2001, "From State to Market: A Survey of Empirical Studies on Privatization," *Journal of Economic Literature*, 39, 321–389.

Metrick, A. and A. Yasuda, 2007, "Economics of Private Equity Funds," Working Paper, Wharton at University of Pennsylvania.

Michaely, R. and W. H. Shaw, 1995, "The Choice of Going Public: Spin-offs vs. Carve-outs," *Financial Management*, 24, 5–21.

Miles, J. A. and J. D. Rosenfeld, 1983, "The Effect of Voluntary Spin-off Announcements on Shareholder Wealth," *Journal of Finance*, 38, 1597–1606.

Mulherin, J. H. and A. L. Boone, 2000, "Comparing Acquisitions and Divestitures," *Journal of Corporate Finance*, 6, 117–139.

Muller, H. M. and F. Panunzi, 2004, "Tender Offers and Leverage," *Quarterly Journal of Economics*, 119, 1217–1248.

Muscarella, C. and M. Versuypens, 1990, "Efficiency and Organizational Structure: A Study of Reverse LBOS," *Journal of Finance*, 45(5, December), 1389–1413.

Myers, S. C. and N. S. Majluf, 1984, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do not Have," *Journal of Financial Economics*, 13, 187–221.

Myers, S. C., 1977, "Determinants of Corporate Borrowing," *Journal of Financial Economics*, 5, 147–175.

Nanda, V., 1991, "On the Good News in Equity Carve–outs," *Journal of Finance*, 46, 1717–1737.

Nanda, V. and M. P. Narayanan, 1999, "Disentangling Value: Financing Needs, Firm Scope and Divestitures," *Journal of Financial Intermediation*, 8, 174–204.

Newbould, G. D., R. E. Chatfield and R. F. Anderson, 1992, "Leveraged Buyouts and Tax Incentives," *Financial Management*, 21, 50–57.

Nikoskelainen, E. and M. Wright, 2007, "The Impact of Corporate Governance Mechanisms on Value Increase in Leveraged Buyouts," *Journal of Corporate Finance*, 13, 511–537.

Nixon, T. D., R. L. Roenfeldt and N. W. Sicherman, 2000, "The Choice between Spin-offs and Sell-offs," *Review of Quantitative Finance and Accounting*, 14, 277–288.

Ofek, E., 1994, "Efficiency Gains in Unsuccessful Management Buyouts," *Journal of Finance*, 49, 637–654.

Ofek, E., 1993, "Capital Structure and Firm Response to Poor Performance: An Empirical Analysis," *Journal of Financial Economics*, 34, 3–30.

Officer, M. S., 2007, "The Price of Corporate Liquidity: Acquisition Discounts for Unlisted Targets," *Journal of Financial Economics*, 83, 571–598.

Officer, M. S., O. Ozbas and B. A. Sensoy, 2008, "Club Deals in Leveraged Buyouts," Working Paper, University of Souther California.

Opler, T. and S. Titman, 1993, "The Determinants of Leverage Buyout Activity: Free Cash Flow vs Financial Distress Costs," *Journal of Finance*, 48, 1985–1999.

Parrino, R., 1997, "Spinoffs and Wealth Transfers: The Marriott Case," *Journal of Financial Economics*, 43, 241–274.

Parsons, C. and S. Titman, 2007, "Capital Structure and Corporate Strategy," in B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 2, Chapter 13 (Elsevier/North-Holland, Handbooks in Finance Series).

Peck, S. W., 1996, "The Influence of Professional Investors on the Failure of Management Buyout Attempts," *Journal of Financial Economics*, 40, 264–294.

Perotti, E. and S. Rossetto, 2007, "Unlocking Value: Equity Carve Outs as Strategic Real Options, *Journal of Corporate Finance*, 13, 771–792.

Perotti, E. C. and K. E. Spier, 1993, "Capital Structure as a Bargaining Tool: The Role of Leverage in Contract Renegotiation," *American Economic Review*, 83, 1131–1141.

Perry, S. E. and T. H. Williams, 1994, "Earnings Management Preceding Management Buyout Offers," *Journal of Accounting and Economics*, 18, 157–179.

Peyer, U. C. and A. Shivdasani, 2001, "Leverage and Internal Capital Markets: Evidence from Leveraged Recaptializations," *Journal of Financial Economics*, 59, 477–515.

Phallipou, L. and O. Gottschalg, 2008, "Governance Effects of LBO Events," Working Paper, Syracuse University.

Phallipou, L. and O. Gottschalg, 2007, "The Performance of Private Equity Funds," Working Paper, University of Amsterdam.

Phillips, G. M., 1995, "Increased Debt and Industry Product Markets: An Empirical Analysis," *Journal of Financial Economics*, 37, 189–238.

Powers, E. A., 2003, "Deciphering the Motives for Equity Carve-outs," *Journal of Financial Research*, 26, 31–50.

Pulvino, T., 1998, "Do Asset Fire-sales Exist? An Empirical Investigation of Commercial Aircraft Transactions," *Journal of Finance*, 53, 939–978.

Pyo, U., 2s007, "Enhancing Managerial Incentives and Value Creation: Evidence from Corporate Spinoffs," *Journal of Empirical Finance*, 31, 341–358.

Rajan, R., H. Servaes and L. Zingales, 2000, "The Cost of Diversity: The Diversification Discount and Inefficient Investment," *Journal of Finance*, 55, 35–80.

Rappaport, A., 1990, "The Staying Power of the Public Corporation," *Harvard Business Review*, 1, 96–104.

Renneboog, L., T. Simons and M. Wright, 2007, "Why Do Firms Go Private in the UK? The Impact of Private Equity Investors, Incentive Realignment and Undervaluation," *Journal of Corporate Finance*, 13, 591–628.

Roden, D. M. and W. G. Lewellen, 1995, "Corporate Capital Structure Decisions: Evidence from Leveraged Buyouts," *Financial Management*, 24, 76–87.

Rosenfeld, J., 1984, "Additional Evidence on the Relation between Divestiture Announcements and Shareholder Wealth," *Journal of Finance*, 39(5), 1437–1448.

Ross, S., 1977, "The Determinants of Financial Structure: The Incentive Signaling Approach," *Bell Journal of Economics*, 8, 23–40.

Scharfstein, D. S. and J. C. Stein, 2000, "The Dark side of Internal Capital Markets: Divisional Rent-seeking and Inefficient Investment," *Journal of Finance*, 55, 2537–2564.

Schipper, K. and A. Smith, 1983, "Effects of Recontracting on Shareholder Wealth: The Case of Voluntary Spinoffs," *Journal of Financial Economics*, 12, 437–467.

Schipper, K. and A. Smith, 1991, "Effects of Management Buyouts on Corporate Interest and Depreciation Tax Deductions," *Journal of Law and Economics*, 34, 295–341.

Schipper, K. and A. Smith, 1986, "A Comparison of Equity Carve-outs and Seasoned Equity Offerings: Share Price Effects and Corporate Restructuring," *Journal of Financial Economics*, 15, 153–186.

Schleifer, A. and R. Vishny, 1992, "Liquidation Value and Debt Capacity: A Market Equilibrium Approach," *Journal of Finance*, 47, 1343–1365.

Schlingemann, F. P., R. M. Stulz and R. A. Walkling, 2002, "Divestitures and the Liquidity of the Market for Corporate Assets," *Journal of Financial Economics*, 64, 117–144.

Seoungpil, A. and D. J. Denis, 2004, "Internal Capital Markets and Investment Policy: Evidence of Corporate Spinoffs," *Journal of Financial Economics*, 71, 489–516.

Servaes, H., 1994, "Do Takeover Targets Invest," *Review of Financial Studies*, 7, 253–277.

Seward, J. K. and J. P. Walsh, 1995, "The Acquisition of Restructured Firms: An Illustration of Market," *Journal of Economics and Management Strategy*, 3, 585–603.

Sicherman, N. W. and R. H. Pettway, 1992, "Wealth Effects for Buyers and Sellers of the Same Divested Assets," *Financial Management*, 21, 119–128.

Slovin, M., M. Sushka and S. Ferraro, 1995, "A Comparison of the Information Conveyed by Equity Carve–outs, Spin–offs, and Asset Sell–offs," *Journal of Financial Economics*, 37, 89–104.

Slovin, M. B., M. E. Sushka and Y. M. Bendeck, 1991, "The Intra-industry Effect of Going-Private Transactions," *Journal of Finance*, 46, 1537–1550.

Smith, A. J., 1990, "Corporate Ownership Structure and Performance," *Journal of Financial Economics*, 27, 143–164.

Stromberg, P., 2007, "The New Demography of Private Equity," Working Paper, Stockholm Institute of Financial Research.

Stulz, R. M., 1990, "Managerial Discretion and Optimal Financing Policies," *Journal of Financial Economics*, 26, 3–27.

Tehranian, H., N. G. Travlos and J. F. Waegelein, 1987, "The Effect of Long–term Performance Plans on Corporate Sell–off–induced Abnormal Returns," *Journal of Finance*, 42, 933–942.

Van de G., M. Linda and W. T. Moore, 1998, "Predicting the Duration and Reversal Probability of Leveraged Buyouts," *Journal of Empirical Finance*, 5, 299–315.

Veld, C. and Y. V. Veld-Merkoulova, 2004, "Do Spin-offs Really Create Value? The European Case," *Journal of Banking and Finance*, 28, 1111–1135.

Vijh, A., 1999, "Long–term Returns from Equity Carve outs," *Journal of Financial Economics*, 54, 273–308.

Vijh, A. M., 2006, "Does a Paren-Subsidiary Structure Enhance Financing Flexibility," *Journal of Finance*, 59, 1337–1360.

Vijh, A. M., 2002, "The Positive Announcement-Period Returns of Equity Carveouts: Asymmetric Information or Divestiture gains?" *Journal of Business*, 75, 153–190.

Vijh, A. M., 1994, "The Spinoff and Merger Ex-date Effects," *Journal of Finance*, 49, 581–609.

Wagner, H. F., 2004, "The Equity Carve-out Decision," Working Paper, University of Munich.

Walker, M., 1998, "Leverage Recapitalization, Operating Efficiency, and Stockholder Wealth," *Financial Review*, 33, 88–114.

Warga, A. and I. Welch, 1993, "Bondholder Losses in Leveraged Buyouts," *Review of Financial Studies*, 6, 959–982.

Williamson, O. E., 1985, *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting* ( Free Press, New York).

Wright, M., S. Thompson and K. Robbie, 1992, "Management-led Leveraged Buy-outs: A European Perspective," *Journal of Business Venturing*, 7, 45–71.

Wruck, E. G. and K. H. Wruck, 2002, "Restructuring Top Management: Evidence from Corporate Spinoffs," *Journal of Law and Economics*, 20, S176–218.

Wruck, K. H., 1994, "Financial Policy, Internal Control, and Performance: Sealed Air Corporation's Leveraged Special Dividend," *Journal of Financial Economics*, 36, 157–192.

Zaima, J. K. and D. Hearth, 1985, "The Wealth Effects of Voluntary Selloffs: Implications for Divesting and Acquiring Firms," *Journal of Financial Research*, 8, 227–236.

*Chapter 17*

# EXECUTIVE COMPENSATION AND INCENTIVES

RAJESH K. AGGARWAL

*Carlson School of Management, University of Minnesota*

## Contents



*Handbook of Empirical Corporate Finance, Volume 2*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

## Abstract

Since the seminal contribution of Jensen and Murphy (1990), our understanding of executive compensation and incentives has greatly improved. Through time, the strength of incentives has increased. In the 1990s, this was accomplished primarily through the use of stock options. The two most important sources of aggregate incentives are holdings of stock and holdings of stock options—these align managers' interests with those of shareholders. Firm size, firm risk, executive aversion to risk, executive productivity, the extent to which executives like or dislike taking certain actions that matter for shareholders, and characteristics of the industry all determine how strong or weak incentives should be. We know that, because shareholders have limited information about what managers do and limited ability to monitor managers, managers need to be provided with some incentives. The key limitation on the provision of incentives is the need to share risk between managers and other shareholders. Actual compensation practices reflect this trade-off. We also know now that boards and shareholders do not have the complete ability to set managerial compensation. Either because of CEO power or a limited supply of managerial talent, at least some CEOs have the ability to extract more compensation than what is dictated by straightforward pay for performance. At this stage, we have a reasonably strong understanding of the determinants of incentives, and incentives seem to work well. We still do not fully understand the level of compensation. A number of well-documented examples are consistent with the rent extraction view of the level of compensation. Instances of excessive compensation and rent extraction seem to be correlated with corporate governance failure, accounting fraud, and poor corporate outcomes.

## Keywords

Executive compensation, incentives, agency, rent extraction

## 1. Introduction

Modern finance theory argues that the proper objective of managers is to maximize the value of the firm, which in general means running the firm in the shareholders' interests. But managers are self-interested individuals. Because managers have effective control over the firm, they can make decisions and take actions that serve their own interests. A common presumption is that the best way to ensure that managers act in the interest of shareholders is to tie their compensation to the company's performance. The higher the sensitivity of executives' pay to firm performance, the closer their interests will be aligned to those of the shareholders. Much of the academic literature has focused on the degree to which managers' interests are aligned with those of shareholders. Understanding this issue and how it informs the optimal design of compensation packages will be the major focus of this survey.

In public discourse about executive compensation, the dominant theme has been the soaring pay packages of chief executive officers (CEOs) and the contribution executive pay has made to rising income inequality in the United States. By 2004, according to *Business Week*, the average CEO compensation package, including bonuses and stock options, was roughly 300 times the average worker's paycheck. (In 2004, *Business Week* revised its calculation method, and so the multiple went down relative to previous years.) This theme has been picked up in recent academic work proposing that the driving force behind CEO compensation is rent extraction by CEOs (Bebchuk and Fried, 2004; Bertrand and Mullainathan, 2001). As this is now the leading alternative to agency explanations for compensation and incentives, this perspective will also be discussed. While earlier discussions of compensation focused on the incentives embodied in compensation and abstracted from concerns about the level of compensation, it is now fair to say that the level of compensation matters as well. For example, in the seminal contribution to the literature on executive compensation, Jensen and Murphy (1990) maintained that political considerations about income inequality limited the amount of incentives that could be provided. Almost 20 years later, after the options explosion, it is difficult to argue that CEO incentives have been limited by such considerations.

This survey begins by examining trends and components in executive compensation in Section 2. Section 3 then discusses what determines incentives, with a particular emphasis on the agency view. Section 4 turns to one of the larger puzzles in the literature: the general absence of the use of relative performance evaluation when compensating and providing incentives to top managers. Section 5 examines what managerial actions incentives should or do influence and, in turn, the associated impact on firm value. Section 6 looks at some alternatives to the agency view, including the increasingly popular skimming or rent extraction view of executive compensation. Section 7 presents conclusions.[1]

[1] Two remarks are in order about the scope of this survey. First, many important papers have necessarily been omitted from this survey, both because of the choice of topics covered and simply out of the need for brevity. Second, this survey focuses relatively more heavily on the literature after 1998, which was written after Kevin Murphy's (1999) earlier and excellent survey of executive compensation.

## 2. Trends in executive compensation

In order to understand the determinants of incentives and, in turn, what incentives are designed to influence, we start by examining the components of executive compensation. As an illustrative example, consider the case of top management at IBM. Figure 1 presents executive compensation data for the five highest paid executives at IBM for fiscal year 2003. The figure includes data on salary, bonus, other annual compensation, restricted stock grants, grants of stock options, long-term incentive payouts (LTIPs), and all other compensation. It does not include data on stock options exercised, sales of stock, or existing holdings of stock or options. The components included add up to the total annual compensation of the executive.

Total annual compensation can be divided into two categories—short-term and long-term components of compensation.

### 2.1. Short-term components of compensation

Short-term components of compensation include salary, bonus, and other annual compensation. Annual salary is fixed in advance and generally does not have an incentive component associated with it. The exception to this statement is that future increases in salary may in part be determined by current firm performance. Jensen and Murphy (1990) show that the present value of current and future increases in salary and bonuses represents a small fraction of total incentives.

Annual bonuses are typically tied to measures of firm performance. Interestingly, the performance measures are often based on accounting information such as earnings, sales, or operating income. Common metrics employed include return on equity (ROE), return

| Name | Title | Salary | Bonus | Other Annual | Restricted Stock Granted | Stock options Granted* | LTIP | All Other Comp. | Total Annual Comp. |
|---|---|---|---|---|---|---|---|---|---|
| S. J. Palmisano | Chmn., Pres., & CEO | 1,550,000 | 5,400,000 | 11,037 | 0 | 12,283,900 | 769,095 | 181,500 | 20,195,532 |
| N. M. Donofrio | Sr. V-P | 637,501 | 1,185,000 | 1,390 | 0 | 2,900,400 | 425,566 | 36,075 | 5,185,932 |
| J. R. Joyce | Sr. V-P & CFO | 615,000 | 1,141,000 | 205 | 1,363,246 | 2,872,100 | 512,730 | 34,950 | 6,539,231 |
| D. T. Elix | Sr. V-P & Group Exec. | 622,501 | 1,010,000 | 764,092 | 1,187,740 | 3,296,100 | 553,748 | 33,925 | 8,098,106 |
| W. M. Zeitler | Sr. V-P & Group Exec. | 487,501 | 995,000 | 2,928 | 908,831 | 2,555,500 | 425,566 | 24,125 | 5,399,451 |
| *Assumes 5% annual stock price appreciation, not a Black-Scholes value. | | | | | | | | | |

Fig. 1. Executive compensation, IBM Corp., fiscal year 2003.
*Source*: IBM Proxy Statement.

on assets (ROA), return on investment (ROI), and economic value added (EVA). Other measures of performance include subjective reports by board members or superiors for lower ranking executives and targets established by the board for investment, product or plant quality (e.g., "zero-defects"), market share, growth rates for income or sales, strategic objectives (e.g., expansion into new lines of business or restructuring of old businesses), and performance relative to that of industry competitors.

There are several points to note about short-term components of compensation. Short-term compensation does have some incentive features, especially the bonus component. Swan and Zhou (2003) argue that bonuses typically have thresholds associated with them (e.g., ROA must exceed 15%) and that performance incentives are quite high-powered around these thresholds. However, bonuses and other forms of short-term compensation are typically not linked to stock performance in the form of stock returns. Given that shareholders presumably care most about stock returns, this might seem surprising. It will become apparent, however, that long-term components of compensation are much more strongly linked to stock returns.

For this reason, the right way to think about annual salary and bonuses is that salary provides the executive with a minimum level of income prior to any performance standards or targets being met. Bonuses typically reflect how well the firm or executive has met nonstock-return-based objectives established by the board. Other annual compensation is usually negligible.[2] When looking at IBM in Figure 1, Chairman, President, and CEO Samuel J. Palmisano's salary and bonus stand out as much higher than those of the other IBM executives.

### 2.2. Long-term components of compensation

Long-term components of compensation include new grants of restricted stock, new grants of stock options, long-term incentive plan payouts, and all other compensation. All other compensation typically includes gross-ups for tax liabilities, perquisites, preferential discounts on stock purchases, contributions to benefit plans, and severance payments and has historically been thought to be relatively unimportant.

Recent work by Yermack (2006), however, suggests that some forms of perquisites—the use of corporate jets—may actually be indicative that management is more interested in rent extraction than in shareholder value maximization.[3] Consistent with this view, Aggarwal, Meschke, and Wang (2007) argue that corporate political donations are also a form of perquisites consumption that may indicate management's greater interest in rent extraction than in shareholder value maximization. In addition, recent disclosures of executive compensation packages given to Jack Welch, former CEO of GE, and

[2] It is the case, however, that other annual compensation can include various forms of perquisites consumption as discussed in the next section.

[3] Rajan and Wulf (2006) offer a contrasting perspective on perquisites such as the use of corporate jets. They state that corporate jets, chauffeur services, and country club memberships may enhance productivity and may therefore be valuable to the firm, and they offer evidence consistent with this view.

Richard Grasso, former CEO of the New York Stock Exchange, suggest that deferred compensation in the form of retirement benefits can be quite substantial. Bebchuk and Fried (2004) call retirement benefits "stealth" compensation. This view is discussed in greater detail later in the chapter.

### *2.2.1. Restricted stock*

Restricted stock grants are restricted in the sense that the executive must remain with the firm for a specified amount of time in order not to forfeit the stock grant. Restricted stock grants with a five-year vesting period are typical. This restriction has two practical implications. First, the executive has potentially a strong incentive to stay with the firm in order to benefit from the grant (see Oyer, 2004, for a discussion of retention). Second, while the vesting period is in effect, the executive cannot sell the stock. She is, in effect, forced to have part of her compensation tied to firm performance over the vesting period. Restricted stock grants clearly align an executive's interests with those of her shareholders.

For this reason, it is interesting that the use of restricted stock grants declined throughout the 1990s. As an incentive mechanism, they were supplanted by stock options. Since the 2000 stock market decline, the use of restricted stock has increased, and the advent of stock option expensing will most likely accelerate the use of restricted stock. Figure 1 shows that of the five highest paid executives at IBM, three received restricted stock in 2003.

### *2.2.2. Stock options*

Throughout the 1990s, stock options were the primary mechanism through which managers' interests were aligned with those of shareholders. Most stock options are granted at the money. A typical stock option grant has a life of 10 years. Interestingly, few options are held to maturity; most options are exercised early, and the stock is sold. Since stock prices on average increase from year to year, over time most stock options will move into the money, although Hall and Knox (2004) argue that a substantial fraction of option grants will be underwater for some period during their life. Stock options usually have a vesting schedule associated with them, such as 25% of an option grant vests every year (with 25% vesting immediately upon grant), so that the full grant vests over three years, with another seven years to maturity.

Historically, for purposes of reported earnings, options granted at the money or out of the money did not count against earnings. Options that are in the money or options that have a variable exercise price (such as indexed options where the option is indexed to aggregate stock market returns) did count against earnings. This explains why most options are issued at the money and why we so rarely see indexed options. Going forward, options will have to be expensed, which decreases the likelihood that we will see as many or as large option grants in the future. Furthermore, option expensing implies that we are more likely to see indexed options and other forms of more tailored option grants to top executives.

For tax purposes, the focus here is on nonqualified options, which are the form of stock options that most executives receive. Nonqualified options have no tax implication at the time that they are issued. When the option is exercised, the executive pays tax on the difference between the stock price and the exercise price at the ordinary income tax rate. The firm deducts the difference between the stock price and the exercise price as compensation expense. If the executive later sells the stock, then the executive pays tax on the difference between the sale price and the market price at exercise of the option at the capital gains tax rate. Because the firm is able to deduct the difference between the stock price at exercise and the exercise price as compensation expense, nonqualified options have favorable tax treatment from the firm's perspective.

Favorable tax treatment is a significant part of the explanation for dramatic increase in the use of stock options. Favorable accounting treatment also explains why the use of stock options increased over the 1990s. For the most part, stock options never show up on accounting statements. In the past, firms were required to disclose grants of stock options but did not have to take an accounting charge for them. As a result, stock options were an excellent way of providing managers with deferred compensation without incurring an accounting liability, even though there is a real economic cost associated with the option grant (see Bulow and Shoven, 2005, for an excellent discussion of how this cost can be recognized dynamically over the life of an option). The fact that favorable accounting treatment only accrues to options that are granted at or out of the money with a prespecified exercise price and date helps to explain the prevalence of granting stock options at the money. Murphy (2002) argues that this favorable accounting treatment leads companies to erroneously view stock options as a low-cost mechanism for compensating executives.

Historically, stock options have been reported in two ways in company proxy statements. The first is to report the value of an option grant assuming the stock price increases by 5 or 10% annually. This is what IBM does, and this is what is reported in Figure 1. This method will generally overstate the value of the option grant. A more satisfactory approach from a researcher's point of view is to use the Black-Scholes formula. The most salient point about using the Black-Scholes model is that it may also overstate the value to the executive of the option grant. Hall and Murphy (2002) argue that because executives are risk averse and hold large, undiversified positions in their own firms, they will value stock options at a lower level than will a well-diversified outside investor.

Conversely, executives have better information than do outside investors about their firms' prospects. Yermack (1997) notes that executives receive option grants before good news is announced and often exercise options in advance of bad news. More recently, Lie (2005) and Heron and Lie (2007) have argued that stock option grants to many top executives have grant dates that are suspiciously the date of the firm's lowest stock price (for the quarter or the year). Since exercise prices are set on the grant date, this exceptionally good timing allows executives to reap windfall gains. They argue that this pattern is consistent with options being backdated—the actual grant date is weeks or months after the chosen grant date.

IBM is quite typical in its grants of stock options. Stock options had become the dominant mechanism through which incentives were provided to managers. Clearly, stock options are tied to the firm's stock price performance. In many situations, they do a good job of aligning a manager's interests with those of the shareholders. This becomes more true as options move into the money. However, a case can also be made that stock options became the dominant mechanism through which executives were able to extract rents from firms. The example of option backdating illustrates one form of rent extraction, as is discussed later.

### *2.2.3. Long-term incentive plans*

Long-term incentive plan payouts are similar to bonuses, but they are awarded for performance over several years. For example, a long-term incentive plan payout may be triggered if ROA is at least 15% for three consecutive years. In general, long-term incentive plans are not that important on a year-to-year basis because they occur only when a long-term target is met. For IBM in 2003, however, long-term incentive plan payouts were quite important for all five executives. As is clear from Figure 1, IBM's top five highest paid executives were quite well compensated in 2003.

## *2.3. Aggregate measure of compensation*

The measure of compensation for incentive purposes that we will use is the total annual compensation of the executive (as in Fig. 1 for IBM), plus the annual change in wealth for the executive based on changes in firm value. Total annual compensation includes salary, bonus, new grants of restricted stock, new grants of stock options, long-term incentive plan payouts, gross-ups for tax liabilities, perquisites, preferential discounts on stock purchases, contributions to benefit plans, and severance payments. The most important elements of total annual compensation given by the firm in terms of magnitude are usually salary, bonus, grants of restricted stock, and grants of stock options. (IBM is a bit of an exception with its large long-term incentive payouts.)

The annual change in wealth for the executive includes changes in the value of existing holdings of stock and existing holdings of stock options; these elements often swamp all of the components of total annual compensation. The annual change in wealth focuses on how large a position the executive holds in the firm. For this reason, it is a good measure of how aligned an executive's interests are with those of other shareholders. At the same time, at least some of an executive's holdings of stock and options are voluntary in the sense that the shares and options are not restricted from being sold. Because some of the holdings are voluntary, the shareholders are not, strictly speaking, providing incentives. While this issue is real, it is also worth noting that the stock market reaction to insider sales is typically negative, implying that there are some limitations on an insider's ability to sell shares. Overall, the annual change in wealth captures the bulk of incentives provided to executives.

### 2.4. Compensation statistics

Figure 2 provides summary information about the elements of compensation. The data are from Standard and Poor's ExecuComp dataset (see Standard and Poor's, 1995). The data cover the 1500 firms included in the S&P 500, S&P MidCap 400, and S&P SmallCap 600 from 1993 to 1997. For each firm, data are reported for the top five executives ranked annually on the basis of salary and bonus. Figure 2 divides executives into four categories based on the classification scheme in Aggarwal and Samwick (2003b). The first category is chief executive officers (CEOs). The second is executives with oversight authority for the firm—for example, executives with titles such as president

| Group and Variable | Mean | Median | Standard Deviation |
|---|---|---|---|
| Total Annual Compensation (in 1997 dollars) | | | |
| CEO | $2,739,000 | 1,494,000 | 5,183,000 |
| Oversight | 1,532,000 | 847,000 | 3,461,000 |
| Divisional | 973,000 | 605,000 | 1,324,000 |
| Neither | 873,000 | 555,000 | 1,254,000 |
| Long-Term Components of Total Annual Compensation (in 1997 dollars) | | | |
| CEO | 1,571,000 | 543,000 | 4,381,000 |
| Oversight | 879,000 | 296,000 | 3,252,000 |
| Divisional | 488,000 | 190,000 | 1,054,000 |
| Neither | 417,000 | 170,000 | 985,000 |
| Holdings of Stock (in 1997 dollars) | | | |
| CEO | 74,606,000 | 5,801,000 | 722,605,000 |
| Oversight | 17,039,000 | 1,079,000 | 117,583,000 |
| Divisional | 2,496,000 | 422,000 | 15,908,000 |
| Neither | 4,898,000 | 475,000 | 98,923,000 |
| Holdings of Options (in 1997 dollars) | | | |
| CEO | 7,968,000 | 2,218,000 | 22,664,000 |
| Oversight | 3,151,000 | 887,000 | 11,320,000 |
| Divisional | 2,018,000 | 583,000 | 7,964,000 |
| Neither | 1,776,000 | 578,000 | 4,514,000 |

Fig. 2. Descriptive statistics on compensation by job classification, 1993–1997.

*Note*: Long-term compensation is comprised of the following components of total annual compensation: grants of restricted stock, grants of stock options, long-term incentive plan payouts, gross-ups for tax liabilities, perquisites, preferential discounts on stock purchases, contributions to benefit plans, and severance payments. *Source:* Aggarwal and Samwick (2003b).

(non-CEO), chief operating officer (COO), chief financial officer (CFO), chairman, and vice chairman. The third is executives with divisional responsibility—CEOs of divisions or subsidiaries, executives in charge of a specific product line or geographical area, and executives with specific production-related responsibilities. The fourth is executives with neither oversight authority nor divisional responsibility—for example, vice presidents with no other information given.

### *2.4.1. Total annual compensation*

Figure 2 highlights a number of important features of executive compensation. Executives, and especially CEOs, seem to be quite well paid. As shown in the figure, in the 1992–1997 period, the average CEO made $2,739,000 in total annual compensation given by the firm. The average is skewed by the fact that some CEOs made enormous amounts of money in some years. Figure 3 lists the 20 highest paid CEOs in 1999, plus Kenneth Lay and Jeffrey Skilling from Enron. The spread in compensation among even

| Name | Company | Total Annual Compensation |
|---|---|---|
| **1. Robert Annunziata** | **Global Crossing, Ltd.** | **193,784,118** |
| **2. Joseph Nacchio** | **Qwest Communication Intl., Inc.** | **172,205,151** |
| **3. L. Dennis Kozlowski** | **Tyco International Ltd.** | **138,331,617** |
| 4. Thomas Siebel | Siebel Systems, Inc. | 134,437,170 |
| 5. Michael Jeffries | Abercrombie & Fitch | 124,513,616 |
| 6. Sanford Weill | Citigroup, Inc. | 109,426,548 |
| 7. Carleton Fiorina | Hewlett-Packard Co. | 101,299,548 |
| 8. Hugh McColl, Jr. | Bank of America Corp. | 84,825,255 |
| 9. John Welch, Jr. | General Electric Co. | 70,366,772 |
| 10. H. Brian Thompson | Global Telesystems, Inc. | 66,548,901 |
| 11. John Chambers | Cisco Systems, Inc. | 64,005,282 |
| 12. Charles Lillis | MediaOne Group, Inc. | 62,515,711 |
| 13. Harvey Golub | American Express Co. | 56,996,770 |
| 14. William McGuire | UnitedHealth Group, Inc. | 51,611,855 |
| 15. John Wren | Omnicom Group | 50,671,565 |
| 16. Robert Davis | Lycos, Inc. | 50,526,628 |
| **17. Bernard Ebbers** | **WorldCom, Inc.** | **50,374,936** |
| 18. William Esrey | Sprint PCS Group | 47,821,117 |
| 19. Michael Bonsignore | Honeywell Intl. Inc. | 45,377,623 |
| 20. Timothy Koogle | Yahoo, Inc. | 44,648,351 |
| **Kenneth Lay** | **Enron** | **22,154,808** |
| **Jeffrey Skilling** | **Enron** | **17,767,941** |

Fig. 3. Twenty highest paid CEOs—1999.
*Source*: Standard & Poor's ExecuComp database.

the 20 highest paid is quite dramatic—from $193,784,118 for Robert Annunziata, the CEO of Global Crossing, to $44,648,351 for Timothy Koogle, the CEO of Yahoo. Of the almost $194 million in compensation Robert Annunziata received, $182.319 million was in the form of a stock option grant (Black-Scholes value). In contrast, Lay and Skilling, though obviously highly compensated, do not particularly stand out.

As shown in Figure 2, the median CEO made $1,494,000 over the period, substantially less than the average CEO. Other executives made smaller amounts, although in all cases executive compensation is quite substantial. Executives with oversight authority made, on average, $1,532,000, with a median of $847,000. Executives with divisional responsibility made, on average, $973,000, with a median of $605,000. Executives with neither oversight authority nor divisional responsibility made, on average, $873,000, with a median of $555,000. While these sums are substantial, they are not as large as is often portrayed in the media. For example, the average worker made about $33,000 in 1997. The median CEO's total annual compensation was 45 times that of the average worker, not the 419 times claimed by *Business Week*. *Business Week* focused on the compensation packages of CEOs at the largest firms, which, as Figure 2 shows, is not representative of all firms. Clearly, politics matters in any discussion of CEO compensation. But comparisons of dollars of compensation cannot tell us much about the strength of incentives or how well aligned managers' interests are with those of shareholders.

### *2.4.2. Holdings of stock and options*

Although the sums paid to executives on an annual basis are large, they pale in comparison to the dollar value of executives' holdings of stock and options, especially, CEOs. On average, the dollar value of CEOs holdings of stock is $74,606,000, while the dollar value of options held is $7,968,000. However, when dealing with estimates of executive compensation, one should always be concerned about the skewness of the data. For example, the median dollar value of holdings of stock for CEOs is $5,801,000. Much of the reason the mean dollar value of stockholdings is so much greater than the median can be attributed to a few outliers. In particular, the holdings of one man, Bill Gates, who was CEO of Microsoft over the sample period, have a pronounced effect on the results. For a brief period in 1999, the dollar value of Bill Gates's holdings in Microsoft exceeded $100 billion. Over the sample period, the dollar value of Bill Gates's holdings averaged about $20 billion. Figure 2 also shows that the dollar value of holdings of stock and options for executives other than the CEO are quite substantial and quite skewed. These holdings of stock and options provide the bulk of incentives for all executives.

When looking at medians, holdings of stock are more important for CEOs and executives with oversight authority than are holdings of options. Conversely, for executives with divisional responsibility and executives with neither divisional responsibility nor oversight authority, options are more important than holdings of stock. This is not true when looking at average dollar values of holdings. Mean holdings of stock are greater than mean holdings of options for all four job groups. Again, medians are more likely to provide a representative picture of the dollar value of stock and option holdings.

One reason holdings of stock are more important for CEOs and executives with oversight authority than for executives with divisional responsibility and executives with neither divisional responsibility nor oversight authority is that the former two groups are more likely to include founders of the firm. Founders typically have large equity positions and may not participate in option programs (a good example of this is Bill Gates of Microsoft). Executives in the divisional and neither groups are less likely to be founders or early employees, and so may have fewer equity holdings. The dramatic extension of stock option grants to employees lower in the firm hierarchy is consistent with this pattern (for a discussion, see Oyer and Schaefer, 2004).

## 3. Incentives and agency

Given these basic facts about executive compensation, one may well wonder how compensation is translated into incentives and how incentives are determined. This section discusses some key issues in the design of incentives.

### *3.1. Pay-performance sensitivities*

Incentives have historically been calculated as pay-performance sensitivities (PPSs). The idea behind the pay-performance sensitivity is to see how much compensation depends on how well the firm performs. Typically, firm performance is measured either in terms of stock returns or accounting returns (e.g., net income divided by the book value of equity or assets). In this section we will focus on firm performance based on stock returns. Pay-performance sensitivities can be calculated in two ways (Murphy, 1999). Under the implicit method, total annual compensation (or its logarithm) is regressed on firm performance (measured either as the dollar change in firm market value or the percentage change in firm market value). The coefficient on firm performance is the PPS. Estimates from the literature using the implicit method find that CEOs receive between \$3.25 (Jensen and Murphy, 1990) and \$5.29 (Hall and Liebman, 1998) for every thousand-dollar increase in shareholder wealth. Aggarwal and Samwick (1999a) show that these estimates are biased downward because they do not take into account the risk executives face from having compensation tied to firm performance. For this reason, the alternative method—the explicit method—will often be preferred to the implicit method.

Under the explicit method, incentives are calculated more directly. An executive's holdings of stock as a fraction of the total equity outstanding provide a measure of explicit pay-performance sensitivity from stockholdings. The number of shares options are written on divided by the total number of shares outstanding multiplied by the options' delta provides a measure of explicit PPS from options. An option's delta measures how the value of the option changes with a change in the price of the underlying stock on which the option is written. For options that are far in the money, the delta will be close to one and the option will have incentive effects similar to that of a share of the

| Group and Variable | Mean | Median | Standard Deviation |
|---|---|---|---|
| Stock and Option PPS | | | |
| CEO | 3.9941 | 1.2769 | 7.1526 |
| Oversight | 1.2504 | 0.3137 | 3.5974 |
| Divisional | 0.3750 | 0.1835 | 0.8289 |
| Neither | 0.4336 | 0.1765 | 1.0910 |
| Stock PPS | | | |
| CEO | 3.1658 | 0.3883 | 7.0067 |
| Oversight | 0.9112 | 0.0622 | 3.5339 |
| Divisional | 0.1626 | 0.0281 | 0.0307 |
| Neither | 0.2245 | 0.0307 | 1.0000 |
| Option PPS | | | |
| CEO | 0.8283 | 0.4153 | 1.2027 |
| Oversight | 0.3393 | 0.1609 | 0.5867 |
| Divisional | 0.2124 | 0.1160 | 0.3805 |
| Neither | 0.2092 | 0.1048 | 0.3641 |

Fig. 4. Descriptive statistics on incentives by job classification.
*Note*: Pay-performance sensitivities (PPSs) reflect the executives' percentage ownership of the firm on a scale of 0 to 100.
*Source*: Aggarwal and Samwick (2003b).

stock. For options that are far out of the money, delta will be close to zero, and the option will have no incentive effects given that it is unlikely to be exercised. An average value of delta for firms in the sample would be around 0.7 (see Hall and Liebman, 1998). The overall explicit PPS for an executive is the sum of the explicit pay-performance sensitivity from stock and the explicit pay-performance sensitivity from options. Summary statistics for these sensitivities are provided in Figure 4.

For CEOs, the average stock and option explicit PPS is 3.99%. That is, we can think of the CEO's holdings of stock and options as being 3.99% of the total equity outstanding in the firm. Pay-performance sensitivities are often reported as the dollars that accrue to an executive from a thousand dollar increase in shareholder wealth. On average, then, a CEO receives \$39.94 from changes in the value of his stock and option holdings for a thousand dollar increase in shareholder wealth. Once again, the medians are smaller than the means. At the median, a CEO receives \$12.77 from changes in the value of his stock and option holdings for a thousand dollar increase in shareholder wealth.

Several other points stand out. First, the skewness of explicit pay-performance sensitivities is much more pronounced for stock than it is for options. Second, CEOs have much greater explicit PPSs than do the other groups of executives. Third, at the median, all four groups of executives have greater explicit PPSs from options than they do from stock. To a large extent, this is a reflection of the fact that options became the incentive device of choice for most companies over the 1990s.

The following figures (Figs. 5 and 6) graphically illustrate the evolution of the use of stock and options as incentive devices from 1993 to 1997.

Although PPSs from stockholdings have remained relatively constant over the sample period, PPSs from options have shown a dramatic increase. From 1993 to 1997, PPSs from options for CEOs more than doubled. The same is true for executives with oversight authority. While the PPSs from options have shown large increases for executives with divisional responsibility and executives with neither responsibility, these increases have not been as large as the increases for CEOs and executives with oversight authority. Aggarwal and Samwick (2003b) argue that this is a reflection of the fact that

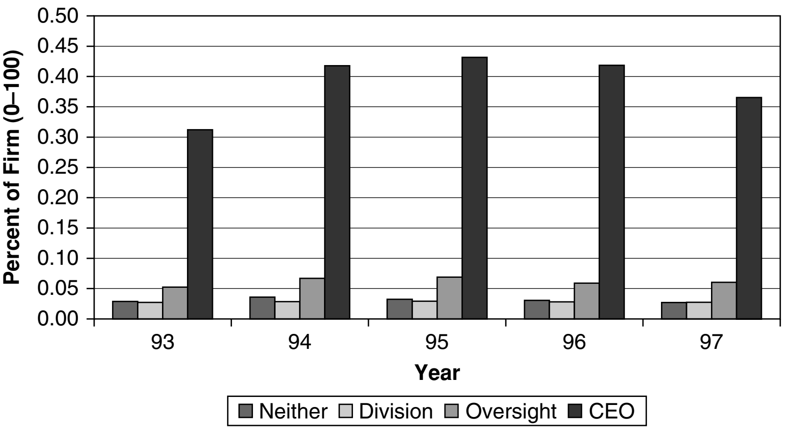



Fig. 5. Median pay-performance sensitivities from stock.

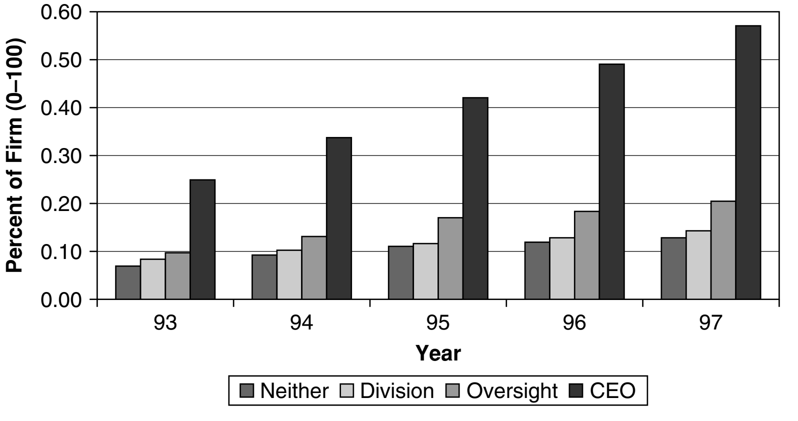



Fig. 6. Median pay-performance sensitivities from options.

overall firm returns are a better measure of a CEO's or an oversight executive's performance than they are of a divisional executive's performance or the performance of an executive with neither divisional responsibility nor oversight authority. In other words, CEOs and oversight executives should have more sensitivity of pay to performance. Barron and Waddell (2003) provide a complementary explanation for this finding. They rank executives by compensation and find that higher-paid executives receive more equity-based incentives. They argue that this is because it is more costly to the firm for a higher ranking executive to make mistakes in project evaluation.

The rise in pay-performance sensitivities from options corresponds to a large rise in the stock market as a whole. Rising stock prices will increase the PPS of existing grants of stock options because the options' deltas increase as options move further into the money. This effect is not driving the full increase in PPSs from stock options.

### *3.2. How should incentives be set?*

Now that we have some sense of the magnitudes involved, how should incentives be set? The answer usually offered by academics is simple to state but fairly hard to implement. Incentives generally ought to be set to maximize the value of the equity held by the shareholders. Berle and Means (1932) noted that the separation of ownership and control in modern corporations creates an agency problem. Since most managers (executives) do not own all or even a substantial fraction of the equity of the firm, they will be more inclined to take actions that are in their own interests rather than in the shareholders' interests. On the other hand, managers are often fairly indispensable, so that eliminating them is not really feasible.

Jensen and Meckling (1976) later gave the agency problem more formal structure. Suppose that there is a single manager—the CEO. At the margin, suppose the manager can take an action that generates an additional dollar for the firm but that costs the manager (almost) a dollar to generate it. If the manager does not get the full dollar generated for the firm because she does not own all of the equity, then she will not take the action in the first place. As the manager's equity ownership decreases, this problem becomes increasingly acute. More and more value-enhancing actions are foregone by the manager. In addition, more and more value-decreasing actions are taken that benefit the manager at the expense of the shareholders.

Examples of this type of behavior include managers choosing not to work hard (shirking), managers choosing to invest too much (overinvestment or empire-building), managers investing too little (underinvestment), managers inefficiently diversifying the firm, managers overspending on corporate headquarters, corporate jets, and other forms of perquisites. Perhaps the best known anecdote describing this behavior is from the takeover of RJR Nabisco. The stories of managers at RJR Nabisco squandering shareholders' cash on corporate jet rides for dogs and celebrity golf tournaments are quite powerful; see Burrough and Helyar's (1990, p. 95). colorful description. This argument has been tremendously influential in the corporate finance and corporate governance literatures.

Jensen (1986, 1993) focuses specifically on situations in which managers choose to overinvest. Jensen (1986) examines overinvestment in the oil and gas industry. He notes that long after oil prices had collapsed in the mid-1980s, oil and gas companies continued to explore and drill for oil. Shareholders would have been better off if development had ceased. Jensen (1993) focuses on value destruction at some of the world's largest corporations. In many cases, shareholders would have been better off if management had not invested in the business and instead, simply distributed the cash flow to the shareholders. In both cases, the inefficiency can be attributed to the fact that the manager does not own all of the equity in the firm.

So why is it that the manager does not own all of the equity in the firm? Generally, having the manager own all of the equity exposes her to too much risk (Mirrlees, 1999; Holmstrom, 1979). As a measure of firm performance, equity returns are subject to random fluctuations that are outside of the manager's control. If the manager is averse to risk, then full ownership of the equity is inefficient. Instead, the manager should share the risk by selling equity to other investors.

If these investors could count on the manager to take the optimal actions that maximize the value of the equity, then the manager should bear no risk and sell all of the equity to investors. In this situation, the manager should be given a fixed wage. Similarly, if investors can perfectly observe what the manager is doing, then there is no reason for the manager to hold equity. Instead, if the manager does not do what the shareholders say she should do, she can be fired.

Of course, as Berle and Means (1932) and Jensen and Meckling (1976) point out, managers are self-interested, and the actions they want to take may substantially differ from the actions shareholders want managers to take. In addition, rarely can shareholders perfectly observe the actions taken by managers (moral hazard). Even if the actions are observable, shareholders may not know enough to judge whether the action taken is the right one. This is especially true in a world in which there are many dispersed shareholders, few of whom own enough shares in any given company to make it worth their while to carefully monitor management (Grossman and Hart, 1980). As a result, management does have to be given incentives to take the right actions.

This trade-off between inducing the manager to take the right actions while not exposing her to too much risk governs how large incentives can be. The implication of this trade-off is that, in general, it will not be optimal for the manager either to have no equity stake or a full 100% equity stake in the firm. The size of incentives (the manager's equity stake or PPS) will be determined by how risk averse the manager is, how volatile the firm's stock is, and how much the manager dislikes taking actions that increase firm value or how much the manager likes actions that decrease firm value. Under a number of assumptions, Holmstrom and Milgrom (1987) derive the pay-performance sensitivity as

$$\alpha^* = \frac{1}{1 + rk\sigma^2}$$

where $r$ is the coefficient of absolute risk aversion of the agent, $k$ is the curvature of the agent's disutility of effort function, and $\sigma^2$ is the variance of the performance measure (frequently, stock price performance).

Jensen and Murphy (1990) present what is still perhaps the best-known estimate of the pay-performance sensitivity for CEOs. They carefully aggregate the pay-performance incentives from a variety of sources into a single PPS. They find that the typical CEO receives approximately \$3.25 of compensation per thousand-dollar increase in shareholder wealth. Of this amount, \$2.50 is due to the median CEO's holdings of stock in the firm, and \$0.15 is due to ownership of stock options. Increases in the present value of current and future compensation and decreases in the probability of dismissal are responsible for \$0.30 each.

Hall and Liebman (1998) show that incentives from stock and particularly stock options have grown substantially since the sample period used by Jensen and Murphy (1990). They calculate that the median CEO has a pay-performance sensitivity of \$5.29 per thousand-dollar increase in shareholder wealth. These estimates are dramatically lower than what is implied by 100% equity ownership—an increase in the manager's wealth of \$1000 when firm value increases by \$1000.

Jensen and Murphy (1990) argue that political forces and concerns with fairness make it quite difficult for managers to have pay that is strongly linked to performance. The implication is that these noneconomic forces cause firms to provide suboptimal incentives to managers. While Hall and Liebman (1998) conclude that pay-performance sensitivities of \$5.29 per thousand-dollar increase in shareholder wealth can lead to large swings in the dollar value of executive stock and option holdings, they do not conclude anything about the optimality of the incentives provided to managers.

Several papers have recognized that full incentives (100% equity ownership) will in general not be optimal for managers. Garen (1994) and Haubrich (1994) both recognize that a manager's risk aversion will put limits on the incentives that can be provided to him. Haubrich (1994) shows that for sufficiently risk-averse executives, Jensen and Murphy's (1990) PPS estimate of \$3.25 per thousand-dollar increase in shareholder wealth can be fully rationalized. Garen (1994) tries to show that the variance of firm returns (the riskiness of stock returns) explains the variation in incentives across CEOs, although his results are not always significant.

Aggarwal and Samwick (1999a) provide evidence that the variance of stock returns explains the variation in incentives across CEOs. They find a strong negative association between pay-performance sensitivities and the variance of dollar stock returns. They point out that earlier estimates of PPS are biased toward zero because they do not take into account the riskiness of firm returns when calculating PPSs. Risk-sharing considerations, they state, do explain why managers do not receive full incentives. Jin (2002) and Garvey and Milbourn (2003) extend this result to show that it is primarily the idiosyncratic component of stock return volatility that is responsible for the negative association between PPSs and the variance of stock returns. These results support one prediction of the Holmstrom and Milgrom (1987) optimal contracting model—that risk-averse agents have incentives that are decreasing in the variance of performance measures.

### *3.3. Executive discretion*

Prendergast (2002) states that the contracting environment has an important influence on incentives. The simplified principal-agent model discussed earlier assumes that risk is in the form of adding noise to executive inputs, thereby making inference more difficult. However, in many settings, what matters is not how hard the executive works, but what task the executive chooses to work. Prendergast argues that in highly uncertain environments, the task decision matters more, and in uncertain environments, the principal is more likely to delegate the task decision (discretion) to the agent. Because of the uncertainty of the environment, the principal cannot monitor inputs and so offers the agent an output-based (high-powered or incentive) contract.

This model has two key implications. The first is that there is a positive relation between environments characterized by task uncertainty and incentives. In the executive compensation literature, the evidence does not generally support this implication (for entrepreneurs, see, e.g., Aggarwal and Samwick, 1999a; Jin, 2002; Garvey and Milbourn, 2003; and Bitler, Moskowitz, and Vissing-Jorgensen, 2005). The second implication is that there is a positive relation between the use of incentives and the delegation of authority. Nagar (2002) finds empirical support for this implication from the banking industry. He finds that there is greater delegation of authority to bank branch managers for high-growth, volatile, and innovative banks. These bank branch managers also receive more incentive-based pay.

### *3.4. Firm size*

A number of other important points have been made in trying to understand the determinants of incentives. The preceding discussion about the merits of using full incentives proceeded with a very solid foundation in theory. But perhaps the most robust conclusion about compensation and incentives is that managers at larger firms generally receive higher compensation in dollar values and have lower ownership in their firms. This finding has been confirmed in papers by Schaefer (1998) and Himmelberg and Hubbard (2000).

Himmelberg and Hubbard observe that it requires more talent to run a large firm than it takes to run a small firm, so managers at large firms are better compensated. Baker and Hall (2004) maintain that CEO marginal products increase dramatically with firm size and that this can explain both the increase in CEO pay as firm size increases and the decrease in incentives. More recently, Gabaix and Landier (2008) show that small differences in CEO productivity or talent can generate dramatic differences in CEO pay by firm size. Section 6 covers this argument in greater detail.

### *3.5. Accounting returns*

Another important issue concerns the use of performance measures other than stock returns. While stock returns are clearly of paramount importance to stockholders, stock returns may not be the best way to provide incentives to managers. To see this, suppose that there exists some other performance measure for managers (say, accounting returns). If accounting returns do a better job of measuring managers' actions (regardless of the outcomes), accounting returns will be a better performance measure than stock returns. The reason for this is that outcomes are influenced by noise or external factors that are beyond the manager's control. The manager should neither be penalized nor rewarded for factors beyond her control. In principle, then, the ideal way to compensate the manager is based on the actions she takes. However, as noted earlier, the moral hazard problem states that shareholders do not directly observe managers' actions. As a result, compensation and incentives have to be based on observable measures, which are frequently outcome-based. If accounting returns, which are outcome-based, do a better job of measuring the actions taken by managers than do stock returns, then accounting returns should be the basis of compensation contracts.

The evidence on the usefulness of accounting returns in understanding incentives is mixed (see Lambert and Larcker, 1987; Janakiraman, Lambert, and Larcker, 1992; and Antle and Smith, 1986). Those papers find that accounting returns have more impact on salary and bonus than do stock returns. This is not surprising given that bonuses are usually tied to nonstock-based measures of firm performance. These papers also find conflicting evidence about the use of accounting measures in relative performance evaluation, which is discussed in greater detail in Section 4.

More generally, any measure that provides unique information about a manager's performance or actions should enter the compensation contract. These measures could be stock returns or accounting returns such as ROA or ROE, as we have already discussed. They could be measures such as divisional sales or profits for a divisional manager. They could be subjective measures such as a report from board members or other superiors for lower ranking executives. They could be other metrics such as EVA. While these measures are undoubtedly important, the evidence on their use and effectiveness is currently limited. Moreover, while accounting measures are important, how much compensation should be based on them is a matter of dispute. Given the possibility of earnings manipulation and gaming of accounting systems, accounting measures are likely to be of secondary importance to stock returns in designing incentive systems for top managers (see, e.g., Bergstresser and Philippon, 2006).

### *3.6. Ability*

Milbourn (2003) argues that CEO ability is an important determinant of the heterogeneity in CEO incentives. In Milbourn's model, more able CEOs have better reputations (the market's assessment of their ability). Since more able CEOs are less likely to be fired, stock prices will be more informative about their efforts, and they will therefore have higher pay-performance sensitivities. Conversely, less able CEOs are more likely to be fired, and therefore stock prices will be less informative about their efforts (as a consequence of the fact that stock prices reflect the cash flows generated by the CEO in place). As a result, less able CEOs will have lower PPSs. Using several different proxies for CEO reputation, Milbourn finds that CEOs with better reputations have higher PPSs.

Frydman (2005) observes that general skills have become more important in the CEO labor market over time relative to firm-specific skills. To the extent that general skills reflect ability that can be measured through both the presence of a general business education and greater occupational mobility, the rising importance of general skills can explain higher CEO pay and greater inequality in top executive pay.

Chang, Dasgupta, and Hilary (2007) present a very nice set of results that suggest ability matters. They show that announcement returns for a top manager's departure are negatively correlated with prior firm returns under that manager, suggesting that managers develop reputations for ability with financial markets. Managers whose departures are perceived to be bad news for the prior firm then go on to do well in subsequent labor market outcomes—they become CEOs at larger firms or earn high salaries in their new CEO positions. Those managers with high salaries at their prior firms (when corporate governance is good) also have good labor market outcomes, suggesting that ability is recognized in salaries. Finally, the loss of a good manager—as measured by negative announcement returns for a departure—is associated with subsequent poor accounting performance. Collectively, these results provide compelling evidence that ability is recognized, rewarded, and impounded into stock returns for top managers.

### *3.7. Incentives within firms*

Interestingly, many companies provide incentives (often stock options) throughout the firm. In many cases, this seems hard to justify given that most employees will not have a first-order impact on firm returns. Oyer (2004) maintains that broad-based incentive pay can be optimal if an employee's future outside opportunities are positively correlated with firm performance. In states of the world in which the firm does well, incentive compensation will be triggered, eliminating the need for costly recontracting when the employee's reservation wage rises. Inderst and Mueller (2006) extend this retention argument by noting that using broad-based incentive pay can protect the employee against ex-post opportunistic behavior by management should the company no longer need the employee due to changes in the firm's strategy or business mix. This protection against ex-post opportunism induces employees to invest in firm-specific human capital.

Oyer and Schaefer (2004) examine three potential explanations for the existence of broad-based stock option plans for employees: incentive provision; sorting–using options to attract employees who are relatively optimistic about a firm's prospects; and employee retention by providing a compensation mechanism (stock options) that will be positively correlated with labor market conditions in the firm's industry to the extent that the firm's returns are positively correlated with industry returns. Oyer and Schaefer find that use of a broad-based stock option plan is associated with smaller firms, more rapidly growing firms, less profitable firms, firms with higher stock returns, New Economy firms, and firms with greater stock return volatility. In calibration exercises, they argue that incentive provision seems unlikely to be the motive for broad-based stock option grants because the cost of compensating the employee for the additional risk of options simply dwarfs the cost of the additional effort generated by the employee. Similar calibrations provide more support for the sorting and retention explanations.

Another strand of the literature has looked at incentives for division managers. Cichello, Fee, Hadlock, and Sonti (2006), in examining the determinants of promotions and turnover for divisional managers, find that better divisional performance is associated with less turnover for divisional managers. Firm performance and performance by other divisions within the firm do not seem to matter, but performance by the division relative to its industry does seem to be important. For promotions, what matters is whether the manager's division outperforms the other division in absolute terms. These results are consistent with the use of tournament incentives (Lazear and Rosen, 1981) within the firm.

Kale, Reis, and Venkateswaran (2007) provide a more direct test for the presence and effectiveness of tournament incentives within firms. Tournament incentives are measured as the gap between the CEO's and the median executive's (non-CEO) pay. The paper finds that tournament incentives are associated with better firm performance, and this association is stronger when the CEO is close to retirement. Conversely, the association is weaker when there is a new CEO who is an outsider or when a succession plan is in place. The results are generally consistent with tournament incentives being both present and meaningful in influencing firm performance.

Wulf (2007) reports that for divisional managers who are also corporate officers pay is more sensitive to firm performance and that the relative weight on firm performance relative to divisional performance is greater than for division managers who are not corporate officers. In a similar vein, Aggarwal and Samwick (2003b) note that for divisional managers, pay-divisional performance sensitivities are positive and increasing in the precision of the divisional performance measure and that the pay-firm performance sensitivity is decreasing in the precision of the divisional performance measure. These papers are consistent with a multisignal principal-agent model in which both divisional and firm performance are informative about an executive's actions.

## 4. Relative performance evaluation

One additional measure that has received extensive attention is the use of industry or market performance as a benchmark for how well a manager has performed, or relative performance evaluation (RPE). In industries where all firms face common external shocks, executive compensation should be inversely related to competitors' performance. This way, an executive is paid when he outperforms his peers, but not when the entire industry happens to be in an upswing. In other words, the optimal compensation of executives is generally believed to rise with a firm's own performance and to decline with that of competitors. While the logic of relative performance evaluation is clear, there is very little evidence supporting its existence and some evidence suggesting that it is counterproductive.

The most important question with respect to relative performance evaluation is what is the appropriate benchmark to use? There are at least two plausible answers to this question—aggregate stock market returns and industry returns. Both have advantages and disadvantages.

### *4.1. RPE using industry returns*

Measuring firm performance relative to industry returns is appealing for a number of reasons. Industry competitors are most likely to face similar conditions and experience common shocks to performance. By comparing a firm's performance to the performances of its industry peers, common industry shocks can be removed while still maintaining a strong baseline of comparability. Interestingly, in their study of relative performance evaluation, Gibbons and Murphy (1990) do not find evidence of the use of relative performance evaluation at the industry level. They find more evidence of the use of relative performance evaluation at the level of the stock market as a whole.

What can explain this finding? Aggarwal and Samwick (1999b) contend that firms in the same industry not only experience common shocks to performance and similar business conditions, but they also directly compete with each other. Relative performance evaluation compensates a manager not only for how well she performs, but also for how badly the benchmark performs. In this case, the industry is the benchmark, and the manager does have some control over how badly the industry performs. For example, the manager can initiate a price war, which will worsen the performance of all the firms in the industry. But this is good news for the manager, since his compensation increases when the industry does worse. For this reason, relative performance evaluation is problematic at the level of the industry. Section 5.2.1 discusses this argument in greater detail.

### *4.2. RPE using market returns*

Measuring firm performance relative to market performance is appealing in that the manager is not rewarded for a generally rising stock market. Clearly, with the large run-up in stock indices from 1994 to 2000, many firms and managers have benefited from marketwide movements in stock prices. Furthermore, managers are unlikely to be able to influence the performance of the market as a whole, so the benchmark is relatively secure from tampering. There are still some problems, however, with using the stock market as a whole for the benchmark.

First, to the extent that firms use relative performance evaluation, they tend to simply net out aggregate market returns. Jin (2002) and Garvey and Milbourn (2003) find that pay-performance sensitivities are not influenced by systematic risk and are decreasing in idiosyncratic risk. Those papers point out that this is consistent either with firms using relative performance evaluation or with executives hedging out the market risk exposure of their compensation packages.

There are other problems associated with the use of relative performance evaluation more generally. For example, to truly use relative performance evaluation, a firm needs to benchmark the most significant sources of a manager's incentives. Let's suppose that the benchmark is aggregate stock market returns. The most significant sources of incentives are holdings of stock and options. Suppose that the firm's stock returns are lower than the aggregate market return. Then under relative performance evaluation, the manager should be required to pay the firm "negative" compensation in proportion to the amount of stock and options she holds. Given that the median CEO holds stock with a dollar value of $5.8 million (from Fig. 2), a 20% underperformance relative to the market would require a transfer from the CEO to the firm of $1.16 million. In a number of circumstances, CEOs would simply choose to quit rather than make the payment or see their base compensation cut.

One area in which the use of relative performance evaluation would seem to be quite sensible is in stock option grants. Specifically, it seems quite reasonable to suppose that the exercise prices of stock options would be indexed to increase or decrease with marketwide changes in stock prices. We do not, however, see this very often. Hall and Murphy (2002) argue that exercise prices are not indexed or set out of the money because managers hold large, undiversified positions in their own firms. As a result, they value stock option grants at a lower level than would an outside investor. Giving executives indexed options or options that are out of the money would simply make this problem worse.

## 5. Do incentives influence firm performance?

With some understanding of the determinants of incentives in hand, let us turn to the question of whether incentives influence firm performance. If incentives are to matter, they must have a significant impact on firm value. Interestingly, the basic question of whether greater incentives improve firm performance has been quite controversial. It is important to note that some simple agency models predict that greater incentives improve firm performance (e.g., Holmstrom and Milgrom, 1987) but others do not. The evidence on this point is also somewhat disjointed.

### *5.1. Some evidence*

Morck, Shleifer, and Vishny (1988) present a well-known result in which greater use of incentives does not improve firm performance. They define incentives as the amount of stock managers' hold in their firms, and they find that firm performance is increasing in managerial ownership for ownership between 0 and 5% of the total equity in the firm. However, they report that firm performance is decreasing in managerial ownership between ownership of 5 and 25%. They then show that firm performance is once again increasing for managerial ownership greater than 25%.

Morck et al. argue that for low levels of ownership (between 0 and 5%), incentives work and greater incentives lead to better firm performance. For intermediate levels of ownership (between 5 and 25%), incentives cease to function. Instead, greater ownership entrenches managers in the sense that they are hard to dismiss when they own a substantial fraction of the equity in the firm. As a result, managers take actions that benefit themselves while decreasing the value of the firm in this intermediate range of ownership. Managers are willing to forgo some firm value, even though this reduces the value of their own equity stakes, because outside shareholders still bear the majority of the value reduction while the managers get all of the benefits. For high levels of ownership (greater than 25%), incentives function once again. Managers own a sufficiently large fraction of the equity that they are no longer willing to take actions that reduce the value of the firm. Greater incentives again lead to better firm performance, although in this region of ownership the relationship is weaker.

One interpretation of these findings is that the full incentives advocated by Jensen and Meckling (1976) and Jensen and Murphy (1990) may not be necessary. Incentives that are sufficiently high (greater than 25% ownership) will have the same effect. An interesting implication of Morck, Shleifer, and Vishny's (1988) findings is that greater incentives are not always better—in the 5 to 25% ownership range, greater incentives are worse.

While this result has important implications for the amount of equity managers should hold in their firms, this result has not gone unchallenged. Himmelberg, Hubbard, and Palia (1999) make the point that we do not know if greater incentives lead to better performance or if firms that perform better happen to give their managers more stock. They much more carefully control for firm-specific factors when estimating the relationship between firm performance and incentives. They find that there is essentially no relationship between incentives and firm performance. One implication of their study is that incentives do not seem to matter much for firm performance.

Aggarwal and Samwick (2006) also examine the relation between firm performance and incentives. Their study looks at many more firms over more years than the Morck, Shleifer, and Vishny (1988) and Himmelberg, Hubbard, and Palia (1999) studies. Their study also incorporates holdings of stock options in addition to ownership of stock, and they find that firm performance is increasing in incentives for all levels of incentives. Interestingly, the implication Aggarwal and Samwick draw from their results is *not* that greater incentives lead to better firm performance. Instead, they argue that every firm is different, and that the amount of incentives each firm provides to its managers is appropriate for that firm.

Aggarwal and Samwick (2006) also point out that the original Morck, Shleifer, and Vishny (1988) paper is based on the disequilibrium idea that many managers have been provided with the wrong level of incentives. To see this, remember that Morck et al. find that firm performance is decreasing in incentives for managerial ownership between 5 and 25%. Firms with managerial ownership in this region could increase firm performance by either lowering managerial ownership to 5% or increasing it past 25%. In practice, it may be easier to lower managerial ownership than to raise it by either having the firm issue equity (thereby diluting managers' ownership stakes) or inducing managers to sell some of their holdings of stock. For some reason, firms do not have managers lower their ownership levels when the firms are in the region of decreasing firm performance. The firms are behaving in what appears to be a value-decreasing fashion.

Aggarwal and Samwick (2006) propose that firms are not, in fact, behaving in a value-decreasing fashion. Instead, firms and managers have very different characteristics that determine how large the level of incentives should be. Managers differ in how risk averse they are, how hard they are willing to work, and how much pleasure they derive from wasteful spending on perquisites such as corporate jets, artwork, celebrity golf tournaments, and expensive corporate staffs. Firms differ in how risky they are, how productive their assets are, and how talented their workers are. All of these differences can influence the level of incentives. Aggarwal and Samwick conclude that firms with "better" characteristics (e.g., less risk-averse managers) have greater incentives and better performance. A firm with worse characteristics (e.g., very risk-averse managers) cannot increase incentives and demonstrate better performance. Increasing incentives would actually result in a worse outcome for shareholders. On average, firms seem to get the level of incentives about right. The conclusion is that incentives do influence firm performance, but the design of incentives relies crucially on firm characteristics.

A different strand of the literature examines whether greater use of incentive compensation can lead to poor firm performance. In effect, can incentives be too high-powered? Recent notable bankruptcies such as Global Crossing, Qwest, WorldCom, and Enron, along with the collapse in share prices since 2000, suggest this possibility. Specifically, the large stock option grants realized by the management teams at the aforementioned firms prior to the bankruptcies raises the question of whether stock-based incentives actually create perverse incentives.

Do stock option grants and other forms of incentive compensation lead to a focus on short-term results, manipulation of accounting statements, tax avoidance, and outright fraud, rather than aligning managers' interests with those of shareholders? Recent papers by Bergstresser and Philippon (2006) and Peng and Roell (2008) provide evidence suggesting that this may be the case. Bergstresser and Philippon (2006) find that there is more earnings manipulation (use of discretionary accruals) at firms with managers who have greater pay-performance sensitivities. Peng and Roell (2008) show that this option-induced earnings manipulation is associated with greater shareholder litigation.

In cases of very large stock option grants (such as those realized by the management teams of Global Crossing, Qwest, and WorldCom), the grants themselves may be symptomatic of deeper problems involving a culture of corporate greed, malfeasance, and failure of corporate governance. Look again at Figure 3, which lists the 20 most highly paid CEOs in 1999, plus Kenneth Lay and Jeffrey Skilling from Enron. One striking feature of this list is that the three highest paid CEOs, Annunziata, Nacchio, and Kozlowski, are from companies that are either bankrupt (Global Crossing and Qwest) or notorious for high-level corporate greed (Tyco). Furthermore, Carleton Fiorina (Hewlett-Packard) was fired after a disastrous merger with Compaq. William McGuire (UnitedHealth) was forced to resign in an options backdating scandal. Bernard Ebbers from WorldCom (bankrupt) also makes the list, and Lay and Skilling from Enron (defunct) are not far behind.

All of these executives received large stock option grants in 1999; in most cases, stock option grants were virtually their entire compensation. Several years after the grants, their companies or the executives themselves were in trouble or bankrupt. The following interesting question arises: are these cases pathological, or do they point to larger problems with the use of stock options? To the extent that extremely disproportionate grants are correlated with deeper firm-level problems, then the existence of such grants can usefully serve as an early warning sign for investors, employees, regulators, and others about impending corporate crises and collapses.

The cases of Enron, WorldCom, and others seem to be characterized by entrenched managers, weak boards, generally poor corporate governance, and cultures of corporate greed. Incentive systems failed at these firms. Instead of using stock options as an incentive device, it now seems clear that these firms used stock options as a tool for management enrichment. These examples are good illustrations of the argument that boards of directors (and, by extension, compensation committees) are either stacked with friends and cronies of the CEO or have been effectively captured by the CEO (Bertrand and Mullainathan, 2001). We return to this argument in Section 6.

The use of stock-based incentives, specifically stock options, is neither inherently good nor bad (see also Holmstrom and Kaplan, 2003). In general, the use of stock options is associated with better firm performance, suggesting that companies use options appropriately. However, in cases of very large option grants that seem disproportionate such as those detailed in Figure 3, executive compensation may be symptomatic of deeper problems within the firm. Corporate governance may well have failed at these firms, and executives may only be constrained by what they can get away with. Thus the current set of corporate crises is not caused by excessive compensation, but excessive compensation may well be a leading symptom of impending collapse.

### *5.2. What actions are incentives designed to influence?*

Given that incentives influence firm performance, it is natural to wonder about the mechanism through which this happens. Clearly, increasing incentives does not directly improve firm performance. Instead, incentives influence something that managers do, which then impacts firm value. What do managers do that is influenced by incentives? The early literature (e.g., Holmstrom, 1979) focused on inducing managers to work hard. Managers disliked working and had to be provided with incentives to do so. While this idea is conceptually appealing, it is not obviously true.

Specifically, can effort provision really be the basis for why managers get incentives? Holmstrom (1992) argues that this is probably not the case. It seems unlikely that incentives are structured to induce managers to work harder. Most top managers are already workaholics, which is perhaps an extreme version of intrinsic motivation. Furthermore, at the margin, it seems unlikely that at very high compensation levels managers would choose to work more. Instead, it seems likely that they would substitute more leisure time when offered higher compensation. If incentives do not influence how hard managers choose to work, what actions do incentives influence?

Several issues need to be considered here. First, managers take actions along many dimensions. They set the strategic direction for the firm, they make investment decisions, they decide to enter or exit different lines of business, they choose how hard to compete within a particular market, they determine how to allocate resources within the firm, they make production decisions, and so on. In short, managers can take many different actions, all of which have a large impact on firm performance. At the same time, for all of these actions, managers' preferences at to which action to take may not coincide with those of the shareholders.

In principle, the shareholders could monitor the actions being taken by the managers and insist that managers take the shareholders' preferred actions. In practice, this will be very difficult to achieve. Instead, shareholders look for a convenient summary measure of all of the actions taken by managers. This summary measure is the firm's stock price. Because compensation tied to the firm's stock performance is designed to influence many actions, it will not necessarily be straightforward to see much impact of incentives on any given type of decision. Nevertheless, a number of papers have examined specific types of decisions to determine whether incentives have any effect. Here we focus on four types of decisions that have been examined in the literature: the setting of prices in product markets; the choice of investment levels; the decision to diversify the firm into new lines of business; and the decision to engage in mergers and acquisitions.

#### *5.2.1. Incentives and pricing policy*

Pricing policy is an important decision for any firm, and it will have a large impact on firm value. Do incentives influence how managers set prices in product markets? Aggarwal and Samwick (1999b) find evidence for this influence in the United States, and Joh (1999) finds evidence of it in Japan. The argument is that shareholders are better off by providing incentives to managers to soften price competition. In effect, incentives can be used to foster tacit collusion within industries.

To see how this works, suppose that firms compensate managers not only for how well the firm but also the industry as a whole performs. Now the manager has incentives to improve not only his own firm's performance, but also that of the industry as well. In such a setting, the manager gets nothing from starting a price war, as both components of his compensation are reduced. Both the industry's and his own firm's performances are worsened. This idea is in stark contrast to what models of relative performance evaluation predict. Relative performance evaluation would compensate a manager if industry performance were to be reduced. Trying to foster collusion can explain why we do not observe relative performance evaluation at the industry level.

Aggarwal and Samwick (1999b) observe that compensation depends positively on both firm performance and industry performance. This clearly suggests that there is no relative performance evaluation. This general finding is confirmed in Himmelberg and Hubbard (2000), and is also found for Japanese firms in Joh (1999). For firms in more competitive industries, the researchers report that how the industry performs is given relatively more weight than how the firm performs. In more competitive industries, the value of collusion is highest but also the hardest to enforce. There is no dominant competitor to enforce discipline, and there are many potential competitors who could cheat on the collusive outcome. In order to induce collusion, firms give managers relatively more incentives based on industry performance than on firm performance.

By contrast, in industries that are not very competitive, firms give managers relatively more incentives based on firm performance than on industry performance. In industries that are not competitive, there is relatively little to gain from fostering collusion. Firms do not compete directly, and so managers are unlikely to find it worthwhile to start a price war in the first place. As a result, little weight need be placed on industry performance.

Aggarwal and Samwick also offer an explanation for why incentives are not necessarily very large. It is not the absolute level of incentives that induces executives to choose the right pricing decisions and avoid price wars. Rather, executives are swayed by the ratio of the weight put on firm performance relative to the weight put on industry performance. In less competitive industries, this ratio will be higher—more weight is put on how the firm performs. In more competitive industries, this ratio should be lower—more weight is placed on how the industry performs. But since what matters is the ratio, the absolute level of incentives based on either firm or industry performance can be low.

The implications of this result for the design of compensation schemes are interesting. Shareholders need to pay attention to the structure of the industry when writing compensation contracts with managers. Attempts to use relative performance evaluation in competitive industries are likely to be counterproductive. Instead, managers should be rewarded for both good firm performance and good industry performance. By contrast, if a firm is in a monopoly position or in a weakly competitive industry, then industry performance is relatively unimportant and should not enter the compensation scheme.

### 5.2.2. *Incentives and investment decisions*

Another important decision made by managers is how much the firm invests from year to year. Investment decisions not only influence current but also future firm performance. Two issues are important with respect to investment and incentives: (1) Do incentives influence investment decisions? (2) If incentives influence investment decisions, how do they influence those decisions?

Aggarwal and Samwick (2006) maintain that incentives are needed to induce managers to invest more. In contrast, Jensen (1986, 1993) quite persuasively argues that managers overinvest because they derive some benefit from doing so, such as empire-building (i.e., managers enjoy controlling more assets and running a larger firm for reasons of prestige; they believe that they will be paid more if they run a larger firm or that their ability to run an even larger firm in the future depends on the size of the firm they run currently). For these reasons, managers may choose to overinvest. Jensen (1986, 1993) contends that shareholders need to curtail overinvestment through the use of debt, dividends, or takeovers. Aggarwal and Samwick (2006) claim that incentives can be used to curtail overinvestment as well, but they find no evidence that incentives are, in fact, used to curtail overinvestment.

Instead, it appears that managers prefer to invest less and for several potential reasons. Investing requires that the manager oversee the investment. When firms expand existing facilities or start new product lines, managers are required to do more work or to spread their talents over a wider array of activities. As a result, managers will generally prefer not to invest more or, in other words, they will underinvest in the sense that they will forgo some positive net present value investment opportunities. Incentives are used to induce managers to invest more than they otherwise would. This result is consistent with the findings of Bertrand and Mullainathan (2003); they show that when managers are insulated from takeover pressure, they tend to let workers' wages rise, slow the closing of old plants, and open fewer new plants. Bertran and Mullainathan view these managers as preferring the quiet life.

The conclusion that incentives are used to induce managers to invest more does not imply that all managers should be given more incentives. Instead, incentives are set according to managers' specific characteristics. Managers who are less averse to risk can be given more incentives and will invest more as a result when compared to managers who are more averse to risk. Managers who are less risk averse will, as a consequence, run better performing firms. These findings stress the importance of choosing the right managers to begin with—those who are not averse to investing in the first place.

### *5.2.3. Incentives and diversification decisions*

Perhaps an even more important managerial decision is to enter an entirely new line of business. Do incentives influence whether managers choose to diversify their firms? Substantial evidence is now available that diversification into new lines of business reduces firm value (Lang and Stulz, 1994; Berger and Ofek, 1995). There is also evidence that managers with greater incentives in the form of stock ownership choose to diversify less (Denis, Denis, and Sarin, 1997). These two pieces of evidence suggest that diversification is an agency problem along the lines of Berle and Means (1932) and Jensen and Meckling (1976), and that shareholders provide managers with incentives to prevent them from diversifying.

If diversification is harmful to firm value, however, shareholders should provide managers with enough incentives so that managers do not diversify at all. Yet this does not happen given the large number of diversified firms that we observe. Clearly, providing managers with even more incentives is costly, but the agency problem argument above does not recognize the existence of those costs.

Aggarwal and Samwick (2003a) examine this question and explicitly consider the possibility that the actions that incentives are designed to influence can be multidimensional. They argue that incentives are designed both to induce managers to take positive actions such as working hard or investing more and to prevent managers from taking negative actions such as diversifying the firm. These authors show how these competing concerns influence the shape of the incentive scheme.

In general, managers who have a stronger preference for diversification will need to be given more incentives. And even though these managers have more incentives, they still diversify more than do managers with less of a preference for diversification (and hence less incentives). The reason for this is that incentives cannot fully offset the preference for diversification. Providing incentives is costly—shareholders give up some of their own returns in exchange for forcing the manager to bear more risk and more of the cost of the preference for diversification. As a result, managers with a stronger preference for diversification will have greater incentives, run firms that are more diversified, and show worse firm performance than will managers with a weaker preference for diversification.

### *5.2.4. Incentives and acquisitions*

Another important decision managers face is whether to engage in a merger or an acquisition. Do incentives induce managers to undertake value-enhancing acquisitions? Here we consider incentives for both target CEOs and acquiring CEOs.

We start by looking at target CEOs. Hartzell, Ofek, and Yermack (2004) show that CEOs of target firms frequently do quite well as a result of the acquisition. While turnover at the time of and subsequent to the merger is quite high for target firm CEOs, they are usually quite well compensated as a result of change in control agreements. Hartzell et al. also show that when target CEOs receive extra compensation in a merger, target shareholders tend to receive lower acquisition premia. Wulf (2004) finds that in mergers of equals, target CEOs are willing to trade merger premia for greater control rights in the merged company. These results suggest that at least some acquisitions have characteristics consistent with rent extraction by the target CEO rather than shareholder value maximization.

For acquiring firm CEOs, the evidence is somewhat mixed. Datta, Iskandar-Datta, and Raman (2001) show a positive relation between equity incentives and merger performance, and Lehn and Zhao (2006) suggest that acquiring firm CEOs are more likely to be replaced after a poor acquisition. On the other hand, Harford and Li (2007) show that CEOs of acquiring firms are rewarded if the acquisition is not wealth destroying but are not penalized if the acquisition is wealth-destroying. In particular, new grants of stock and options offset the effects of poor stock price performance. This result is similar in spirit to Garvey and Milbourn's (2006) asymmetric benchmarking result. The Harford and Li (2007) result is also in contrast to more general forms of investment such as capital expenditures, where CEOs benefit if the investment is wealth-creating and are penalized if it is wealth destroying. Consistent with the idea that acquisitions either insulate managers or are a form of rent extraction for acquiring CEOs, Grinstein and Hribar (2004) find that more powerful acquiring CEOs are able to extract larger M&A bonuses from their boards.

This section has considered several actions that managers can take to influence firm performance. Another mechanism through which incentives might work is capital structure decisions. Although some progress has been made in this area (see, e.g., Chang, 1993, and Qiu, 2006), we do not yet have a fully satisfactory theory supported by evidence. This remains an important area for continuing research. Importantly, however, we have been able to establish some of the other mechanisms through which incentives work. In doing so, we also have been able to determine the types of considerations that should influence the design of incentives. In looking at some of the actions that incentives are designed to influence, we have taken it for granted that some representative of the firm or the shareholders sets incentives. We now examine more carefully who actually sets incentives and executive pay.

## 6. Alternatives to the agency view

In general, compensation is determined by a compensation committee. The compensation committee is a subset of the board of directors that often includes outside directors and sometimes the CEO as well. A popular view in the press and now in the literature is that CEOs effectively set their own pay. The argument is that boards of directors (and, by extension, compensation committees) are either stacked with friends and cronies of the CEO or have been effectively captured by the CEO. When the CEO sits on the compensation committee, this view certainly seems quite plausible. In this view, compensation is divorced from firm performance. Instead, CEOs are able to skim firm profits and are constrained only by what they think they can get away with.

Gillan, Hartzell, and Parrino (2006) document several interesting facts that may shed light on possibilities for rent extraction, as well as the nature of contracting with top executives. They find that fewer than half of the CEOs in the S&P 500 have an explicit contract or employment agreement. In cases where CEOs do have an explicit contract, it appears that the contract is designed to protect the CEO from expropriation by the firm rather than to foster rent extraction by the CEO. This also leaves open the possibility that rent extraction by CEOs occurs when implicit contracts are in place—less formal agreement ex ante may foster or be a prerequisite for informal collusion ex post. In this section, we consider the skimming view as well as several other alternatives to the agency view.

### *6.1. The skimming view and rent extraction*

Bertrand and Mullainathan (2001) suggest an interesting approach to see if skimming is a valid argument. They argue that if managers are able to skim firm profits, they will be rewarded for “lucky” events that reflect nothing about the manager’s efforts or talent. We already know that the effects of common stocks are not removed through the use of relative performance evaluation. Bertrand and Mullainathan (2001) look at specific stocks in certain industries. For example, they examine exchange rate movements and changes in the price of oil. They argue that it is unlikely that managers have much influence over these stocks, so these are truly lucky or unlucky events for the manager. For the oil price example, they find that managers at oil companies often see their pay fluctuate with oil prices. When oil prices are up, compensation goes up, and when oil prices are down, compensation goes down. This suggests that CEOs are being rewarded and penalized for random events. However, Bertrand and Mullainathan also find that in a reasonable number of cases, CEOs are shielded from the consequences of random but downward movements in oil prices—oil prices decrease but compensation does not. This suggests that managers are rewarded for good luck but not always penalized for bad luck (see also Garvey and Milbourn, 2006).

Bertrand and Mullainathan (2001) draw several conclusions. First, they argue that in poorly governed firms, managers essentially set their own pay and are able to skim firm profits in the form of pay for luck. Second, skimming is less prevalent in firms that are well governed. Well-governed firms are those in which the board of directors and compensation committees are independent of the CEO. Third, there is evidence that managers are not overpaid in firms that are well governed in the sense that stock options are substitute compensation for other forms of compensation. By contrast, in firms that are not well governed, stock options seem to be strictly additional compensation. These findings suggest that board independence is an important check on CEO compensation.

### *6.2. Pay without performance and stealth compensation*

Bebchuk and Fried (2004) state that managerial power explains the rise in executive compensation. In their view, boards are controlled by CEOs, and so there is no arms-length bargaining or efficient contracting between boards and top executives. Instead, CEOs dictate the pay-setting process. As a result, CEO compensation is often independent of firm performance. In particular, the more powerful the CEO, the less sensitive CEO pay is to firm performance. They further argue that the only limits to CEO rent extraction come from outrage costs. Compensation that is too high will draw scrutiny and criticism from the media, shareholders, and other activists.

Bebchuk and Fried extend this argument by showing that many forms of compensation are hidden from the view of shareholders by effectively skirting SEC disclosure requirements. As an example, retirement benefits are not disclosed at the time that they are paid because the executive is no longer with the company. Historically, companies did not disclose the present value of supplemental executive pensions; instead, they estimated the annual payment based on number of years of service. As a result, an executive’s pension could be quite large in relation to their actual compensation while serving. The case of Richard Grasso, the former CEO of the NYSE, whose pension rose to $140 million in value, is instructive in this regard.

There are numerous other examples as well. Yermack (1997) presents evidence that stock returns are abnormally high after stock option grants. Lie (2005) and Heron and Lie (2007) also present evidence that stock returns are abnormally low prior to stock option grants. Furthermore, the frequency with which stock option grants occur when stock prices are at their lowest within some interval (usually a quarter) is strikingly anomalous. Lie (2005) argues that this result is strongly suggestive of stock option backdating, in which executives choose grant dates after the fact. Subsequent confirmed examples of backdating at companies such as United Health have borne this hypothesis out. Because backdating implies an immediate transfer to the executive from the shareholders, it is perhaps the clearest example of stealth compensation. It appears that often only the executive and the board or some members of the compensation committee are even aware of it.

Stock option backdating can be viewed as the most compelling evidence for widespread rent extraction by top executives. Stock option backdating can explain, for example, the Bertrand and Mullainathan (2001) and Garvey and Milbourn (2006) results that executives are not penalized for poor stock returns. Indeed, poor stock returns can become an opportunity for rent extraction. Lie and Heron (2007) show that the pattern of abnormal negative returns prior to stock option grants and abnormal positive returns after stock option grants largely disappears after 2002, when the SEC required that stock option grants be reported within two business days. This suggests that at least one channel for rent extraction has been shut down.

Kuhnen and Zwiebel (2007) model hidden or stealth compensation and show that it emerges naturally in a setting in which executives have discretion over their pay. They emphasize that if an executive takes too much stealth compensation, then the shareholders' assessment of the executive's ability will become more negative, increasing the likelihood of termination. As a result, the use of stealth compensation is more likely in environments with greater noise in either output or managerial ability.

### *6.3. The shortage view*

An alternative interpretation of how CEO pay is established is that it is set in the managerial labor market. Himmelberg and Hubbard (2000) have noted the insufficient supply of talented top managers. As a result, those top managers with talent are able to command high wage premiums. One potential implication is that compensation is divorced from firm performance because firms must pay CEOs what they demand. Himmelberg and Hubbard argue that this is what explains the lack of relative performance evaluation and not the idea that CEOs skim as much profit as possible. They examine the possibility that the supply of talented CEOs is thin by arguing that the size of the firm a CEO manages is a good proxy for how talented the CEO is. More talented CEOs will work at larger firms where their talents are presumably better employed.

The authors then examine how the sensitivity of CEO pay to firm performance is influenced by macroeconomic shocks. The idea is that favorable macroeconomic shocks should increase the demand for talented CEOs without increasing the supply of talented CEOs (at least, in the short run). They find that macroeconomic shocks disproportionately increase the pay of CEOs for large firms.

Gabaix and Landier (2008) develop a competitive matching model that implies that the most talented CEOs will manage the largest firms. As firms increase in size, CEO pay will grow as well, as a reflection of CEO productivity. Calibrations of the model can explain both the increase in CEO pay in the United States and cross-country dispersion in CEO pay. Intriguingly, Gabaix and Landier contend that the increase in CEO pay over the period 1980–2003 can be explained by the increase in firms' market capitalizations. In addition, their calibrations suggest that small cross-sectional differences in CEO talent drive large differences in CEO pay. One concern with their paper is that they compare CEO pay (a flow variable) with firm market value (a stock variable). A more appropriate comparison might be between changes in CEO firm-specific wealth and changes in firm value. Because much of compensation occurs in the form of stock options, which have convex payoffs, the change in CEO firm-specific wealth might well be nonlinear in relation to changes in market capitalization.

Two competing ideas—that there is an insufficient supply of talented CEOs and that CEOs skim firm profits—make it clear that it is hard to disentangle how compensation is being set. Nonetheless, one feature that does stand out from the Bertrand and Mullainathan (2001), Himmelberg and Hubbard (2000), and Gabaix and Landier (2008) studies is that compensation does appear to be strongly linked to firm performance. Much of the literature also shows that compensation committees do not use relative performance evaluation when they set CEO compensation, presumably because CEOs do not wish to be evaluated on a relative basis. Is this because CEOs have a strong bargaining position or because they have co-opted the committee? As an illustration of this point, consider what happens when a CEO is separated from her firm.

### *6.4. The value of termination*

In this section, we examine what happens when CEOs and other executives are terminated. In general, CEO dismissals are newsworthy events (e.g., Carleton Fiorina's termination at Hewlett-Packard). Furthermore, the stock market reaction to such dismissals is often quite positive (Hewlett-Packard's market capitalization was up $4 billion upon Fiorina's termination). To address the causes and consequences of termination, researchers typically estimate models of the likelihood of CEO turnover as a function of different variables, including firm performance. The assumption is that most turnover represents firings unless there is evidence to the contrary, such as the manager reaching retirement age or becoming CEO elsewhere. Not surprisingly, poor performance is associated with a much higher likelihood of being terminated. Above-average performance is associated with promotions to CEO at other firms (Fee and Hadlock, 2003). Intriguingly, stock performance matters more than does accounting performance. A recent example of this is the termination of Robert Nardelli, former CEO of Home Depot, who generated consistent accounting performance but poor stock return performance.

In addition, performance relative to industry benchmarks seems to be important, suggesting a role for relative performance evaluation. Jenter and Kanaan (2006) show that CEOs are much more likely to be terminated after poor industry or market performance, even after controlling for firm-specific performance. This result seems to be at odds with relative performance evaluation, since arguably both industry and market performance are outside of the CEO's control. Going a step further, they also find that within a peer group of CEOs, those whose firms perform the worst are most likely to be fired. Thus, there appears to be a form of asymmetric relative performance evaluation in firing decisions. When there is a positive market or industry shock, the probability of being fired is low even with poor performance in a relative sense. When there is a negative shock, the weakest performers within the group are more likely to be fired.

The turnover–performance relation is stronger for firms with outside dominated boards, suggesting a role for monitoring (Weisbach, 1988). The turnover–performance relation is weaker when managers have a sizeable equity stake, which supports the view that equity ownership can serve to entrench management for some levels of ownership. The turnover–performance relation is weaker when the CEO is also chairman of the board, again suggesting a role for management entrenchment. There is evidence that turnover rates have increased sharply over time, perhaps suggesting heightened shareholder scrutiny (Hadlock and Lumer, 1997; Huson, Parrino, and Starks, 2001). This increased pressure on executives may lead them to focus more on increasing shareholder value, but the increased pressure may also enhance incentives to mislead capital markets (Peng and Roell, 2008).

Turnover–performance sensitivities in the United States are relatively high compared to those of other countries. These sensitivities are relatively higher in firms where performance is a relatively more precise signal of managerial ability. If a firm is performing particularly poorly, it is likely to choose an outsider as the replacement CEO. If it is doing moderately poorly, it is likely to hire an inside replacement. In addition, many other senior executives leave when the CEO is dismissed (Fee and Hadlock, 2004).

What are the consequences of CEOs being terminated? In general, new managers make many changes at the firm, and accounting performance improves. For the dismissed CEOs, there is both good news and bad news. They typically get large severance payments on the order of two times their salary. But they usually have poor labor market prospects, especially if their dismissal occurs in bankruptcy or in conjunction with a scandal (Fee and Hadlock, 2004).

Given that CEOs typically get paid quite well upon departure, one could ask why CEOs are rewarded for performance that induces termination. The skimming view is that the CEO writes his own compensation contract (including payments for termination). On the other hand, an insufficient supply of talented managers should not imply that a CEO who is about to be terminated would receive a large payment. Presumably, a CEOs bargaining power is weakest when the CEO is about to be terminated. However, this presumption need not be correct. In order to induce the CEO to depart, the board may have to pay the CEO quite handsomely. The threat of not departing may give the CEO quite a lot of bargaining power.

As an example, consider the case of Mattel, Inc. Mattel's CEO, Jill Barad, made \$39,130,708 in compensation in 1997. In 1999, Mattel suffered enormous losses attributable to its acquisition of the Learning Company. Barad resigned from Mattel on February 3, 2000. Barad's severance package was worth more than \$37 million, with some estimates as high as \$50 million. It is hard to argue that Barad's severance package reflects pay for performance. If the board's intention was to allow Barad to skim some value from the firm, her severance package suggests that the board succeeded. On the other hand, the negative publicity and controversy generated by the severance package were certainly not good news for Mattel and the board. The alternative interpretation is simply that it took \$37 million to \$50 million to induce Barad to leave. If the general point of incentives is to induce the manager to take the right actions, then termination incentives may also have to be strong. While the board presumably could have simply fired her, contractually they may not have been able to avoid most of these payments. In any event, voluntary departure is typically viewed as better for the firm.

### *6.5. Common agency and boards*

Aggarwal and Nanda (2006) argue that boards seem to take into account the objectives of a diverse group of stakeholders. They maintain that this is true even in the United States, where nominally boards have a fiduciary responsibility to shareholders. In their model, shareholders may grant access to the board to constituencies in exchange for noncontractible assets that these constituencies possess. As an example, the threat of hold-up by a union (e.g., a strike) can be ameliorated by giving board representation to the union. Here the asset is the union's right to strike, which is traded for the ability to influence how employees are treated through participation on the board. Aggarwal and Nanda state that because the different constituencies (and their board members) have divergent interests, each will try to influence the actions taken by management. As a consequence of this dissonance, managerial incentives are weakened, and management does less of every task that matters to the various board members.

Aggarwal and Nanda test this theory and find evidence consistent with the argument—firms with more board members pursue a greater variety of objectives, provide their executives with fewer stock-based incentives, and exhibit weaker share-price-based performance. Importantly, these outcomes are still consistent with ex-ante efficiency: accepting weaker ex-post incentives is the price of getting the various constituencies to contribute their assets to the firm. Thus, their argument can be viewed as an alternative to the skimming or board capture view. Boards may tolerate (and, in fact, encourage) weaker incentives as part of a larger optimization that incorporates other factors of production but is still consistent with shareholder value maximization.

### 6.6. *Executive compensation and executive beliefs*

Executives' decisions about whether to sell shares that they hold can provide insight into whether executives believe that their stock is overvalued or undervalued (Jenter, 2005). Jenter finds that executives in firms with high book-to-market ratios add to their personal holdings of their firms' stock, while executives in low book-to-market firms reduce their personal holdings. He also finds that firm-level capital structure decisions mirror the decisions of executives in their personal portfolios—executives selling stock from their personal portfolios run firms that issue equity. Intriguingly, after controlling for book-to-market and size, Jenter finds that executives do not earn abnormal returns from their personal portfolio trades. This suggests that the information executives have about their firm's stock is already publicly available. This result also suggests that book-to-market is indicative of market mispricing rather than risk. It is unlikely that executives would choose to increase their exposure to their own firm's idiosyncratic risk rather than purchase a portfolio of high book-to-market stocks.

Malmendier and Tate (2005) contend that an executive's option exercise decisions can be used as a metric for how overconfident the executive is. Their argument is that executives are overexposed to firm-specific risk because of vesting periods on options and restricted stock grants. Thus, executives should exercise stock options early to reduce their exposure to their own firms. Executives who, for example, hold their options to expiration are deemed to be overconfident. Malmendier and Tate find that overconfident executives have higher investment–cash flow sensitivities. Baxamusa (2007) suggests that this result is also consistent with the possibility that instead of being overconfident, executives believe that their firms are undervalued by the market, as argued by Jenter (2005). It is fair to say that executives' decisions about their holdings of stock and options in their own firms do inform us about divergences between their own beliefs and those of the market. What is less clear is whether the executives are right or the market is right, and in what situations one is more likely to be right than the other.

## 7. Conclusion

What can we conclude about the appropriate design of incentives? Since 1990, when Jensen and Murphy made their seminal contribution to the literature on executive compensation, our understanding of executive compensation and incentives has greatly improved. After almost two decades of research, we know that, through time, the strength of incentives has increased. During the 1990s, this was accomplished primarily through the use of stock options. While the other components of compensation (salary, bonus, long-term incentive plan payouts, and restricted stock grants) are important, the most important source of new incentives has been stock options. The two most important sources of aggregate incentives are holdings of stock and holdings of stock options; they are what align managers' interests with those of shareholders.

We know that firm size, firm risk, executive risk aversion, executive productivity, the extent to which executives dislike or like taking certain actions that matter for shareholders, and characteristics of the industry all determine how strong or weak incentives should be. We know that, because shareholders have limited information about what managers do and limited ability to monitor managers, managers need to be provided with some incentives. The key limitation on the provision of incentives seems to be the need to share risk between managers and other investors (shareholders). Actual compensation practices when it comes to granting executives stock and options seem to reflect this trade-off.

There are strong reasons to believe that firm performance should be evaluated relative to an industry or marketwide benchmark. Yet in practice we do not see this happening. While the industry would appear to be the most informative benchmark, it is also the benchmark that is most susceptible to tampering by a firm's managers. The market as a whole is also potentially a useful benchmark, but several conceptual problems are associated with its use. Indexed options might provide a reasonable way to achieve some benchmarking, but they may not provide sufficient incentives to risk-averse managers. In addition, indexed options have historically had unfavorable accounting treatment associated with them, which has limited their use. With the advent of option expensing in the United States, this situation may well change.

We also know that boards and shareholders do not have the complete ability to set managerial compensation. Either because of CEO power or a limited supply of managerial talent, at least some CEOs have the ability to extract more compensation than what is dictated by straightforward pay for performance. At this stage, we have a reasonably strong understanding of the determinants of incentives, and incentives seem to work well. We still do not have a full understanding of the level of compensation, however, although the recent paper by Gabaix and Landier (2008) represents a good start. A number of well-documented examples are consistent with the rent extraction view of the level of compensation. Instances of excessive compensation and rent extraction seem to be correlated with corporate governance failure, accounting fraud, and poor corporate outcomes. The extent of such instances has not yet been fully documented. A better understanding of issues related to the level of compensation is the next critical task for researchers in the field of executive compensation.

## References

Aggarwal, R. K., F. Meschke and T. Y. Wang, 2007, "Corporate Political Contributions: Investment or Agency?" Working Paper, University of Minnesota.

Aggarwal, R. K. and D. Nanda, 2006, "Access, Common Agency, and Board Size," Working Paper, University of Minnesota and Duke University.

Aggarwal, R. K. and A. A. Samwick, 2006, "Empire-Builders and Shirkers: Investment, Firm Performance, and Managerial Incentives," *Journal of Corporate Finance,* 12, 489–515.

Aggarwal, R. K. and A. A. Samwick, 2003a, "Why Do Managers Diversify Their Firms? Agency Reconsidered," *Journal of Finance,* 58, 71–118.

Aggarwal, R. K. and A. A. Samwick, 2003b, "Performance Incentives within Firms: The Effect of Managerial Responsibility," *Journal of Finance*, 58, 1613–1649.

Aggarwal, R. K. and A. A. Samwick, 1999a, "The Other Side of the Tradeoff: The Impact of Risk on Executive Compensation," *Journal of Political Economy,* 107, 65–105.

Aggarwal, R. K. and A. A. Samwick, 1999b, "Executive Compensation, Relative Performance Evaluation, and Strategic Competition: Theory and Evidence," *Journal of Finance,* 54, 1999–2043.

Antle, R. and A. Smith, 1986, "An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives," *Journal of Accounting Research*, 24, 1–39.

Baker, G. P. and B. J. Hall 2004, "CEO Incentives and Firm Size," *Journal of Labor Economics*, 22, 767–798.

Barron, J. M. and G. R. Waddell, 2003, "Executive Rank, Pay, and Project Selection," *Journal of Financial Economics*, 67, 305–349.

Baxamusa, M., 2007, "Managerial Perception and Corporate Decisions," Working Paper, University of Minnesota.

Bebchuk, L. and J. Fried, 2004, *Pay Without Performance*, Harvard University Press Cambridge, MA.

Berger, P. G. and E. Ofek, 1995, "Diversification's Effect on Firm Value," *Journal of Financial Economics*, 37, 39–65.

Bergstresser, D. B. and T. Philippon, 2006, "CEO Incentives and Earnings Management," *Journal of Financial Economics*, 80, 511–529.

Berle, A. A. and G. C. Means, 1932, *The Modern Corporation and Private Property*, Macmillan, New York.

Bertrand, M. and S. Mullainathan, 2003, "Enjoying the Quiet Life? Corporate Governance and Managerial Preferences," *Journal of Political Economy*, 111, 1043–1075.

Bertrand, M. and S. Mullainathan, 2001, "Are CEOs Rewarded for Luck? The Ones Without Principals Are," *Quarterly Journal of Economics*, 116, 901–932.

Bitler, M. P., T. J. Moskowitz and A. Vissing-Jorgensen, 2005, "Testing Agency Theory with Entrepreneur Effort and Wealth," *Journal of Finance*, 60, 539–576.

Bulow, J. and J. Shoven, 2005, "Accounting for Stock Options," *Journal of Economic Perspectives*, 19, 115–134.

Burrough, B. and J. Helyar, 1990, *Barbarians at the Gate: The Fall of RJR Nabisco*, Harper & Row, New York.

Chang, C., 1993, "Payout Policy, Capital Structure, and Compensation Contracts When Managers Value Control," *Review of Financial Studies*, 6, 911–933.

Chang, Y. Y., S. Dasgupta and G. Hilary, 2007, "Managerial Reputation in Financial and Labor Markets," Working Paper, Hong Kong University of Science and Technology.

Cichello, M. S., C. E. Fee, C. J. Hadlock, and R. Sonti, 2006, "Promotions, Turnover, and Performance Evaluation: Evidence from the Careers of Division Managers," Working Paper, Michigan State University.

Datta, S., M. Iskandar-Datta and K. Raman, 2001, "Executive Compensation and Corporate Acquisition Decisions," *Journal of Finance*, 56, 2299–2336.

Denis, D., D. Denis and A. Sarin, 1997, "Agency Problems, Equity Ownership, and Corporate Diversification," *Journal of Finance*, 52, 135–160.

Fee, C. E. and C. J. Hadlock, 2004, "Management Turnover Across the Corporate Hierarchy," *Journal of Accounting and Economics*, 37, 3–38.

Fee, C. E. and C. J. Hadlock, 2003, "Raids, Rewards, and Reputations in the Market for Managerial Talent," *Review of Financial Studies*, 16, 1315–1357.

Frydman, C., 2005, "Rising through the Ranks: The Evolution of the Market for Corporate Executives," 1936–2003, Working Paper, MIT.

Gabaix, X. and A. Landier, 2008, "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics*, 123, 49–100.

Garen, J., 1994, "Executive Compensation and Principal-agent Theory," *Journal of Political Economy*, 102, 1175–1199.

Garvey, G. and T. Milbourn, 2006, "Asymmetric Benchmarking in Compensation: Executives Are Rewarded for Good Luck but not Penalized for Bad," *Journal of Financial Economics*, 82, 197–225.

Garvey, G. and T. Milbourn, 2003, “Incentive Compensation When Executives Can Hedge the Market: Evidence of Relative Performance Evaluation in the Cross Section,” *Journal of Finance*, 59, 1557–1581.

Gibbons, R. and K. J. Murphy, 1990, “Relative Performance Evaluation for Chief Executive Officers,” *Industrial and Labor Relations Review*, 43, 30S–51S.

Gillan, S. L., J. C. Hartzell, and R. Parrino, 2006, “Explicit vs. Implicit Contracts: Evidence from CEO Employment Agreements,” Working Paper, Texas Tech and University of Texas.

Grinstein, Y. and P. Hribar, 2004, “CEO Compensation and Incentives—Evidence from M&A Bonuses,” *Journal of Financial Economics*, 73, 119–143.

Grossman, S. and O. Hart, 1980, “Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation,” *Bell Journal of Economics*, 11, 42–64.

Hadlock, C. J. and G. B. Lumer, 1997, “Compensation, Turnover, and Top Management Incentives,” *Journal of Business*, 70, 153–187.

Hall, B. J. and T. A. Knox, 2004, “Underwater Options and the Dynamics of Executive Pay-to-Performance Sensitivities,” *Journal of Accounting Research*, 42, 365–412.

Hall, B. J. and J. B. Liebman, 1998, “Are CEOs Really Paid Like Bureaucrats?” *Quarterly Journal of Economics*, 113, 653–691.

Hall, B. J. and K. J. Murphy, 2002, “Stock Options for Undiversified Executives,” *Journal of Accounting and Economics*, 33, 3–42.

Harford, J. and K. Li, 2007, “Decoupling CEO Wealth and Firm Performance: The Case of Acquiring CEOs,” *Journal of Finance*, 62, 917–949.

Hartzell, J. C., E. Ofek and D. Yermack, 2004, “What’s in It for Me? CEOs Whose Firms Are Acquired,” *Review of Financial Studies*, 17, 37–61.

Haubrich, J. G., 1994, “Risk Aversion, Performance Pay, and the Principal-Agent Problem,” *Journal of Political Economy*, 102, 258–276.

Heron, R. A. and E. Lie, 2007, “Does Backdating Explain the Stock Price Pattern around Executive Stock Options Grants?” *Journal of Financial Economics*, 83, 271–295.

Himmelberg, C. P. and R. G. Hubbard, 2000, “Incentive Pay and the Market for CEOs: An Analysis of Pay for Performance Sensitivity,” Working Paper, Columbia University.

Himmelberg, C. P., R. G. Hubbard and D. Palia, 1999, “Understanding the Determinants of Managerial Ownership and the Link between Ownership and Performance,” *Journal of Financial Economics*, 53, 353–384.

Holmström, B., 1992, *Comments, in Contract Economics*, edited by Lars Werin and Hans Wijkander, Blackwell Publishers, Oxford and Cambridge, MA.

Holmström, B., 1979, “Moral Hazard and Observability,” *Bell Journal of Economics*, 10, 74–91.

Holmström, B. and S. Kaplan, 2003, “The State of U.S. Corporate Governance: What’s Right and What’s Wrong?” *Journal of Applied Corporate Finance*, 15, 8–20.

Holmström, B. and P. Milgrom, 1987, “Aggregation and Linearity in the Provision of Intertemporal Incentives,” *Econometrica*, 55, 303–328.

Huson, M. R., R. Parrino and L. Starks, 2001, “Internal Monitoring Mechanisms and CEO Turnover: A Long-Term Perspective,” 56, 2265–2297.

Inderst, R. and H. Mueller, 2006, “Specific Human Capital and Broad-based Incentive Pay,” Working Paper, New York University.

Janakiraman, S., R. Lambert and D. Larcker, 1992, “An Empirical Investigation of the Relative Performance Evaluation Hypothesis,” *Journal of Accounting Research*, 30, 53–69.

Jensen, M. C., 1993, “The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems,” *Journal of Finance*, 48, 831–880.

Jensen, M. C., 1986, “Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers,” *American Economic Review*, 76, 323–329.

Jensen, M. C. and W. H. Meckling, 1976, “Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure,” *Journal of Financial Economics*, 3, 305–360.

Jensen, M. C. and K. J. Murphy, 1990, “Performance Pay and Top Management Incentives,” *Journal of Political Economy*, 98, 225–264.

Jenter, D., 2005, "Market Timing and Managerial Portfolio Decisions," *Journal of Finance*, 60, 1903–1949.

Jenter, D. and F. Kanaan, 2006, "CEO Turnover and Relative Performance Evaluation," Working Paper, MIT.

Jin, L., 2002, "CEO Compensation, Diversification, and Incentives," *Journal of Financial Economics*, 66, 29–63.

Joh, S. W., 1999, "Strategic Managerial Incentive Compensation in Japan: Relative Performance Evaluation and Product Market Collusion," *Review of Economics and Statistics*, 81, 303–313.

Kale, J., E. Reis and A. Venkateswaran, 2007, "Rank Order Tournaments and Incentive Alignment: The Effect on Firm Performance," Working Paper, Georgia State University.

Kuhnen, C. M. and J. Zwiebel, 2007, "Executive Pay, Hidden Compensation, and Managerial Entrenchment," Working Paper, Northwestern and Stanford,

Lambert, R. and D. Larcker, 1987, "An Analysis of the Use of Accounting and Market Measures of Performance in Executive Compensation Contracts," *Journal of Accounting Research*, 25, 85–125.

Lang, L. H. P. and R. M. Stulz, 1994, "Tobin's q, Corporate Diversification, and Firm Performance," *Journal of Political Economy*, 102, 1248–1280.

Lazear, E. P. and S. Rosen, 1981, "Rank-order Tournaments as Optimal Labor Contracts," *Journal of Political Economy*, 89, 841–864.

Lehn, K. and M. Zhao, 2006, "CEO Turnover after Acquisitions: Are Bad Bidders Fired?" *Journal of Finance*, 61, 1759–1811.

Lie, E., 2005, "On the Timing of CEO Stock Option Awards," *Management Science*, 51, 802–812.

Malmendier, U. and G. Tate, 2005, "CEO Overconfidence and Corporate Investment," *Journal of Finance*, 60, 2661–2700.

Milbourn, T., 2003, "CEO Reputation and Stock-based Compensation," *Journal of Financial Economics*, 68, 233–262.

Mirrlees, J. A., 1999, "The Theory of Moral Hazard and Unobservable Behaviour: Part I," *Review of Economic Studies*, 66, 3–21.

Morck, R., A. Shleifer and R. Vishny, 1988, "Management Ownership and Market Valuation—An Empirical Analysis," *Journal of Financial Economics*, 20, 293–315.

Murphy, K. J., 2002, "Explaining Executive Compensation: Managerial Power versus the Perceived Cost of Stock Options," *University of Chicago Law Review*, 69, 847–869.

Murphy, K. J., 1999, "Executive Compensation," In Orley Ashenfelter and David Card (Eds.), *Handbook of Labor Economics*, Vol. 3B, Elsevier/North-Holland, Amsterdam, 2485–2563.

Nagar, V., 2002, "Delegation and Incentive Compensation," *Accounting Review*, 77, 379–395.

Oyer, P., 2004, "Why Do Firms Use Incentives that Have No Incentive Effects?" *Journal of Finance*, 59, 1619–1649.

Oyer, P. and S. Schaefer, 2004, "Why Do Some Firms Give Stock Options to All Employees? An Empirical Examination of Alternative Theories," *Journal of Financial Economics,* 76, 99–133.

Peng, L. and A. Roell, 2008, "Executive Pay, Earnings Manipulation, and Shareholder Litigation," *Review of Finance*, 12, 141–184.

Prendergast, C., 2002, "The Tenuous Tradeoff Between Risk and Incentives," *Journal of Political Economy*, 110, 1071–1102.

Qiu, H., 2006, "The Role of Debt and Managerial Compensation in Resolving Agency Problems," Working Paper, Hong Kong University.

Rajan, R. and J. Wulf, 2006, "Are Perks Purely Managerial Excess?" *Journal of Financial Economics*, 79, 1–33.

Schaefer, S., 1998, "The Dependence of Pay-Performance Sensitivity on the Size of the Firm," *Review of Economics and Statistics*, 80, 436–443.

Standard and Poor's, 1995*, S&P Compustat ExecuComp: The Executive Compensation Database*, McGraw-Hill, Englewood.

Swan, P. L. and X. Zhou, 2003, "Performance Thresholds in Managerial Incentive Contracts," *Journal of Business*, 76, 665–696.

Weisbach, M. S., 1988, "Outside Directors and CEO Turnover," *Journal of Financial Economics*, 20, 431–460.

Wulf, J., 2007, "Authority, Risk, and Performance Incentives: Evidence from Division Manager Positions inside Firms," *Journal of Industrial Economics*, 55, 169–196.

Wulf, J., 2004, "Do CEOs in Mergers Trade Power for Premium? Evidence from Mergers of Equals," *Journal of Law, Economics, and Organization*, 20, 60–101.

Yermack, D., 2006, "Flights of Fancy: Corporate Jets, CEO Perquisites, and Inferior Shareholder Returns," *Journal of Financial Economics*, 80, 211–242.

Yermack, D., 1997, "Good Timing: CEO Stock Option Awards and Company News Announcements," *Journal of Finance*, 52, 449–476.

*Chapter 18*

# MANAGING CORPORATE RISK

CLIFFORD W. SMITH JR.

*William E. Simon Graduate School of Business Administration*
*University of Rochester*

## Contents



* I would like to thank Mike Barclay, Neil Doherty, John Graham, Ludger Hentschel, Leslie Marx, David Mayers, Erwan Morellec, Deana Nance, Charles Smithson, Robert Schwartz, René Stulz, Sykes Wilford and Lee Wakeman; their collaboration on various aspects of risk management has enriched my understanding of this topic materially.

*Handbook of Empirical Corporate Finance, Volume 2*
*Edited by B. Espen Eckbo*

*DOI: 10.1016/S1873-1503(06)01003-8*

## Abstract

Recent developments in corporate risk management include an expansion of the available instruments, a material reduction in the costs of risk management products, and a more sophisticated understanding of their benefits. The chapter examines the underlying theory of how risk management increases firm value, and it summarizes the evidence on the use of risk management instruments.



## Keywords

risk management, derivatives, insurance, taxes

## 1. Introduction

Risk management by firms has expanded substantially over the past two decades. This expansion has produced both a more sophisticated understanding of the benefits of an appropriate risk management program and a material reduction in the cost of risk management products. Much of the risk management literature focuses on the use of derivatives—forwards, futures, swaps, and options—in hedging corporate exposures to interest rates, foreign exchange rates, and commodity prices. But the array of risk management instruments is much broader. Both financially engineered hybrid instruments and engagement in specific real production activities represent important alternative methods of managing risk.

Devising and implementing an effective risk management strategy involves several steps: (1) the identification and quantification of risk exposures, (2) the design of potential risk management instruments and an assessment of their respective effectiveness, (3) an assessment of the potential benefits and costs of risk management, and (4) the selection and implementation of the strategy.

Much of this discussion focuses on the underlying theory of the mechanisms through which risk management can increase the value of the firm. This is a critical step in the design of an effective corporate risk management strategy. For example, there is apparent disagreement on how one should measure a firm's risk exposures: Should management focus on cash flows, firm value, or reported earnings? Discovering *why* a firm hedges has direct implications for *how* one should measure these corporate exposures as well as *what* instruments the firm should employ to hedge.

## 2. Risk exposures and hedging

Some of the risks to which firms are exposed affect only individual firms, while others affect a broad cross section of firms in the marketplace. Figure 1 arrays these risks along a spectrum. At one end of this risk spectrum are marketwide risks; these risks—for example, the impact of unexpected changes in interest rates, FX rates, or oil prices—are not localized to a specific firm or industry. At the other end are firm-specific risks; these include fires, lawsuits, outcomes of research and development projects, and outcomes of exploration and development activities for firms in natural resource industries.

### *2.1. Risk management instruments*

An advantage of arraying the sources of risk as in Figure 1 is that it illustrates the fact that different risks are managed with different hedging instruments. In the second column of the figure, for example, insurance policies are employed to manage firm-specific risks like fires. Marketwide risks, such as exposures to interest rates, can

| Risk Exposures | | Financial | | Real Production |
|---|---|---|---|---|
| | | Specialized | Hybrid | |
| Firm-specific | Fire | Insurance | | Loss prevention program |
| | Lawsuit | | | Joint venture |
| | Research & development projects | Warrants | Convertible bonds/ convertible preferred stock | |
| | Exploration & development projects | | | |
| | Commodity prices | | Oil-indexed notes | Technology choice |
| | FX rates | Forwards/ futures/ swaps/ options/ | Dual currency bonds | Plant siting |
| Market-wide | Interest rates | | Reverse floating rate notes | Vertical integration |

Fig. 1. Corporate Risk Spectrum.

be managed with specialized derivative instruments, such as forwards, futures, swaps, and options.[1]

Over the past two decades, many new financially engineered securities have been introduced to provide customized solutions to corporate risk management problems.

Since these hybrid securities are structured around bonds or preferred stock, they normally are carried on the firm's balance sheet. In creating these hybrids, financial engineers operate much like General Motors in producing automobiles to meet specific customer demands: GM customizes its cars by assembling various combinations of

[1] Although forward, futures, and simple swap contracts differ in administration of the contract, liquidity, and settlement terms, all three instruments have similar exposure profiles. In buying a forward, futures, or swap contract, the value of the contract appreciates with unexpected increases in the underlying asset price, and it falls with unexpected reductions. Writing a forward, futures, or swap contract produces the opposite exposure; the value of the position falls with unexpected increases in the underlying asset price (see Smith, Smithson, and Wilford, 1989).

off-the-shelf components—frames, engines, trim packages, interior appointments, and so on. Similarly, hybrid securities are customized, but the components that make up these instruments are themselves fairly basic off-the-shelf debt instruments, preferred stock contracts, swaps, forwards, and options.[2]

As illustrated in Figure 1, the firm's choice of real production activities also can be used to manage its risk exposures. For example, moving production overseas changes a firm's foreign-exchange exposure. But producing in a new market with new suppliers, new workers, and different labor laws is a major strategic decision. One material advantage offered by financial risk management products is that they allow more effective separation of production and risk management activities. Moreover, financial contracts are more liquid, so if market conditions and exposures change, this added flexibility permits more rapid adjustments.

### *2.2. Risk exposure*

In analyzing a firm's hedging incentives, it is important to understand the relation between an underlying risk and firm value. Some relations are straightforward; for example, an uninsured casualty loss directly reduces firm value. However, other exposures can be more subtle.

Figure 2 illustrates the exposure profile for an oil company. Because this corporation owns substantial oil reserves, higher oil prices raise revenues and increase firm value. The exposure profile relating the unexpected change in firm value that results from an unexpected change in oil prices thus has a positive slope. (For simplicity, this relation is shown as a straight line.)

But the exposure profile would look quite different for a petrochemical firm: higher oil prices would raise the costs of a major input. Thus, the exposure profile relating the

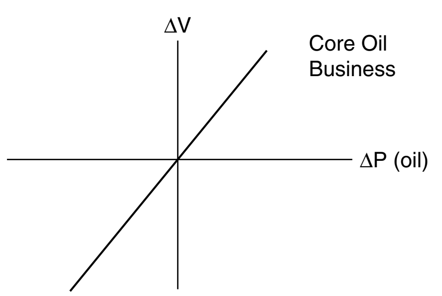



Fig. 2. Exposure profile for an oil company relating the unexpected change in firm value resulting from an unexpected change in oil prices.

[2] For example, an oil producer might issue bonds that include a forward contract on oil, a silver-mining firm might issue bonds incorporating an option on silver, and a copper producer might issue a bond giving investors a strip of copper options, one at every coupon payment (see Smith and Smithson, 1990).

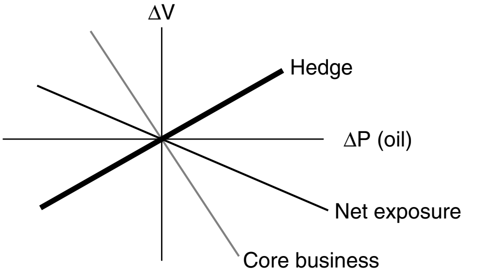



Fig. 3. Exposure profile for a petrochemical company. The Core Business curve relates the unexpected change in firm value resulting from an unexpected change in oil prices. The Hedge curve illustrates the payoff to an oil swap which receives cash-flow based spot oil prices. The net exposure curve reflects the modification in exposure to oil prices from the hedge.

unexpected change in firm value that results from an unexpected change in oil prices has a negative slope. This exposure profile is illustrated in Figure 3.

With the firm's exposure identified, I now can illustrate a basic impact of risk management on firm value. For example, to hedge its exposure to oil prices, the chemical company in Figure 3 must employ a hedging instrument that appreciates in value if oil prices increase. Because gains on such a hedge would offset losses in the firm's core business, risk management reduces the volatility of firm value.

Richard Breeden, former Securities and Exchange Commission chairman, notes that "derivatives are the moving vans of risk—they shift risk from place to place by substituting one type of risk for another." Yet this analysis suggests that such a characterization of derivatives ignores a critical aspect of these instruments. The differences in exposures across firms offer potentially important gains from the use of derivatives. For example, an oil swap with our petrochemical firm as one end-user and our oil company as the other allows both firms to hedge. Therefore, derivatives do more than just shift risk from place to place; they also can reduce the total risk in the system.

Although this hedging activity reduces the exposure of firm value to oil prices of our chemical firm, hedging generally will not affect the firm's optimal pricing or production decisions. Basic economic theory implies that optimal production and pricing occur where marginal cost equals marginal revenue. Relevant costs reflect opportunity costs, and the opportunity cost of the oil is its spot price. For this reason, hedging affects neither the firm's relevant marginal costs nor marginal revenues. Therefore, pricing and production decisions and their decisions to use financial instruments to hedge their exposures generally are separable.

## 3. Benefits of risk management

Because financial risk management reduces the volatility of firm value, one might presume that all firms would want to engage in risk management. Yet there is wide variation in the use of risk management instruments across firms, even among firms that have similar exposures. There are firm characteristics that can provide firms with strong economic incentives to hedge.

### *3.1. Ownership structure*

In analyzing risk management benefits, it can generally be assumed that the objective of the firm is to maximize its current market value. In their personal affairs, risk-averse individuals have incentives to manage risk because doing so lowers the required rate of return in order to engage in a particular risky activity. Thus, for example, an insurance company has a competitive advantage over most individuals in bearing risk and hence is willing to do so at a price lower than the individual would demand. Similarly, for an individual proprietorship, a partnership, or a closely held corporation, the risk aversion of the firm's owners is sufficient to motivate the firm to engage in risk management activities. But for a widely held corporation this logic fails. Portfolio theory implies that a corporation's required rate of return does not depend on total risk but on the systematic risk of its cash flows. Thus, a hedging instrument that works primarily on diversifiable risk does not provide a lower discount rate for a widely held firm whose owners hold well-diversified portfolios. And even if risk management affects systematic risk, as long as the investment is appropriately priced, risk management still will not affect firm value. To illustrate, consider a capital asset pricing model framework. To increase firm value, the firm must acquire an asset that plots above the security market line. But a fairly priced asset will plot on the line. Thus, even if hedging changes the firm's beta, a fairly priced hedge would simply move the firm along the security market line—it would not increase firm value.

Risk management increases the value of a widely held firm by increasing the firm's expected net cash flows—not by reducing its required rate of return. To understand how this might occur, recall the Modigliani-Miller proposition: The firm's financing decisions, including its risk management activities, will not affect firm value, assuming the firm's investment decisions are fixed and as long as there are no taxes and no contracting costs. For my purposes here, it is useful to restate this proposition in a logically equivalent way that emphasizes its managerial implications: If financial decisions, including risk management decisions, affect firm value, they must do so through their effect on investment decisions, taxes, or contracting costs.

### *3.2. Risk shifting within the firm*

Thus far, I have viewed the firm from the perspective of its investors. Of course, the corporation is a vast network of contracts between various parties with conflicting as well as common interests. In addition to bondholders and stockholders, a corporation has other constituencies, such as employees, managers, suppliers and customers. All have vested interests in the company's success.

Like the owners of private or closely held companies, the firm's managers, employees, suppliers, and customers may not be able to diversify their risks; if not hedged, these risks can affect their future payoffs in their respective relationships with the firm. Because they are also risk averse, these groups are likely to require extra compensation to bear any risk that is not assumed by the owners or transferred through hedging to a counterparty (see Smith and Stulz, 1985; Stulz, 1990).

Employees will demand higher wages (or reduce their loyalty or perhaps their work effort) at a company where the probability of layoff is greater. Managers with alternative opportunities will demand higher salaries (or maybe an equity stake in the company) to run firms where the risks of insolvency and financial embarrassment are significant. Suppliers will be more reluctant to enter into long-term contracts, and trade creditors will charge more and be less flexible with companies whose prospects are more uncertain. And customers concerned about the company's ability to fulfill warranty obligations or service their products in the future may be reluctant to buy those products.

Because of limited liability, the amount of risk that can be allocated to stockholders is limited by the capital stock of the company. Companies in service industries, for instance, often employ limited capital. And for such companies, where the claims—and thus the risks—of managers and employees are likely to be large relative to the claims of investors, substantial benefits may be gained from hedging those risks.

Note, however, that one important aspect of achieving these potential risk-management benefits has received little attention—a firm's ability to pre-commit to a hedging strategy. This is less of a problem with some firm-specific risks: supplier, employment, and customer contracts have long stipulated levels of insurance coverage. But it is rare to see a supplier contract that specifies that interest rate risk be managed on an ongoing basis.

Without an ability to pre-commit to hedge, the realized gains to a firm in these dimensions will be lower. It is difficult to rely on implicit reputational effects to support an ongoing hedging policy because of potential incentive incompatibility problems. In circumstances where these claimholders might value hedging quite highly, the firm's stockholders face big incentives to unwind the hedge. (Morellec and Smith, 2007, examine conditions under which shareholders have incentives to maintain the firm's hedging policy after fixed claims have been issued.)

Consideration of comparative advantage in risk-bearing also has implications for the design of compensation contracts. Effective compensation plans achieve an appropriate balance between two potentially conflicting goals—strengthening employees' performance incentives and insulating them from risks beyond their control. Incentive considerations dictate that firms link compensation to performance measures such as share price changes or earnings. Yet a potential problem with such performance proxies is that they contain significant variation that is unrelated to employees' actions. Because financial price risks are a potential source of such noise, companies may also achieve economies in risk-bearing by excluding them more effectively from performance measures that serve as the basis for employee evaluations and bonuses. (See also DeMarzo and Duffie, 1991, 1995; Breeden and Viswanathan, 1998).

### 3.3. Taxes

With progressivity in the tax structure, after-tax payoffs are concave; thus, hedging reduces the expected tax liability, increases after-tax liability, and increases after-tax cash flows and value (Mayers and Smith, 1982; Smith and Stulz, 1985). In their analysis of more than 80,000 COMPUSTAT firm-year observations, Graham and Smith (1999) find that in approximately 50 % of the cases, corporations face convex effective tax functions and thus have tax-based incentives to hedge. In approximately 25% of the cases, firms face linear tax functions and thus have no tax-related incentives to hedge. The remaining firms face concave effective tax functions, which provide a tax-based disincentive to hedge. Of the cases with convex tax functions, roughly one-quarter of the firms have potential tax savings from hedging that appear material; in extreme cases savings exceed 40% of the expected tax liability. For the remaining firms, the tax savings are fairly small. Thus, the distribution of potential tax savings from hedging appears quite skewed.

Firms most likely face convex tax functions when (1) their expected taxable incomes are near the kink in the statutory tax schedule (i.e., taxable income near zero), (2) their incomes are volatile, and (3) their incomes exhibit negative serial correlation (hence the firm is more likely to shift between profits and losses).

The Graham/Smith methods also allow them to decompose the basic structure of the tax code to examine the incremental impact of statutory progressivity, net operating loss carrybacks and carryforwards, investment tax credits, the alternative minimum tax, and uncertainty in taxable income. They find that most of the convexity is induced by the asymmetric treatment of profits and losses in the tax code. Carryback and carryforward provisions effectively allow firms to smooth their losses, thereby reducing tax function curvature at its most convex points but making the function convex over a broader range of taxable income. In contrast, the alternative minimum tax and investment tax credits have only modest effects on the convexity of the tax function.

### 3.4. The underinvestment problem

Although well-diversified stockholders and bondholders may not be concerned about the prospect of unhedged losses per se, they will become concerned if such losses raise the likelihood of insolvency. For example, companies that wind up in Chapter 11 face considerable involvement by the bankruptcy court in their operating decisions as well as substantial direct costs of administration and reorganization. And short of the formal bankruptcy process, financial difficulty can impose large indirect costs. One such cost is the underinvestment problem identified by Myers (1977).

If a company's effective leverage is high enough, management can have incentives to reject an available positive net present value project. As Myers demonstrates, if enough of the value of the new investment is captured by the fixed claimholders so that what is left for the shareholders fails to provide them a normal return given the capital employed and the risk, then the stock price will fall. Taking the project would generate a wealth

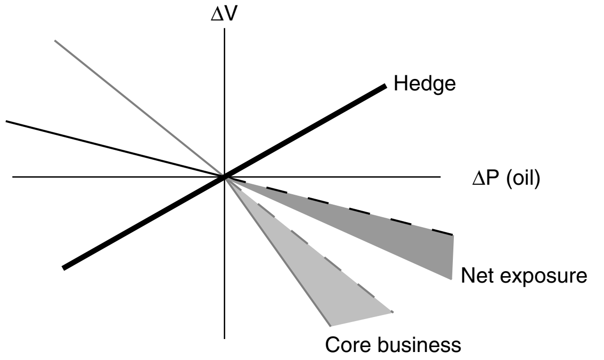



Fig. 4. Underinvestment and Hedging. For a petrochemical firm, an increase in oil prices raises costs and lowers firm value which induces an increase in leverage, an exacerbation of underinvestment problems, and a further fall in value—the shaded wedge. Hedging reduces these underinvestment costs.

transfer from stockholder to lenders. To illustrate, again consider our petrochemical firm from Figure 3. Without hedging, an unanticipated increase in oil prices would raise costs, lower profits, and reduce firm value. But this reduction in firm value causes an induced increase in leverage; higher leverage exacerbates the underinvestment problem and further reduces firm value. This is depicted by the shaded wedge below the core business line in Figure 4. (One can think of the original "core business" line as one that holds investment policy fixed and the steeper line reflects the underinvestment costs.)

Now if this petrochemical firm were to hedge its oil price exposure, the reduction in operating cash flows from an unexpected increase in oil prices would be offset by the increased value of the hedge. Thus, the induced increase in leverage and the exacerbation of the underinvestment problems would be smaller. In Figure 4, this is illustrated by the smaller shaded area associated with the firm's net exposure. Note that this benefit of hedging is that this wedge is reduced, not that the curve is flatter.

Hedging can be an important mechanism for controlling underinvestment costs; it can be a more effective method than reducing leverage. So, an additional benefit of hedging is that it can increase the firm's debt capacity (see Stulz, 1984). This benefit of controlling underinvestment problems should be more pronounced for firms whose value lies primarily in its growth opportunities.

### *3.5. Information problems*

Froot, Scharfstein, and Stein (1993) note that a similar result is obtained when one considers the information asymmetry issues raised by Myers and Majluf (1984). Froot et al. state that raising external capital is costly because of this information asymmetry, and thus a firm like our petrochemical firm in Figure 4 might hedge. Without hedging, higher oil prices would lower firm value, raise leverage, and thus induce management to raise expensive external equity. By reducing the fall in firm value when oil prices rise, hedging reduces the induced increase in leverage and thus the likelihood that the firm would have to access external capital markets. Note, however, that these information asymmetry costs are likely to be small in the specific cases where hedging opportunities are greatest. Because investors can observe events such as a fire, a lawsuit, or a fall in oil prices, informational asymmetries are smaller and managers who raise external capital in these circumstances face more limited costs. This benefit of hedging should be greatest for firms with substantial informational asymmetries between managers and external investors.

### *3.6. Free cash flow problems*

Hedging also can control the free cash flow problem. Jensen (1986) defines free cash flow as cash flow generated by the firm in excess of that required to fund available positive net present value projects. He argues that financing a firm generating substantial free cash flow with debt allows managers to make believable promises to distribute the free cash flow. If we again return to our petrochemical firm in Figure 3, however, unexpectedly lower oil prices reduce costs, increase firm value, decrease the firm's effective leverage, and thus exacerbate the free cash flow problem. The shaded area in Figure 5 illustrates this free cash flow. (Again, one can think of the original "core business" line as one that holds investment policy fixed, and the flatter line reflects the induced overinvestment because of the free cash flow incentives.)

Morellec and Smith (2007) state that hedging can control this free cash flow problem. For example, if this petrochemical firm now hedges its oil price exposure, the increase in operating cash flows from an unexpected reduction in oil prices is offset by the reduced value of the hedge. Thus, both the induced reduction in leverage and the exacerbation of the free cash flow problems are smaller, as illustrated by the

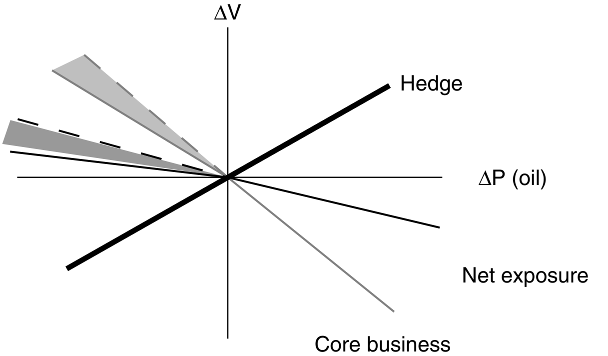



Fig. 5. Free Cash Flow and Hedging. For a petrochemical firm, a reduction in oil prices reduces costs and raises firm value which induces a reduction in leverage and less effective control of the free cash flow problem and a fall in firm value—the shaded wedge. Hedging reduces the free cash flow costs.

smaller shaded area associated with the firm's net exposure in Figure 5. This benefit should be more pronounced for firms whose value is comprised primarily of assets in place.

### *3.7. Hedging motives and methods*

Understanding the motives for risk management is a critical step in designing an effective hedging program for a firm. If the primary consideration for a particular firm in hedging is taxes, this firm should focus on hedging its taxable income. If hedging is prompted by risk-sharing concerns, then a firm where the bonus is linked to accounting returns might focus on hedging accounting earnings. If the cost of financial distress and the underinvestment problem are the primary factors that motivate hedging, the firm should hedge firm value.

In general, hedging value and hedging earnings are not the same thing. FASB rules have evolved to a point where it is typically difficult to obtain hedge-accounting treatment for an off-balance-sheet hedge. Most firms that use standard derivatives are thus required to mark the hedge to market in each accounting period. Yet accounting rules also generally prohibit the firm from marking to market the value of its core assets or liabilities that give rise to the exposure. This means that a firm can engage in risk management activities that, while reducing the volatility of firm value, increase the volatility of reported earnings.

Because access to hedge-accounting treatment for derivatives has been restricted, there has been an increase in the use of structured notes and other hybrid securities. This has occurred in part because accounting rules generally do not require that a risk management contract bundled with a loan or preferred stock issue be marked to market.

Finally, note that with three independent instruments, three different targets can be achieved. Therefore, in principle, with the appropriate choice of hedging instruments a firm could simultaneously manage the impact on its value, reported earnings, and taxable income.

## 4. The costs of risk management

It is important to identify those aspects of the risk management transactions that represent real costs. Basically, the relevant cost of hedging is the sum of any out-of-pocket fees, the implicit cost of the bid-ask spread, and the opportunity cost of management's time in the administration of the program. For standard swaps, many of these costs have fallen dramatically over the past two decades. In the early 1980s, the bid-ask spread for swaps at times exceeded 100 basis points. In 2006, it could be as low as two basis points for a standard interest rate swap. Thus, profound reduction in hedging costs, which reflects the material increase in the liquidity of these markets, makes the net benefit from accessing the market greater and explains part of the observed growth in these markets. Moreover, standardization and increased familiarity with these instruments and their uses have lowered the administrative costs.

In addition to this variation in cost over time, there is also important variation in costs across hedging instruments at a given date. In general, the costs of hedging will be lower: (1) the greater the volume of transactions in a given market, (2) the lower the volatility of the underlying asset price, and (3) the less private information is relevant for pricing the underlying asset. Therefore, hedging costs are generally lower for derivatives on interest rates and major currencies but higher for more customized hedging instruments.

## 5. Evidence on corporate hedging

Derivative instruments can be used either to speculate on financial price movements or to hedge. Thus, a basic question to be addressed is: do firms use derivatives to hedge or to speculate? Although the evidence is still preliminary, the answer appears to be, for the most part, to hedge.

An early survey of the corporate use of derivatives was conducted by Dolde (1993). The overwhelming majority of the 244 Fortune 500 companies that responded to Dolde's questionnaire reported that their policy was to use derivatives primarily to hedge their exposures. At the same time, however, only about 20% of the responding firms reported that they attempted to hedge their exposures completely. Moreover, smaller firms—those likely to have lower credit ratings and, hence, greater default risk—reported hedging larger percentages of their exposures than big companies.

About 90% of the firms in Dolde's survey also said that they sometimes took a view on the market direction of interest rates or exchange rates. And although roughly one in six of even these companies hedged its exposures completely, the rest claimed to modify their positions to accommodate their view.

For example, if they expected rates to move in a way that would increase firm value, they might hedge only 30% of their exposure; but if they expected rates to move in a way that would reduce firm value, they might hedge 100% of their exposure. Only two firms said that they would use hedge ratios outside the 0–100% range. In effect, this means that less than 1% of the firms said they would use derivatives to speculate and enlarge an existing exposure.

Of course, surveys present a number of problems. For example, some companies might be reluctant to admit that they use derivatives to speculate. Yet other evidence on hedging also bears out this corporate propensity to hedge. For example, a mounting number of studies find that firms with operating characteristics that theory suggests should make hedging more valuable appear to use more derivatives. If derivatives were used primarily to speculate, no such associations should be expected.

Before examining those empirical results, however, I must note one caveat about the data examined by these studies. There are important limitations in our current ability to judge whether one firm hedges more than another. These limitations are of three basic varieties.

First, some firms hedge using derivatives, whereas others employ hybrid debt and preferred stock issues. Many empirical studies of corporate hedging focus on hedging using derivatives but ignore the risk management implications of hybrid securities issued by the firm.

Second, over the past decade there has been wide variation in the disclosure of corporate hedging activities. Prior to the adoption of SFAS 105, disclosure by firms was generally required only if a hedging activity was material. Yet, some firms voluntarily disclosed more than was required. This means that firms with essentially equivalent hedging policies might appear different simply because of different disclosure policies. The adoption of SFAS 105 reduced this problem but did not eliminate it.

The third problem is more fundamental. Even with complete access to hedging data, if two firms employ different risk management instruments, judging which firm hedges more can be difficult. For example, assume that firm A has \$10 million (notional principal) of three-year interest rate swaps, firm B has \$20 million of three-year swaps, and firm C has \$10 million of seven-year swaps. Firm A clearly hedges less than either B or C, but comparing B with C is more difficult. For the next three years, B hedges more than C, but for the following four years firm C hedges more.

If we turn to options, the problems become dramatically more difficult—attempting to compare firms with contracts of different size and different exercise prices is quite difficult. In principle, one could estimate the contracts' deltas, but deltas depend on the prices at which they are evaluated. These data problems limit the power of all the empirical work in this area.

### *5.1. Investment policy*

Geczy, Minton, and Schrand (1997), Nance, Smith, and Smithson (1993), and Mian (1996) examine whether firms with more growth opportunities in their investment opportunity sets are more likely to hedge. Nance et al., who examine hedging activity by 169 Fortune 500 firms, conclude that firms with more growth options hedge more. Mian, analyzing data for 3022 firms listed on Compustat, finds conflicting evidence across different measures. Geczy et al., who examined the use of currency derivatives by 372 large firms, find no significant relation. Morellec and Smith (2007) suggest that one reason studies fail to find a robust relation between the firm's investment opportunities and hedging is that hedging can control both underinvestment and free cash flow problems. Thus, firms with both substantial assets in place and growth options can have important incentives to hedge.

### *5.2. Financing policy*

Several studies (e.g., Block and Gallagher 1986, and Geczy, Minton, and Schrand, 1997) examine the association between hedging and leverage. Most report no significant association between the two. This result potentially reflects a fundamental statistical problem. Both leverage and hedging decisions are endogenous. Thus, simply putting leverage on the right-hand side of an ordinary-least-square (OLS) regression to explain hedging choices introduces a potential simultaneous-equation bias in the reported coefficients. At our current state of knowledge, this will be a difficult problem to solve. It is not clear that our theory is sufficiently rich to identify structural equations so that simultaneous equation estimation methods can be employed.

Booth, Smith, and Stolz (1994), Wall and Pringle (1989), and Mayers and Smith (1990) examine the impact of the probability of financial distress on the incentive to hedge. Booth et al. study the use of interest rate futures by 238 financial institutions; Wall and Pringle, hedging by 250 swap users from the NAARS database; and Mayers and Smith, reinsurance purchases for a sample of 1276 property-casualty insurance companies. Wall and Pringle find that firms with lower credit ratings use more swaps, Mayers and Smith report that insurers with lower Best's ratings reinsure more, and Booth, Smith, and Stolz report that S&Ls hedge more than banks. These results suggest that with a higher probability of financial distress, firms have stronger incentives to hedge.

### *5.3. Managerial incentives*

Tufano (1996) examines managerial incentives to hedge and concludes that firms that compensate managers with more stock hedge more, although firms that use more stock options hedge less. He argues that with more restricted stock, managerial risk aversion increases the incentives to hedge. But the impact of volatility on option values (see Black and Scholes, 1973) implies that managers who receive options should hedge less. Geczy et al. (1997) also look at managerial option ownership. The evidence suggests that firms that use currency derivatives grant more options to their managers than nonusers.

Geczy et al. are appropriately concerned about simply inserting a measure of managerial compensation as an explanatory variable in a hedging equation. Compensation is endogenous and thus doing so would introduce simultaneous equation problems. Interestingly, Tufano's analysis is less likely to suffer from such problems. He focuses on 48 firms in the gold-mining industry. Because our theories are unlikely to be able to explain the observed variation in compensation structure across this reasonably homogeneous population of firms, any simultaneous equation bias is likely to be small in his analysis.

### *5.4. Firm size*

Booth et al., Block and Gallagher, Nance et al., Mian, Geczy et al., and Tufano (1996) all report that large firms are more likely to hedge than are small firms. This finding is consistent with the proposition that there are significant transaction and information costs as well as scale economies. Hedging instruments frequently are viewed as sophisticated products, and large firms are more likely to hire managers with the requisite expertise to manage a hedging program.

The analysis in Mayers and Smith, however, indicates that small insurers reinsure more. This result is consistent with size-related tax and financial distress incentives.

Moreover, the information cost issues that are associated with derivative instruments for industrial or financial firms are likely to be less important for insurance companies' reinsurance purchases.

### *5.5. Taxes*

Geczy et al., Graham and Rogers (2002), Mian, Nance et al., and Tufano test the proposition that statutory progressivity of the tax function provides an incentive for firms to hedge. Some find that the greater the likelihood that a firm's pretax income falls in the progressive region of the tax schedule, the more likely the firm is to hedge. The effective tax schedule can be convex because of limitations on the use of tax credits. Mian finds more hedging by firms with more foreign tax credits, and Nance, Smith, and Smithson document more hedging by firms with more investment tax credits. They report inconsistent evidence on the impact of tax-loss carryforwards.

Note, however, that these variables may proxy for things other than a firm's tax status. For example, the presence of tax-loss carryforwards also might proxy for financial distress; similarly, ITCs may proxy for aspects of a firm's investment opportunities; finally, foreign tax credits potentially proxy for a foreign currency exposure. (Both Mian, 1996, and Houston and Mueller, 1988, find that firms with foreign operations are more likely to hedge.)

To the extent that these variables proxy for firm characteristics other than the progressivity of the firm's effective tax schedule, we have a potentially important identification problem. More powerful tests of these tax hypotheses will require proxies that better isolate the firm's tax status.

### *5.6. Ownership structure*

There has been little analysis of the use of forwards, futures, swaps, or options by closely held firms, largely because of data limitations. Within their sample of property-casualty insurance companies, however, Mayers and Smith examine closely held insurance firms. In their analysis of hedging through reinsurance contracts, they find that closely held insurance firms buy more reinsurance. This finding is consistent with the proposition that firms with owners whose portfolios are more ill-diversified have stronger incentives to hedge.

## 6. Conclusion

Derivative instruments represent a material addition to the corporate financial officer's tool kit. These instruments provide incredible flexibility in structuring a customized risk management strategy for the firm.

To realize their potential requires a detailed understanding of the instruments and their uses. Used appropriately, they reduce risk and increase firm value. But used inappropriately, they have caused firms to collapse. As noted above, implementation of a risk identification strategy involves several steps: (1) exposure identification, (2) instrument design, (3) net benefit assessment, and (4) strategy implementation. Heretofore, the academic community has focused substantially all its attention on the first three steps. Yet the implementation step is critical to an effective strategy. As with other aspects of organizational design, it involves three critical aspects: (1) the assignment of decision rights—who has the authority to structure and implement the policy and who has the responsibility to monitor these decisions, (2) the methods of rewarding these key individuals, and (3) the structure of systems to evaluate the effectiveness of the policy, including the details of its implementation (see Brickley, Smith, and Zimmerman, 2004). To date, this last step has received little academic attention, yet it may be the single most important in terms of creating firm value.

## References

Black, F. and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–659.

Block, S. B. and R. J. Gallagher, 1986, "The Use of Interest Rate Futures and Options by Corporate Financial Managers," *Financial Management*, 15, 73–78.

Booth, J. R., R. L. Smith and R. W. Stolz, 1984, "The Use of Interest Rate Futures by Financial Institutions," *Journal of Bank Research*, 15, 15–20.

Breeden, D. and S. Viswanathan, 1998, "Why Do Firms Hedge: An Asymmetric Information Model" Working Paper—Duke University.

Brickley, J. A., C. W. Smith and J. L. Zimmerman, 2004, *Managerial Economics and Organizational Architecture*, McGraw-Hill/Irwin.

DeAngelo, H. and R. W. Masulis, 1980, "Optimal Capital Structure under Corporate and Personal Taxation," *Journal of Financial Econornics*, 39, 1199–1206.

DeMarzo, P. M. and D. Duffie, 1995, "Corporate Incentives for Hedging and Hedge Accounting," *Review of Financial Studies*, 8, 743–771.

DeMarzo, P. M. and D. Duffie, 1991, "Corporate Financial Hedging with Proprietary Information," *Journal of Economic Theory*, 53, 261–286.

Dolde, W., 1993, "The Trajectory of Corporate Financial Risk Management," *Journal of Applied Corporate Finance*, 6(3), 33–41.

Froot, K. A., D. S. Scharfstein and J. C. Stein, 1993, "Risk Management: Coordinating Corporate Investments and Financing Policies," *Journal of Finance*, 5, 1629–1658.

Geczy, C., B. A. Minton and C. Schrand, 1997, "Why Firms Use Currency Derivatives?" *Journal of Finance*, 52, 1323–1354.

Graham, J. R. and C. W. Smith, 1999, "Tax Incentives to Hedge" *Journal of Finance*, 54, 2241–2262.

Graham, J. R. and D. A. Rogers, 2002, "Do Firms Hedge in Response to Tax Incentives?" *Journal of Finance*, 57, 815–839.

Guay, W., 1999, "The Impact of Derivatives on Firm Risk: An Empirical Examination of New Derivatives Users," *Journal of Accounting and Economics*, 26, 319–362.

Guay, W. and S. P. Kothari, 2002, "How Much Do Firms Hedge with Derivatives?" *Journal of Financial Economics*, Forthcoming.

Houston, C. O. and G. G. Mueller, 1988, "Foreign Exchange Rate Hedging and SFAS No. 52—Relatives or Strangers?" *Accounting Horizons*, 2, 50–57.

Jensen, M. C. 1986, "Agency Costs of Free Cash Flow and Corporate Finance Takeovers" *American Economics Review*, 76, 323–329.

Mayers, D. and C. W. Smith, 1990, "On the Corporate Demand for Insurance: Evidence from the Reinsurance Market," *Journal of Business*, 63, 19–40.

Mayers, D. and C. W. Smith, 1982, "Corporate Insurance and the Underinvestment Problem," *Journal of Risk and Insurance*, 54, 45–54.

Mayers, D. and C. W. Smith, 1987, "On the Corporate Demand for Insurance, *Journal of Business*, 55, 281–296.

Mian, S. L., 1996, "Evidence on Corporate Hedging Policies," *Journal of Financial and Quantitative Analysis*, 31, 419–439.

Morellec, E. and C. W. Smith, 2007, "Agency Conflicts and Risk Management" *Review of Finance*, 11(1), 1–23.

Myers, S. and N. Majluf, 1984, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have" *Journal of Financial Economics*, 13, 187–221.

Myers, W., 1977, "Determinants of Corporate Borrowing," *Journal of Financial Economics*, 5, 147–175.

Nance, D. R., C. W. Smith and C. W. Smithson, 1993, "On the Determinants of Corporate Hedging," *Journal of Finance,* 48, 267–284.

Smith, C. W. and C. W. Smithson, 1990, *Handbook of Financial Engineering*, Harper and Row, New York.

Smith, C. W., C. W. Smithson and D. S. Wilford, 1989, "Managing Financial Risk," *Journal of Applied Corporate Finance*, 1(4), 27–48.

Smith, C. W. and R. Stulz, 1985, "The Determinants of Firms' Hedging Policies," *Journal of Financial and Quantitative Analysis*, 20, 391–405.

Stulz, R. M. 1984, "Optimal Hedging Policies" *Journal of Financial and Quantitative Analysis*, 19(2), 127–140.

Stulz, R. 1990, "Managerial Discretion and Optimal Findings Policies" *Journal of Financial Economics*, 26(1), 3–26.

Tufano, T., 1996, "Who Manages Risk? An Empirical Examination of Risk Management Practices in the Gold Mining Industry," *Journal of Finance*, 51, 1097–1137.

Wall, L. D. and J. J. Pringle, 1989, "Alternative Explanations of Interest Rate Swaps: An Empirical Analysis," *Financial Management*, 18, 119–149.

# AUTHOR INDEX

**C**

## G



## H

**T**



**U**



**V**



**W**

# SUBJECT INDEX

## C

## M

## T