THE EFFECT OF THE SARBANES-OXLEY ACT OF 2002 ON CORPORATE GOVERNANCE

by

ILIANA DIMITROVA

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2010

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Professor Devra Golbe

Date — Chair of Examining Committee

Professor Merih Uctum

Date — Executive Officer

Professor Devra Golbe

Professor Ingmar Nyman

Professor Wim Vijverberg

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

THE EFFECT OF THE SARBANES-OXLEY ACT OF 2002

ON CORPORATE GOVERNANCE

by

Iliana Dimitrova

Advisor: Professor Devra Golbe

The Sarbanes-Oxley Act of 2002 (SOX) introduced strict governance rules for public companies in the US to prevent diversion of resources by corporate insiders. Another governance mechanism that serves to align the interests of managers and directors of the firm with outside shareholders is ownership structure as discussed in Chapter 1. The equity structure of dual-class firms is characterized by multiple classes of shares with differential voting rights, which allows insiders to hold disproportionate dividend and voting rights. Dual-class status may serve as a protection from value-reducing takeovers and help the realization of growth options. At the same time, since managerial cash-flow rights have an incentive alignment effect, while voting control reduces the discipline of the market for corporate control, dual-class firms may protect large private benefits of insiders.

Chapter 2 evaluates the effect of SOX on incentive alignment between managers and shareholders. The value-increasing effect of managerial cash-flow rights is expected to be smaller in magnitude after SOX. Similarly, the negative effect of voting control should be less pronounced after SOX. Using dual-class firms, these effects can be separately estimated in a panel data model. Results indicate a smaller positive effect of cash-flow rights which is evidence that the governance provisions of SOX improve managerial incentive alignment.

Chapter 3 analyses differences in returns on inferior and superior shares of dual-class firms, which are expected to reflect reduced value of voting control after SOX. The market models in the event study methodology account for non-synchronous trading due to thin market in the superior shares. There is no evidence of improved performance of inferior shares vs. superior shares in the market's expectations around key SOX dates.

Insiders would value control less if SOX limits private benefits of control and may want to divest their holdings by recapitalizing into single-class. At the same time, if dual-class status is beneficial to outside shareholders, the new rules would reduce the cost in terms of diversion of resources. Chapter 4 empirically tests the effect of SOX on recapitalization decisions of dual-class firms into single-class by formulating a survival model. Results indicate that the propensity of firms to recapitalize into single class is significantly higher after the passage of the Act.

Acknowledgements

I would like to express my deep gratitude to Professor Devra Golbe for providing me with invaluable support and guidance through every stage of this dissertation. Her contributions were extremely helpful in all aspects of the project, from formulating hypotheses to modeling to editorial remarks. I am also thankful for the independence she gave me in working on my thesis.

I am indebted to Professor Ingmar Nyman for his comments on corporate governance theory, encouraging me to see the big picture. Professor Wim Vijverberg contributed with numerous suggestions for the estimations and provided a reality check for several issues in financial economics.

I would like to thank participants in the 2010 PhD meetings at City University London and the Graduate Center Economics seminar at City University of New York for helpful comments and suggestions.

Paul Gompers, Joy Ishii and Andrew Metrick, the authors of “Extreme Governance” (2008) generously provided me with their dual-class data.

I would also like to thank Professor Thom Thurston for supporting my academic career and extracurricular work at the Economics department, and Professor Michael Grossman for his genuine interest in my progress.

Finally, I am grateful to my husband Hristo Dimitrov for his help with corporate law issues, and for his endless support throughout my graduate studies.

Table of Contents

List of Tables

List of Figures

CHAPTER 1. GENERAL CONCEPTS

1. Introduction

The Sarbanes-Oxley Act (SOX) was adopted in July 2002 in response to scandals like Enron and WorldCom, which involved accounting and governance malpractices. The legislation introduced strict corporate governance rules for public companies in the US in order to prevent fraud and diversion of resources by corporate insiders. There may be a conflict of interest between managers and outside shareholders due to the separation of ownership and control in a corporation. This agency problem includes misalignment of incentives as well as extraction of private benefits. The governance provisions of SOX are meant to improve incentive alignment and limit extraction of private benefits by corporate insiders. This dissertation examines the effect of SOX on various dimensions of governance though the lens of dual-class ownership. The equity structure of dual-class firms is characterized by multiple classes of shares with differential voting rights. Thus dual-class firms are a useful empirical tool to study the effect of SOX because dual-class shares allow for disproportionate dividend and voting rights of insiders and separate identification of their effects on firm value.

Chapter 2 of this thesis evaluates the effect of SOX on incentive alignment between managers and shareholders. Managerial equity ownership helps resolve agency problems by aligning incentives but the voting control that comes with large shareholdings allows the extraction of private benefits by insiders. If SOX is successful in improving incentive alignment, the value-increasing effect of managerial cash-flow rights is expected to be smaller in magnitude after SOX. Similarly, if SOX limits extraction of private benefits, the negative effect of voting control should be less pronounced after SOX.

The voting control of managers can insulate them from the market for corporate control and allow them to divert resources at the expense of outside shareholders. The corporate governance provisions of SOX are expected to reduce such extraction of private benefits and hence the value of voting control. Chapter 3 analyses differences in returns on inferior and superior shares of dual-class firms due to changes in the value of voting control. Compliance costs as well as contemporaneous macroeconomic factors affect the returns on both classes of shares of the firm in the same way, so only changes in the value of control will be reflected in the difference in returns. Thus improved governance due to SOX is expected to cause better performance of inferior vs. superior shares.

Chapter 4 looks at the decision of dual-class firms to recapitalize into single-class. If the governance provisions of SOX limit extraction of private benefits, then insiders would want to divest control by recapitalizing into single-class. However, the choice of dual-class structure may be dominated by factors creating shareholder value, such as protection from value-reducing takeovers and realization of growth options. In that case, the new governance rules would allow outside shareholders to reap the benefits of dual-class structure longer by reducing the cost of insiders' control. I empirically test the effect of SOX on recapitalization decisions by collecting a survival-time dataset and formulating a duration model of the time spent in dual-class status until recapitalization into single-class.

As the previous discussion indicates, there are several concepts essential to the analysis and recurring in the subsequent chapters: agency problems, managerial equity ownership and dual-class structure. The purpose of the present chapter is to explain in detail these general concepts. This chapter also includes a section on the main provisions of Sarbanes-Oxley.

2. Agency theory

2.1. Agency problems

An agency relationship is a contract under which the principal engages the agent to perform some service on his behalf whereby delegating some decision making authority to the agent [Jensen and Meckling (1976)]. The agent is hired as an expert who knows more than the principal, i.e. there is an intrinsic and unavoidable information asymmetry in the agency relationship. Agency problems arise when the agent does not act in the principal's interest, for example by choosing suboptimal effort level which is unverifiable by the principal. The principal can limit divergences from his interest by establishing appropriate incentives for the agent and by incurring monitoring costs. The agent may incur bonding costs to guarantee that she will not take actions to harm the principal. The monitoring and bonding costs are called agency costs.

Agency theory helps explain the choice of capital structure in terms of equity versus debt financing of projects at a corporation. Optimal capital structure is achieved by maximizing firm value and hence minimizing the agency costs associated with the choice between equity and debt financing. Similarly, optimal ownership structure is defined by minimizing the agency costs associated with the choice between inside and outside equity.

2.2. Solutions to agency problems

Thus one of the corporate governance mechanisms to alleviate agency problems is inside equity ownership. Share ownership by managers helps align managerial incentives with the interests of outside equity holders. The possibility of takeovers if share value is not maximized under current management also provides a disciplining force for insiders. The market for corporate control is closely related to managerial share ownership. They are negatively related:

the higher the inside share ownership and hence voting control, the weaker the disciplining force of takeovers.

There are a variety of other mechanisms for monitoring and motivating managers, such as the discipline imposed by the capital market, the board of directors, incentive contracts based on performance, and regulation[1]. It is important to consider them because they, like managerial equity ownership, may affect firm value.

Capital structure is an important mechanism for motivating managers in agency problems. For example, in a model where managerial effort is a problem, share ownership can be a solution but debt is better in a sense that it provides the highest upside potential for managerial reward [Jensen and Meckling (1976)]. When a project is financed with debt, all of the upside potential accrues as a payoff to the entrepreneur and he has an incentive to choose high effort level. In fact, in some agency models, such as the free cash flow theory, only debt works as it limits the unproductive use of free cash flows [Jensen (1986)]. On the other hand, debt financing can lead to value-decreasing asset substitution, which involves undertaking high-risk projects and reaping the upside potential in case of success by the owner-manager, while the debtor bears the cost in case of failure.

The board of directors is a mechanism that has attracted a lot of attention recently. A relatively new anti-takeover device such as the staggered board has a negative effect on firm valuation [Bebchuk and Cohen (2005)]. Staggered boards make it extremely difficult for a bidder to gain control by requiring at least a year between the takeover bid and taking control of the board, and two elections far apart in time. On the other hand, outside directors can serve as a

[1] Another mechanism for aligning the incentives of managers and shareholders that I am not going to address is the labor market. Although this was recognized in a discussion of the firm as a "nexus" of contracts, the labor market is out of focus in recent research. Managers have firm-specific human capital and incentives to invest in it should be value-increasing. The labor market for their skill should regulate their behavior [Fama (1980)].

countervailing force to insiders with a lot of voting power. Under SOX, for example, public companies are required to have an audit committee of the board composed entirely of independent directors. The degree of independence of the board can have a positive effect on value, according to Bebchuk and Cohen (2005).

Performance pay for managers and designing an incentive contract is a method for providing appropriate incentives for managers. To some extent, performance pay overlaps with share ownership by managers but it can also be based on certain performance targets or indicators [Fahlenbrach (2008)]. Thus a performance pay scheme may include meeting performance targets as well as compensation in the form of shares or options.

Regulation such as SOX is also a potential solution to agency problems. Research shows that regulation can serve as substitute governance mechanism. La Porta et al (2002) find that the effect of managerial cash-flow ownership as an incentive device differs across countries depending on their corporate governance laws. The positive affect of cash-flow ownership in a cross-section of firms implies that firms with low managerial holdings have lower value relative to otherwise identical firms. This effect is less pronounced in countries with strong investor protection and governance laws so that regulation compensates as a governance mechanism.

3. Managerial equity ownership

Ownership structure, i.e. the choice between inside and outside equity, has been studied extensively as a corporate governance mechanism. In theory, managerial ownership of equity can alleviate the agency problem by aligning incentives [Jensen and Meckling (1976), La Porta et al (2002)]. When managers own a share of cash-flow rights, they have an incentive to maximize firm value which is in the interest of shareholders.

At the same time, the voting rights associated with inside ownership create the potential of entrenchment due to insiders' ability to extract private benefits [Grossman and Hart (1988), Harris and Raviv (1988)]. In these models, the disciplining force of the market for corporate control is weaker, hence there is a conflict between controlling shareholders and minority shareholders. The way that entrenched managers can engage in value-reducing activities include empire building, sales growth maximization, perquisite taking, and maximizing employees' welfare. Large outside shareholders can also engage in value-reducing activities such as dealing on preferential terms with the company—for example, borrowing at low interest rate.

There are two other important factors associated with managerial equity ownership that must be included in the analysis, however. First, there is a cost associated with concentrated ownership, which may be value-decreasing. Higher managerial ownership levels imply less portfolio diversification for risk-averse managers [Fama (1980)]. Portfolio theory states that for any given level of expected return, the efficient portfolio minimizes the variance of returns by choosing appropriate weights for the stocks. Under certain conditions, the optimal portfolio is in fact the market portfolio. The constraint that managers should hold a large share in their company prevents the optimal diversification of risk. If shareholders design incentive contracts for the managers that force them to hold a higher percentage of the cash-flow rights, this aligns incentives and is value-increasing. At the same time, it entails a cost to managers so shareholders may not be able to retain the best possible team in terms of human capital which is value-decreasing [Gomes (2000)]. The optimal contract involves a trade-off between diversification and incentives for performance.

Second, there may be a value-increasing effect of control because of the possibility of value-reducing takeovers. In some cases, managers receive a bonus for acquiring another

company, and their compensation and prestige may be tied to company size. Thus consolidating control allows managers to protect the company from hostile value-decreasing takeovers. Consolidation of voting control can be valuable in firms where it is difficult for insiders to communicate information about performance or investment opportunities [Stein (1988)]. Managers, to avoid being replaced in a takeover, may undertake costly methods of signaling, including investing in projects that have observable payoffs even though these projects have lower expected returns than others with less observable payoffs. Thus voting control as takeover protection can be value-increasing.

4. Factors determining dual-class status

4.1. Dual-class firms

Dual-class firms in the US comprise about six percent of the number of public companies and eight percent of the market capitalization [Gompers et al (2008)]. Dual-class companies are more common in Europe and East Asia [Cronqvist and Nilsson (2005), Nenova (2003)]. About two-thirds of East Asian firms are controlled by a single shareholder with higher control rights than cash-flow rights [Claessens et al (2002)].

This study uses dual-class firms in the US in the period 1994-2002 as an empirical tool. The defining feature of dual-class firms is that they have multiple classes of shares with differential voting rights. Usually there are two classes, with ten votes attached to the superior voting stock and one vote attached to the inferior stock. The superior class may also elect a higher proportion of directors on the board. Dividends may differ between the classes, but usually the inferior class pays the higher dividend. In companies with a one share-one vote structure voting and dividend rights bestowed by the shares are proportionate. In contrast,

insiders of dual-firms owned 37% of the cash-flow rights but 59% of the voting rights on average in 2002. In terms of inside holdings, officers and directors held 80% of the outstanding number of superior shares, while holding 23% of the inferior shares. This seems high relative to share holdings of managers at single-class firms, which according to Morck et al (1988) are about 10% on average for Fortune 500 firms. The percentage is 18% in Himmlberg et al (1999), which is higher because the sample is representative of all Compustat firms including smaller firms. Thus the particular mix of dividend and voting rights of insiders at dual-class firms is achieved through ownership of not only superior but also inferior shares.

Although dual-class firms are used mostly as an empirical tool in this thesis, it is important to understand the factors determining the choice of dual-class status, which is a deviation from the one share-one vote structure. The literature on this subject parallels the discussion about managerial equity ownership and the pros and cons of insider cash-flow ownership and voting control, but with some additional considerations. The central question is whether dual-class status is optimal for outside shareholders, or whether the private benefits of insiders outweigh the benefits to outside shareholders, leading to suboptimal security value.

4.1. Optimality of one share-one vote

In the case of widely-held firms, Grossman and Hart (1988) argue for the optimality of one share-one vote rule in general. They focus on takeover bids as a mechanism for allocating control. There are two types of control benefits—benefits to security holders which determine firm value and private benefits to the controlling party. One share-one vote maximizes the importance of benefits to security holders relative to benefits to the controlling party and hence encourages the selection of an efficient management team.

Grossman and Hart (1988) note that the agency problem caused by the delegation of control does not bear directly on the security-voting structure--it implies only that management should receive performance-based compensation. It is clear that shareholders as a group have an incentive to monitor management and hence tying votes to shares may be desirable to allow them to act on this incentive. In practice, however, monitoring by individual small shareholders is limited because of free-rider problems. Then monitoring of management will be effective when a single party becomes large enough to internalize the externalities of collective action by making a takeover bid. This suggests that the main impact of a firm's security voting structure will be in its influence on the market for corporate control. When the rival's private benefits are lower than the incumbent's but their security benefits are higher, the one share-one vote structure increases value for shareholders because the bidder can use some their security value to pay for control in the takeover bid. Sufficient conditions are derived for one share-one vote to be optimal overall, under which deviations such as dual-class shares are value-reducing.

Shleifer and Vishny (1986) observe than even without the possibility of takeovers, the one share-one vote structure allows outside shareholders control that is commensurate with their cash-flow claims and motivates them to monitor managers. Monitoring of managers is a public good and its provisions can be made by parties who already own a large share of the firm. As the largest consumer of the public good, large shareholders may pay for it themselves and bring about value-increasing changes in corporate policy in widely-held corporations.

In the case of closely-held firms where a controlling shareholder owns 51% or more of the equity, one share-one vote is optimal according to the model in Burkart and Lee (2008). A control transfer can only take place with the consent of the controlling shareholder and is therefore best viewed as the outcome of a bilateral negotiation between the controlling

shareholder and the potential acquirer. A mandatory one share-one vote rule would improve the allocation of takeover premium, while deviations are value-decreasing for minority shareholders.

4.2. Optimal deviations from one share-one vote

Grossman and Hart (1988) find that under certain conditions, deviations from one share-one vote may be desirable. Particularly in the case of widely-held companies with significant private benefits, dual-class status can be value-increasing. For example, if both the rival's public and private benefit are higher than those of the incumbent, dual-class shares help shareholders extract more of the rival's private benefit in the competition for control.

More recently, Kalay and Pant (2008) claim that one share-one vote is neither optimal nor enforceable. It is not optimal because by making votes "more expensive" in terms of cash-flows, dual-class structure helps transfer value from private benefit to security value. It is not enforceable because in the presence of derivatives markets, shareholders can choose their desired mix of cash-flows/votes and vary it through time.

In closely-held corporations, private benefits are a motivation for the choice of dual-class status at IPO when the ownership wishes to retain control [DeAngelo and DeAngelo (1985)]. This is not necessarily value-reducing for outside shareholders, because in a perfect market, investors will factor it into the share price. Thus the owner is willing to sell under-priced securities in exchange for keeping control, and the loss of value at the IPO is compensated with extracting private benefits.

There are factors other than private benefits that can motivate the choice of dual-class structure as well. Concentration of voting control requires that insiders hold a large share of the equity in the firm. While the incentive-alignment effect of managerial equity ownership is value-

increasing, there is a cost for managers from holding an undiversified portfolio in terms of risk [Fama (1980)]. The constraint that managers should hold a large share in their company prevents the optimal diversification of risk and imposes a cost to managers so that the best team may not be retained. A dual-class structure can alleviate this problem and help achieve optimal management for the company.

Another factor is related to the asymmetric information in the market for corporate control. Consolidation of voting control can be valuable in firms where it is difficult for insiders to communicate information about increased investment opportunities [Stein (1988)]. Managers may undertake costly methods of signaling in order to avoid being replaced in a takeover if the firm is currently undervalued. This includes investing in projects that have observable payoffs even though these projects have lower expected returns. Thus voting control can serve as protection from value-reducing takeovers.

Firms with growth opportunities are particularly susceptible to value-reducing takeovers as they may be currently undervalued in a market with asymmetric information. Lehn et al (1990) find that dual-class status is more common among high-growth firms as it helps the realization of growth options. In addition, firm growth implies the need for financing. When issuing equity is the optimal choice of financing, there would be a bigger loss of control when votes are proportional to cash-flows as with single-class shares relative to dual-class shares. Entrepreneurs may resort to inferior forms of financing as they try to avoid the risk of losing control by issuing equity. Therefore, firm growth and the corresponding need for financing are a reason to deviate from one share-one vote structure in Burkart and Lee (2008).

Another potential efficiency gain is suggested by DeAngelo and DeAngelo (1985). Consolidation of voting control may encourage managers to invest in firm-specific human capital

or otherwise the returns on such investment can be appropriated by outside bidders. If shareholders wish to insulate managers from the threat of takeovers by consolidating voting control among the managers, they may not require the insiders to increase their ownership of residual claims commensurately since the costs of foregone diversification may be large [Lehn et al (1990)].

Another reason to deviate from one share-one vote for the firm may be the fact that private benefits serve as efficient compensation for managers [Burkart et al (1997)]. Managers make valuable contributions to the firm for which it may be difficult to compensate them due to contractual incompleteness. The role of outside shareholders may be counterproductive in this case because there is a trade-off between monitoring and managerial initiative. Over-monitoring by large outside shareholders may be value-reducing because they do not internalize the private benefit that is the reward for managerial initiative. This is the "alternative hypothesis" to Shleifer and Vishny (1986), where large outside shareholders serve as a value-increasing monitoring device for managers.

Further, Zingales (1995) proposes a model in which an owner may take the firm public in order to divest his interest in the firm and diversify. In that case, dual-class shares can serve as a means to control the timing of his ultimate divestment rather than unduly entrench the manager.

In the case when voting control by insiders is beneficial, there may be counterbalancing governance mechanisms to prevent extraction of inefficient private benefits. These include the discipline imposed by the capital market and the board of directors. Studies find that dual-class firms are significantly more levered than single-class firms [Cronqvist and Nilsson (2005)]. The choice of debt financing post-IPO may be due to insiders not willing to dilute their control through seasoned equity offerings but it can also serve as counterbalancing governance

mechanism. In addition, there is evidence of contractual, institutional and personal mechanisms which are designed to mitigate the entrenchment effect, e.g. coat-tail provisions and strong board of directors [Moyer et al (1992), Taylor and Whittred (1998)].

4.3. Empirical evidence

Empirical evidence includes Amoako-Adu and Smith (2001) who find that the private benefit of a large shareholder related by family with the founders of the firm is significant for the choice of dual-class status at IPO. Family-owned firms tend to choose dual-class structure more often for reasons such as employing family members in positions of senior management and having the option of passing on control to the next generation. At the same time, Amoako-Adu and Smith (2001) find that in the general case of a controlling shareholder without family interest, dual-class shares are used to minimize the cost of foregone diversification and control the timing of divesting control, rather than unduly entrench the manager.

Empirical studies on stock price reactions surrounding dual-class status changes show mixed results. Jarrell and Poulsen (1988) estimate negative abnormal returns at the announcement of dual-class recapitalization in a sample of 97 firms between 1976 and 1987 consistent with entrenchment. On the other hand, Cornett and Vetsuypens (1989) find that dual-class recapitalization of 70 firms between 1962 and 1986 leads to abnormal price increases consistent with optimality. The two studies employ very similar methods but event studies are very sensitive to choice of event dates. The older study uses the proxy mailing date as the announcement date while the later one identifies announcement dates from the news. In that regard, Cornett and Vetsuypens (1989) may be correct in identifying a positive reaction at the

announcement of dual-class recapitalization. The evidence in Dimitrov and Jain (2006) also supports value-enhancing effect of dual-class recapitalization.

Overall, theoretical and empirical evidence on dual-class shares is mixed. It seems fair to say that if dual-class status is chosen by a widely-held firm, it is optimal choice determined by factors such as high private benefit of the rival. Dual-class shares may also be value-increasing in the case of closely-held firms where insiders have control, if they serve as protection from hostile takeovers and there are countervailing mechanisms, such as coat-tail provisions. Possible exceptions are family-controlled firms where family interest adversely affects valuation.[2]

4.4. Efficient vs inefficient private benefits

The central question regarding the efficiency of dual-class status is whether it is optimal for outside shareholders or whether the private benefits of insiders come at a cost of reducing security value for outside shareholders. In the case when dual-class status is optimal, the benefits and costs of insiders' control are balanced at the margin. This does not necessarily imply the absence of private benefits for insiders but that the outcome in terms of firm value cannot be improved by restricting these private benefits so that they are "efficient." They may exist to allow the raider to benefit from the takeover and thus allow value-increasing takeovers to take

[2] In the 1980s, dual-class companies and their governance were subject to some controversy which led to changes in the regulatory environment relevant to dual-class shares as described in Jarrell and Poulsen (1988). The NYSE was the only exchange maintaining the one share-one vote rule until 1984, when the NYSE imposed a moratorium on its enforcement of this policy. The reasons for the moratorium were possible loss of listed securities to AMEX or NASD, requests by NYSE-listed companies to be allowed to use dual-class recapitalization as defensive tactic against hostile takeovers, and allowing flexibility in NYSE-listed firms in choosing their own capital structure. The SEC which proposed a rule that allows newly public companies to choose dual-class structure at IPO or already dual-class companies to do seasoned offerings. Otherwise, companies are allowed to issue new classes of equity with lower voting rights through a public offering but are not allowed to change the voting rights of existing securities. Burkart and Lee (2008) observe that regulatory attitudes towards dual-class shares have swung from restrictive to relaxed, but the current trend especially in Europe is back to restrictions, exemplified by the ongoing debate about a EU-wide prohibition on deviations from one share-one vote.

place as in Grossman and Hart (1988) or to compensate managers in the case of contractual incompleteness as in Burkart et al (1997).

On the other hand, there are conditions under which dual-class status can be suboptimal. The existence of dual-class shares can protect an incumbent with high private benefits and low security value, which is an inefficient outcome. Similarly, if counterbalancing governance at a controlled firm fails and private benefits are higher than what was reflected in security value at the IPO, outside shareholders are expropriated.

Empirical studies of the effect of insider's cash-flow rights and voting control on firm value do not make a distinction between efficient and inefficient private benefits [Claessens et al (2002), Gompers et al (2008)]. The reason is that efficient private benefits would not have a value-decreasing effect in a cross section of firms. On the contrary, firms with inefficient private benefits of insiders will have lower valuation than otherwise identical firms. Therefore, for the empirical tests in Chapters 2 and 3, any negative effect of voting control is due to inefficient private benefits. In Chapter 4 that distinction does not matter because the analysis focuses on the decision of dual-class firm to recapitalize into single-class. If the additional restriction that SOX is not expected to reduce efficient private benefits is imposed, then again the results will be driven by inefficient private benefits of insiders.

5. Sarbanes-Oxley

Enacted in July 2002, SOX applies to exchange-listed firms and its provisions fall under three general categories: corporate governance, financial reporting, and auditor oversight. It was introduced in response to scandals like Enron, Global Crossing, Adelphia, Tyco and WorldCom,

which involved improper accounting practices and off-balance sheet transactions, loans and other improper payments to insiders[3]. Specific corporate governance provisions include:

- audit committee of the board of directors
- internal control audits
- CEO and CFO certification of financial statements
- higher penalties on officers who are charged with forging documents
- prohibition on insider loans
- prohibition on related party transactions involving large outside shareholders

In terms of financial reporting provisions, there are expanded SEC reporting requirements and more timely disclosure of equity transactions by corporate insiders, and in terms of auditor oversight, there is the formation of a quasi-public institution to oversee auditing (PCAOB) [Coates (2007), Mitchell (2003), Lins (2003)].

The SEC collaborated with the major stock exchanges to develop a stricter set of exchange listing requirements for publicly traded firms, and NYSE and NASDAQ changed their listing rules around the same time SOX was passed [Leuz (2007)]. The rules include:

- majority of board members must be independent
- independent audit committee of the board with at least one financial expert
- independent board committees that nominate directors and compensate managers
- a code of ethics

The effect of SOX cannot be identified separately from the effect of the related listing rules except in event studies looking at specific dates and announcements. Therefore studies of valuation effects refer to the Act and the listing rules collectively as SOX [Chhaochharia and Grinstein (2007)].

[3] Among these, Adelphia was a dual-class firm where a family controlled the company through its ownership of supervoting class of stock (65% of votes) but had a minority of the cash-flow rights (11%).

CHAPTER 2. SARBANES-OXLEY AND MANAGERIAL OWNERSHIP AS ALTERNATIVE GOVERNANCE MECHANISMS: AN EVALUATION USING DUAL-CLASS FIRMS

1. Introduction

Sarbanes-Oxley (SOX) introduced strict corporate governance provisions for public companies. Its effect on managerial incentive alignment and extraction of private benefits has not been established empirically in the literature. This chapter examines whether the new corporate governance rules help resolve such agency problems. I estimate the impact of SOX via the effect of equity ownership by mangers on firm value, particularly the positive incentive alignment effect of their cash-flow rights, and the negative entrenchment effect of their voting control.

In general, agency costs can be resolved through mechanisms such as ownership structure, capital structure, or regulation. Managerial equity ownership in particular helps to align the interests of manager and shareholder because the cash-flow rights granted by the shares provide an incentive to the manager to act in the shareholder's interest. When insiders own equity in their firm, the incentive alignment has a positive effect on firm valuation. However, the voting rights included with equity ownership create managerial entrenchment, characterized by the extraction of private benefits. When insiders have voting control in the firm, they are insulated from the market for corporate control, which has a negative effect on firm value.

Romano (2005) notes that the corporate governance mandates of SOX are "substantive" and "unprecedented." The Act and the related exchange listing rules include provisions whose purpose is to ensure alignment of incentives of corporate insiders with those of investors, and to reduce the likelihood of corporate misconduct and fraud. Thus, the relationship between managerial equity ownership and firm value is expected to change due to SOX and the purpose of this chapter is to establish this effect empirically. If SOX is successful in improving incentive alignment, the value-increasing incentive effect of managerial cash-flow rights is expected to be

smaller in magnitude after SOX. Similarly, if SOX improves corporate governance and limits extraction of private benefits by insiders, then the negative entrenchment effect of voting control should be less pronounced after SOX.

Dual-class firms are a useful empirical tool to study the effect of SOX on corporate governance as they allow for disproportionate cash-flow and voting rights for insiders and the separate identification of their effects on firm value. In firms with only a single class of shares, both effects are measured with the same variable—managerial equity ownership—because cash-flow and voting rights are proportionate. Therefore, dual-class firms provide a tool to test the positive incentive alignment effect and the negative entrenchment effect and to evaluate the effect of SOX on these corporate governance relationships.

2. The effect of SOX on the role of managerial ownership

2.1. Empirical studies of managerial ownership

As discussed in Chapter 1, in theory managerial ownership of equity can alleviate the agency problem by aligning incentives [Jensen and Meckling (1976), La Porta et al (2002)]. When managers own a share of cash-flow rights, they have an incentive to maximize effort which is in the interest of shareholders. At the same time, the voting rights associated with inside ownership create the potential of entrenchment due to insiders' ability to extract private benefits [Grossman and Hart (1988), Harris and Raviv (1988)]. There are two other important factors associated with managerial equity ownership: cash-flow ownership by managers may be value-reducing because of the cost it entails for managers, while voting rights may be beneficial as protection from value-decreasing takeovers.

If ownership structure is optimally chosen in response to the firm's environment, there will be no empirical relationship between observed levels of ownership and firm value in a cross-section of firms. In other words, two identical firms will have equal levels of inside equity ownership. If, however, managers own a suboptimal amount of equity in a firm, its value will be lower, controlling for all other factors. Thus in a cross-section of firms, there will be a positive relationship between valuation and managerial ownership.

Numerous empirical studies of the effect of managerial ownership on firm valuation find a significant nonlinear relationship between managerial ownership and value as measured by Tobin's Q [(Morck et al (1988), McConnell and Servaes (1990)]. Over a range of inside equity ownership, there is a positive and decreasing effect on firm value, which is consistent with incentive alignment outweighing the negative effect of entrenchment. At higher levels the relationship turns negative, which is consistent with managerial entrenchment outweighing the positive effect of incentive alignment. Specific ranges differ across studies, but approximately there is a strong positive effect when managers own 0-5% equity, much smaller effect when managers own 5-30%, and negative at about 40%. Empirical evidence that does not find a relationship between managerial share holdings and firm value indicating that these are optimally determined include Himmelberg et al (1999), Gomes (2000) and Dahya et al (2008). Studies like these which take a more structural approach controlling for the interrelatedness of governance choices or unobserved factors impose high demands on the model and estimation.

2.2. Effects of SOX in the literature

Chapter 1 listed the main corporate governance provisions of SOX, including majority of independent directors on the board, an audit committee of the board of directors, and an internal

controls audit. The costs and benefits of SOX are hard to quantify. Some costs that have been recognized in the literature are the “out-of-pocket” compliance costs as well as certain opportunity costs. Block (2004) discusses the higher cost of being public after SOX due to higher fees paid to outside directors, higher premiums for officers and directors’ insurance and more expensive audits. In addition, there is the time pressure on officers due to the need to personally certify financial statements, oversee audits, and attend committee meetings of the board to ensure SEC compliance.

In terms of corporate governance benefits, SOX aims to better align incentives between managers and shareholders and to limit the extraction of private benefits by insiders [Coates (2007)]. To the extent that agency problems are minimized due to SOX, there should be a positive effect on firm valuation. Keeping costs constant, we can expect a positive effect for those companies that did not operate at the value-maximizing point of corporate governance [Wintoki (2007)]. The ultimate effect of SOX is an open empirical question.

Several studies focus on stock market reactions to SOX and establish that the valuation of some firms increased because of the stricter governance provisions of SOX. Chhaochharia and Grinstein (2007) differentiate firms based on the degree of compliance with the provisions of SOX prior to their implementation and then compare stock price reactions. Being less compliant with the provisions would be an indication of higher agency costs, and indeed these firms experienced an increase in value relative to the more compliant firms. For example, firms that had many related party transactions before SOX had a positive reaction in terms of valuation after SOX. Other papers that find positive relative effect of SOX in terms of stock price reaction are Hochberg et al (2007), Jain and Rezaee (2006), and Li et al (2008).

While some firms can benefit from SOX, it is possible that other firms are made worse off because compliance costs outweigh the benefits. Certain costs are fixed and since there are no exemptions for small firms, compliance is relatively more costly for small firms [Block (2004)]. These firms may also find it harder to attract independent directors. Chhaochharia and Grinstein (2007) find that small firms experience a negative relative effect in terms of stock price.

The estimation of the effect of SOX on aggregate stock market valuation net of all costs and benefits is complicated by the need to isolate confounding contemporaneous events. Since SOX applies to all public companies, there is no readily available control group of firms that are not affected. Litvak (2007) uses foreign firms listed in the US as controls and finds a negative overall effect of SOX on US firms. Zhang (2007) compares the stock market reaction of all US firms to foreign firms and also finds that there was an absolute negative reaction but the results are criticized because foreign markets do not adequately control for contemporaneous events [Leuz (2007)].

2.3. The effect of SOX on managerial incentive alignment and entrenchment

Apart from looking at stock price reactions, it is important to see how several of the most important corporate governance relations are affected due to SOX, which is the subject of this study. Since cash-flow rights have been found to be a mechanism for aligning incentives and empirically there is a positive relation with firm value, it is important to see how this relation changes with the legal environment. La Porta et al (2002) find that the magnitude of the positive effect of managerial cash-flow ownership differs across countries depending on their corporate governance laws. In the US, there has not been a study of how SOX and related rules affected the role of managerial share ownership in optimal market valuation. We can hypothesize that any

positive effect of insiders' cash-flow rights and any negative effect of their voting rights on Tobin's Q are smaller in absolute value after SOX due to better incentive alignment decreased ability of insiders to extract private benefits.

To state these relationships in a slightly more formal way, we can think of governance net of associated cost as the input in a production function of firm value. Let *V(G; L)* be the value function which depends on the firm's choice of governance G and an environmental variable *L*, which stands for corporate law or other legislation concerning governance. Common assumptions regarding production functions are relevant here as well. These are $V_G>0$, $V_L>0$, $V_{GG}<0$ and $V_{GL}<0$. The first two derivatives state that firm valuation is increasing in governance and in legislation as they help resolve agency problems. The third one means that the marginal effect of governance is decreasing. The final one states that the marginal effect of governance is decreasing in legislation because legislation and governance are substitutes so that the marginal effect of governance decreases as legislation increases.

The optimal choice of governance *G** is the one that maximizes firm value, *max V(G, L)*. The first-order condition is $V_G(G, L)=0$. To see how *L* can affect *G** in equilibrium, we can employ standard comparative statics. Differentiating the first order condition $V_G(G, L)=0$ with respect to *L*:

$$V_{GL} + V_{GG}\frac{dG^*}{dL} = 0$$

$$\frac{dG^*}{dL} = -\frac{V_{GL}}{V_{GG}}$$

Thus in equilibrium, the optimal choice of governance is inversely related to legislation because $V_{GG}<0$ and $V_{GL}<0$. For example, the better the corporate governance laws, the less is the optimal share of cash-flow rights held by managers.

Although in equilibrium optimal cash-flow rights change in response to legislation, there will be a short-term effect in which the firm's stock market valuation changes but managers take longer to rebalance their portfolios. In other words, in the short term V_G will fall as L increases because $V_{GL}<0$ but G^* is fixed. The change in the marginal effect of governance can be used to evaluate whether SOX was successful in improving governance, which is the purpose of the empirical tests in this chapter. Specifically, the effectiveness of SOX can be established by looking at the change of the marginal effects of cash-flow rights and voting control immediately after SOX was enacted. Because legislation and governance are substitutes, we expect a decrease in the positive marginal effect of cash-flow rights and a decrease in the negative marginal effect of voting control in absolute value as a result of better regulation due to SOX.[4]

3. Dual-class firms

In Chapter 1, theoretical factors determining the choice of dual-class status and the associated managerial control include private benefits of insiders, their cost in terms of undiversified risk, protection from value-reducing takeovers and realization of growth options of the firm, and human capital investment by management. This section looks more closely at some characteristics of dual-class firms, which are used in the estimation. As already indicated, dual-class firms have multiple classes of equity with differential voting rights. Dividends may also differ between the classes, and usually the inferior class pays the higher dividend. This equity structure allows the separate identification of insiders' share of cash-flows and their voting control.

[4] It would be interesting to test whether SOX causes a decrease in managerial equity ownership in equilibrium but the dual-class data used for the analysis ends in 2002. Extending the dataset to include inside ownership of cash-flow and voting rights at dual-class firms post-2002 is a task left to future research.

The dual-class sample includes all publicly-traded dual-class firms in the US in the period 1994-2002. It was provided by the authors of a study of dual-class firms and a detailed explanation of the construction of the dataset is available in Gompers et al (2008)[5]. Panel A in Table 1 presents the number of firms by year and information regarding trading of the classes. Superior and inferior classes are defined so that the number of votes attached to the superior shares is greater than the number of votes attached to the inferior shares. If the votes are equal, the superior class elects a higher proportion of directors on the board. Typically in the US, only one or two of the multiple classes trade publicly. This is usually the inferior voting class as shown in the table. When two classes trade, one is usually superior and the other inferior, but it is possible for both traded classes to be inferior to a closely-held class.

Although multiple classes of shares are possible ranging from two to four, most often there are two classes. There is a wide variety of arrangements in terms of votes, proportion of directors elected, and dividends. Panel B in Table 1 summarizes some of these characteristics when there are two--one superior and one inferior--classes of shares. Most often, the number of votes attached to the superior class is 10 versus 1 for the inferior class. The superior class also elects a higher proportion of directors on the board. It is commonly assumed that the inferior class pays higher dividends. Although on average this is the case, it is possible for the superior class to pay the higher dividend. In terms of inside holdings, officers and directors hold on average 80% of the outstanding number of superior shares, while holding 23% of the inferior shares.

[5] To build a comprehensive set of dual-class firms, the authors first constructed a candidate sample using data from the Securities Data Company (SDC), Compustat, CRSP and the Investor Responsibility Research Center (IRRC). Then they examined the proxy statements and/or 10-Ks available online since 1994 to determine if every candidate was dual-class. Once the dual-class sample was completed, the next step was to determine the insider ownership for each class of stock as of the proxy record date for each firm-year. They parsed the tables and footnotes to compute the common-stock ownership, excluding all options and other rights. Since the share classes sometimes have differing cash-flow rights, they also collected dividend data for all firms.

Panel C of Table 1 contains some well-known companies which have dual-class shares. Industries with higher concentration of dual-class firms include publishing, food, apparel, cosmetics, media, and telecoms.

Empirically, the most important feature of dual-class shares is that they allow for disproportionate holding of cash-flow rights and voting rights by insiders. Thus the positive incentive alignment effect and the negative entrenchment effect on firm value can be identified separately in a cross-section of dual-class firms. Empirical studies that identify these effects separately when examining managerial equity ownership use international companies with pyramidal equity structures. Claessens et al (2002) and La Porta et al (2002) find a positive effect of cash-flow rights on firm value as measured by Tobin's Q. Claessens et al (2002) also find a negative effect of the wedge when voting rights exceed cash-flow rights, a result similar to Lins (2003). Gompers et al (2008) study US dual-class firms and find strong empirical evidence of the value-increasing effect of managerial cash-flow ownership and a more ambiguous role for the wedge between voting and cash-flow rights.

After identifying the separate effects of cash-flow ownership and voting rights by using dual-class shares, the impact of SOX on corporate governance can be evaluated through its effects on incentive alignment and extraction of private benefits according to my hypothesis. While estimations can provide useful results for the sample of dual-class firms, the ability to draw inference for all firms may be clouded by the possibility that dual-class firms are different from single-class firms. Sample selection bias can arise if dual-class companies are not representative of all companies for the relationship between ownership structure and firm value. Gompers et al (2008) correct for the possibility of such sample selection bias by using the methods of Heckman (1979). The selection equation aims to explain the choice of dual-class

status with measures of private benefits, takeover probability, etc., but the estimated coefficients of the ownership variables are not sensitive to such correction. Another study using the same dual-class data to examine the effects of the wedge between cash-flow rights and voting rights of insiders on investment and other decisions also uses the Heckman method to correct for sample selection bias [Masulis et al (2008)]. The inverse Mills ratio is insignificant indicating that dual-class firms are not fundamentally different from single-class. Therefore, there is little reason to believe that dual-class firms have different relationships between ownership and firm value from single-class firms so I do not make sample selection corrections in what follows.

4. Empirical model

The basic model defines the relationship between firm value and managerial ownership measures:

$$Q_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + \alpha + \varepsilon_{it} \quad (1)$$

Q_{it} is firm value, X_{1it} contains the managerial ownership variables, and X_{2it} contains other control variables. The managerial ownership measures of interest here are insiders' cash-flow ownership, and the wedge between insiders' cash-flow and voting rights. All public companies are required to disclose the share ownership of each member of the board of directors, as well as for all executive officers and directors as a group. This group of managers and directors is called the "insiders" in the literature.

Q_{it} and X_{2it} are measured over the fiscal year or at fiscal year end. Data on insiders' share holdings is not available as of fiscal year end and generally comes from proxy statements. Since ownership is a determinant, X_{1it} is measured as of the nearest proxy record date prior to fiscal

year end. For example, for a fiscal year end on December 31, 1999, we would expect the nearest prior proxy date to be in April 1999.

4.1. Data

In addition to a comprehensive sample of dual-class firms in the US in the period 1994-2002, the data from Grompers et al (2008) contains information by class on the following: insider share ownership, shares outstanding, votes and dividends. It is measured as of the proxy record date in each year (or as of the fiscal year end if the 10-K was used). There are a total of 738 firms and 3718 firm-years in 1994-2002. I have supplemented the dual-class data with financial data from Compustat, 1993-2003. Compustat data is measured over the fiscal year or at fiscal year end. The data was merged so that each fiscal year end matches the nearest prior proxy record date. Proxy data older than a year was not allowed. For the estimation, nominal variables were adjusted for inflation using annual CPI.

Financial firms and regulated utilities are excluded from the sample. For financial firms, valuation data is difficult to calculate and to compare with other sectors. For regulated utilities, valuation can be influenced by government regulations [Claessens et al (2002)]. Utilities and financial firms are classified according to the Fama-French 49-industry definitions from French's website, so that SIC 4900-4942 are in Utilities (31) and SIC 6000-6799 are in Banking (45), Insurance (46), Real Estate (47), and Trading (48). The number of excluded firms is 101 out of 738 dual-class firms, so that 637 firms are left in the sample.

Summary statistics on insider ownership and financial variables are given in Table 2. Panel A shows that insiders of dual-firms owned 37% of the cash-flow rights but 59% of the voting rights in 2002. This illustrates that on average, voting control is higher than cash-flow

ownership by insiders at dual-class firms producing a positive difference, or wedge, between them. The maximum wedge is 83% in the sample. It is possible, however, for voting rights to be disproportionately smaller than cash-flow rights, so that the wedge between them is negative, as indicated by the minimum equal to -22%.

Panel B of Table 2 gives summary statistics on key financial variables of dual-class firms in 2002. Average book value of assets is $4023 million. Sales growth is 19%, with ratio of capital expenditures to sales of 0.09. Leverage measured by the ratio of book value of long-term debt to book value of assets is 0.23. The ratio of research and development to sales and the ratio of advertising to sales are 0.19 and 0.02 respectively. The number of observations for R&D and advertising is smaller because data is missing for many firms which may indicate these expenses are insignificant and hence are not reported.

Summary statistics for both dual-class and single-class firms for the representative year of 2000 reported in Gompers et al (2008) indicate that these groups have about the same average level of assets and market capitalization. Significant differences between dual-class and single class firms arise only in terms of leverage and age. Dual-class firms are more levered than single-class firms, with mean debt-to-assets ratio of 0.23 vs. 0.17. Dual-class firms are also older than single-class firms, where age is defined as years from the firm's CRSP listing date. The average age of dual-class firms is 13 years vs. 10 for single-class firms.

4.2. Variables

The dependent variable *Q* is firm valuation measured by Tobin's Q, the ratio of market value of assets to their replacement cost, as of fiscal year end in year *t*.

$$Q = \frac{market\ value\ of\ assets}{replacement\ cost\ of\ assets}$$

Market value of assets is equal to the market value of equity plus the market value of debt as an accounting identity. Market value of equity is calculated as the number of common shares outstanding (*CSHO*) multiplied by share price (*PRCC_F*). There is no direct estimate of the market value of debt so it is calculated as the book value of assets (*AT*) minus the sum of the book value of common equity (*CEQ*) and balance sheet deferred taxes (*TXDB*). Replacement cost of assets is approximated with their book value [Kaplan and Zingales (1997)]. Thus

$$Q = \frac{AT + CSHO \times PRCC\_F - CEQ - TXDB}{AT}. \quad (2)$$

Table 3 contains all variable definitions and data sources.

Tobin's Q is meant to be a measure of the surplus value of the firm above replacement cost, which is higher the higher the growth options and the higher the intangibles. When calculating Tobin's Q, the market price of shares enters the numerator. When two classes trade, the price of superior shares reflects a control premium. However, the purpose is to understand how the value of the firm belonging to minority shareholders is affected by the ownership variables. Thus, the value belonging only to controlling shareholders should be excluded and all shares valued at the price of inferior shares, which is *PRCC_F* in (2).

For example, La Porta et al (2002) compute the market value of equity for firms with multiple classes of common stock by multiplying the total number of outstanding shares other than preferred stock by the price per share of the most widely traded class of common stock. This procedure typically prices equity using prices of lower-voting shares. The authors claim that this is correct conceptually since the model's predictions concern the value of equity to outside minority shareholders without the voting premium that reflects the power to divert.

Lease et al (1983) demonstrate that the effect of valuing all shares at the inferior price is likely to be small. When a company has two classes of common stock, which differ only in their

voting rights, superior voting shares generally trade only at a small premium (5%) above their inferior counterparts. It should also be noted that the price of any common stock includes a vote component, but in most publicly traded companies, this value cannot be directly measured [Zingales (1995)].

Estimates of Tobin's Q are notoriously subject to measurement error. The accounting data used to estimate replacement cost of assets is of particular concern since book values of some intangible assets are often quite different from their true replacement cost. Measurement error in the denominator of equation (2) causes errors in *Q* to be right-skewed, with some very extreme outliers. Figure 1 illustrates this point. Fig. 1a plots *Q* with a break in the axis to show both the bulk of the data and the outliers, where an extreme value of 212 is very unlikely. To reduce the weight of outliers, researchers censor extreme values at 1% or 5%, or use the natural logarithm of Tobin's Q [Dahya et al (2008), La Porta et al (2002)]. I use the latter strategy so that the dependent variable *Q* in most of the specifications is formulated as *lnQ*. Fig. 1b plots *lnQ* and we can see how the influence of outliers is substantially reduced. A similar effect is achieved by censoring at 1% both ends of the distribution as shown in Fig. 1c.

Cash-flow ownership by insiders (*INSIDERCF*) and the wedge between cash-flow rights and voting rights (*WEDGE*) are the main independent variables. *INSIDERCF* measures the incentive alignment effect of managerial equity ownership. Its coefficient represents the marginal effect of insiders' cash-flow rights and is expected to be positive in a cross-section of firms. Claims on the cash-flows of the firm for each class are calculated in accordance to the dividends paid by each class or the correct proportions of cash-flow rights if no dividends are paid. When no actual dividends are paid, cash-flows are based on the proportion of dividends to be paid in the future, as stipulated by incorporation charters. If the classes are non-dividend paying, cash-

flow rights are proportionate. *INSIDERCF* is thus formulated as the percentage of total claims on firm cash-flows held by insiders, and Table 3 shows how I calculate it.

Voting rights is the percentage of total votes held by insiders (*INSIDERV*). The difference between *INSIDERV* and *INSIDERCF* is the *WEDGE*. The *WEDGE* measures the degree of divergence of insiders' voting rights from their cash-flow rights, and hence the entrenchment effect. Since *WEDGE = INSIDERV – INSIDERCF*, when *INSIDERCF* and *WEDGE* are included together in the regression, they identify the cash-flow and voting rights of insiders separately. The coefficient of *WEDGE* measures the effect of deviations of voting rights from cash-flow rights, but since cash-flow rights are kept constant via *INSIDERCF*, the effect of *WEDGE* is driven by voting rights. The coefficient of *WEDGE* is expected to be negative because, keeping *INISDERCF* constant, the higher the voting power of insiders, the lower the firm value due to extraction of private benefits. Thus *INSIDERCF* and *WEDGE* serve to un-bundle the cash-flow and voting rights of insiders, while including *WEDGE* rather than *INSIDERV* in the regression is more appropriate in the case of dual-class firms. Both *INSIDERCF* and *WEDGE* are measured as of the proxy record date prior to each fiscal year end.

Other independent variables include firm controls for growth because the numerator of Tobin's Q reflects the capitalized value of growth options [(Morck et al (1988)]. Firm controls are identified from literature and include size, capital expenditures, and sales growth [La Porta et al (2002), Claessens et al (2002)]. Small firms are expected to grow faster because they have more growth options and the estimated coefficient is expected to be negative. Empirically, size is also found to be a significant indicator for growth prospects. A pick up in sales may also indicate growth opportunities, so sales growth is used as a control. Finally, the firm may undertake increased capital expenditures if future growth is predicted.

SIZE is defined as the log of book value of assets at end of fiscal year *t. SALESGROWTH* is sales growth during fiscal year *t.* Since Tobin's Q is a forward-looking valuation which includes growth options, the most recent indicators of growth should be used. In fact, La Porta et al (2002) make the point that past sales growth may be a poor measure of investment opportunities so future sales growth should be more appropriate. That is why at the very least both *SIZE* and *SALESGROWTH* should be measured in the most recent fiscal year. *CAPEX* equals the ratio of capital expenditures to sales in fiscal year *t-1.* Unlike size and sales growth as indicators of growth options, capital expenditures must pick up in advance to meet expected growth, so *CAPEX* is measured in the previous fiscal year consistent with existing literature.

4.3. Estimation using fixed effects model (FE)

Unobservable firm characteristics that could determine both managerial ownership and firm value present a problem in estimations involving Tobin's Q because omitting them from the regression leads to biased coefficients. One approach to solve the unobserved firm heterogeneity problem is to estimate FE [Himmelberg et al (1999)]. For example, there may be an unobservable monitoring technology available to the firm so that under optimal contracting the owners will choose a lower level of managerial ownership to align incentives. If measures of the quality of monitoring technology are omitted from the specification, a regression of firm value on managerial ownership will show a spurious negative relation because ownership is a negative proxy for the quality of monitoring technology. The example is from Himmelberg et al (1999) where, after controlling for unobservable firm characteristics, there is no positive effect of managerial equity holdings on firm value.

In general, there could be a “good firm quality” that remains constant over time and determines both high valuation and high managerial cash-flow ownership leading to the false conclusion that high *INSIDERCF* increases firm value in a cross-section of firms. By transforming the regression variables into deviations from their means in the estimation of the coefficients, FE eliminates the unobserved effect and hence the source of bias. Similarly, the FE estimation would control for a “bad firm quality” that can simultaneously cause low valuation and entrenched managers who fear for their jobs leading to spurious negative relation between *WEDGE* and firm value. These qualities are unobservable but as long as they remain constant through time, the FE estimation eliminates the bias without identifying empirically measurable proxies. As these are firm-specific factors, it is necessary to include firm-level fixed effects, rather than more aggregated industry dummies.

Other than unobservable firm quality, there is reason to believe that some omitted corporate governance factors are also practically time-invariant. Performance pay schemes discussed in Chapter 1 usually consist of a combination of specific targets for managers as well as a requirement for share ownership in the company. If meeting the targets is an effective motivation device, the existence of a performance pay scheme could simultaneously determine high *INSIDERCF* and high firm value. Performance pay, once instituted at the firm, is unlikely to change over time however, so that the effect would be subsumed in the fixed effect. Another important governance mechanism that may be treated as relatively time-invariable is the board of directors. A stronger and more independent board may have a positive effect on value but at the same time allow higher voting power for insiders because it serves as countervailing governance mechanism. The coefficient of *WEDGE* would be biased toward zero in this case, but the FE estimation helps eliminate the bias.

Therefore, I estimate

$$\ln Q_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + \alpha_i + \varepsilon_{it} \quad (3)$$

where α_i is unobserved firm-specific characteristic that is time-invariant and correlated with at least one of the variables in X_{it}. Estimation of β is based on the standard fixed effects model. The independent variables are

$$X_1 = [INSIDERCF \;\; WEDGE] \quad (4)$$

$$X_2 = [SIZE \;\; SALESGROWTH \;\; CAPEX] \quad (5)$$

where *INSIDERCF* is the percentage of firm cash flows owned by officers and directors and *WEDGE* is the difference between voting and cash-flow rights held by insiders measured at the nearest proxy record date to each fiscal year end. *SIZE* is the log of book value of assets at fiscal year end, *SALESGROWTH* is sales growth during the fiscal year, and *CAPEX* equals the ratio of capital expenditures to sales during the previous fiscal year.

It is necessary to include the effect of SOX to see whether firms that have a higher degree of agency problem experience a positive effect on firm value due to the provisions of SOX and related rules. In a cross-section, this will translate into a less positive coefficient on cash-flow rights and a less negative coefficient on the wedge. In the specification including the effect of SOX,

$$X_1 = [INSIDERCF \;\; WEDGE \;\; AFTERSOX{*}INSIDERCF \;\; AFTERSOX{*}WEDGE] \quad (6)$$

where *AFTERSOX* is an indicator equal to one if the timing of the observation is after SOX was signed into law on July 30, 2002, and zero otherwise. As already noted, the sample ends in 2002 but the expected effect is realized though immediate adjustment of stock market prices and can be identified in a cross-section by the end of 2002. Writing out the full equation,

$$\ln Q_{it} = b_0 AFTERSOX_t + b_1 INSIDERCF_{it} + b_2 WEDGE_{it} + b_3 INSIDERCF_{it} * AFTERSOX_t + b_4 WEDGE_{it} * AFTERSOX_t + b_5 SIZE_{it} + b_6 SALESGROWTH_{it} + b_7 CAPEX_{it} + \alpha_i + \varepsilon_{it} \quad (7)$$

The introduction of *AFTERSOX* into the model implies an effect of time on firm value and its relationship with inside ownership variables. *AFTERSOX* by itself is equivalent to a dummy variable for the year 2002. All fiscal year ends that are after the introduction of SOX in July 2002 until the end of 2002 are treated as after SOX, and since most of the firms have fiscal years ends that coincide with the end of the calendar year, *AFTERSOX* is practically a dummy for the year 2002. The coefficient of *AFTERSOX* therefore captures macroeconomic effects on firm value that are not necessarily attributable to SOX. However, I do not attempt to say anything about the aggregate effect of SOX on firm valuation as measured by Tobin's Q, only about the interaction of *AFTERSOX* with the ownership variables. I therefore do not consider the interpretation of the coefficient of *AFTERSOX* by itself.

AFTERSOX is interacted with the ownership measures *INSIDERCF* and *WEDGE*, and the interacted terms are the focus of this study. In that case, we can be relatively certain that any observed effects are only attributable to SOX as major corporate governance legislation. The reason is that the relationship between the ownership variables and Q is a cross-sectional relationship. The only factor with a macroeconomic component is growth, and it is already controlled for in the equation. Other factors that could affect this relationship include idiosyncratic risk, private benefits, takeover probability, dividend policy, pay-for-performance schemes, and board of directors as already discussed. It is difficult to imagine that any of these governance factors could change for all firms at the same time between July 2002 and end of 2002[6]. Thus regulation is definitely a dominant factor that can affect corporate governance and change the cross-sectional relationship.

[6] Possible exception is takeover probability which is determined in the market of corporate control. For example, there could be a spike in takeover activity so that higher insider voting power would be beneficial. However, it is difficult to control for this empirically and even takeover activity tends to be industry-concentrated. There is no evidence of a "wave" of takeovers in 2002 as there was in the 1980s.

Therefore I focus on how the relationships between firm value, insider cash-flow rights and voting control change due to SOX and evaluate the effect of the regulation via these coefficients. In the context of the panel dataset, this practically compares the cross-section of firms in all years before SOX to those after SOX. The original study of the relationship between governance and firm value in fact used only one cross-section of firms--Fortune 500 firms in 1980 [Morck at al (1988)]. Having more years in a panel dataset, though, contributes to the precision of coefficient estimates. There is no need for time dummies and previous studies of these relationships do not include them [Cleassens et al (2002), La Porta et al (2002), Gompers et al (2008)]. In Section 6, I perform an additional econometric check to confirm that *AFTERSOX* measures the effect of SOX and not other macroeconomics factors on the relationship between *Q, INSIDERCF* and *WEDGE*.

When examining repeat cross-sections, or panel data, there may be heteroskedasticity and/or autocorrelation in the error terms. Ignoring the covariance structure will not bias the coefficients but may lead to incorrect inferences by underestimating the standard errors of the coefficients. In the case of Tobin's Q, we can expect the errors to be correlated for each firm across time because we are dealing with a measure of assets in place, not returns. Therefore, for the same firm, even after accounting for fixed effects, there could be serial correlation that dies out over time [Wooldridge (2002)]. Wooldridge proposes a clustered covariance matrix estimator which allows for an arbitrary covariance structure within an observation unit in panel data. Recent papers suggest clustering as well [Petersen (2006), Thompson (2006)]. Therefore, I estimate clustered standard errors allowing for arbitrary covariance structure within firms. This does not affect the estimation of coefficients, only inference via the standard errors.

5. Empirical results

Table 4 presents the basic estimation results. I start with a model based on previous literature estimating the relationship between managerial ownership and firm value with OLS. I then estimate the FE model defined in equations (3), (4), and (5) which examines the effects of insiders' cash-flow rights and voting control on firm value without accounting for SOX. The model is similar to those estimated in previous literature but is more restrictive in controlling for unobserved factors by including firm fixed effects. To account for serial correlation within firm, I estimate clustered standard errors. I then explore the effect of SOX by adding a dummy variable as defined in equation (7).

In Model 1 in Table 4, I use an approach similar to Claessens et al (2002) but utilizing US dual-class data[7]. The relationship between firm value and the ownership measures is estimated with OLS, including controls for size, sales growth, and capital expenditures. There are also 11 industry dummies, defined on the basis of two-digit SIC codes. Financial firms and regulated utilities are excluded. I get results that are similar to Specification 1 in Table 4 in Claessens. The coefficients of cash-flows and wedge are positive and negative respectively, although the significance levels in my results are lower at 10% and 5%, respectively, whereas in Claessens they are 1%. The coefficient of sales growth and size are positive and significant at 1%, as in Claessens. I get a positive and significant at 1% coefficient on capital spending, unlike Claessens where this variable is insignificant.

Results from the estimation of Model 2 defined in (3), (4) and (5) are presented in the second column of Table 4. In terms of the ownership variables, the coefficient of *INSIDERCF* is positive and significant, while the coefficient of *WEDGE* is not different from zero. The

[7] Claessens et al (2002) study the effects of cash-flow rights and the wedge in East Asian economies. They use the market-to-book ratio of assets as a dependent variable which is only slightly different from Tobin's Q and the same independent variables as I do with the exception of firm age since establishment.

specification allowing for unobserved firm effect related with the explanatory variables is supported by the Hausman test of fixed versus random effects[8]. As to the controls for growth prospects, *SIZE*, *SALESGROWTH* and *CAPEX* all have the expected signs, with *SIZE* and *SALESGROWTH* being statistically different from zero.

Model 3 allows for clustered standard errors by firm, in which point estimates remain the same but standard errors increase in general. The main results carry over with the exception that *SALESGROWTH* is no longer significant, while *CAPEX* becomes important as an indicator of growth opportunities. Economically this makes more sense since sales growth is not as forward-looking as capital expenditures.

The positive and significant coefficient of *INSIDERCF* indicates that cash-flow ownership by insiders is value-increasing. When insiders own a larger share of the firm's cash-flows, they have an incentive to maximize effort. The incentive alignment between managers and outside shareholders has a positive effect on firm value as measured by Tobin's Q. This result is consistent with a multitude of studies finding a positive effect of managerial share ownership on firm value, starting with Morck et al (1988) and McConnell and Servaes (1990), who measure the incentive effect with a single variable, to Claessens et al (2002) and Lins (2003), where voting rights are controlled for by using pyramidal ownership structures. The positive coefficient of *INSIDERCF* in a cross-section of firms is evidence that insider ownership of cash-flows helps resolve agency problems because increasing inside ownership increases firm value. Valuation of firms where insiders hold a small share of cash-flows is lower than valuation of otherwise similar firms where insiders hold a large share of cash flows. The value-increasing effect is obtained even after accounting for unobserved firm characteristics with fixed effects. This result is

[8] Chi-sq(5)=124.21 (p-value=0.00)

contrary to the findings in Himmelberg et al (1999)[9] who, after adjusting for unobserved firm characteristics, find no significant effect of managerial equity ownership on firm value. It is possible that the ability to separate the cash-flow and voting rights here leads to better identification of the incentive alignment effect.

The coefficient on the *WEDGE* is insignificant, contrary to the entrenchment hypothesis and the empirical findings in Claessens et al (2002). With the inclusion of firm fixed effects instead of industry dummies, the significance of *WEDGE* goes away. The coefficient of WEDGE measures the effect of deviations of voting rights from cash-flow rights of insiders as discussed in Section 4.2. Since cash-flow rights are kept constant in the regression framework, the result implies an insignificant effect of insider's voting control per se. It seems that there is an unobserved firm quality that simultaneously lowers value and causes entrenched management, so that there is no effect of control on firm valuation after including fixed effects in the estimation. The coefficient of the *WEDGE* is consistent with some results in Gompers et al (2008) who, after accounting for endogeneity, fail to uncover a negative relationship between voting control and firm value in all specifications.

At the same time, the insignificance of the *WEDGE* variable may be due to a misspecification in the way this variable enters the regression. While it is logical to expect positive *WEDGE* when insiders own more voting rights than cash-flow rights, negative and zero *WEDGE* firms may be different (7% and 8% of the observations in the sample, respectively). Insiders in negative and zero *WEDG*E firms own about the same fraction of cash flows as do their peers in the positive *WEDGE* sample (36% vs. 40%) but a substantially smaller fraction of

[9] Himmelberg et al (1999) use m, managerial share holdings as percentage of total shares outstanding, and m-squared to study the effect of ownership on firm valuation. The approach is rather different from mine as they also include several variables designed to control for managerial discretion. Managerial share holdings are significant in a model without firm fixed effects, but not when fixed effects are included.

the voting rights (33% vs. 65%). When I estimate model 2 for only positive *WEDGE* firms, or when I specify squared terms in the ownership variables, I still find that *WEDGE* is insignificant.

The lack of a negative relationship between the *WEDGE* and firm value is consistent with the more ambiguous role for voting control in the literature. It is possible that the value-decreasing effect of private benefits is counterbalanced by the positive effect of protection from value-reducing takeovers. The benefits that come from consolidated control such as investment in firm-specific human capital and realization of growth options in the face of asymmetric information in the market of corporate control compensate the negative effect of insider's ability to extract private benefits. The negative result about the coefficient of the WEDGE points to the fact that deviations from one share-one vote do not affect firm value because the entrenchment effect is counterbalanced by the benefits of insiders having voting power.

SOX is expected to change the relationships between measures of managerial ownership and firm value. The corporate governance provisions are meant to improve incentive alignment and limit extraction of private benefits. Model 4 in Table 4 is estimated based on the specification in (7) which includes the variable *AFTERSOX* . These interaction terms between the ownership variables and *AFTERSOX* estimate the change in the coefficients of *INSIDERCF* and *WEDGE* due to SOX. There are 292 observations from the passage of SOX in July until the end of the sample in 2002 that allow the estimation of the interaction terms. Even if managerial equity holdings adjust only slowly, the effect takes place immediately when SOX is enacted though the adjustment of stock market prices and hence firm value as measured by Tobin's Q.

The interaction term between *INSIDERCF* and *AFTERSOX* is negative and significant at 10%. There is no effect of *AFTERSOX* by itself, and Model 5 in Table 4 omits *AFTERSOX* as an

independent variable. While there is no significant change in the coefficient estimates, omitting *AFTERSOX* raises the significance level of the interaction term with *INSIDERCF* to 1%.

The negative and significant coefficient of the interaction term translates into a decrease in the positive coefficient of *INSIDERCF* after SOX. Thus the total effect of insider cash-flow rights becomes significantly smaller after the legislation was enacted. This is evidence that the incentive alignment effect of managerial cash-flow ownership is weaker as a result of SOX. The effect is consistent with findings that the magnitude to the positive effect of cash-flow rights depends on the legal environment and investor protection [La Porta et al (2002)]. Stricter governance rules due to SOX improve incentive alignment as a substitute for managerial ownership so that valuation of firms where insiders hold a small share of cash-flows rises relative to valuation of similar firms where insiders hold a large share of cash-flows. As a result we observe a significantly smaller marginal effect of *INSIDERCF* after SOX.

6. Robustness of the results

Table 5 explores some alternative models to see if the results hold up to different specifications. For convenience I include the baseline specification (Model 5 in Table 4) as Model 1 in Table 5. Model 2 in Table 5 does not adjust for outliers by taking the natural log of Tobin's Q so that the dependent variable is just *Q*. The important result on the coefficient of *INSIDERCF* does not hold up to this change. While measurement error in the dependent variable does not cause bias, it does inflate the residuals and standard errors, making inference more difficult. The standard errors in Model 2 indicate that this is indeed the case. Since the most extreme outliers tend to be driven by firm-specific measurement error, the log transformation is a natural way to reduce the influence of outliers. The alternative is to censor extreme values at 1%

and 99%. Fig. 1c shows that the effect of censoring on the distribution of Tobin's Q is similar to that of taking the log but leads to the loss of some observations. There is no substantial change in the main conclusions of the estimation as shown in Model 3 in Table 5 so the natural log of Q is the preferred specification.

Since firm value may be sensitive to capital structure, we can include debt [Lins (2003)]. As already discussed in Chapter 1, debt plays a major role in agency. In a stylized setting where the entrepreneur needs to finance a project, there is a choice between equity and debt. The dual-class structure is by definition equity, which may sell at a discount because of agency costs that the owners are willing to accept in exchange for preserving control. The discount may be mitigated by the insider's decision to keep a higher proportion of the cash-flows. For debt to have a place in the regression of Tobin's Q on ownership measures, there must be a clear mechanism through which subsequent projects and the choice of debt financing influence the effect of ownership structure on value. Hence many regressions involving Tobin's Q and managerial equity holding do not include debt [La Porta et al (2002), Claessens et al (2002)].

In the case of dual-class firms, however, the choice of debt financing post-IPO may be due to insiders not willing to dilute their control through seasoned equity offerings or to its role as a counterbalancing governance mechanism. Studies find that dual-class firms have low leverage at IPO [Taylor and Whittred (1998)], but debt is important in the post-IPO capital structure of dual-class firms [Cronqvist and Nilsson (2005)]. To check if the coefficients on cash-flow rights and wedge are sensitive to capital structure, debt is included in Model 4 in Table 5. Results indicate that the inclusion of debt does not affect estimated coefficients.

Other controls in regressions involving Tobin's Q and ownership measures include advertising and R&D, although they can have an ambiguous role as controls for growth options.

Also, data is missing for many firms in Compustat which would reduce the size of the sample significantly. To avoid this difficulty, Gompers et al (2008) assume zero values for advertising and R&D whenever data is missing. Model 5 in Table 5 uses the same approach but the coefficients are not significant. These variables are best left out of the regression as the number of non-zero values of advertising, for example, is only about 30% of the total sample. In addition, imputing zeros introduces measurement error which biases estimates towards zero.

Some researchers specify a nonlinear relationship between managerial ownership and firm valuation [Morck et al (1988)]. Tobin's Q is found to first increase and then decrease with managerial equity holdings. This nonlinear relationship exists because the two opposing effects of incentive alignment and entrenchment are measured with a single variable, managerial ownership. The relationship is positive because of incentive alignment but concave because the negative effect of voting control is stronger at higher levels of ownership. However, when it is possible to measure *INSIDERCF* and *WEDGE* separately as in the case of dual-class firms, there is less reason to expect a concave relationship. In any case, the squared terms of *INSIDERCF* and *WEDGE* are included in the specification of Model 6 in Table 5, but the coefficients are not different from zero. The joint Wald test of the coefficient of *WEDGE* and its quadratic term gives $F(2,561)=1.03$ (p-value=0.36).

Finally, I perform some econometric checks that *AFTERSOX* measures the effect of the SOX legislation and not other macroeconomic factors. As already explained in Section 4.3, apart from economy-wide regulation, the dominant factors that affect the relationship between *Q*, *INSIDERCF* and *WEDGE* are related to firm governance and estimated in a cross-section. Growth options determine firm value, and there is definitely a macroeconomic component in firm growth, but that is already accounted for in the model. Nevertheless, Table 6 presents the

econometric check that macro factors do not affect the slopes of the ownership variables. The model is estimated prior to SOX, in the years 1994-2001. It includes year dummies for each year except 1994, and terms that multiply the year dummies with *INSIDERCF* and *WEDGE*[10]. The F-test of the joint significance of all year dummies and interacted terms gives F(21, 2033)=4.61 (p-value=0.00), which indicates that these terms are important for firm value as measured by Q. The 1996-98 year dummies are significant, indicating that business cycle effects matter. This is plausible because some macro factors, like the "dot-com" bubble, could have an effect apart from the one already reflected in *SALESGROWTH* and *CAPEX*. What matters for the validity of the effect of *AFTERSOX* on the relationship between firm value and ownership structure are the interacted terms of the year dummies with *INSIDERCF* and *WEDGE*. The F-test of joint significance of these interacted terms gives F(14, 2033)=1.25 (p-value=0.24), which does not reject the null that they are jointly zero. We can infer that SOX and no other macro factors is the dominant reason for the change in the coefficient of *INSIDERCF* in the cross section after SOX was introduced in July 2002 until the end of 2002.

7. Endogeneity and instrumental variables

7.1. The problem of endogeneity

If the ownership variables are endogenous, coefficients produced by the fixed effects estimation may be biased. Although "endogenous" traditionally means "determined from within the model," endogeneity is commonly defined as the problem of correlation between any one of the independent variables and the error term leading to biased coefficients [Wooldrige (2002)]. The problem of endogeneity in reduced-form regressions of firm value on ownership structure is recognized as very common and hard to fix in corporate governance research [Demsetz and

[10] I also use GDP growth in place of the year dummies but it is not significant by itself or interacted.

Villalonga (2001), Coles et al (2007)]. In the present context, there are two sources of possible endogeneity: omitted variables and reverse causality.

Studies find significant substitutability among different governance choices so that they are negatively related, including incentive contracts based on performance, the discipline imposed on managers by the capital and labor markets, and the board of directors [Agrawal and Knoeber (1996)]. For example, a firm with a large shareholder who may collude with managers to divert corporate resources would also have a board with more independent directors [Dahya et al (2008)]. With dual-class firms in particular studies find strong countervailing governance such as outside directors and coat-tail provisions [Lehn et al (1990), Moyer et al (1992)].

Thus, in a regression involving firm value in a cross-section of firms, we need to control for all governance mechanisms, otherwise the excluded ones will be absorbed by the error term and will be correlated with the explanatory variable, leading to biased coefficients. This is an extremely difficult task empirically because it is hard to measure many corporate governance factors. However, the governance mechanisms most likely to be related with *INSIDERCF* and *WEDGE* such as performance pay and the structure of the board of directors are unlikely to change significantly in the time span over which I estimate my model, so that they are subsumed in the fixed effect.

The second source of endogeneity is reverse causality [Kole (1996)]. Suppose managers have inside information about high growth prospects not currently reflected in the firm's market value and increase their equity holdings. Subsequently, market valuation rises to arrive at the correct level. The reverse is also possible. If the firm becomes overvalued, managers divest their holdings, while valuation subsequently falls in a market correction. A spurious positive relation between valuation and insider ownership is found due to reverse causality. Managers rebalance

their portfolios in response to temporary over- and under-valuation of their firm's stock so that value is determining managerial ownership rather than ownership determining value.

It is not very likely, however, that firms can change their ownership structures quickly and frequently in light of temporary over- and under-valuations. La Porta et al (2002) find that ownership structures of firms in some East Asian economics tend to be very stable, which may be the case for US firms as well. More generally, Claessens et al (2002) note that regression results are based on cross-sectional relationships. The possibility of reverse causality would thus lead to a bias only if insiders change their cash-flow rights frequently in light of temporary changes in valuation, while maintaining their control rights and did so systematically across many corporations. Such behavior seems unlikely, they say.

Yet another reason why such behavior seems unlikely is insider trading laws. Insiders are not allowed to trade on the basis of material non-public information. There are strict reporting requirements with the SEC concerning insider trading. All managers and directors of public firms are required to file Form 4 with the SEC every time they trade shares, regardless of how many shares they own or how many are traded. Financial analysts follow closely insider trading while class action suits are often filed alleging insider trading on the basis of material non-public information. Managers of course can take a long-term view of the prospects of the firm using public information and trade their shares on the basis of that (or get away with trading based on inside information). Nevertheless, I think that to negate the positive coefficient on *INSIDERCF* on the basis of reverse causality implies that managers are consistently willing and able to evade insider trading prohibitions across many companies, which is unrealistic.

Despite the likelihood that endogeneity is not be a problem, I will address the issue for completeness. A remedy for the possibility that ownership variables are not orthogonal to the

error term in a regression of firm value is to use instrumental variables. Instruments should identify exogenous variation in managerial cash-flow ownership and voting rights and at the same time be independent of firm valuation, the dependent variable. I use two strategies to attempt to deal with endogeneity using instrumental variables: the traditional 2SLS approach and the control function approach.

7.2. Instruments

Instruments should identify exogenous variation in *INSIDERCF* and *WEDGE* and at the same time be uncorrelated with the dependent variable *lnQ*. Plausible candidates for instruments in the present context are difficult to find. For example, large firm size is found to be a deterrent for takeovers so that managerial ownership at large firms may be low [Morck et al (1988)]. However, size is also a proxy for growth opportunities and hence is positively related with Tobin's Q so that it is not a good instrument for managerial ownership.

Several instruments have been used in regressions involving firm value and managerial ownership in the literature, including stock price volatility, CEO compensation, and lagged ownership. The most plausible case can be made for using stock price volatility as an instrument [Demsetz and Lenn (1985), Himmelberg et al (1999), Amit and Villalonga (2006)]. Because higher managerial cash-flow rights imply less portfolio diversification for risk-averse managers, the optimal contract involves a trade-off between diversification and incentives for performance. Thus the higher the volatility of the firm's stock price, the lower the optimal *INSIDERCF*. For risk to be considered a good instrument, however, it must not be correlated with firm value.

According to the CAPM, asset risk can be decomposed into a systematic and an idiosyncratic component. The expected return on a stock is given by

$$E(r_i) = r_f + \beta_i (E(r_m) - r_f)$$

where r_i is the return on stock *i*, r_f is the risk-free rate, and r_m is the market return. The CAPM in regression form is

$$r_i^e = \alpha_i + \beta_i r_m^e + \varepsilon_i$$

Thus the excess return on a stock has two components: the systematic or market risk ($\beta_i r_m^e$) and unsystematic or idiosyncratic risk (ε_i). The variance can be decomposed accordingly:

$$\text{var}(r_i^e) = \beta_i^2 \sigma_m^2 + \sigma_i^2$$

where the first term is the market component and the second is the idiosyncratic component. Firm value depends inversely only on the systematic component because the standard implication of the CAPM is that only market risk is priced. Managerial cash-flow ownership, however, depends on idiosyncratic risk as well.

As an empirical proxy for idiosyncratic risk, I use the standard deviation of the idiosyncratic component of daily stock prices (*SIGMA*). The idiosyncratic component is constructed as the residual in the market model in (9) using daily returns for each firm. Since ownership variables are measured as of proxy date, the period for the regression covers the current year of the proxy as in Himmelberg et al (1999) as well as the prior year to allow for slower response of managerial ownership to changes in volatility. Idiosyncratic risk is the estimated root mean standard error, also called standard error of the regression. Pricing data from the period 1993-2002 is obtained from CRSP.

Although *SIGMA* is a viable instrument for *INSIDERCF*, there is one more possibly endogenous variable, *WEDGE*. Since *WEDGE* measures the effect of the divergence between cash-flow and voting rights, the variable is driven by private benefits. In the dual-class data, a proxy for private benefit that can also be used as an instrument in a regression of firm value is

the indicator variable *NAME*, which is equal to one of a person's name appears in the company's name at IPO. It is found to significantly predict dual-class status as a proxy for private benefits but is unlikely to be related with valuation [Gompers et al (2008)].

The last candidate for instrument for *WEDGE* is CEO compensation (*CEOCOMP*). Zinagles (1995) finds that entrenched managers have higher salaries than those with lower voting power in comparable firms. There may be problem with the exogeneity of *CEOCOMP*, however. On one hand, the base salary of CEOs should be determined by the labor market and be independent of firm characteristics. On the other hand, even base CEO compensation may be determined by firm value so that higher-valued firms pay their executives higher salaries, making *CEOCOMP* endogenous. Because of the fact that CEO compensation has been found to be higher in firms where managers have higher voting power, *CEOCOMP* is included as an instrument but it may not be a good one. Data on *CEOCOMP* is obtained from Execucomp by fiscal year for the period 1994-2002. It represents the salary and bonus for the top executive of each dual-class firm. Unfortunately, data is only available for about a third of the 640 firms used for the FE estimation which reduces the size of the sample.

7.3. Instrumental variables model (IV)

In order to evaluate the performance of the instruments in terms of identifying exogenous variation in the endogenous variables, I start with a general IV model based on (1). I then add *AFTERSOX* since it was found to matter in the equation by affecting the coefficient of *INSIDERCF*. The general model is simpler to evaluate in terms of the explanatory power of the instruments because the interacted terms with *AFTERSOX* would also be endogenous.

The general model from (1) can be written as

$$\ln Q_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + \alpha + \varepsilon_{it} \quad (8)$$

where Q_i is a T_i by 1 vector representing the number of observation in a cross section of the panel dataset. Consistent with the discussion in Baum et al (2003, 2007), let X_1 be the set of endogenous variables and X_2 be the set of exogenous variables. Let $X = [X_1 \;\; X_2]$, so that there are a total of 5 variables in X.

$$X_1 = [INSIDERCF \;\; WEDGE] \quad (9)$$

$$X_2 = [SIZE \;\; SALESGROWTH \;\; CAPEX] \quad (10)$$

Let the full set of exogenous variables be Z. Z can be split into a set of instruments (Z_1) and a set of exogenous variables (Z_2) so that $Z = [Z_1 \;\; Z_2]$ and $X_2 = Z_2$.

$$Z_1 = [SIGMA \;\; NAME \;\; CEOCOMP] \quad (11)$$

$$Z_2 = [SIZE \;\; SALESGROWTH \;\; CAPEX] \quad (12)$$

AFTERSOX has been found to matter in the equation by affecting the coefficient of *INSIDERCF*. Therefore it belongs in the model to account for this structural break. Note that since *INSIDERCF* and *WEDGE* are endogenous, their interactions with *AFTERSOX* are also endogenous, and the instruments need to be interacted with *AFTERSOX* as well. In addition, the model should properly be estimated with firm fixed effects, as already argued in Section 4.3. Unfortunately, including firm FE does not allow dummy variables to be included as instruments. The FE model including *AFTERSOX* is:

$$\ln Q_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + \alpha_i + \varepsilon_{it} \quad (13)$$

$$X_1 = [INSIDERCF \;\; WEDGE \;\; INSIDERCF*AFTERSOX \;\; WEDGE*AFTERSOX] \quad (14)$$

$$X_2 = [SIZE \;\; SALESGROWTH \;\; CAPEX \;\; AFTERSOX] \quad (15)$$

$$Z_1 = [SIGMA \;\; CEOCOMP \;\; SIGMA*AFTERSOX \;\; CEOCOMP*AFTERSOX] \quad (16)$$

$$Z_2 = [SIZE \;\; SALESGROWTH \;\; CAPEX \;\; AFTERSOX] \quad (17)$$

There are two conditions for identification. First the order condition states that the number of instruments in Z_1 has to be greater than or equal to the number of endogenous variables in X. In this case, this is satisfied with 3 instruments in Z_1 and 2 variables in X_1. This is a necessary but not sufficient condition. The second is the rank condition which states that the matrix $Q_{XZ} = E(X_i'Z_i)$ has to be full rank. The rank condition can be interpreted in terms of correlations: the endogenous variables X_1 must be correlated with the instruments Z_1. These two conditions allow for identification and estimation of the system of equations. However, if the correlations are weak, there may be a weak identification problem which causes bias in the coefficients that is increasing in the number of instruments as discussed in Bound et al (1995).

2SLS allowing for clustered standard errors is efficiently estimated with Generalized Method of Moments (GMM). Feasible efficient GMM uses an optimal weighting matrix to minimize the J-function representing the moment conditions, and an important feature is that the estimate of the variance-covariance matrix of the error term allows for clustering by firm. The moment conditions state that the instruments should be orthogonal to the error term, that is $E(Z_i\varepsilon_i) = 0$. Apart from identification or weak identification problems, violation of the orthogonality conditions presents another complication in the IV estimation. That is, if *SIGMA, NAME* and *CEOCOMP* are related to the dependent variable, firm value, and not orthogonal to the error term in the first-stage regression, there may still be a source of bias in the coefficients.

The control function approach (two-stage residual inclusion) is the second strategy I use to address endogeneity. It is an IV method for correcting for bias in nonlinear models [Basu et al (2008)]. In the first stage of the control function approach, the endogenous variables are regressed on the instruments and the residuals are generated. The first-stage residuals are then included as additional repressors in the second-stage regression, instead of using the predicted

values of the endogenous variables. Other than having better statistical properties in the general case of nonlinear IV, the residual inclusion strategy is particularly useful when the first stage is nonlinear and the predicted values of the endogenous variables cannot be directly obtained.

There is reason to believe that the first stage should be estimated by a nonlinear model because *INSIDERCF* is formulated as a proportion and bounded between zero and one, while *WEDGE* is bounded between negative one and one. Figure 2 presents statistics and histograms of the ownership variables. The formulation of the first stage must take into account the proportional nature of *INSIDERCF* so the model is correctly specified. The logistic transformation can be used to model proportions so that the transformed variables *INSIDERCF'* and *WEDGE'* are:

$$INISDERCF' = ln(INSIDERCF) - ln(1 - INSIDERCF)$$

$$WEDGE' = ln(1 + WEDGE) - ln(1 - WEDGE).$$

The intuition behind the control function approach is that the residual from the first stage serves as a proxy variable that conditions out the part of the endogenous variable that is related with the error term in the first stage. The remaining variation in endogenous variable is independent of the error and consistent estimates are obtained [Petrin and Train (2010)]. The approach uses instrumental variables to condition out the variation due to unobserved factors. The model is estimated in two steps. First, the endogenous variable is regressed on observed instruments. The residuals of this regression are retained and used to create the control function, which in its simplest form represents the residuals themselves. Second, the model is estimated with the control function entering as an extra variable. This variable in effect contains the endogenous variation in the explanatory variable and conditions it out in the estimation. The remaining variation is exogenous and does not lead to biased coefficients. The asymptotic

variance of the second-step estimator needs to take this extra source of variation into account so I use the bootstrap method to estimate standard errors.

7.4. Results from IV estimation

All instruments identified in Section 7.2 are used together with the exogenous variables in the IV to estimate the model defined in (8)-(12). Model 1 in Table 7 gives the first-stage results of the 2SLS estimation allowing for clustered standard errors. *SIGMA*, *NAME* and *CEOCOMP* are the instruments in the first stage, and *INSIDERCF* and *WEDGE* are the endogenous variables in the second stage.

We need to examine the performance of the IV estimation. As already discussed, the system needs to be identified to allow estimation (rank condition). The only significant instrument in the first-stage results for Model 1 given in Table 7 is *CEOCOMP*, while *SIGMA* and *NAME* appear to be irrelevant, which points to the lack of explanatory power of the instruments. The F-test of joint significance of the instruments[11] in the equation for *INSIDERCF* is F(3,166)=0.61 (p-value=0.61), which does not reject the null that they are not jointly different from zero. The situation is similar in the equation for *WEDGE*, although *CEOCOMP* is significant there: F(3,166)=1.82 (p-value=0.15). Under-identification is formally confirmed by the Keibergen-Paap (2006) rk statistic which does not reject the null hypothesis of less than full rank of the matrix of reduced-form coefficients (chi-sq(2)=0.34, p-value=0.84).

To model *INSIDERCF* and *WEDGE* as proportions in the first stage, the transformed variables are used in the control function approach with residual inclusion. The results are not

[11] The F-statistic in the case of clustering is defined as F=W/q, where W is the Wald test statistic and q is the number of restrictions. The F-stat is compared to the critical F(q,d) where d is the number of clusters minus one.

reported as there is almost no difference from the results in Model 1 in Table 7. Thus there is no improvement in the fit and under-identification is still a problem.

The fact that the *WEDGE* has a pile-up of zero values as shown by Fig. 2c may be a cause for misspecification of the first-stage regression. However, the fit does not improve when zero values of the *WEDGE* are omitted. The regression is also estimated without *CEOCOMP* because missing data for that variable significantly reduces the sample, as well as with the log of CEOCOMP because this variable too is highly skewed, but there is no improvement.

Model 2 in Table 7 presents the full model defined in (13)-(17), including firm fixed effects and the effect of SOX. The instruments are interacted with AFTERSOX, but NAME could not be used in the FE estimation because it is a categorical variable. Although there is a slight improvement in overall fit of the equations, the explanatory power of the instruments is still a problem. Although papers that use instrumental variables for managerial ownership rarely report first-stage results, a problem of weak instruments for insider cash-flow and voting rights is also encountered in Gompers et al (2008). Therefore, the results from the IV estimation are inconclusive.

While I could not estimate a model accounting for endogeneity, other authors have ignored the issue as well. Causation that runs from firm value to managerial ownership is often assumed away in reduced-form analysis. For example, the model in La Porta et al (2002) regarding the effect of cash-flow rights on firm value assumes that managerial share ownership is exogenous because ownership patters are very stable with respect to temporary fluctuations in firm value. Claessens et al (2002) also do not attempt to address the problem after noting that the likelihood of endogeneity due to reverse causation is small.

8. Conclusion

The unprecedented corporate governance requirements of SOX aimed at reducing conflicts of interest between managers and outside shareholders and limiting corporate fraud. This paper examines the effect of SOX on managerial incentive alignment and extraction of private benefits though the role of managerial share ownership. By using dual-class firms, the separate forces of insider cash-flow rights and voting control are identified in the determination of firm value. The possibility of biased coefficients due to omitted factors is most effectively addressed with firm fixed effects as the models using instrumental variables encounter a problem of weak instruments. The fixed effects estimation indicates a positive effect of managerial cash-flow rights that is robust to various specifications. On the other hand, there is no evidence of an effect on firm value of disproportionate cash-flow and voting rights of insider due the deviation from one share-one vote. This negative result suggests that dual-class status is on average an efficient choice given the firm's environment. SOX significantly decreases the positive effect of managerial cash-flow rights after its introduction in July 2002, acting as substitute governance mechanism. One limitation of the model is that the estimated effect of SOX may not represent long-term equilibrium because data on inside ownership at dual-class firms is not currently available post-2002. However, even if managerial equity holdings adjust only slowly, the expected effect is realized though immediate adjustment of stock market prices and can be identified in a cross-section of dual-class firms by the end of 2002. The evidence suggests that SOX was effective in improving managerial incentive alignment as the value-increasing effect of insider cash-flow rights is smaller in magnitude after SOX.

Tables and Figures

Table 1
Dual-class firms

Panel A: Number of firms in the merged dual-class/Compustat data by year of proxy record date

	1994	1995	1996	1997	1998	1999	2000	2001	2002
Number of dual-class firms	106	375	405	457	464	450	452	397	352
Only inferior class trades	61	264	286	342	335	320	330	284	249
Only superior class trades	12	33	40	41	53	53	50	45	43
Superior and inferior classes trade	30	73	74	69	71	71	63	59	52
Two inferior classes trade	2	3	4	4	3	3	4	4	5

Total number of dual-class firms: 738

Panel B: Characteristics of the superior and inferior class when there are two classes

	5th %ile	50th %ile	95th %ile	99th %ile	Mean
Number of votes of superior class	1	10	15	50	43
Number of votes of inferior class	0	1	1	1	0.8
Proportion of directors elected by superior class	0	0.62	0.91	1	0.48
Proportion of directors elected by inferior class	0	0.25	0.61	0.81	0.23
Dividends paid by superior class	0	1	1	1.36	0.73
Dividends paid by inferior class	0.06	1	1.28	3	0.78
Inside holdings of superior class	10%	94%	100%	100%	80%
Inside holdings of inferior class	0.3%	16%	72%	100%	23%

Panel C: Examples of dual-class firms

New York Times	Estee Lauder	Nextel
Washington Post	Revlon	Viacom
Readers Digest	Ralph Lauren	Comcast
Scholastic	Timberland	Adelphia
American Greetings	Kenneth Cole	
Playboy	Abercrombie&Fitch	Berkshire Hathaway
		Goldman Sachs
Kraft Foods	FOX Entertainment	
Coca Cola	Blockbuster	Ford Motor Co
Nabisco	Martha Stewart	
Hershey Foods	Loews Cineplex	Continental Airlines
Pepsi		Hawaian Airlines
Ben&Jerrys		

Not-so-famous: Coinmatch Laundry Co, Apple Orthodontix Inc, Badger Meter Inc, Falcon Building

Superior and inferior classes are defined so that the number of votes of the superior is greater than the inferior. If the classes have equal voting rights, the superior class is the one that elects the higher proportion of directors. When number of votes and proportion of directors differ in ranking, the case is eliminated.

Table 2
Ownership and financial variables in 2002*

Panel A: Insiders' cash-flow and voting rights

	Mean	Min	Max	Obs
Insiders cash-flow ownership	37%	1.7%	93%	311
Insiders voting rights	59%	1.0%	100%	302
Wedge	23%	-22%	83%	300

Panel B: Financial variables for fiscal year end in 2002

	Mean	Obs
Book value of assets, in $mil	4022.65	301
Salesgrowth	16%	297
Ratio of capital expenditures to sales	0.09	296
Book ratio of debt to assets	0.23	301
Ratio of R&D to sales	0.19	297
Ratio of advertising to sales	0.02	297

*Excluding regulated utilities and financial firms.

Table 3
Variable definitions

Variable	Description	Source
Q	Tobin's Q, the ratio of market value of assets to their replacement cost; Calculated as Q = (AT + CSHO*PRCC_F - CEQ - TXDB) / AT	
lnQ	Equals the natural log of Q	
AT	Book value of assets	Compustat Item 6
CSHO	Common shares outstanding	Compustat Item 25
PRCC_F	Price fiscal year close	Compustat Item 199
CEQ	Book value of common equity	Compustat Item 60
TXDB	Balance sheet deferred taxes	Compustat Item 74
INSIDERCF	Percentage of cash-flow rights owned by officers and directors Calculated as INSIDERCF= (DIVAMTA*INSIDERA+…+DIVAMTD*INSIDERD)/(DIVAMTA*SHROUTA+…+DIVAMTD*SHROUTD)	
DIVAMTA-DIVAMTD	Equals dividends per share for classes A-D if dividends are paid; Equals the proportion of dividends across classes if paid in the future; Equals one if dividends not paid	Dual-class data*
SHROUTA-SHROUTD	Shares outstanding of classes A-D	Dual-class data*
INSIDERA-INSIDERD	Number of shares of classes A-D owned by insiders	Dual-class data*
INSIDERV	Percentage of votes held by officers and directors Calculated as INSIDERV= (VOTEA*INSIDERA+…+VOTED*INSIDERD)/(VOTEA*SHROUTA+…+VOTED*SHROUTD)	
VOTEA-VOTED	Number of votes attached to classes A-D	Dual-class data*
WEDGE	Equals the difference between INSIDERV and INSIDERCF	
SIZE	Equals *ln*(AT)	
SALESGROWTH	Growth of sales	Compustat Item 12
CAPEX	Ratio of capital expenditures to sales	Compustat Items 128, 12
DEBT	Book ratio of long-term debt to assets	Compustat Items 9, 6
ADVERT	Ratio of advertising expenditures to sales	Compustat Items 45, 12
R&D	Ratio of research and development expense to sales	Compustat Items 46, 12
AFTERSOX	Equals one if observation is after SOX signed into law on July 30, zero otherwise	
SIGMA	Idiosyncratic risk calculated as the standard error of regression of daily stock return on market return	CRSP, Fama-French factors
NAME	Indicator if firm named after the founder	Dual-class data*
CEOCOMP	Salary and bonus of top executive	Execucomp TOTAL_CURR

*Dual-class data is from Gompers, Ishii and Metrick, "Extreme Governance," 2008.

Table 4
The effect of SOX using FE estimation

	Model 1 (OLS)	Model 2 (FE)	Model 3 (FE)	Model 4 (FE)	Model 5 (FE)
	lnQ	lnQ	lnQ	lnQ	lnQ
AFTERSOX				-0.067	
				(0.062)	
INSIDERCF	0.046*	0.586***	0.586***	0.529***	0.545***
	(0.025)	(0.096)	(0.154)	(0.155)	(0.152)
WEDGE	-0.116**	0.137	0.137	0.118	0.150
	(0.057)	(0.136)	(0.163)	(0.162)	(0.162)
INSIDERCF*AFTERSOX				-0.194*	-0.293***
				(0.110)	(0.066)
WEDGE*AFTERSOX				0.054	-0.044
				(0.123)	(0.089)
SIZE	-0.034***	-0.223***	-0.223***	-0.219***	-0.221***
	(0.007)	(0.018)	(0.041)	(0.040)	(0.040)
SALESGROWTH	0.041***	0.012*	0.012	0.010	0.010
	(0.009)	(0.006)	(0.011)	(0.011)	(0.011)
CAPEX	0.010***	0.002	0.002**	0.002**	0.002**
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Constant	0.636***	1.324***	1.324***	1.344***	1.338***
	(0.058)	(0.116)	(0.243)	(0.240)	(0.239)
Observations	2621	2621	2621	2621	2621
Obs after SOX				292	292
R-squared	0.09	0.12	0.12	0.13	0.13
Number of gvkey		562	562	562	562

Model 1 is estimated with OLS including 11 industry dummies (coefficients not reported).
Models 2, 3, 4 and 5 are estimated with firm-level fixed effects.
Model 2 assumes homoskedastic errors and does not account for the effects of SOX.
Models 3, 4, and 5 allow for clustered standard errors by firm.
Model 4 examines the effects of SOX by including the indicator variable AFTERSOX.
Model 5 includes AFTERSOX only in the interaction terms with INSIDERCF and WEDGE.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5
Robustness of the FE results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	lnQ	Q	Q censored	lnQ	lnQ	lnQ
INSIDERCF	0.545***	1.023	1.338***	0.538***	0.549***	0.122
	(0.152)	(1.034)	(0.340)	(0.155)	(0.152)	(0.422)
WEDGE	0.150	1.330	0.177	0.171	0.154	0.531
	(0.162)	(1.504)	(0.379)	(0.160)	(0.161)	(0.389)
INSIDERCF*AFTERSOX	-0.293***	-0.693***	-0.598***	-0.293***	-0.295***	-0.282***
	(0.066)	(0.206)	(0.129)	(0.067)	(0.067)	(0.067)
WEDGE*AFTERSOX	-0.044	-0.244	-0.049	-0.048	-0.044	-0.055
	(0.089)	(0.374)	(0.168)	(0.088)	(0.089)	(0.089)
SIZE	-0.221***	-1.517***	-0.439***	-0.223***	-0.220***	-0.219***
	(0.040)	(0.552)	(0.082)	(0.040)	(0.041)	(0.040)
SALESGROWTH	0.010	0.067	0.014	0.010	0.011	0.011
	(0.011)	(0.079)	(0.018)	(0.011)	(0.011)	(0.011)
CAPEX	0.002**	-0.002	-0.026	0.002**	-0.001	0.002**
	(0.001)	(0.007)	(0.056)	(0.001)	(0.021)	(0.001)
R&D					0.000	
					(0.002)	
ADVERT					-0.156	
					(0.193)	
DEBT				-0.017		
				(0.120)		
INSIDERCF^2						0.528
						(0.485)
WEDGE^2						-0.542
						(0.509)
Constant	1.338***	9.334***	3.592***	1.349***	1.336***	1.351***
	(0.239)	(2.985)	(0.503)	(0.242)	(0.240)	(0.246)
Observations	2621	2621	2561	2605	2621	2621
R-squared	0.13	0.10	0.10	0.13	0.13	0.13

For comparison purposes, Model 1 is the same as the baseline specification of Model 5, Table 4.
In Model 2, the dependent variable is just Tobin's Q, not the natural log of Tobin's Q.
In Model 3, extreme values of Tobin's Q are censored at 1% and 99%
Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6
Year dummies to test for business cycle effects

Dependent variable: lnQ			
INSIDERCF	0.388***	InsiderCF_95	0.109
	(0.125)		(0.151)
WEDGE	0.165	InsiderCF_96	-0.052
	(0.154)		(0.135)
SIZE	-0.212***	InsiderCF_97	0.035
	(0.018)		(0.127)
SALESGROWTH	0.008	InsiderCF_98	0.142
	(0.006)		(0.123)
CAPEX	0.001	InsiderCF_99	0.214*
	(0.001)		(0.114)
year95	0.122	InsiderCF_00	-0.005
	(0.077)		(0.118)
year96	0.223***	InsiderCF_01	0.162
	(0.069)		(0.133)
year97	0.211***	Wedge_95	-0.206
	(0.064)		(0.165)
year98	0.132**	Wedge_96	-0.265*
	(0.062)		(0.147)
year99	0.01	Wedge_97	-0.1
	(0.063)		(0.139)
year00	0.075	Wedge_98	-0.071
	(0.061)		(0.139)
year01	0.047	Wedge_99	0.099
	(0.061)		(0.138)
		Wedge_00	0.001
			(0.135)
		Wedge_01	-0.166
Constant	1.213***		(0.136)
	(0.122)		
Observations	2621		
R-squared	0.17		

The FE model is estimated over the years 1994-2001 including year dummies (1994 is omitted) and interactions between the year dummies and the ownership variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7
First-stage regressions for IV models

	Model 1		Model 2			
	INSIDERCF	WEDGE	INSIDERCF	WEDGE	INSIDERCF* AFTERSOX	WEDGE* AFTERSOX
SIGMA	-0.359	-0.364	-0.256	-0.682	-0.099	-0.103
	(1.127)	(1.101)	(0.615)	(0.645)	(0.218)	(0.227)
NAME	0.034	0.03				
	(0.032)	(0.037)				
CEOCOMP	0.000	0.001**	0.000	0.001**	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SIGMA*AFTERSOX			-1.513	0.760	-1.056*	0.654
			(1.578)	(1.657)	(0.561)	(0.582)
CEOCOMP*AFTERSOX			0.000	0.000	0.000	0.001**
			(0.000)	(0.000)	(0.000)	(0.000)
SIZE	-0.017	-0.015	-0.018***	-0.015***	-0.002	-0.003
	(0.012)	(0.014)	(0.005)	(0.006)	(0.002)	(0.002)
SALESGROWTH	0.014	0.004	0.012	0.003	-0.002	0.000
	(0.013)	(0.013)	(0.015)	(0.016)	(0.005)	(0.005)
CAPEX	0.079**	0.011	0.078***	0.011	0.007	0.003
	(0.038)	(0.034)	(0.026)	(0.027)	(0.009)	(0.009)
AFTERSOX			0.027	-0.037	0.303***	0.187***
			(0.053)	(0.055)	(0.019)	(0.019)
Constant	0.391***	0.313***	0.399***	0.334***	0.016	0.021
	(0.089)	(0.100)	(0.041)	(0.043)	(0.015)	(0.015)
Observations	868	868	868	868	868	868
Centered R-sq	0.03	0.02	0.04	0.03	0.65	0.57

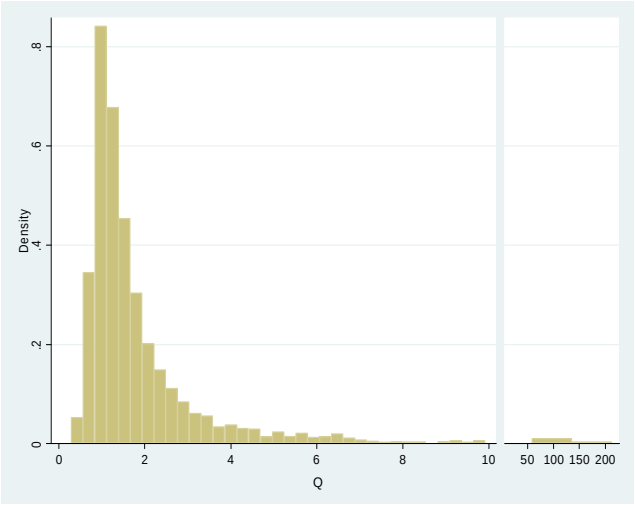
Model 1 is 2SLS allowing for clustered standard errors (efficient GMM).
Model 2 is the same as Model 1 but includes firm fixed effects and SOX.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 1
The effect of outliers in Tobin's Q

Fig. 1a

Tobin's Q

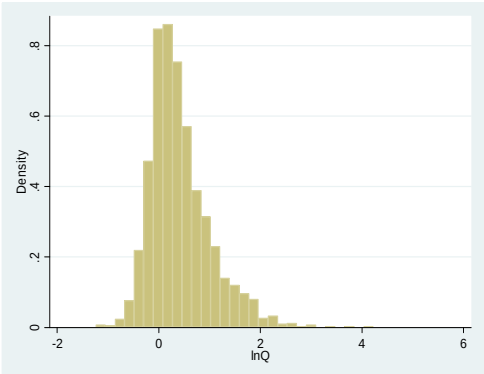
	Percentiles	Smallest	
1%	0.552	0.291	
5%	0.723	Largest	
10%	0.827	212.315	
25%	1.016		
50%	1.356	Obs	3634
		Mean	2.043
75%	2.043	Std. Dev.	4.504
90%	3.419	Variance	20.285
95%	5.071	Skewness	31.171
99%	9.923	Kurtosis	1344.587



lnQ

	Percentiles	Smallest	
1%	-0.593	-1.235	
5%	-0.325	Largest	
10%	-0.189	5.358	
25%	0.016		
50%	0.304	Obs	3634
		Mean	0.427
75%	0.714	Std. Dev.	0.615
90%	1.229	Variance	0.378
95%	1.623	Skewness	1.385
99%	2.295	Kurtosis	7.017

Fig. 1b



Q censored at 1% and 99%

	Percentiles	Smallest	
1%	0.624	0.561	
5%	0.744	Largest	
10%	0.845	9.916	
25%	1.025		
50%	1.357	Obs	3551
		Mean	1.803
75%	2.016	Std. Dev.	1.326
90%	3.300	Variance	1.759
95%	4.581	Skewness	2.660
99%	7.116	Kurtosis	11.642

Fig. 1c

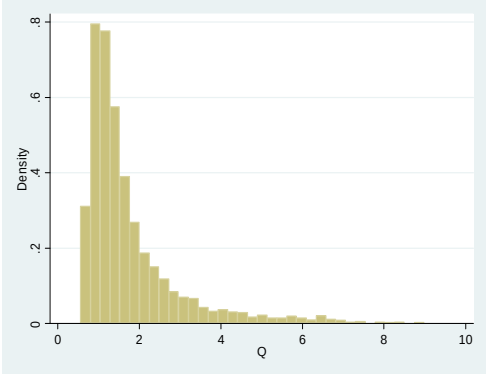
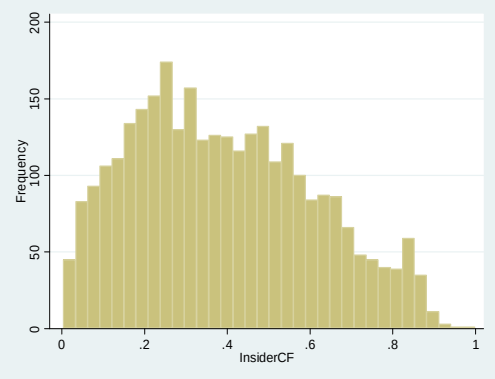


Figure 2
The distributions of INSIDERCF, INSIDERV and WEDGE

INSIDERCF

	Percentiles	Smallest	
1%	0.025	0.004	
5%	0.070	Largest	
10%	0.113	0.999	
25%	0.216		
50%	0.367	Obs	3012
		Mean	0.391
75%	0.549	Std. Dev.	0.218
90%	0.696	Variance	0.047
95%	0.793	Skewness	0.338
99%	0.867	Kurtosis	2.259

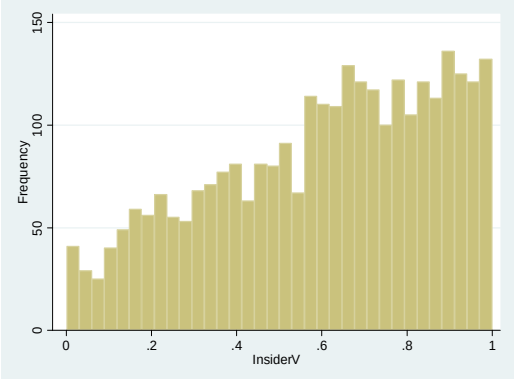
Fig. 2a



INSIDERV

	Percentiles	Smallest	
1%	0.023	0.002	
5%	0.126	Largest	
10%	0.203	1.000	
25%	0.398		
50%	0.644	Obs	2927
		Mean	0.603
75%	0.827	Std. Dev.	0.265
90%	0.931	Variance	0.070
95%	0.967	Skewness	-0.414
99%	1.000	Kurtosis	2.143

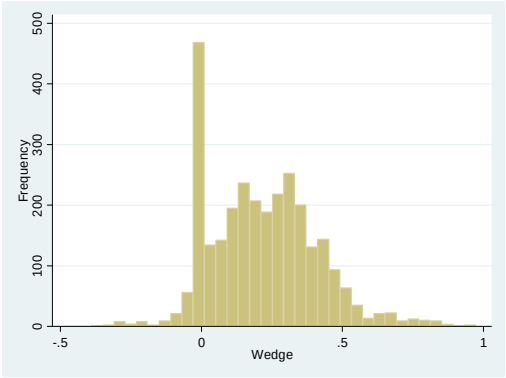
Fig. 2b



WEDGE

	Percentiles	Smallest	
1%	-0.135	-0.390	
5%	-0.014	Largest	
10%	0.000	0.974	
25%	0.056		
50%	0.205	Obs	2919
		Mean	0.216
75%	0.338	Std. Dev.	0.189
90%	0.453	Variance	0.036
95%	0.524	Skewness	0.485
99%	0.753	Kurtosis	3.320

Fig. 2c



CHAPTER 3. THE EFFECT OF SARBANES-OXLEY ON THE PERFORMANCE OF INFERIOR VS. SUPERIOR SHARES OF DUAL-CLASS COMPANIES

1. Introduction

The Sarbanes-Oxley Act (SOX) applies to exchange-listed firms and a major category in its provisions concerns corporate governance, including the power exercised by corporate insiders. Managers and directors of firms, otherwise called "insiders," can have voting rights by owning shares in their company. Voting control insulates insiders from the disciplining force of the market for corporate control and allows the extraction of private benefits. To the extent that extraction of private benefits and entrenchment is reduced due to the corporate governance provisions of SOX, there should be a reduction in the value of voting control of insiders. The effect of SOX on corporate governance has not been studied in this context. I exploit differences in the returns of dual-class shares to examine the effect of SOX on the value of control.

Dual-class firms are a useful empirical tool to study changes in the value of voting control when both classes are traded. A defining feature of dual-class shares is that they allow for disproportionate voting rights and sometimes disproportionate rights in electing the board of directors. Superior shares have more votes and/or elect more directors on the board than inferior shares. As a result of these differences, there is a voting premium reflected in the prices of shares so that superior shares command a higher price. The private benefit of insiders who hold the superior class for the purpose of maintaining voting control is one of the most important factors determining the magnitude of the voting premium. The corporate governance provisions of SOX are expected to reduce the value of control and hence the voting premium by limiting extraction of private benefits and transferring value from insiders to outside shareholders. Since the value of voting control is reflected in the prices of superior and inferior shares of dual-class companies,

the reduction of the value of control will translate into better performance of inferior vs. superior stock returns around key SOX dates.

Empirical studies of the effect of SOX often use the market price of shares. Several studies focus on stock market reactions to SOX and establish that the valuation of some firms increased because of the stricter governance provisions of SOX. Chhaochharia and Grinstein (2007) differentiate firms based on the degree of compliance with the provisions of SOX prior to their implementation. Being less compliant with the provisions would be an indication of higher agency costs, and indeed these firms experienced an increase in value relative to the more compliant firms. Li et al (2008) examine the effect of SOX on firms with sub-optimal governance before SOX evidenced by a lot of earnings management and find a positive reaction. Other papers that find positive relative effect of SOX in terms of stock price reaction are Hochberg et al (2007), Jain and Rezaee (2006), and Li et al (2008).

While some firms can benefit from SOX relative to others, it is possible that some or all firms can be made worse off because compliance costs outweigh the benefits. Some costs associated with SOX that have been recognized in the literature are the “out-of-pocket” compliance costs as well as certain opportunity costs. Block (2004) discusses the higher cost of being public after SOX due to higher fees paid to outside directors, higher premiums for officers and directors’ insurance and more expensive audits. In addition, there is the time pressure on officers due to the need to personally certify financial statements, oversee audits, and attend committee meetings of the board to ensure SEC compliance. Certain costs are fixed and since there are no exemptions for small firms, compliance is relatively more costly for small firms [Block (2004)]. These firms may also find it harder to attract independent directors.

Chhaochharia and Grinstein (2007) find that small firms experience a negative relative effect in terms of stock price.

An important advantage of dual-class shares is that we can identify whether there was any benefit of SOX in terms of corporate governance regardless of the costs. Changes in the voting premium translate into differences in the returns on inferior and superior shares. Since compliance and other costs will affect the returns on both classes of shares, only the positive governance effect of SOX will be present in the difference of returns on the inferior and superior class. The inferior class is expected to perform better than the superior class due the reduction of private benefits of control reflected in the prices of the two classes of shares.

The estimation of the effect of SOX on firm value net of all costs and benefits is complicated by the need to isolate confounding contemporaneous events. Since SOX applies to all public companies, there is no readily available control group of firms that are not affected. Litvak (2007) and Smith (2007) use foreign firms listed in the US as controls and find a negative overall effect of SOX on US firms. Zhang (2007) tests the overall effect of SOX on US firms relative to foreign firms cross-listed in the US not subject to SOX and also finds a negative effect of SOX but the results are criticized because foreign markets do not adequately control for contemporaneous events [Leuz (2007)].

There has not been a study that examines how the price ratio of superior and inferior shares in the US, or the voting premium, changes as a result of SOX. This approach would control for concurrent market events much better. The fact that the shares were issued by the same company means that it is possible to control for contemporaneous macroeconomic factors that affect the returns of both classes of shares of the company. Macroeconomic factors, for

example, would affect both classes in the same way. Thus, country, market, industry and some firm factors are effectively washed out.

The way SOX affects the voting premium is by limiting the ability of insiders to expropriate minority shareholders, or through changes in private benefits. The challenge is to control for other governance factors that specifically affect the premium. For instance, there may be simultaneous changes in the market for corporate control affecting the probability of control contest. However, other governance factors are unlikely to change in a narrow event window around key SOX dates.

Therefore, using an event study approach, I test the hypothesis that SOX reduced the ability of insiders to extract private benefits. The reduced value of voting control is empirically reflected in higher returns on inferior shares relative to superior shares around key SOX events. The event study methodology assumes that markets are efficient and prices of stocks immediately adjust to reflect new information. Thus the effect of SOX can only be seen in the market reaction of inferior versus superior shares of dual-class firms, and hence in the market expectation of this effect. It is important to identify SOX events that reveal new information, otherwise the political process is already reflected in prices.

2. SOX event history

Previous studies have performed event history analysis of SOX identifying dates when new information about the probability of passage of the Act and the nature of the rules is revealed to the market. Zhang (2007) compiles a comprehensive history of the events starting with the Enron scandal in late 2001 until the passage of the legislation on July 30, 2002. None of the events until the announcement of fraud by WorldCom and the SEC filing suit on June 25,

2002 are found to trigger significant market reactions by Zhang (2007). The reason is that, although much of the content of the reform had been discussed prior to June 25, it lacked support in Congress. However, news of massive fraud at WorldCom changed the political environment for reform. Therefore Li at al (2008) collect eight critical events during the SOX legislative process starting on June 25, 2002, and I follow their event taxonomy. The event history is developed on the basis of a legislative history of SOX, the Library of congress's Advanced Bill Summary and Status Search for the 107th Congress, and an extensive search of articles in the press using Factiva and Lexis-Nexis databases. The eight critical events released new information to the market relating to the probability that SOX would become law and/or the extent of financial reporting reforms in the legislation.

Thus I consider the event periods in Li et al (2008): June 25, June 26-27, July 8-12, July 15-17, July 24, July 25-26, July 29-30, August 14-15. The only difference is that I combined the first two events into one on June 25-27, as in Zhang (2007), because I do not have the intraday data they use for the June 25 event. I also consider an additional event on August 16 related to the changes in listing requirements. As discussed in Chapter 1, NYSE and NASDAQ changed their listing requirements around the same time SOX was passed to include new corporate governance requirements such as majority of independent directors on board, independent audit committee, and independent board committees that choose new directors and compensate managers. Table 1 lists the events.

All of the developments relating to SOX took about a year to play out. The final legislation combined the reform bills of Sen. Sarbanes and Rep. Oxley. The two bills, together with other legislative proposals towards corporate reforms, were triggered directly by the collapse of Enron, which exposed an unprecedented accounting scandal and allegedly seriously

corrupted governance system. There was no significant development in securities rule making in 2001. The first signal of regulatory overhaul was reported in January 2002 when SEC Chairman Pitt announced a reform plan to create an independent regulatory organization. The Bush Administration unveiled their response to the Enron scandal in February and March, while Congress moved ahead with several proposals towards accounting reforms. Republican Rep. Oxley's bill was introduced in the House in February. Democrat Sen. Sarbanes' tougher reform bill passed the Senate Banking Committee on June 18, but the media did not expect it to become law at that time.

The WorldCom scandal on June 25 boosted rulemaking activities, and the process accelerated after President Bush delivered a speech regarding accounting reforms on July 9. The Senate started debate on Sarbanes' bill on July 8, and the bill was passed in the Senate on July 15. House GOP leaders allegedly sought to dilute Sarbanes' bill after its passage. However, they soon retreated from such efforts and offered minor changes to complete the legislation. The House and Senate formed a Conference Committee and started final negotiations to merge the bills. The final rule was agreed upon on July 24 when the Conference Committee Report was issued. SOX was passed in Congress on July 25, and signed into law on July 30. The implementation of SOX started soon after its passage and the rulemaking activities continued afterwards. Specifically, the SEC collaborated with major exchanges to produce new listing requirements that were proposed by the NYSE in August 2002.

The underlying reasoning behind using event studies to test the effect of legislation is that financial markets are efficient. When markets are "informationally" efficient, prices reflect all publicly available information and instantly change to reflect new information. The choice of the eight SOX events is exactly because the information as to the probability of the passage of the

Act and how tough or easy the rules are expected to be released on these dates is new to the market and is hence not yet incorporated in prices. Otherwise, the political negotiation process would already be incorporated in price expectations and the focus on these dates would be self-defeating. The SOX events used in this study are such that new information is released to the market which leads to the updating of expectations. This also assumes that market agents know precisely how SOX will affect the corporate world so that expected effect is the same as actual effect of the legislation.

Such a model relies on the assumption of market efficiency. Accordingly, I test a variety of event windows as markets may take time to process the implications of SOX. I look at cumulative returns over adjacent event windows as well as over all events in the study. In addition, the downside to using short event windows is that market expectations are volatile and results may be sensitive to selection of event dates. Selection of event dates and interpretation of market signals varies across studies. Chhaochharia and Grinstein (2007) avoid this problem by considering a “buy-and-hold” period that captures all belief updates from beginning of November 2001 to end of October 2002. Return volatility is estimated in the 100 trading days before the event window. Following this approach, I also include wide event windows calling them “buy-and-hold” periods in Table 1 to distinguish them from the short events.

Thus one of my buy-and-hold periods is as in Chhaochharia and Grinstein (2007), which is a year long. With a wide event window other factors, such as a pickup in takeover activity in market for corporate control affecting specifically the voting premium, can change concurrently as well. Therefore I also consider a shorter buy-and-hold period given Table 1. The most important developments leading up to SOX happened during about a month before the final adoption, so that buy-and-hold period is from June 25, 2002 to July 30, 2002. The downside to

wide event windows and cumulative returns over adjacent trading days, however, is that it becomes difficult to identify when new information is released and what the market reaction is.

3. The voting premium

Theoretical research summarized in Chapter 1 shows that private benefits of insiders are an important factor for insiders' control. Other factors determining the choice of dual-class status include the managers' cost due to undiversified holdings in the company, protection from value-reducing takeovers and realization of growth options of the firm, and human capital investment by management.

Zingales (1995) defines the voting premium as the ratio between the prices of superior and inferior shares. Thus the voting premium is $VP = P_S/P_I$, where P_S and P_I are the prices of superior and inferior shares of the same company. In his model, voting premium is determined by the magnitude of private benefits of control (relative to ability to produce income), probability of control contest, and proportion of superior shares outstanding. The voting premium also depends on differential voting rights between the two classes, differential dividend rights, and market liquidity. Corporate charter provisions prohibit the preferential treatment of superior shares in terms of other capital distributions. Thus after adjusting for any difference in dividends, voting rights, and liquidity, the voting premium will reflect the value of control due to private benefits. Depending on the definition and adjustments, studies find significant voting premium in the US of 4.5% [Nenova (2003)], 5.4% [Lease et al (1983], and 10.5% [Zingales (1995)].

Changes in the voting premium will translate into a difference in returns on inferior and superior shares. Taking logs, $lnVP = ln\ P_S - ln\ P_I$. Taking derivates, $dVP/VP = dP_S/P_S - dP_I/P_I$. Therefore, the percentage change in the voting premium is equal to the difference in the

percentage changes in the prices of superior and inferior shares. In other words, changes in the voting premium are reflected in the difference in returns on superior and inferior shares. Since an important factor in the voting premium is private benefits, if SOX limits diversion of resources from outside shareholders, the legislation will affect the returns of the superior vs. inferior class.

Other governance factors such as takeover activity or dividend policy are unlikely to change around the SOX event dates, so that only changes in private benefits will affect the voting premium. In addition, compliance costs affect both classes and would not cause difference in returns. Suppose, for example, that the dual-class company had minimized its agency costs and was operating at the efficient value-maximizing point before SOX. After SOX, compliance costs affect equally both classes and their respective returns, but there is no benefit and no difference in the returns. If, on the other hand, there was scope for improvement in governance before SOX, which was not undertaken due to private benefits for insiders, then SOX is value-increasing for outside shareholders. Keeping costs constant, there is a positive effect in the security value for those companies that did not operate at the value-maximizing point of corporate governance [Wintoki (2007)]. We can see this effect in the difference in returns on inferior and superior classes net of the costs that affect both classes.

An example here may serve to illustrate the expected effect of SOX. Assume private benefits are $100 and that they can be transformed into security value (i.e. shared by all shareholders) at cost to shareholders of $60, which is the compliance cost associated with SOX. The compliance cost and security benefits are shared equally between holders of inferior and superior shares. Then after compliance, the inferior shareholders are better of by $20 (because they paid $30 in costs and received $50 of formerly private benefits). The superior shareholders are worse off by $80 (the $50 in foregone private benefits and $30 in compliance costs). In this

example, SOX is not value-increasing overall because there is a $60 loss in value due to compliance costs. However, SOX affects how private benefit is transferred into security value. In other words, the effect of SOX is the $50 that gets redistributed from the superior to the inferior shareholders, who are better off even after compliance costs.

This is not to say that all of the benefit of SOX can be seen in the difference of returns between inferior and superior classes. Many of the provisions of SOX will affect both classes in terms of costs *and* benefits. For example, better financial reporting will raise returns on both classes. However, since the pricing difference between superior and inferior class is driven by corporate governance, some provisions of SOX will affect one of the classes more than the other. This way it is possible to observe if the legislation was effective in terms of restricting private benefits and improving governance regardless of the compliance costs.

4. Data

4.1. Description of sample

Typically in the US, only one or two of the multiple classes trade publicly, which is usually the inferior class. When two classes trade, they are usually superior and inferior class, but it is possible both traded classes to be inferior to a closely-held class. Superior and inferior classes in this study are defined so that the number of votes attached to the superior shares is always greater than the number of votes attached to inferior shares. If the votes are equal, the superior class elects a higher proportion of directors on the board. I eliminate the cases when two inferior classes trade, and when votes and proportion of directors elected produce different rankings. For each event, the sample includes all dual-class firms whose superior and inferior classes are traded and share price data is available from CRSP.

Data on number of votes, number of existing classes, number of trading classes, votes and proportion of directors elected by class comes from the dual-class data in Gompers et al (2008). The number of firms varies between 54 and 56 according to the event date. Data on market return and benchmark factors were obtained from Kenneth French's webpage. Compustat and Execucomp were used for financial statistics on dual-class firms.

4.2. Sample selection issues

As already discussed in Section 3, the voting premium reflected in the prices of superior and inferior shares is determined by private benefits of control, probability of control contest, and proportion of superior shares outstanding. The voting premium also depends on differential voting rights between the two classes, differential dividend rights, and market liquidity. Thus the main driver of the voting premium relevant for corporate governance regulation is the existence of private benefits of the controlling shareholders. SOX is expected to reduce extraction of private benefits, which translates into better performance of inferior shares relative to superior shares around key SOX events.

The sample includes only firms that trade both the inferior and the superior class, because otherwise prices would not be available for the superior class. Hence, there may be some non-randomness in which firms do not have both classes trading, i.e. selection issues. Only about 15% of all dual-class firms in 2002 have both classes trading as shown in Panel A of Table 2. The reason why the superior class is usually closely held is that it provides a mechanism through which insiders maintain voting control of the firm. For example, in firms where only the inferior class trades in 2002, insiders' voting rights are on average 60% (Table 2, Panel B). In contrast, when both superior and inferior classes trade, voting control is only 49% and the difference is

significant at 5%. Thus voting control is significantly lower at firms with two traded classes which may be a sign of low private benefits for insiders.

Although the lower voting control of insiders at firms with two traded classes may indicate private benefits are not so important, what matters really is the allocation of voting and cash-flow rights among insiders and outside shareholders and the factors that determine this allocation. Panel B of Table 2 explores differences between dual-class firms that trade both the superior and inferior class from those with only inferior traded class in 2002. I will discuss financial indicators, measures of inside ownership, and finally idiosyncratic risk. Panel C of the table gives some examples of dual-class firms, although the "famous" ones with a superior and inferior traded class are not so many, probably because of the small number of these firms to begin with.

In terms of financial indicators, sales growth, which can be used as a proxy for growth opportunities, is significantly lower for firms with two traded classes. They are also older, on average 25 versus 13 years, and the difference is significant. They are bigger based on book value of assets but the difference is not significant at 5%. There are no significant differences in terms of capital expenditures and leverage. So far the picture that emerges is that dual-class firms with two traded classes are older and have fewer growth options than firms with only inferior traded class. Growth seems to be a factor for choosing dual-class structure at IPO, which is the case for the firms trading only one class. Firms with two traded classes generally recapitalize into dual-class class much later than the IPO.

Along the dimensions of ownership and private benefits, the significantly lower voting control of insiders at firms where superior and inferior classes trade (49% vs. 60%) was already discussed. However, cash-flow ownership is also significantly lower (31% vs. 38%). In fact,

Claessens et al (2002) argue that the difference between voting and cash-flow rights (WEDGE) is a better proxy for private benefits. In a sample of East Asian firms, they find that high-wedge firms have low value relative to low-wedge firm. This is because at high-wedge firms, the negative effect of insiders' voting control is stronger than the positive incentive alignment effect of holding cash-flow rights. Although the wedge between voting and cash-flow rights at firms where both superior and inferior classes trade is lower than those where only the inferior class trades, the difference is not significant. Therefore, using this proxy of private benefits, there is no evidence of difference among the two categories of firms.

Zingales (1995) finds that CEO compensation can be used as a proxy for private benefits because CEO compensation is higher at firms with high voting premium. There is no evidence of that here since the difference in CEO compensation is not significant.

Gompers et al (2008) propose yet another proxy for private benefits, which is whether the firm is named after one of its founders. They collect a variable equal to one if a person's name appears in the company name at IPO. Surprisingly, there is a significant difference in this variable between firms that trade two classes and those that trade only one. It appears that firms where both inferior and superior classes trade are named after the founders *more* frequently. If this variable is a good proxy for private benefits, it would mean that they are in fact higher at the firms used for estimation.

Finally, a significant difference emerges in terms of idiosyncratic risk. Idiosyncratic risk is measured as the root mean standard error (RMSE) in a regression of the inferior return on the market return during 2001 and 2002. Idiosyncratic risk can be a factor for low cash-flow holdings by managers because of the cost of holding an undiversified portfolio. In a diversified portfolio, idiosyncratic risk is eliminated and only systematic risk remains. However, since

managers hold an undiversified portfolio, the higher the cash-flow rights in the company, the higher the forgone-diversification cost.

In sum, firms that trade both classes are older and have lower growth. They also have lower idiosyncratic risk, which may be related to lower growth options. Although insiders' voting rights are lower on average, they are enough to ensure effective control. Trading in the superior class would provide the benefit of a more liquid market for the controlling shares. The wedge between the voting control and cash-flow rights is lower but not significantly different. CEO compensation is also not significantly different. In fact, these companies appear to be named after their founders more often.

Even if the sub-sample of dual-class firm used in the estimation is representative of all dual-class firms, an important question is whether it is representative of all firms. Along the dimensions of financial indicators, Gompers et al (2008) find that dual-class firms are significantly different from single-class firms in terms of only age (13 years for dual-class firms vs. 10 year for single-class firms) and leverage (0.23 debt-to-asset ration for dual-class firms vs 0.17 for single-class firms). In the specific sample of dual-class firms trading two classes, I find that they are even older (23 years) while leverage is still significantly higher (0.2). In terms of sales growth, firms with two traded classes are comparable to single-class firms.

For the event study results, however, what really matters for sample selection are factors that affect the voting premium, i.e. those that pertain to corporate governance. I find that 39% of firms that trade two classes are named after their founder, but this is the case for only 12% of single class firms. Both insider cash-flow rights and voting rights are higher than inside holdings at single-class firms of 10-18% [Morck et al (1988), Himmelberg et al (1999)]. Gompers et al (2008) account for the possibility of such sample selection bias when examining dual-class

firms' governance by using the methods of Heckman. The selection equation includes single-class firms and aims to explain the choice of status with measures of private benefits, takeover probability, etc., but firm value is not sensitive to such correction. Despite the differences described between dual-class firms with two traded classes and those with one traded class, and more generally, between dual-class and single-class firms, the evidence in Gompers et al (2008) suggests that firm value is not sensitive to the distinction between dual-class and single-class status. Thus I propose that my event study results are representative of the wider class of firms.

5. Empirical model

5.1. Matched pairs

Evaluating the effect of SOX requires controlling for contemporaneous factors. One way to do that is to compare firms subject to SOX to firms similar in all relevant aspects that are not subject to SOX, or less compliant with the provisions of SOX. Studies like these devise methods to match companies based on certain characteristics to control for contemporaneous effects.

In Litvak (2007), "pair return" is the difference in returns of two companies—one subject to SOX and the other not. The companies are "matched" on certain characteristics and the matching controls for country, industry and size factors. The expected difference in returns is zero because the matching produces similar pairs of companies. Any statistically significant observed difference is attributed to SOX. In Chhaochharia and Grinstein (2007), companies are matched on industry, size and market-to-book ratio. They claim that the matching portfolios have similar risk characteristics so that their expected returns are the same. Because of the risk/return trade-off, expected difference in returns is zero.

In the case of dual-class companies, matching is not necessary since the classes belong to the same company. In a sense, the two classes are perfectly matched so that many events, such as macroeconomic or industry, would affect both classes equally. Therefore, any observed difference in returns should be due only to changes in the factors determining the voting premium. Provided that there is no difference in risk between the two classes of shares, the benefit of the matched pairs approach is that not only market but also industry and even some firm factors that affect both classes are controlled for.

Let R_{Sjt} be the daily total return on the superior class of common stock of dual-class company j, and R_{Ijt} is the daily return on the inferior class of the same company. In the case when regulation affects a large number of firms at the same point in time there are statistical reasons for using portfolio returns instead of analyzing the returns to each individual asset in associations with a regulatory change. Since returns are contemporaneously correlated, probability statements based on the analysis of several individual asset returns for the same time period are not independent, and there is no simple way to combine the single-asset tests into a joint probability statement. The portfolio return, however, directly incorporates the cross-sectional dependence, facilitating joint tests of significance [Schwert (1981)]. Equally-weighted portfolios are used as in Brown and Warner (1985) and Chhaochharia and Grinstein (2007). Therefore, let R_{St} be the equal-weighted return on a portfolio of superior shares, and R_{It} be the equal-weighted return on the matching portfolio of inferior shares. That is, $R_{St} = \sum R_{Sjt}/N$ and $R_{It} = \sum R_{Ijt}/N$. Let the difference in daily portfolio returns be $\Delta_t = R_{It} - R_{St}$. The matching portfolios have the same risk characteristics so that $E(R_{It}) = E(R_{St})$, and $E(\Delta_t) = 0$. The t-statistics of observed Δ_t on SOX event days will show if reactions were significant.

5.2. Models of risk-adjusted expected return

Using a market model is another way to control for contemporaneous effects and examine abnormal returns associated with certain events, as in a traditional event study. The benefit of this approach is that it allows for differential risk characteristics of the superior and inferior shares. The CAPM is a standard market model that is used for explaining expected returns and can be used as a benchmark to calculate abnormal returns. The OLS version of the CAPM involves a regression of portfolio returns on market return which is used to predict returns over the event period and calculate excess returns.

The two benchmarks—market model and matching companies--are not completely different, given that both try to control for contemporaneous effects. Litvak (2007) mentions that the "pair return" of matched companies is like the market-adjusted return in a standard event study where the market index consists of non-SOX firms. The benefit of the traditional event study based on abnormal returns over the market is that the two classes of shares could have different risk.

As in the matched-pairs approach, let R_{Sjt} be the daily total return on the superior class of common stock of dual-class company j, and R_{Ijt} is the daily return on the inferior class of the same company. Let R_{St} be the equal-weighted return on a portfolio of superior shares, and R_{It} be the equal-weighted return on the matching portfolio of inferior shares.

Using a market model:

$$R_{St} = \alpha_S + \beta_S R_{Mt} + \varepsilon_{St} \quad (1)$$

$$R_{It} = \alpha_I + \beta_I R_{Mt} + \varepsilon_{It} \quad (2)$$

We can first estimate the equations using daily return data to obtain the parameter estimates. Zhang (2007) does similar regressions in 100 days prior to December 28, 2001. Using

these parameter estimates, we can predict returns absent the impact of SOX and calculate abnormal returns on SOX event days:

$$AR_{St} = R_{St} - \hat{\alpha}_S - \hat{\beta}_S R_{Mt} \quad (3)$$

$$AR_{It} = R_{It} - \hat{\alpha}_I - \hat{\beta}_I R_{Mt} \quad (4)$$

My hypothesis is that the abnormal returns will vary by class, specifically that the inferior class will have a higher abnormal return than the superior class. It is possible for both returns to be negative if markets expect SOX to have net costs, which will affect both classes of dual-class firms. But the return on the inferior class will be less negative due to the positive corporate governance effect. With a short event window around SOX events, we can be relatively certain that there are no other governance changes that affect the voting premium as well.

The magnitude of the t-statistic of the observed difference in returns on superior and inferior shares indicates its significance. Let the observed difference in abnormal returns on any event day t be DAR_t so that $DAR_t = \mathrm{AR}_{It} - \mathrm{AR}_{St}$. For event periods longer than a day, define "difference in cumulative abnormal return"

$$DCAR(t_1, t_2) = \sum_{t1}^{t2} DAR_t \quad (5)$$

We can form the t-statistic for any event lasting from date t_1 to date t_2 as

$$\frac{DCAR(t_1, t_2)}{\sqrt{(t_2 - t_1 + 1)S_t^2}}, \quad (6)$$

where S_t is the standard deviation of the difference in errors in the estimation period [Schwert (1981), MacKinlay (1997)].

Fama and French (1993) find that in addition to the market factor, there are two other benchmark factors that are significant in explaining expected returns. Small-minus-big (SMB) is

the return on a portfolio of small stocks relative to big stocks. Size is measured by the market value of equity, which is defined as price multiplied by shares outstanding. The breakpoint for the Small and Big portfolios is the median NYSE market equity. Portfolios are rebalanced quarterly to reflect changes in market values. Thus SMB summarizes the performance of small stocks relative to big stocks, with small stocks outperforming big stocks on average. The three-factor model treats this difference in performance as a "risk" factor earning additional return and includes it in the regression together with the market return.

High-minus-low (HML) is the return on a portfolio of value stocks relative to growth stocks. Value stocks are those that have high book value of equity relative to market value of equity. Book value of equity is defined as the book value of stockholders' equity from the balance sheet plus deferred taxes minus the book value of preferred stock. Growth stocks have low book-to-market equity because growth options are reflected in stock price which enters the denominator of the book-to-market ratio. The breakpoints for the three portfolios growth, neutral and value stocks are the 30^{th} and 70^{th} percentiles of the NYSE book-to-market equity ratio. Thus HML summarizes the performance of value stocks relative to growth stocks. The price performance of value stocks is found to be better than that of growth stocks. Similarly to HML, this "risk" factor is included in the three-factor model.

The equations for the two portfolios of superior and inferior shares are:

$$R_{St} = \alpha_{St} + \beta_{S1} R_{Mt} + \beta_{S2} SMB_t + \beta_{S3} HML_t + \varepsilon_{St} \quad (7)$$

$$R_{It} = \alpha_{It} + \beta_{I1} R_{Mt} + \beta_{I2} SMB_t + \beta_{I3} HML_t + \varepsilon_{It} \quad (8)$$

These are estimated in addition to (1) and (2) over the period of 100 trading days prior to December 28, 2001, and the parameters are used to predict excess returns on event days. These

are then compared to actual returns and the significance of abnormal reaction is tested using the t-test as in the traditional event study.

5.3. Non-synchronous trading

Thin trading occurs when stocks do not trade every period or if they do, trades happen only occasionally during the day. Total returns used in the event study are based on closing prices. The closing price is the last transacted price, and when stocks trade infrequently, the transaction may have happened earlier in the period. Thin trading is especially a problem for the superior class of shares of dual-class companies. There are two main consequences of thin trading when using daily returns. First, the return of the thinly traded stock may not correspond to the market return in terms of timing because it is calculated based on a price that is too old. Using such non-synchronous returns can lead to biased estimates of beta in the market model. Second, thin trading can lead to non-normality of excess returns. Thus the standard t-test used in event studies to detect abnormal returns on event days may be misspecified in the case of superior shares of dual-class firms.

Table 3, Panel A shows differences in superior and inferior shares in terms of trading during the event study estimation period of 100 trading days prior to December 28, 2001. There are 59 dual-class firms trading two classes of shares during this period. Although in general it is possible for multiple share classes to exist, in this sample the existing classes are always two so that the proportions of shares of the superior and inferior class add up to one. The number of shares outstanding and the market value of the superior class is about half of the inferior class on average in this period. The average daily volume of the superior class is also about half of that of the inferior class. Thus in terms of average daily volume there is no indication of thin trading in

the superior class. However, this statistic mixes trading frequency and size of trade. In other words, it is possible to have less frequent trading in the superior class coupled with high volumes traded.

Panel B illustrates thin trading of the superior class better. Since there are 59 firms with a maximum of a 100 trading days in the sample, the statistics are given in firm-trading days. There are a total of 5,815 firm-trading days because a few firms enter the sample after the start of the period or exit before the end. Closing prices based on last transaction are available for more than 90% of the firm-trading days in the inferior class, but for only 68% in the superior class. This means that often there are no trades in superior shares happening at all during the day. And as already pointed out, even when closing price is available for the superior class, the transaction may have happened earlier in the day because the average daily volume is much lower than the inferior class. Thus the superior class is characterized by thin trading which has implications both for the estimates of beta in the market model and for the distribution of excess returns and test statistics in the event study.

In the case of very infrequent trading there is the problem of zero volume on some days. When no trading happens at all during the day, CRSP uses the average of the bid-ask spread instead of last transacted price in calculating total return. In general, it is possible to have a missing observation in the price field when trading status is halted, suspended or unknown. This never happens in the current sample, so the rest of the prices are based on the bid-ask average (10% of the firm-trading days for the inferior class and 32% for the superior class). Note that it is possible to have non-zero volume but a bid-ask average instead of a transaction price. This happens because for shares traded on the NASDAQ, volumes of after-hours trades are included in the current day while the price quote is given the next day.

Thus total returns calculated by CRSP are based on the trading price, or if unavailable, on the bid-ask average quote. Research shows that, other than using bid-ask averages, there can be different methods to treat missing prices [Maynes and Rumsey (1993)]. One way is to exclude these observations. However, this approach can severely reduce the size of the sample, especially when there is a limited selection of stocks, as in the case of dual-class firms. For example, if the stock trades every second day, it would be completely ignored.

Another approach is to "fill in" missing returns by assuming no change in price if no trade happens so that all the multiperiod return is allocated to the day the stock trades and the return over the intervening periods is equal to zero. Maynes and Rumsey (1993) call this "lumped return." This would produce a larger than normal number of zero returns and when trading does occur, the realized return would represent return over more than one day. Another approach would be to calculate returns on a "trade-to-trade" basis and regress these on market returns using the same trade-to-trade time intervals. Although this is found to be the optimal approach statistically, it requires additional information as to the exact timing of the trades.

Finally, a more complicated approach for filling in missing returns is the one in Heinkel and Kraus (1988). It involves modeling two processes. The first is the return generating process, which can be the CAPM. The second is the process generating the bid-ask spread. Trading prices reflect a bid-ask spread which is caused by the specialist earning a profit from uninformed parties and/or the specialist costs and rents. In other words, the trading price may be higher or lower than the true price, sometimes even equal to the bid or ask, depending on the specialist position as monopolist in trading. Because of the bid-ask spread, the variance of the price is upward-biased, and both processes need to be modeled when filling in missing returns.

The CRSP approach of using the bid-ask average when no trading price is available is a straightforward method that does not require additional information on actual trades or complicated modeling. It is also better than the "lumped" return approach because bid-ask averages update even when no trading happens. Admittedly, it introduces measurement error in the estimation of the true price[12] [Himmelberg at al (1999)]. However, an argument can be made that using the closing price also introduces measurement error [Venkatesh (1992)]. Since transacted prices reflect a bid-ask spread, CRSP returns based on closing price differ from true returns because of the bid-ask effect increasing volatility estimates. In other words, the bid-ask average may be a better approximation of true price as opposed to closing transaction price. Otherwise, there is bid-ask noise in transaction price. However, market model properties such as beta and commonly used test statistics are found to be unaffected when using closing prices.

Apart from measurement errors involved in using either the closing price based on last transaction, or the bid-ask average, both measures suffer from the problem of not being synchronous with the market return when there is thin trading. Especially for the superior class where trading volumes are very low, closing prices based on transactions may be unevenly spaced. When closing prices are unavailable, bid-ask averages update more frequently than if using the "lumped return" approach with no change in price. However, it also happens for the bid and ask to remain the same several days in a row if there is no trading, producing a larger number of zero-return values. The number of zero-return zero-volume trading days in the superior class is about twice that of the inferior class, as shown in Table 3, Panel B. The existence of a higher number of zero returns, as well as large returns, is further investigated below in the discussion about the normality of the distribution of returns.

[12] Again, the true price differs from observed prices because of the bid-ask spread. The trading price may be higher or lower than the true price, sometimes even equal to the bid or ask, if the specialist is a monopolist in trading.

Scholes and Williams (1977) demonstrate that non-synchronous trading leads to errors-in-variables bias when daily returns are used in the calculation of the coefficients in a market model. With thin trading, prices recorded at the end of time period tend to represent the outcome of a transaction which occurred earlier in the period. Therefore the return measured at any time is not the true one that corresponds to the market return at that time but is based on prices that may be too old. Errors-in-variables occur when measured returns are used as proxies for the true unobserved returns. As a result of this measurement error, beta is biased downward in the CAPM market model.

The market return is based on an index which with thinly traded stocks becomes an average of temporarily ordered underlying values of shares. Consequently, positive serial correlation is introduced into returns which are calculated from the index and the estimated variance of the index is biased downwards. The covariance with the returns of the less frequently traded stock is also biased downwards but the bias is smaller because the index consists of stocks with a variety of trading frequencies. If there were two securities only, the discrepancy between measured covariance from true covariance is greatest when one security trades on average very infrequently while the remaining security trades very frequently.

In other words, the covariance of an individual stock with the market return is positively related with the trading frequency. As a result, the beta of a thinly traded stock is biased downward. One approach would be to calculate returns on a trade-to-trade basis as discussed above. Dimson (1979) proposes a method for measuring beta that eliminates the bias and does not require additional information as to the exact timing of the trades. The method involves including non-synchronous plus synchronous market returns as explanatory variables in the regression (aggregated coefficients method). The regression is

$$R_t = \alpha + \sum_{k=-n}^{n} \beta_k M_{t+k}$$
$$\beta = \sum_{k=-n}^{n} \beta_k \quad (9)$$

Thus true systematic risk can be obtained from security price data which is subject to thin trading. All that needs to be done is to run a regression on lagged, matching and leading market returns. A consistent estimate of beta is obtained by aggregating the slope coefficients from this regression.

5.4. Non-normality of returns

Other than biased estimates of beta in the case of thinly traded stocks, there is the problem of non-normality of returns. As an extreme example, if a stock trades only once at the opening of the first day and once at the closing of the second day, the return calculated from these transacted prices would practically represent a two-day return. The reverse is also possible, with trades being spaced very close to each other. Thus thinly traded stocks are more likely to be characterized by numerous zero and large non-zero returns, resulting in non-normal return distribution with fat tails, and less often skewness.

Even in the general case of thickly traded stocks, individual returns are found to be highly non-normal in daily data, which makes the t-test inappropriate. If the disturbances have a normal distribution, the t-statistic has a t-distribution. If individual returns are non-normal, the distribution of t is asymptotically a t-distribution, but in small samples it can deviate from the t-distribution. Brown and Warner (1985) observe that the daily returns for an individual security are non-normal. However, event studies often deal with the cross-section sample mean of excess return of N securities. In the present case, there are portfolios of 59 superior and inferior shares,

so that excess returns are cross-sectional averages. According to the Central Limit Theorem, if the excess returns in a cross section are iid, the distribution of the sample mean converges to normal as N increases. Therefore, the more securities there are in the portfolio, the more likely the distribution of mean excess returns, or portfolio returns, is to be normal. Cable and Holland (2002) find that normality of stock returns is restored with portfolios of 50-60 stocks.

In the case of thinly traded stocks, there are significant departures from normality for individual securities [Maynes and Rumsey (1993)]. The question is whether normality is indeed restored in the portfolios of 59 inferior and superior shares used in the event study. Figure 1 graphically explores normality of the distribution of returns of superior and inferior shares. In addition, the skewness/kurtosis tests for normality give individual statistics as well as a joint statistic for normality (sktest). In the case of the superior and inferior shares of one random dual-class company, it is obvious that the distributions are highly non-normal. The sktest also shows that we reject the null hypothesis of normality of returns. In the case of equally-weighted portfolios of ten random companies, there is improvement in terms of kurtosis and skewness, especially in the inferior portfolio, but the null is still rejected at 5%. Finally, with a portfolio of 59 dual-class companies, normality is restored but only in the case of inferior shares, where the null cannot be rejected at 5%. However, in the superior portfolio, we still reject normality at 5%.

The non-parametric test designed by Corrado (1989) can help in the testing for the existence of abnormal return in event studies when the underlying distribution is non-normal. It is appropriate for skewed and fat-tailed distributions and offers improved specification under the null of no abnormal returns and more power to detect abnormal returns under the alternative. The non-parametric rank test involves transforming each security's time series of market model excess returns during the event study estimation period into their respective ranks. In the case of

the superior and inferior portfolios of shares, the procedure involves transforming the difference in excess returns of the superior and inferior portfolios into its respective rank over the event study estimation period. Let K_t=rank(DAR_t). Since there are 100 trading days in the event study estimation period, there are 100 ranks and K_t ranges from 1 to 100. The ranking transforms the distribution of excess returns into a uniform distribution. Therefore, the sample mean and standard deviation of the ranks are

$$\overline{K} = \frac{1}{2} + \frac{100}{2} = 50.5$$
$$S(K) = \sqrt{\frac{1}{100}\sum (K_t - 50.5)^2} \quad (10)$$

The nonparametric rank test statistic is

$$\frac{(K_t - 50.5)}{S(K)} \quad (11)$$

The cumulative pdf of the rank test statistic is the same as the normal in this case because the sum of uniformly distributed variables converges to normal for more than twenty stocks in the portfolio. Thus the percentiles are the same as the normal distribution in the traditional event study using the t-statistic. The nonparametric rank test performs well for most thinly traded stocks as long as there is no variance increase on the event date.

6. Results

First, I estimate the simplest model which is the matched-pairs approach. The companies in the inferior and superior portfolios are “matched” since the shares belong to the same companies. Assuming they have the same risk characteristics, this effectively controls for country, industry and firm factors other than the changes in governance. Returns on inferior shares are expected to be higher than returns on superior shares around events that signal tough

rules, and returns on superior shares are expected to be higher around events that reveal rules would not be as tough as expected by the market. The events have been identified from the literature so that they convey new information to the market, while up to the event all available information has been incorporated in prices. Any updating of beliefs regarding the passage of the Act or the how tough the rules would be are reflected in the difference in abnormal returns of inferior and superior shares.

Results are presented in Table 4. Panel A lists the eight specific events and tests for the difference in returns of the inferior vs. superior portfolio. Significant differences are found for two events, E4 on July 24 (inferior return higher than superior return) and E6 on July 29-30 (superior return higher than inferior return). Testing for cumulative returns over the contiguous trading days in E2-E3 and E4-E6 finds no significant difference in Panel B. There is no difference in returns over all events either. Panel C gives the test results from wide event windows, or buy-and-hold periods. None of them shows a difference in the performance of the inferior portfolio relative to the superior portfolio. Thus the marched pair results are dependent on the width of the event window. However, this approach suffers from several weaknesses, so I turn to the other tests described in Section 5.

There are two reasons why we need to estimate abnormal reactions based on models of risk-adjusted expected return in addition to the matched-pairs approach. First, the risk characteristics of the two classes may be different. Second, there is the problem of non-synchronous trading which requires an adjustment that cannot be incorporated in the matched pairs approach. The estimated coefficients of the three models—market model, 3-factor, and market model adjusted for thin trading—are given in Table 5.

If the two classes of shares are not the same in terms of risk, mean returns should be higher for the riskier portfolio over the event study estimation period because of the risk/return trade-off. I tested if the expected difference in return is zero over the 100 trading days prior to December 28, 2001. Mean superior and inferior portfolio returns are not different as indicated by both a t-test and a Wilcoxon signed-rank test[13] in Table 5.

At the same time, the hypothesis that the coefficient of market return, or beta, in the market model for superior shares is the same as the beta in the market model for inferior shares is rejected at 5% (p-value=0.045)[14]. The significant difference in betas of the inferior and superior portfolios indicates differences in risk. There is no evidence of that based on the three-factor model though. There is an improvement in the fit of the regression since both SMB and HML are significant, but neither of the coefficients of the three factors in the inferior model are different from the corresponding coefficients in the superior model at 5%.

The lack of difference in mean returns over the estimation period and in the three-factor coefficients may be evidence that the superior and inferior portfolios have the same risk characteristics. However, the significant difference in betas in the market model as well as the evidence for asynchronous trading and the potential associated bias call for an estimation based on separate market models for the inferior and superior portfolios and an adjustment for thin trading. Indeed, the beta of the superior portfolio is always smaller than that of the inferior

[13] Wilcoxon signed-rank test is a nonparametric test for the difference between two related samples. It is the alternative to the paired t-test when the population is not normal. It involves ranking the differences between measurements from the two samples, summing the ranks separately for the positive and negative differences, and forming the test statistic on the basis of the smaller sum, which is distributed normally for n>20.

[14] The Chow test using the F-distribution can be used to test for difference in coefficients in the superior and inferior models if the observations in both portfolios were i.i.d. However, this is guaranteed not to be the case here so a pooled model with robust standard errors is estimated and a t-test is used to detect structural break in the betas.

portfolio in the market models. The last model in Table 5 includes one lead and one lag of the market return in the regression as in Scholes and Williams (1979) and Dimson (1979)[15].

Thus the model most appropriate in this case is the market model adjusting for thin trading in two aspects[16]. First, the downward bias in beta is corrected and the results are estimated using the standard t-test for abnormal returns. In addition, the nonparametric rank test is used to identify abnormal returns after the correction for bias in the beta estimates. Table 6 presents the results from the event study. There is no substantial difference in the results from the matched pairs approach and the significant events are again E4 and E6.

The nonparametric rank test statistic from Corrado (1989) is given in Table 6 in additional to the standard t-statistics. There is no change in the main findings. The test statistics are higher than the standard t-statistics except on the days with significant abnormal results which is testament to increased power under the alternative. However, they are not high enough to detect abnormal returns other than those already significant by the t-test. It seems that the distributions of the superior and inferior portfolio returns are sufficiently close to normal, satisfying the conditions for a standard t-test.

In sum, on the basis of both the matched-pairs approach and the market model adjusted for thin trading, one of the events associated with significant market reactions of inferior shares relative to superior shares is the issuance of the Conference Committee's report on July 24. Abnormal returns on inferior shares were significantly higher than abnormal returns on superior shares by 1.2% (t=2.22). The market return was 5.2%.

---

[15] CRSP also uses a three-day moving-average market window with one lead and one lag of the market return to calculate annual Scholes-Williams betas. For NYSE stocks, only trading prices are used and there must be at least half trading days in a year to produce a beta for the stock. For NASDAQ stocks, both trading prices and bid-ask averages are used since using only trading prices would severely limit the sample (CRSP Stock and Indices Databases Data Description Guide, Version 200901, p. 111).

[16] It is possible to estimate the 3-factor model with Dimson betas. However, the coefficients of the leads and lags of SMB and HML are insignificant, and the implications are not substantially different from the market model.

The interpretation of this event in the literature is conflicting. Li et al (2008) claim that it represented the first unambiguous indication that the more demanding reforms of the proposed bills would form the core of SOX. They find market returns associated with this event of 5.4%, which is the incremental return related to the event over the average market return for 234 nonevent days in 2002. News about contemporaneous events is unfavorable so the positive reaction cannot be due to confounding factors. Thus they find evidence supporting the hypothesis that SOX improve corporate governance. Zhang (2007) compares returns of the US market to foreign markets and also finds a positive abnormal return of the US market associated with this event of 4.5%. However, the interpretation of the positive reaction in Zhang is not that there are governance benefits, but that the market realized that the costs would be lower than already expected.

Thus both Li et al (2008) and Zhang (2007) find a positive market reaction to the Conference Committee's report on July 24. Li attributes it to more-beneficial-than-expected reforms, while Zhang attributes it to lower-than-expected costs. The source of this conflicting interpretation is that neither study has a research design able to distinguish between expected costs from expected benefits. In my study, only the expected governance effect is reflected in the differential return on inferior and superior shares since costs and even other provisions of the Act would affect both classes. Therefore the higher return on inferior shares is consistent with the interpretation in Li et al (2008) and shows market expectations of improved corporate governance associated with the event on July 24.

In addition to the issuance of the Conference Committee's report, another major governance-related event on July 24 was the arrest of Adelphia executives, the founder and his two sons, around 6 a.m. on July 24, for governance violations such as major off-balance sheet

loans to the controlling family and practically looting the company. Adelphia was in fact a dual-class firm where the family controlled the company through its ownership of super-voting class of stock (65% of votes) but had a minority of the cash-flow rights (11%). The Conference Committee reached an agreement in the morning and the content of the conference report was revealed during the day, although the exact timing is not available. The arrest signaled a change of attitude of the government and an increase in the likelihood that a tough rule would pass. Therefore this event likely reinforced the overall expectation of substantial governance reform, in some ways probably even more than the issuance of the Conference Committee report.

The second significant event is July 29-30 when abnormal returns on *superior* shares were significantly higher than abnormal returns on inferior shares by 2.3% (t=-2.93). The market return was 5.8%. This event included the announcement by the SEC that it would publicly identify CEOs who fail to certify financial reports and the signing of the bill into law. The final uncertainty regarding President Bush signing the law is resolved prior to the actual signing on July 30 during a speech he delivered on July 29 on welfare reform in which he explicitly indicated that he would sign the accounting and corporate governance reform bill into law. The fact that abnormal returns on inferior shares were lower than superior shares may be consistent with investors expecting tougher governance rules than what was finally announced. However, even if some uncertainty regarding the passage of the law was resolved, the nature of the rules was known to the market already. Thus it is possible that the 2-day excess return is spurious and that wider event windows may be more reliable.

Testing over contiguous trading days and all individual SOX events shows no significant difference in cumulative returns on inferior and superior shares during the legislative period. As already discussed, different interpretations as to whether the information release on the chosen

event dates increased or decreased the likelihood the law would be passed in previous studies can be avoided with a wider event window. Treating the whole period from June 25 to July 30 as one event, or buy-and-hold period, does not show different reactions of inferior vs. superior shares. In fact, the results from both wide buy-and-hold periods are essentially the same as those from the matched-pairs event study, with no significant difference in abnormal returns.

Results are also robust when the traditional market model and the three-factor model are the benchmarks for calculating excess returns, as shown in Table 7. According to Brown and Warner (1985), excess returns based on the returns of thinly traded stocks without the correction would be biased. However, if the average bias in the portfolio of securities is zero, the event study methodology is not necessarily misspecified. This can happen of the securities in the portfolio are representative of a wide range of trading frequencies. This seems to be the case here since excess returns are calculated in portfolios of superior and inferior shares, so that after the correction, there is no substantial difference in results.

7. Conclusion

This chapter analyses differences in returns on inferior and superior shares of dual-class firms in response to SOX events that reveal new information to the market about the passage of the Act. The analysis assumes efficient financial markets and a voting premium reflected in the prices of inferior and superior shares, so that any identified price reactions reflect the market's expectation of the effects of SOX. Strict governance rules are expected to reduce diversion of resources by corporate insiders at the expense of outside shareholders, and hence the value of voting control. A decrease in the value of control is in turn reflected in better performance of inferior shares relative to superior shares around key SOX dates. No other factors should affect

this differential since compliance costs and macroeconomic factors affect the returns on both classes of shares of the firm in the same way. The model adjusts for non-synchronous trading in the beta estimates and includes a nonparametric test to allow for non-normality in returns. Results show that inferior shares did not perform better than superior shares overall, indicating that SOX was ineffective in reducing private benefits. There was a significant market reaction to the arrest of Adelphia executives and issuance of the Conference Committee's report on July 24, which signaled that tough rules would be implemented. Inferior shares gained on superior shares, but the reaction was reversed at the final resolution of uncertainty and passage of the law on July 29-30. There is no substantial change in this conclusion when a traditional market model or a three-factor model is used. It is possible that the lack of difference in returns may be due to small value of control when both inferior and superior classes trade publicly as opposed to the general case of only inferior traded class. However, the main difference between these types of dual-class firms appears to be in terms of lower growth prospects for firms with two traded classes, rather than in terms of private benefits. The event study results are conditional on the assumption of efficient markets. I have considered wider event windows to address the fact that markets may take time to digest the implications of SOX. Although again no significant effects are found, the wider event windows make it difficult to control for other events and identify reactions specifically to news about SOX. Overall, there is no evidence of improved performance of inferior shares over superior shares of dual-class companies in the market's expectation of the effects of SOX.

Tables and Figures

Table 1
Events relating to the passage of SOX and SEC exchange-listing requirements

Panel A: Event history

Event	Event window	Description	Li et al (2008)	Zhang (2007)
E1	June 25-27	WorldCom admitted to understating expenses by $3.8 billion; SEC filed suit at the WorldCom announcement of fraud; Introduction of Sarbanes bill in Senate; SEC requiring CEO/CFO certifications of financial reports	E1&E2	E13
E2	July 8-12	Senate consideration of Sarbanes bill; Bush delivered speech on corporate reforms	E3	E14
E3	July 15-17	Senate passage of Sarbanes bill and House deliberation and passage of Oxley bill	E4	E15
E4	July 24	Issuance of House-Senate Conference Committee report; Arrest of Adelphia executives	E5	E17
E5	July 25-26	House and Senate Passage of Conference Committee's report; Congress passed SOX	E6	E17
E6	July 29-30	SEC to post names of CEO/CFOs who fail to certify financial reports; Bush signs SOX into law July 30	E7	
E7	August 14-15	Due date of CEO/CFO certifications with SEC	E8	
E8	August 16	NYSE filed proposed rules with SEC regarding exchange listing requirements		

Panel B: Buy-and-hold periods

B&H1	6/25/02-7/30/02	Wide event window to capture the most important rule-making activites until the passage of SOX
B&H2	11/01/01-10/31/02	All SOX-related events as in Chhaocharia&Grinstein (2007); Enron restated financial results in November 2001; NYSE and similar NASDAQ listing rules were approved by the SEC in October 2002

Table 2
Characteristics of firms with two traded classes of shares

Panel A: Number of firms in the merged dual-class/Compustat data by year of proxy record date

	2001	2002
Number of dual-class firms	397	352
Only inferior class trades	284	249
Only superior class trades	45	43
Superior and inferior classes trade	59	52
Two inferior classes trade	4	5

Panel B: Characteristics of firms in 2002

	Superior and inferior classes trade	Only inferior class trades	Difference (t-stat)
Asset size (in $millions)	5,910	3,276	2,635 (0.92)
Age of firm (years since first CRSP date)	25	13	12* (6.33)
Sales growth	2.94%	16.41%	13.47%* (2.45)
Capex as a ratio of sales	0.08	0.09	0.01 (0.15)
Leverage as ratio of debt to assets	0.2	0.23	0.03 (0.73)
CEO compensation (in $thousands)	2,000	1,575	426 (0.96)
Cash-flow ownership of insiders	31%	38%	7%* (2.08)
Vote ownership of insiders	49%	60%	11%* (2.50)
Wedge	20%	23%	3% (1.11)
Founder's name in company name	0.39	0.26	0.13* (2.01)
Idiosyncratic risk	0.0295	0.0406	0.0111* (3.02)

Panel C: Examples of firms

Superior and inferior classes trade	Only inferior class trades
Readers Digest	American Greetings
Continental Airlines	Adelphia
	Coca Cola
	Dow Jones
	Ford Motor
	Hershey Foods
	Nabisco
	New York Times
	Revlon
	Scholastic
	Washington Post

Accounting statistics are measured as of fiscal year end in 2002. Source: Compustat
Inside ownership is measured as of the proxy record date. Source: Dual-class data from Gompers et al (2008)
CEO compmensation represents salary and bonus for the top executive from Execucomp.
Idiosyncratic risk is measured as RMSE of a regression of inferior return on market return during 2001 and 2002.
*Indicates significance at 5%

Table 3
Non-synchronous trading in the event study estimation period 08/01/2001-12/27/2001

Panel A: Differences between superior and inferior classes

	Superior class	Inferior class
Shares outstanding as % of total shares	32%	68%
Market value of shares as % of total equity	34%	66%
Average daily trading volume	113,412	278,341

Panel B: Trading statistics in firm-trading days

	Superior class	Inferior class
Total number of firm-trading days	5815 (100%)	5815 (100%)
Number of days when price based on last transaction	3953 (68%)	5270 (90%)
Number of days when price based on bid-ask average quote	1862 (32%)	545 (10%)
Number of zero-volume trading days	1858 (32%)	542 (10%)
Number of zero-volume, zero-return trading days	938 (16%)	437 (8%)

Total shares outstanding and market value of equity are based on all existing classes of shares, which in this sample are always two per firm.
There are 59 dual-class firms trading during all or some part of the estimation period.
Trading status is never halted, suspended or unknown in this sample: either actual trading price or bid-ask average are available in the price field.
On NASDAQ, volumes of after-hours trades are included in the current day, while the price is quoted the next day. Therefore, it is possible to have bid/ask quotes paired with nonzero volumes.

Table 4
Matched pairs event study

Panel A: Testing over individual SOX events

Event	Event Window	Number of firms	Inferior Portfolio Return	Superior Porfolio Return	Difference	t-stat
E1	June 25-27	56	0.0168	0.0075	0.0093	0.94
E2	July 8-12	56	-0.0323	-0.0263	-0.0060	-0.27
E3	July 15-17	56	-0.0172	-0.0167	-0.0006	-0.06
E4	July 24	56	0.0255	0.0070	0.0185	3.23*
E5	July 25-26	56	0.0152	0.0213	-0.0061	-0.76
E6	July 29-30	56	0.0279	0.0421	-0.0142	-1.75*
E7	August 14-15	54	0.0392	0.0314	0.0078	0.97
E8	August 16	54	0.0014	0.0006	0.0008	0.14
Panel B: Cumulative returns						
E2-E3	July 8-17	56	-0.0495	-0.0431	-0.0065	-0.87
E4-E6	July 24-30	56	0.1117	0.1338	0.0222	0.89
	All events	56	0.0284	0.0381	0.0096	0.38
Panel C: Testing over wide event windows (buy-and-hold periods)						
B&H1	6/25/02-7/30/02	56	-0.0286	-0.0439	0.0172	0.63
B&H2	11/01/01-10/31/02	59	0.1569	0.1515	0.0054	0.06

Inferior portfolio return is the equally-weighted return of inferior shares, and superior portfolio return is the equally-weighted return of matching superior shares of the same company.

* indicates that the difference is positive or negative at significance level of 5% (one-tailed critical |t|=1.645).

Table 5
Market models of superior and inferior portfolios

	Market model		Three-factor model		Dimson betas	
	Superior Ret	Inferior Ret	Superior Ret	Inferior Ret	Superior Ret	Inferior Ret
Market Return	0.492***	0.634***	0.536***	0.693***	0.480***	0.615***
	(0.050)	(0.042)	(0.052)	(0.039)	(0.052)	(0.427)
Market Return_t-1					0.059	0.067
					(0.051)	(0.042)
Market Return_t+1					0.02	0.051
					(0.051)	(0.042)
SMB			0.515***	0.551***		
			(0.115)	(0.087)		
HML			0.280**	0.366***		
			(0.125)	(0.095)		
Constant	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	100	100	100	100	100	100
R-squared	0.49	0.7	0.58	0.79	0.49	0.7

t-test of difference in mean superior and inferior portfolio returns: t-stat=0.39 (p-value=0.69)
Wilcoxon signed-rank test of difference in mean returns: z=-0.007 (p-value=0.99)

Market return is the daily value-weighted return on all NYSE, AMEX and NASDAQ stocks.
The benchmark factor SMB summarizes the performance of small versus big stocks.
HML summarizes the performance of value stocks relative to growth stocks.

Estimation period: 100 trading days prior to Dec. 28, 2001
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6
Excess returns based on market model adjusted for thin trading

Panel A: Testing over individual SOX events

Event	Event Window	Market Return	Abn. Return Inferior Portfolio	Abn. Return Superior Porfolio	Difference	t-stat	Nonparametric rank test stat
E1	June 25-27	-0.0023	0.0164	0.0066	0.0098	1.05	1.61
E2	July 8-12	-0.0638	0.0078	0.0022	0.0055	0.46	1.30
E3	July 15-17	-0.0132	-0.0080	-0.0110	0.0030	0.32	0.85
E4	July 24	0.0524	-0.0052	-0.0172	0.0120	2.22*	1.75*
E5	July 25-26	0.0091	0.0015	0.0110	-0.0095	-1.23	-1.51
E6	July 29-30	0.0579	-0.0142	0.0083	-0.0225	-2.93*	-1.68*
E7	August 14-15	0.0472	0.0070	0.0057	0.0013	0.17	0.49
E8	August 16	0.0001	-0.0016	-0.0018	0.0001	0.02	0.10

Excess returns are calculated based on a market model with one lead and one lag of the market return.
The nonparmetric rank test detects abnormal returns without assuming normal or symmetric distribution.
* indicates that the difference is positive or negative at significance level of 5% (one-tailed citical |t|=1.645).

Table 7
Traditional event study

Panel A: Market model

Event	Event Window	Market Return	Abn. Return Inferior Portfolio	Abn. Return Superior Porfolio	Difference	t-stat
E1	June 25-27	-0.0023	0.0162	0.0060	0.0102	1.08
E2	July 8-12	-0.0638	0.0046	0.0006	0.0040	0.33
E3	July 15-17	-0.0132	-0.0110	-0.0129	0.0019	0.20
E4	July 24	0.0524	-0.0084	-0.0197	0.0113	2.08*
E5	July 25-26	0.0091	0.0080	0.0150	-0.0070	-0.92
E6	July 29-30	0.0579	-0.0102	0.0118	-0.0220	-2.86*
E7	August 14-15	0.0472	0.0079	0.0063	0.0016	0.20
E8	August 16	0.0001	0.0002	-0.0006	0.0009	0.16

Panel B: Three-factor model

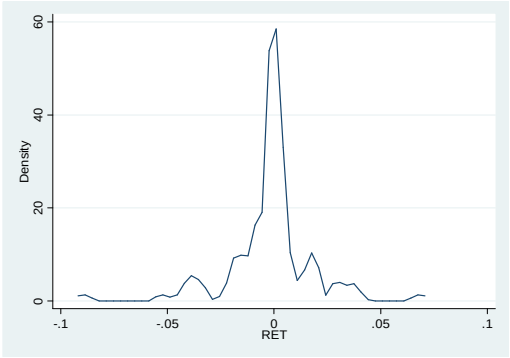
Event	Event Window	Market Return	SMB	HML	Abn. Return Inferior Portfolio	Abn. Return Superior Porfolio	Difference	t-stat
E1	June 25-27	-0.0023	0.0170	-0.0178	0.0144	0.0032	0.0112	1.19
E2	July 8-12	-0.0638	0.0121	-0.0008	0.0035	-0.0012	0.0047	0.39
E3	July 15-17	-0.0132	0.0077	-0.0154	-0.0079	-0.0111	0.0031	0.33
E4	July 24	0.0524	-0.0171	-0.0164	0.0042	-0.0083	0.0125	2.31*
E5	July 25-26	0.0091	-0.0048	0.0062	0.0084	0.0159	-0.0075	-0.98
E6	July 29-30	0.0579	-0.0155	-0.0041	-0.0030	0.0190	-0.0220	-2.87*
E7	August 14-15	0.0472	-0.0181	-0.0114	0.0198	0.0174	0.0024	0.32
E8	August 16	0.0001	0.0124	-0.0030	-0.0052	-0.0059	0.0007	0.13

Abnormal returns are calculated from a model regressing portfolio return on market return in Panel A,
and from a model regressing portfolio return on market return, SMB and HML in Panel B.
* indicates that the difference is positive or negative at significance level of 5% (one-tailed citical |t|=1.645).

Figure 1
Normality of returns in the 100-day event study estimation period 08/01/2001-12/27/2001

One random company

Fig. 1a Superior class

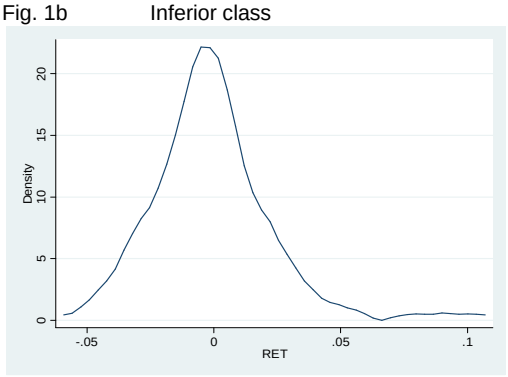


Mean	-0.0017	Skewness	-0.6496
Std. Dev.	0.0194	Kurtosis	8.0902

Skewness/Kurtosis tests for Normality

Pr(Skew.)	Pr(Kurt.)	adj chi2(2)	Pr>chi2
0.009	0.000	20.12	0.000

Fig. 1b Inferior class



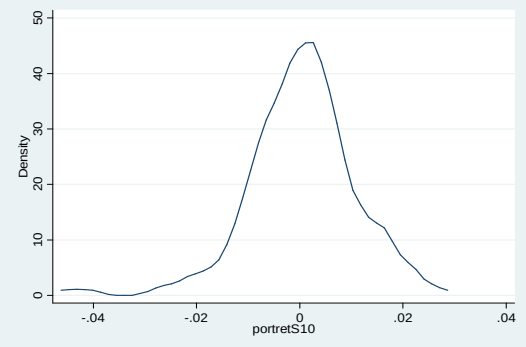
Mean	-0.0009	Skewness	1.1731
Std. Dev.	0.0234	Kurtosis	6.5991

Skewness/Kurtosis tests for Normality

Pr(Skew.)	Pr(Kurt.)	adj chi2(2)	Pr>chi2
0.000	0.000	24.48	0.000

Equally-weighted portfolio of 10 random companies

Fig. 1c Superior portfolio

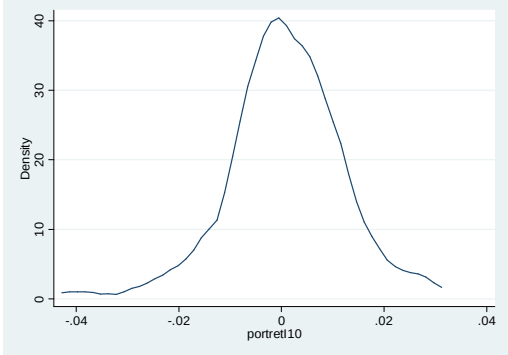


Mean	0.0004	Skewness	-0.6770
Std. Dev.	0.0103	Kurtosis	5.6225

Skewness/Kurtosis tests for Normality

Pr(Skew.)	Pr(Kurt.)	adj chi2(2)	Pr>chi2
0.007	0.001	14.61	0.007

Fig. 1d Inferior porfolio



Mean	0.0008	Skewness	-0.3800
Std. Dev.	0.0111	Kurtosis	4.4296

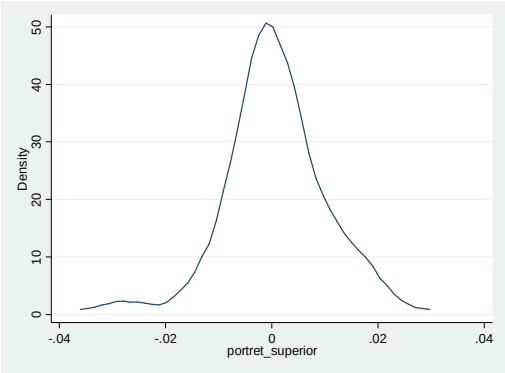
Skewness/Kurtosis tests for Normality

Pr(Skew.)	Pr(Kurt.)	adj chi2(2)	Pr>chi2
0.108	0.019	7.4	0.025

Figure 1, continued

All 59 dual-class companies in the sample

Fig. 1e Superior portfolio

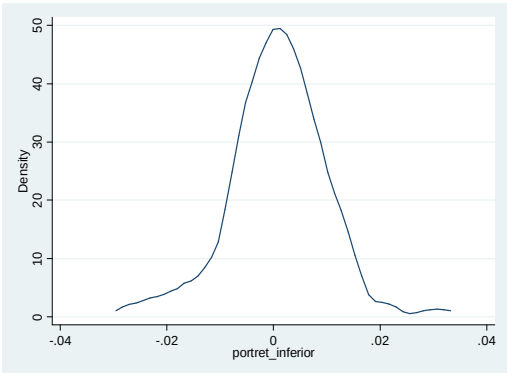


Mean	0.0006	Skewness	-0.3302
Std. Dev.	0.0097	Kurtosis	4.5136

Skewness/Kurtosis tests for Normality

Pr(Skew.)	Pr(Kurt.)	adj chi2(2)	Pr>chi2
0.161	0.015	7.22	0.027

Fig. 1f Inferior portfolio



Mean	0.0008	Skewness	-0.1793
Std. Dev.	0.0090	Kurtosis	4.4006

Skewness/Kurtosis tests for Normality

Pr(Skew.)	Pr(Kurt.)	adj chi2(2)	Pr>chi2
0.439	0.021	5.78	0.066

CHAPTER 4. DOES SARBANES-OXLEY AFFECT THE DECISION OF DUAL-CLASS FIRMS TO RECAPITALIZE INTO SINGLE-CLASS

1. Introduction

An important category in the provisions of Sarbanes-Oxley (SOX) concerns corporate governance, which can be studied in the context of agency theory. There may be a conflict of interest between managers and outside shareholders due to the separation of ownership and control in a corporation. This agency problem includes misalignment of incentives between managers and shareholders, as well as extraction of private benefits by corporate insiders. The governance provisions of SOX are meant to improve incentive alignment and limit extraction of private benefits. Since private benefits are an important factor in the choice of dual-class structure, this chapter examines the effect of SOX on the decision of dual-class firms to recapitalize into single-class.

Dual-class shares allow for disproportionate dividend and voting rights for insiders. Cash-flow ownership by managers helps align their interests with those of outside equity holders by providing an incentive to maximize effort. At the same time, voting control by managers reduces the possibility of takeovers if share value is not maximized under current management and allows insiders to extract private benefits. Therefore a dual-class structure where insiders' cash-flow rights are disproportionately smaller than voting rights creates the potential for a large agency problem.

On the other hand, there are reasons why a dual-class structure may be an efficient choice by the firm. The voting control of insiders holding super-voting stock can serve as protection from value-reducing takeovers in the presence of imperfect information in the market for corporate control. In the case of a growing firm which may be undervalued in the market, this protection helps realize growth options. In addition, the voting control of managers encourages

their investment in firm-specific human capital. Since control does insulate managers from the disciplining force of the market for corporate control, there may be counterbalancing governance mechanisms such as strong board of directors. The SOX legislation is another such substitute governance mechanism that may help resolve agency problems at dual-class firms.

If SOX is effective in limiting the extraction of private benefits, we can expect dual-class firms where extraction of private benefits by insiders is the dominant reason for dual-class status to recapitalize into single-class. If the ability of insiders to extract private benefits is limited due to SOX, they would value control less and may want to divest their holdings. Becoming single-class would be advantageous for two reasons. First, a single-class recapitalization eliminates disproportionate holding of cash-flow and voting rights, which may increase security value. This will happen in the case when insiders' private benefits come at a cost to outside shareholders in terms of security value after SOX, i.e. in the case of inefficient private benefits above and beyond the reduction in private benefit due to SOX. Second, superior shares are usually not traded, or traded very thinly, and selling ordinary shares would allow insiders to monetize holdings gradually and in a more liquid market.

At the same time, if dual-class structure is beneficial for outside shareholders because it adds to security value, the fact that SOX serves as substitute governance mechanism may cause firms to remain dual-class longer. If the protection from value-reducing takeovers provided by insiders' voting control is important for realization of growth options, the fact that shareholders need to worry less about extraction of private benefits by insiders means dual-class firms are less likely to recapitalize into single-class. Thus if SOX is effective in reducing private benefits while dual-class status is an optimal choice in terms of shareholder value, we can expect fewer recapitalizations into single-class after SOX.

Therefore the effect of SOX on the decision of dual-class firms to recapitalize into single-class is an open empirical question. It can be tested by modeling the time it takes a dual-class company to recapitalize into single-class and the effect of SOX on this duration. The empirical model to examine recapitalization decisions of dual-class firm is based on survival analysis. Results indicate that dual-class firms are more likely to recapitalize into single-class after SOX.

2. The effect of SOX on single-class recapitalizations of dual-class firms

A single-class recap may be induced by a change in any of the factors that are found important for the choice of dual-class status. Since Chapter 1 describes these theoretical factors in detail, the purpose of this section is to highlight the variables which have been shown to be important empirically and how they relate to the theoretical factors identified before.

2.1. Empirical studies of recapitalization decisions

The general factors involved in dual-class governance models identified in Chapter 1 are again private benefits, protection from value-reducing takeovers and realization of growth options, diversification of risk and divestment, and firm-specific human capital. Dual-class status can be chosen at IPO or the company can recapitalize into dual-class status post-IPO. An existing dual-class firm can subsequently recapitalize into single-class. Explaining the choice of dual-class structure and modeling transitions of status involves identifying empirically measurable factors that influence these decisions.

Amoako-Adu and Smith (2001) formulate several observable factors that may influence the choice to become dual-class at IPO of Canadian firms. These are the amount of votes of the largest shareholder, whether it is a family firm, size, age, and whether the largest shareholder

was a controlling shareholder of a previously listed dual-class firm. The first two are related to the benefits of voting control and private benefits. The last one relates to the preference for control of the largest shareholder and also their familiarity with the use of dual-class shares.

Age controls for firm-specific human capital of the controlling manager: the longer the time since establishment, the more the founder has developed skills specific to the business. According to Taylor and Whittred (1998), the most noticeable aspect among companies utilizing dual-class shares is the extent to which firm value is dependent on skills specific to the founder. They study Australian firms at IPO and observe that the value of growth opportunities of dual-class firms is highly dependent on the human capital of the founding shareholders.

Morck et al (1988) find that large size deters hostile takeovers. Thus, as a takeover defense, dual-class capitalization may be more important for smaller firms. Consistent with this argument, Taylor and Whittred (1998) find that Australian dual-class firms tend to be smaller at IPO than their single-class counterparts. Post-IPO, Gompers et al (2008) find no significant difference in the size of dual-class versus single-class firms in the US.

As already discussed in theoretical factors for dual-class status, firm growth is important for the choice of dual-class status. Dual-class stock serves as a temporary device against value-reducing hostile takeovers. Empirically, firms with growth opportunities are more likely to choose dual-class structure at IPO [Taylor and Whittred (1998)]. High-growth firms are also more likely to recapitalize into dual-class post-IPO [Lehn et al (1990)]. The anti-takeover protection of dual-class stock is put in place to protect a firm expecting growth in the future but temporarily undervalued in the market because of asymmetric information [Dimitrov and Jain (2006)].

There have been fewer studies of dual-class companies recapitalizing into single-class. Amoako-Adu and Smith (2001) study all dual and single-class firms that went public over the period 1983-1998 on the Toronto Stock Exchange. They follow the dual-class firms post-IPO and observe the reasons why some of them reclassify into single-class. Information was gathered from proxy statements and press releases. Approximately half of the reclassifications were attributable to three reasons: meet the terms of a debt restructuring agreement, facilitate the sale of control block, and increase institutional appeal for stock prior to seasoned offering. In the case when financial distress led to a debt restructuring plan, the control by the largest shareholder was normally lost and there was a need to reduce the leverage of the firm by issuing new equity. To make this equity more attractive to outside investors, a single-class structure was established. Similarly, in the cases where a controlling shareholder sold a block of shares, reclassification to a single class was used to increase the appeal of the shares to outside investors. In the case of seasoned equity offerings, greater liquidity was a desirable feature for investors, particularly institutions. Other reasons for reclassification into single class of equity include increasing institutional investor interest even in the absence of a seasoned offering, adoption of shareholder rights plan, the scheduled expiry of the dual-class structure as established by prior shareholder's agreement, and a plan to list on US exchanges where interest in dual-class shares was lower than in Canada.

While collecting data for my statistical analysis, I reviewed the SEC filings of about 550 dual-class companies and found evidence of recapitalization into single-class in 176 cases. Although I did not follow up by reading press releases, the reasons for single-class recapitalization, whenever given in the actual proxy, were consistent with those above. The method of recapitalization most often is the exchange method. The superior shares are

exchangeable into inferior shares and the company can redeem all of the outstanding superior shares. The elimination of the superior class is particularly easy if it is held only among a few individuals. Less often the Articles of Incorporation may require shareholder approval of the recapitalization. The immediate change of voting power of the largest shareholder usually is not enough to lose the control position, but further divesting is expected gradually over time. In some cases, compensation is given to the holders of the superior class.

Exhibit 1 contains excerpts from an actual proxy statement announcing the single-class recap of Commonwealth Tel Co (CTE) in 2003. Level 3 Communications is a company that is the largest shareholder in CTE and owns superior voting stock allowing 15 votes per share. Level 3 has voting power of 29%, with which it effectively controls CTE, while cash-flow rights are 4%. However, Level 3 wishes to monetize the position which entails giving up a large control block. After the single-class recap, Level 3 would own about 5% of the common shares. Although the exact reasons for such change in control are unclear from the proxy, given this decision, the single-class recap is advantageous for Level 3 and for the company in several ways. It increases the liquidity and trading volume of the common stock. It aligns cash-flow and voting rights so that voting influence is proportionate to economic interest. Better liquidity attracts institutional investors, which is also beneficial in the case of a future seasoned equity offering. Finally, if there is a control change, the single-class structure would allow for the takeover premium to accrue to all shareholders equally.

It is interesting to note that one of the reasons for the single-class recap is that the new capital structure would conform to that of most other publicly traded corporations. There is a “perceived” negative impact on stock price due to having two classes of traded stock with generally similar characteristics except voting rights. This perception that dual-class structure is

not very shareholder-friendly is mentioned in Dimitrov and Jain (2006), who talk about the "stigma" assigned to dual-class firms because of perceived high agency costs. They speculate this attitude is changing in recent years, due to dual-class companies like Berkshire Hathaway and Google.

2.2. The effect of SOX on recapitalization decisions

In terms of corporate governance benefits, SOX aims to better align incentives between managers and shareholders and to limit the extraction of private benefits by insiders [Coates (2007)]. Thus we can expect that if SOX was effective in improving corporate governance by aligning incentives between managers and shareholders, dual-class firms harboring high private benefits of insiders will recapitalize into single-class. If the new rules prevent insiders from extracting private benefit, the cost of holding undiversified share of the company equity becomes relatively higher so insiders may want to divest their position. A recapitalization into single-class would be advantageous for reasons such as more liquid market as well as better value for shares when economic interest and voting influence become aligned at firms with inefficient private benefits of insiders. Exhibit 2 is an example of a proxy where a recapitalization because of SOX is explicitly discussed, as well as the compliance requirement for independent audit committee of the board.

Another reason to recapitalize is SOX-related compliance costs, which can be seen as a corollary to private benefits. Some of the exchange-listing requirements in the regulatory application of SOX include independent audit committee of the board, majority of independent directors, and independent nominating and compensating committees. A survey in Block (2004) quotes as one of the reasons for the higher cost of being public after SOX the higher fees paid to

outside directors. Based on this and the perception that dual-class structure is not very shareholder-friendly in Dimitrov and Jain (2006), we can speculate that outside directors' compensation would be particularly high for dual-class firms although a formal test of this is left for future research. Dual-class companies may decide to reclassify into singe-class if some of these costs can be alleviated. There is evidence that compliance costs drove some public companies, especially small ones, to go private [Kamar et al (2008)]. There has not been a study of the effect of SOX on dual-class decisions.

On the other hand, SOX may serve as substitute governance mechanism at firms where dual-class status is optimal in terms of shareholder value. As already discussed, the reasons include protection from value-reducing hostile takeovers, realization of growth options and investment in firm-specific human capital. Research shows that regulation can serve as substitute governance mechanism. For example, La Porta et al (2002) find that the effect of managerial cash-flow ownership as an incentive device differs across countries depending on their corporate governance laws. Thus we can hypothesize that if SOX acts as counterbalancing governance, firms where dual-class status is efficient will remain dual-class longer because of the reduced extraction of benefits by insiders at the expense of outside shareholders. The ultimate effect of SOX on the decision of dual-class firms to recapitalize into single-class is an open empirical question, which can be examined in the framework of survival analysis.

3. Empirical model

3.1. Survival analysis

A natural first step in formulating an empirical model in the present case is to think of OLS where the dependent variable is time. Let T be the time it takes a dual-class firm to

recapitalize into single-class, or the difference between the single-class date and the dual-class date. If we want to know how this duration is explained by a set of variables X, we could formulate the following regression model:

$$\log T = X\beta + e \quad (1)$$

The log formulation may be appropriate because the distribution of durations is usually skewed. However, T is based only on firms that undergo recap into single-class, while the remaining dual-class firms are ignored. The problem when episode length is unknown exactly is defined as "censoring." With right censoring, at the time of observation the relevant event has not yet occurred, which is the problem here [Garson (2009)]. Observing duration only for firms that did recapitalize during a given observation period may bias the estimate of duration spent in dual-class status. The reason is that we only observe the "short durations," or the firms that recapitalize relatively quickly, while the ones that take longer to become single-class are not observed. The longer the observation period, the more "long durations" we observe, but if there are any remaining dual-class firms, their single-class date is still unknown. This idea is discussed further and illustrated in Section 4.1.

Thus OLS is not appropriate because it cannot handle censoring leading to severe bias and loss of information as the censored observations are thrown away [Jenkins (2005)]. The second problem with OLS is that if X includes time-varying covariates, their values change during the duration T so that the current values are misleading.

Survival analysis (or event history analysis) involves modeling duration, where data units move along a finite series of states [Kiefer(1988)]. In this case there are two possible states--dual-class and single-class status. The time between the dual-class capitalization and the single-

class recapitalization is the episode duration. The “failure” event is the recapitalization into single-class, which may or may not occur.

Duration times are modeled via the hazard rate, which is the chance of making transition out of the current single-class state at each time period conditional on survival as a dual-class firm up to that point. Let time in dual-class state, or episode duration, be the realization of a continuous random variable T with c.d.f. F(t) and p.d.f. f(t). In continuous time, f(t) is the slope of F(t),

$$f(t) = \frac{\partial F(t)}{\partial t} \tag{2}$$

F(t) is called the failure function, while S(t)=1-F(t) is the survivor function. In discrete time, the p.d.f. for t=5, for example, can simply be interpreted as the proportion of the sample that recapitalizes into single-class after 5 years. The failure function would be the proportion of the sample that recapitalizes in less than 5 years, while the survivor function is the proportion of sample with durations over 5 years. The hazard rate is defined as the instantaneous rate of leaving the state at duration t given it was not left before duration t [Petersen (1986)].

$$h(t) = \frac{f(t)}{S(t)} \tag{3}$$

In discrete time, it can be interpreted as the conditional probability of having an episode length of exactly t conditional on survival up to time t. In other words, it is the number of firms that recapitalize in 5 years as a proportion of the sample of firms with durations of 5 or more years. There is a one-to-one relationship between the hazard, failure, and survivor functions.

$$S(t) = \exp(-\int_0^t h(u)du) \tag{4}$$

The shape of the hazard function is empirically relevant for examining the duration dependence and the effect of covariates on time to recapitalization. To account for differences in hazard rates between firms, a functional form of the hazard rate that depends on some covariates is formulated. In general, the hazard rate can be expressed as

$$h(t)=h(t, X(t)) \tag{5}$$

The coefficient estimates are chosen to maximize the probability of observing what has been observed in a procedure called maximum likelihood estimation (ML). We need to express the probability of observed data as a function of unknown coefficients and then maximize the joint probability which is the likelihood function [Allison (1984)]. The likelihood function to be maximized in the estimation of the parameters is

$$L = \prod_{i=1}^{N} f(T_i) \tag{6}$$

With K right-censored observations, at time of observation the relevant event has not yet occurred so that the episode has length T>t. The likelihood function in that case can be written as

$$L = \prod_{j=1}^{J} f(T_i) \prod_{k=1}^{K} S(T_k) \tag{7}$$

Thus uncensored cases define the failure density function. Both censored and uncensored cases define the survival function and in this way information from the censored cases is also used.

Explanatory variables whose values change over the course of episodes are conceptually straightforward to handle in the hazard function framework [Jenkins (2005)]. Two types of time-varying covariates can be distinguished. One type stays constant for finite sub-periods of time. The other type changes all the time [Petersen (1986)]. In the current context, all financial and ownership variables are of the first type, i.e. they are reported as fixed for certain periods of time such as fiscal year. In the estimation, it is not necessary to integrate over the entire path of the

covariate. For each interval where covariate stayed constant, a separate record is created and treated as a separate observation (episode splitting). The model is then estimated using ML provided the survivor function has a closed form as in parametric models.

3.2. Duration models

Continuous parametric models specify a closed-form function of the hazard rate in (5), which has convenient mathematical properties. Semi-parametric methods (Cox model) do not specify the shape of the hazard function with respect to time, i.e. duration dependence, but only estimate the effect of covariates. Nonparametric methods are estimated based on empirical data but such methods cannot handle multivariate controls. I am going to consider two popular parametric models, and the Cox model. Choosing which model best describes the data involves estimating them and evaluating their predictions and fit.

In the class of parametric models, there are proportional hazards (PH) and accelerated failure time (AFT) models. In PH, the effect of covariates is that they scale the hazard function. For example, dual-class firms with high private benefits of insiders are expected to take longer to recapitalize into single-class. In PH, the baseline hazard function describing duration dependence is scaled by a proportionality factor which does not depend on time. The PH assumption means that a 10-year-old high-private benefit firm is expected to recapitalize slower than a 10-year-old low-private benefit firm, and the scale factor is the same if the firms were 2-years-old.

The Weibull model is a popular PH model. It allows for the hazard rate to increase or decrease with time. The Weibull model can be expressed as

$$\begin{aligned} h(t, X) &= pt^{p-1}\exp(X\beta) \\ h(t, X) &= h_o(t)\exp(X\beta) \end{aligned} \qquad (7)$$

The shape parameter p answers the question whether risk is constant over time. If p>1 the hazard rate rises monotonically with time (positive duration dependence), if p<1 it falls with time (negative duration dependence), and if p=1it is constant (special case of exponential model).

Thus the baseline hazard function $h_0(t)$ is separable from and scaled by $\exp(X\beta)$, which is the essential characteristic of PH models. X contains the covariates affecting the hazard rate. The marginal effect of the covariates is given by

$$\beta_k = \frac{\partial \log h(t, X)}{\partial X_k} \tag{8}$$

The coefficients do not depend on t and summarize the proportional effect on the hazard function of an absolute change in a covariate [Kalter (2008)].

The hazard ratio in PH models is the ratio between the hazard rates of two groups if the variable is categorical, or the ratio between the hazard rates for 1 unit increase in a continuous covariate. For example, the hazard ratio for a dummy independent variable X_1 is defined as

$$\frac{h(t, X_1 = 1, X_2, \ldots)}{h(t, X_1 = 0, X_2, \ldots)} = \exp(b_1) \tag{9}$$

In AFT models, the effect of covariates is to scale survival time itself, i.e. multiply the time scale. Thus a proportional change in an explanatory variable would lead to a proportional change in survival time. In other words, the single-class recapitalization clock ticks faster for firms with low private benefits of insiders relative to firms with high private benefits. Thus the parameter estimates for AFT and PH models have opposite signs.

The AFT model can be expressed in regression form as

$$\log T = X\beta + e \tag{10}$$

which is the same as (1). If there are only fixed covariates and no censoring, AFT models can be consistently estimated by OLS [Allison (1984)]. Otherwise, ML estimation must be used.

Assuming a normal distribution of e gives rise to a lognormal distribution of T, which is the AFT lognormal model.

The estimated coefficients in AFT models are

$$\beta_k = \frac{\partial \log(T)}{\partial X_k} \tag{11}$$

and the time ratio is defined as

$$\exp(\beta_k) = \frac{T_i}{T_j} \tag{12}$$

The expressions for the hazard rate in AFT models are more complicated. In general,

$$\begin{aligned} h(t, X) &= \Psi h_o(t\Psi) \\ \Psi &= \exp(-X\beta) = \exp(-\mu) \end{aligned} \tag{13}$$

In the AFT lognormal model, the hazard rate is

$$h(t, X) = \frac{\frac{1}{t\sigma\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{\log(t)-\mu}{\sigma})^2)}{1 - \Phi(\frac{\log(t)-\mu}{\sigma})} \tag{14}$$

The functional form of (14) allows the hazard rate to first increase and then decrease with time via the parameter sigma, which may be an appropriate behavior for the recapitalization rate. Young dual-class firms are more likely to recapitalize up to a certain point in time. After that, the likelihood to recapitalize into single-class given it has been dual-class for so long may start to decrease because the dual-class structure was not meant as a temporary device.

Finally, the semi-parametric Cox model is a model which assumes proportional hazards like the PH Weibull model. However, the shape of the baseline hazard function is not specified so that $h_0(t)$ in (7) is the nonparametric part of the hazard function. Thus the Cox model does not require assumptions about the hazard function with respect to time but is inappropriate when duration dependence is of particular interest. Whether the hazard function increases or decreases

with time is fitted from the data, and there is no constant. Cox regression is estimated by a procedure called maximum partial likelihood (PL) where only the order of failure times is important, not the actual survival times [Garson (2009)].

In sum, I am going to consider the PH Weibull model defined in (7)-(9) and the AFT lognormal model defined in (10)-(14). In addition, I will estimate the Cox PH semi-parametric model.

3.3. Variables

Time is the dependent variable in survival analysis. Table 1, Panel A shows that a firm that recapitalizes into single-class does so on average 8 years after becoming dual-class (the uncensored observations). The minimum can be less than a year, while the maximum is 45 years. Figure 1a is a histogram of the uncensored observations. The duration spent as dual-class by the firms that did not recapitalize by 2009 (the censored observations) is on average 14 years, with the oldest firm being 83. Figure 1b is the histogram of the censored observations. The fact that the average time before recapitalization is only 8 years, while the average age of the censored observations is 14 years is evidence that observing only the recaps would lead to censoring bias. If indeed a dual-class firm recapitalizes into single-class on average after 8 years, we should not observe average age higher than 8 years. The reason for the discrepancy is that we observe more of the short durations of the sample of firms in 1994-2002 since we stop observation in 2009.

In terms of explanatory variables, the theoretical factors identified as important for the decision of dual-class firms to recapitalize into single-class include private benefits. Thus some of the empirical covariates in X are based on measures of inside ownership as a proxy for the importance of private benefits for insiders.

Cash-flow ownership by insiders (INSIDERCF) can be defined as the percentage of total firm cash-flows owned by insiders. Claims on the cash-flows of the firm for each class are calculated in accordance to the dividends paid by each class or the correct proportions of cash-flow rights if no dividends are paid. When no actual dividends are paid, cash-flows are based on the proportion of dividends to be paid in the future, as stipulated by incorporation charters. If the classes are non-dividend paying, cash-flow rights are proportionate. INSIDERCF can be used as a proxy for the positive incentive alignment effect of managerial equity ownership [Gompers et al (2008)].

Similarly, voting rights are the percentage of total votes held by insiders (INSIDERV). INSIDERV can be used as a proxy for the negative entrenchment effect of insiders owning voting rights, which insulates them from the disciplining force of the market for corporate control. As a result of the disproportional cash-flow and voting rights of dual-class shares, there is a difference between the INSIDERV and INSIDERCF (WEDGE). Claessens et al (2002) find that firms with high discrepancy between INSIDERCF and INSIDERV experience a negative effect on firm value. Thus the WEDGE measures the degree of divergence of insiders' voting rights from their cash-flow rights, and hence is a good proxy for private benefits.

Therefore the WEDGE is an empirical covariate that can explain the decision of dual-class firms to recapitalize into single class. Keeping INSIDERV constant, high-WEDGE firms are less likely to recapitalize than low-WEDGE firms because private benefits are more important. At the same time the time spent in the dual-class state is expected to be longer if the WEDGE is high. Since the coefficients of the PH and AFT models have opposite signs, the expected effect of WEDGE on the hazard rate in the PH models is negative, and the expected effect on duration in the AFT model is positive.

INSIDERV can have an ambiguous effect on the duration as dual-class firm and the hazard rate. Private benefits may be important at firms where insiders control a higher percentage of the voting rights so that they are less likely to recapitalize into single-class. Therefore as a measure of entrenchment, INSIDERV should have a negative effect on the hazard rate. On the other hand, if insiders have a high share of equity, recapitalizing into single-class would be advantageous for them when they ultimately divest control because of more liquid market and better value for their shares. Therefore, after controlling for private benefits with WEDGE, firms with high INSIDERV may be more likely to recapitalize into single-class. This means that the coefficient of ISNIDERV is expected to be positive in the PH models, while in the AFT model it is negative.

Another control for private benefits in the dual-class data is NAME which is an indicator variable equal to one if the firm's name at IPO includes the founder's name. If private benefits such as perquisite-taking, empire building, and passing control to the next generation are important to the founder, the firm is more likely to be named after the founder. In that case the firm is less likely to recapitalize into single-class and the duration in dual-class status is longer. Thus NAME is expected to have a negative coefficient in the PH models, and a positive coefficient in the AFT model.

Other independent variables in X include controls for growth options since dual-class structure can be used by a growing firm that is undervalued in an imperfect market as a defense against value-decreasing takeovers. Variables identified from the literature include size, capital expenditures, and sales growth [Claessens et al (2002), La Porta et al (2002),]. Small firms are expected to grow faster because they have more growth options and the effect on the hazard rate of single-class recapitalization is expected to be negative. Large size may by itself be a deterrent

for hostile takeovers, so that large firms may not need the protection of dual-class shares as an anti-takeover device. Thus large dual-class firms are expected to recapitalize more often into single-class because of lower growth prospects and smaller threat of hostile takeovers. SIZE is defined as the log of book value of assets and the expected effect is positive in terms of hazard rate in the PH models and negative in terms of duration in the AFT model.

A pick up in sales may also indicate growth opportunities, so sales growth is used as a control (SALESGROWTH). SALESGROWTH is expected to have a negative effect on the risk of becoming single-class (negative expected coefficient in the PH models) and a positive effect on the time spent as dual-class firm (positive expected coefficient in the AFT model). Capital expenditure are also a proxy for future expected growth, but using capital expenditures as a control for growth leads to the loss of many observations in the survival-time dataset because of missing data in Compustat. Therefore, capital expenditures are not used in the regression because the sample is too small to afford to lose observations. Finally, asset growth can be used as a measure of current growth. However, asset growth is insignificant when used together with SALESGROWTH as well as by itself in the estimation, so SALESGROWTH is preferred because it is more forward-looking signal for the decision to recapitalize into single-class.

Financial indicators such as assets and sales vary by fiscal years. The total period for each firm is split into episodes. The beginning of each episode is defined by fiscal year ends as available, except in the first episode when it is the dual-class date. The episode ends are defined by subsequent fiscal year ends, except in the last episode. In the last episode, either the recap event happens or the observation is censored. If the single-class date or last observation date is before the next available fiscal year end, the financials are a weighted average of the adjacent

fiscal year ends. If there is no more financial data after the event date which happens in cases like bankruptcy, the variables are assumed equal to the previous fiscal year end.

Insider ownership variables are also added to the episodes in the period 1994-2002, matching the proxy record date to the prior fiscal year end. For the years outside the dual-class data, the ownership variables are assumed fixed as of the first available record for the years prior, and as of the last available record for the years after. Since ownership structure tends to be stable through time, this assumption may be reasonable. Details about episode splitting and variable definitions are given in the Appendix.

A more formal test whether the ownership variables can be assumed constant is to compare the variation within firms to the variation across firms in the years where these variables are available. The total variation in share ownership by insiders can be decomposed into within-firm variation and between-firm variation by means of ANOVA[17]. Statistical significance is tested with the F-stat which gives the ratio of the between-variance to the within-variance. The closer it is to one, the more likely it is that they are equal. In the present case F=44.04 (p-value=0.00), which rejects the null of equal variances and indicates that the between-firm variation is much higher than the within-firm variation.

Table 1, Panel A lists summary statistics of the explanatory variables as of the last episode. INSIDERCF is on average 39%, while INSIDERV is 57%. The average WEDGE is 18% which illustrates the disproportionate holding of cash-flow and voting rights by insiders. About 23% of the firms in the sample are named after their founders. The financial variables

---

[17] Let the insider share ownership be X_{it}. The total variation can be decomposed as $\sum_i \sum_t (X_{it} - \overline{X}) = SS_{within} + SS_{between}$, where $SS_{within} = \sum_i \sum_t (X_{it} - \overline{X})^2$ $SS_{between} = T\sum_i (\overline{X}_i - \overline{X})^2$

For dual-class firms in the years 1994-2002, SS_{within}=22.73 and $SS_{between}$=199.01.

book value of assets and sales growth are measured as of the last episode before recap or final observation. Because these dates occur at different times, the variables are adjusted for inflation so that real book value of assets and sales growth after inflation are given in the table.

It is more meaningful to compare the average values of the empirical covariates by groups of firms differentiated by the decision to recapitalize. Both INSIDERCF and INSIDERV are higher for the firms that did not recapitalize into single-class during the observation period, as shown in Panel B of Table 1. The WEDGE, which is a proxy for private benefit, is significantly higher as well. These firms are also more likely to be named after the founder. Although these averages highlight important differences, the effects of these variables on the propensity to recapitalize can only be estimated in the duration model, including all observable factors for single-class recap. Real book value of assets and sales growth adjusted for inflation as of the last episode are not significantly different between the groups.

Finally, the effect of SOX is included with the variable AFTERSOX. It is an indicator equal to one if the episode ends after SOX was introduced on July 30, 2002, and zero otherwise. There is no lag because recapitalizations can be effected quickly once management makes a decision. In general, the shares of the superior class are automatically exchangeable into the shares of the inferior class and are held mostly by insiders. Thus AFTERSOX is formulated as a time-varying dummy variable, taking zero value in the episodes ending before SOX, and unit value in the episodes after SOX. The expected effect of SOX is ambiguous, as already stated in Section 2.2. If the new rules limit the extraction of benefits by insiders and a single-class structure is preferred, we can expect a positive coefficient of AFTERSOX in the PH model and a negative coefficient in the AFT model. The expected effects are reversed if dual-class status is beneficial for outside shareholders so that firms remain dual-class longer after SOX.

4. Data

4.1. The structure of survival-time data

In order to evaluate the effect of SOX on recapitalization decisions using a duration model, it is necessary to construct a survival-time dataset containing dual-class firms that recapitalize into single-class, as well firms that do not recap. The date of becoming dual-class (dual-class date) and the date of recapitalizing into single-class (single-class date) define the time spent in dual-class status until the event of recap occurs. For the firms that are still dual-class at the end of the observation period, or those that undergo bankruptcy or merger during the observation period, the duration is censored.

The idea of right-censoring is illustrated in Figure 2. Suppose we take a sample of the stock of firms in 1994 and observe their recaps until 2009. We should observe long durations such as A as well as short durations such as C. Duration B is censored as of 2009 since that firm continues to exist as dual-class. Suppose now that in addition to the existing firms in 1994, we have all new dual-class firms in the period 1994-2002, as in the present case. Observing their recapitalizations until 2009 would only produce short durations such as D and E. F would be censored and omitting it would lead to underestimation of the duration from becoming dual-class to single-class recap. Taking into account such censored observations is an important aspect of survival analysis. There is information in the fact that a dual-class firm has not yet recapitalized as of 2009. Thus it is important to include the right-censored observations to avoid bias. Right-censoring can be handled by duration models via the estimation method of maximum likelihood, which is described in Section 3.1.

4.2. Data collection

A comprehensive sample of dual-class firms in 1994-2002, together with panel data on inside ownership, is available from the authors of a study of dual-class firms in the US [Gompers at al (2008)]. The data collection involves identifying which firms in the dual-class data recapitalize into single-class until 2009. Right-censored observations are also included. These are firms that are still dual-class at the end of the observation period, as well as firms which underwent merger or bankruptcy at some point during the observation period. The mergers and bankruptcies are treated as censored observations because the event of single-class recap does not occur and hence duration, already defined as the time spent in dual-class status until single-class recap, is unknown. These events are not treated as competing risks because there is no reason to expect a relationship between the choice between single and dual-class status, and mergers/bankruptcies. Detailed explanation about the data collection process is given in the Appendix.

Table 2 describes the sample of dual-class firms and the results from data collection on single-class recapitalizations. Panel A of Table 2 shows the number of firms by year in the dual-class data, as well as the number of firms that leave the sample. One reason for leaving the dual-class sample is recapitalization into dual class, so these exits constitute the candidate recapitalizations in the period 1994-2001. However, it cannot be automatically assumed that these firms undergo a recapitalization. Other reasons for the firm leaving the dual-class sample include merger, bankruptcy or missing proxy data in that year. Therefore it is necessary to identify what happens to the firms that leave the dual-class sample.

On average, about 20% of the firms trade both the inferior and superior class. In these cases, it is possible to identify a recapitalization based on the CRSP delisting code signifying

exchange. As a result, about 16% of the recapitalizations are based on an exchange code in CRSP. The remaining exits are either eliminated based on delisting codes signifying other events such as merger of bankruptcy within a year of the last proxy record date in the dual-class data, or checked in the SEC filings via EDGAR. 2002 is the last year in the dual-class sample so the whole sample of firms, except those that could be eliminated based on CRSP delisting codes, was observed until 2009. As a result of this process, 176 recapitalizations were identified in the period 1994-2009 out of the total 739 firms in the dual-class data. 344 firms experienced other events such as merger or bankruptcy, and 219 firms remained dual-class as of the last observation date. This date was chosen to be 12/31/2008 since at the time of data collection, CRSP data was only available until end of 2008.

The dual-class date is very important for survival analysis. It is the beginning of analysis time and a missing dual-class date would mean the observation is omitted. That is why I need to identify the dual-class date of the observations in the sample. For this purpose I use CRSP listing codes for the firms that trade two classes and hand-check the remaining firms in the SEC filings.

When two classes are traded, the listing date of each class is available in CRSP (first listing date). Then the date the firm capitalized into dual-class can be identified. After examining these dates, I found out that a firm with two traded classes becomes dual-class much later than the IPO in general. Panel B of Table 2 shows that there are 112 observations of firms with two traded classes in the survival-time dataset. Of these, only 16% establish a dual-class structure at the IPO so that the dual-class date is the same as the first listing date. In the remaining 84% of the cases, the difference between the dual-class date and the first listing date is on average 14 years, with a minimum of a couple of months to a maximum of 48 years. This shows that dual-

class firms with two traded classes are usually not set up as dual-class at IPO but recapitalize into dual-class later.

The dual-class date cannot be established from the first listing date when only one class trades. I checked the dual-class date of some of these firms in the SEC filings and other company sources. Specifically, I checked the dual-class date of the firms that recapitalize in the period 1994-2003, which are 111 observations. I found that the situation with firms that trade only one class is reversed. In the vast majority of cases (93%) the firm established a dual-class structure at the IPO. Therefore assuming that the dual-class date is the same as the first listing date when the dual-class date is not available would lead to an error of only 7%. Since identifying the dual-class date is time-consuming, I assumed it is the same as the first listing date for the 549 observations where the dual-class date is missing.

As a result, the duration of each period begins with the time the firm became dual-class, or the first listing date if the dual-class date is not available. It ends with the time the firm recapitalized into dual-class, or 12/31/2009 if the firm is still dual-class, or the delisting date if merger/bankruptcy (treated as censored). Thus each episode ends in a single-class recapitalization or the observation is censored. This data structure can be used to examine the decision of firms to recapitalize into single class and the effect of SOX. Panel C of Table 2 gives the number of observations before and after SOX was introduced on July 30, 2002 and shows that there are 60 recaps and 259 censored observations after SOX.

Additional data necessary for the analysis includes time-varying insider ownership variables available in the original dual-class data, as well as financial indicators, available from Compustat. After adding these variables to the survival-time dataset, 702 firms remain in the sample used in the estimation.

5. Results

5.1. Choosing the model

I first estimate the PH Weibull, AFT lognormal the Cox semi-parametric PH models. The purpose is to understand which model performs best in terms of assumptions about the hazard rate, expected parameter estimates, and goodness of fit measured by the Akaike Information Criterion (AIC). I will then focus on the interpretation of the effect of SOX on the hazard rate of single-class class recap and duration as dual-class based on the chosen model. The explanatory variables in X are INSIDERV, WEDGE, SIZE, SALESGROWTH, and AFTERSOX. The parametric models include a constant, while the Cox model does not because it is subsumed in the non-parametric part of the hazard function.

Model 1 in Table 3 shows the coefficient estimates from the PH Weibull model. The coefficients indicate the marginal effect of the covariate on the hazard rate of singe-class recap (not the hazard ratios). The log-likelihood ratio test can be used to evaluate the importance of covariates. If the null model is the constant-only model and the full model includes predictor variables, the difference in fit between the two is the basis for computing the log-likelihood ratio. The form of the test statistic is LR=2(-lnL(null model)+lnL(full model)), which is distributed chi-square with k degrees of freedom for k covariates. The LR tests for this and all subsequent models indicate that the models with covariates are better than constant-only models.

In Model 1, the coefficient of WEDGE representing the marginal effect on the hazard rate is negative and significant at 1%. This indicates that dual-class firms where the discrepancy between the cash-flow rights and voting rights of insiders is high indicating high private benefit are less likely to recapitalize into single-class. The coefficient of NAME is also negative and

significant at 1% confirming the expectation that firms named after their founders recapitalize into single-class less often. AFTERSOX is positive and significant at 1% indicating that dual-class firms are more likely to recapitalize into single-class after SOX was introduced. None of the other coefficients in the model is different from zero at 5%.

The ancillary parameter ln(p) is significant at 1% indicating that the hazard rate is not constant with time. The corresponding value of p is 0.80 and, since it is less than one, it indicates that the hazard rate decreases with time. Figures 3a and 3b show the graphs of the hazard rate and survival function.

In other words, after accounting for other factors, the effect of time on the hazard rate is negative so that the longer a dual-class firm remains dual-class, the less likely it is to recapitalize into single-class. Duration dependence can be interpreted as the effect of firm age as in Amoako-Adu and Smith (2001) and Taylor and Whittred (1998). Age in these studies controls for firm-specific human capital of the manager. The effect of age here cannot be separated from duration dependence because age as a dual-class firm is the same as duration. I find that the longer the time since becoming dual-class, the less likely the firm is to recapitalize which may be due to the fact that the founder has developed skills specific to the business.

The AFT lognormal model allows the hazard rate to first increase and then decrease with time. This may be a more reasonable form of the hazard of single-class recap. Young firms where dual-class status is a temporary anti-takeover device may be more likely to recap with time. At some point however, the longer the firm remains dual-class, the less likely it is to recap because of human capital investment by the founder. The AFT lognormal model also does not assume a proportional effect of the covariates through time. This model is presented as Model 2 in Table 3.

The coefficients of WEDGE, NAME and AFTERSOX behave as in the PH Weibull model. Note that the signs are reversed because they give the marginal effects on the time to single-class recapitalization as opposed to the marginal effects on the hazard rate. Thus high-WEDGE firms take longer to recap, and the effect of NAME on duration in dual-class status is also positive. On the other hand, AFTERSOX has a negative effect on duration so that firms recapitalize faster after SOX. In addition, the coefficient of SIZE is negative and significant at 5%, indicating that the larger the firm, the shorter the time spent in dual-class status.

The ancillary parameter sigma is significant at 1%. The value of 1.78 allows the hazard rate to first steeply increase with time and then decrease with time, as shown in Figure 3c. The survival function in Figure 3d is similar to the one for the Ph Wibull model in Figure 3b, which stems from the fact that over most of the life of the firms, the hazard rate is decreasing. However, the hazard rate in the AFT model is allowed to first increase with time for very young dual-class firms, and then decrease, which contributes to the better fit of the AFT lognormal model. The AIC is lower than the PH Weibull model, indicating improved fit.

Finally in Table 3, Model 3 is the semi-parametric Cox model. It assumes proportional hazards but does not postulate a specific form of the hazard function. Thus the baseline hazard and survival functions are not solved for parametrically but fitted from the data. Thus we cannot make conclusions about the exact shape of the hazard with respect to time but the model is less restrictive in terms of assumptions. Figures 3e and 3f show that plots of the hazard and survival functions as they are fitted from the data. The main conclusions regarding the parameter estimates remain unchanged, but the AIC is much larger as can be expected.

Even through no assumptions are made about the form of the hazard function in the Cox model, it does assume proportional hazards. A simple way to check the validity of this

assumption is to plot the hazard rates for a categorical variable, such as NAME. If the graphs are parallel, the effect of NAME on the hazard rate would be the same regardless of how old the firms are. Figure 3g indicates that the PH assumption may not be appropriate since there is a wider difference in the graphs in the beginning than in the end. Since the irregular shape of the non-parametric baseline hazard prevents good visual examination, another test that can be used is the "log-log" plot. The test plots the –ln(-ln(survival)) curve versus ln(analysis time) for each category of a variable. The PH assumption is violated when the curves are not parallel. Figure 3h shows that this is indeed the case.

Overall it appears that the AFT lognormal model is best in terms of assumptions about the hazard rate, expected parameter estimates, and goodness of fit. Therefore I use the AFT model to interpret the effect of SOX.[18]

5.2. The effect of SOX in the AFT lognormal model

Table 4 focuses on the AFT lognormal model because it is used for the interpretation of the effect of SOX. Column 1 is the same as Model 2 in Table 3. In addition, Column 2 presents the exponentiated coefficients, or time ratios. To summarize the effect of the control variables in Column 1, dual-class firm where there is a large WEDGE between the voting and cash-flow rights of insiders, indicating high private benefit of control, take longer to recapitalize into single class as it makes cash-flow and voting rights proportional. Another measure of the importance of control for the founder is the NAME variable, and it has a positive effect on duration so that firms named after their founders remain dual-class longer. Size is also a significant determinant of recapitalization and large firms become single-class faster than small firms. The effects of

[18] The effect of SOX in the PH Weibull and Cox models is essentially the same as in the AFT lognormal.

sales growth and insiders' voting control by itself are not found to be different from zero in the estimation.

AFTERSOX has a significant negative effect on duration as dual-class firm. This means that dual-class firms recapitalize into single-class faster after SOX was introduced. The corporate governance provisions of the Act aimed at aligning incentives of corporate insiders and outside shareholders as well as restricting diversion of resources. Dual-class firms recapitalize into single-class more quickly after SOX. This empirical result is consistent with the hypothesis that private benefits are an important determinant of dual-class status, and that insiders inferred that the passage of SOS would reduce their ability to extract private benefits. Insiders value control less after SOX and want to divest their holdings to minimize the under-diversification cost of concentrated holdings. Becoming single-class may help this process by eliminating disproportionate holding of cash-flow and voting. In addition, there is a more liquid market in ordinary shares which allows insiders to monetize holdings gradually and more efficiently.

Column 2 in Table 4 gives the coefficients in the time ratio form defined in equation (12). The time ratio of AFTERSOX indicates that the time firms spend in dual-class status after SOX is about 80% shorter than the duration before SOX. Consequently, dual-class firms are more likely to recapitalize into single-class after the legislation was introduced. Figure 4 graphically illustrates the effect of SOX on the hazard and survival functions. The hazard rate is higher after SOX in Figure 4a, and survival is lower. In terms of the control variables, the time ratio of high-WEDGE firms indicates that they stay dual-class about 10 times longer, and those named after their founders 2 times longer. The duration of large firms is about 8% shorter, indicating that they are more likely to recapitalize into single class.

6. Conclusion

This chapter estimates a duration model of the time if takes a dual-class firm to recapitalize into single-class. If the corporate governance provisions of SOX are effective in limiting extraction of private benefits of insiders, they would no longer value control and would want to divest their position. At the same time, if dual-class structure is optimal in terms of outside shareholder value, then SOX would cause firms to remain dual-class by reducing the cost of extraction of private benefits. I find significant evidence that the duration firms spend in dual-class status is shorter after SOX. SOX limits the extraction of private benefits and since private benefits are an important factor in the choice of dual-class status, firms are more likely to recapitalize into single class after the legislation was passed. The overall effect of SOX on recapitalization decisions of dual-class firms is that the legislation serves to limit extraction of private benefits of insiders evidenced by the significantly shorter time spend in dual-class status and the higher likelihood of single-class recapitalization after SOX.

Appendix: Data construction

The purpose of the data construction is to build a survival-time dataset consisting of single-class recapitalizations of dual-class firms as well as firms which did not recapitalize in the observation period. It is also necessary to have the dates of firms' becoming dual-class and recapitalizing into single-class, or the last observation date if the recap event did not occur. Single-class recaps can be identified on the basis of information from CRSP when two classes trade publicly, or in about 20% of the dual-class firms. However, data on single-class recaps is not readily available when only one class trades. In this case, the event can be identified from firm filings with the SEC, which include financial reports (10-K) and proxy statements (DEF 14A). This appendix describes the process of hand-collecting data from the individual firms' SEC filings and building the survival-time dataset. Once the event of single-class recap is identified, explanatory variables such as insider ownership and financial indicators are included to examine the decision of dual-class firms to become single-class.

The starting point for the data construction is the dual-class data from Gompers et al (2008) which contains a comprehensive sample of dual-class firms in the period 1994-2002. It is a panel dataset on insider ownership variables measured as of the proxy record date of each firm in the period 1994-2002.[19] The number of firms per year varies from a minimum of 109 in 1994 to a maximum of 504 in 1998. Table A1 lists some of the variables in the dataset that are relevant for examining recap decisions. Using the structure of the dual-class data helps narrow down the search for recaps in the SEC filings, which for some firms extend over decades.

[19] To build a comprehensive set of dual-class firms, the authors first constructed a candidate sample using data from the Securities Data Company (SDC), Compustat, CRSP and the Investor Responsibility Research Center (IRRC). Then they examined the proxy statements and/or 10-Ks for every candidate to determine it was dual-class. Once the dual-class sample was completed, the next step was to determine the insider ownership for each class of stock as of the proxy record date for each firm-year. They parsed the tables and footnotes to compute the common-stock ownership, excluding all options and other rights. Since the share classes sometimes have differing cash-flow rights, they also collected dividend data for all firms.

Table A1: Variables in the original dual-class data

RECORDYEAR:
GVKEY: Compustat gvkey
NAME: Company name
RECORDDATE: Proxy record date
PERMNOA-PERMNOD: CRSP permno of "A" class, CRSP permno of "B" class, etc. Recorded as missing when the class does not trade publicly. The designation of A versus B need not be the same as the company's designation.
CLASSNUM: Number of classes of stock
VOTEA-VOTED: Number of votes attached to class A shares, number of votes attached to class B shares, etc.
DIVCODE: Equals "NA" when no dividends were paid. Equals "NTD" when no dividends were paid but dividend rights will differ if they are paid in the future. Equals "value" otherwise when dividends were paid on some class.
DIVAMTA-DIVAMTD: Dividends per share for class A, dividends per share for class B, etc. Equals 1 across classes when divcode is "NA." Defined to provide the correct ratio of dividends across classes should dividends be paid in the future when divcode is "NTD". Equals dividends paid when divcode is "value".
INSIDERA - INSIDERD - Number of shares of class A common stock owned by insiders, number of shares of class B common stock owned by insiders, etc.
SHROUTA-SHROUTD - Shares outstanding of class "A" stock, shares outstanding of class "B" stock, etc.

Table A2 gives a small part of the data for better understanding of the structure. I have randomly chosen several observations in 1997 and 1998. Wrigley and Abercrombie&Fitch are well-known companies, the others are not so famous. Suppose these are all the companies in 1997 and 1998. Observe that Alpharma and Wrigley remain in the sample both years. Thus there cannot be a recapitalization into single-class for these firms. On the other hand, several firms in 1997 are no longer in the sample in 1998, such as Graphic Industries and Abercrombie&Fitch. It is possible that these firms became single-class but we cannot automatically assume a recapitalization. The reason they are no longer in the sample could be a merger, bankruptcy, or proxy data was missing in 1998. Finally, there are some new firms in 1998, such as Hertz.

Table A2: Examples of observations in the dual-class data

YEAR	GVKEY	NAME	RECORDDATE	PERMNOA	PERMNOB	PERMNOC	CLASSNUM	VOTEA	VOTEB	INSIDERA	INSIDERB
1997	1034	ALPHARMA	4/1/1997	.	65832	.	2	4	1	9500000	563981
1997	5275	GRAPHIC INDUSTRIES	4/17/1997	.	39670	.	2	10	1	4478092	1432785
1997	11609	Wrigley	1/15/1997	.	15472	.	2	10	1	12913621	22307666
1997	12270	COSMETIC & FRAGR	1/28/1997	10394	77410	84724	2	1	0	854205	972401
1997	21158	GENEVA STEEL	12/30/1997	76116	.	.	2	1	1	409556	19031348
1997	61595	BIG FLOWER PRESS	5/9/1997	82632	.	.	2	1	0	6942166	0
1997	63643	ABERCROMBIE & FITCH	3/28/1997	.	83976	.	2	3	1	43000000	60976
1998	1034	ALPHARMA	3/24/1998	.	65832	.	2	4	1	9500000	1740031
1998	5600	Hertz Corp	3/25/1998	.	84663	.	2	5	1	67310167	20612339
1998	6116	INTERNAT SPEEDWAY	2/5/1998	.	84226	.	2	1	0.2	23269475	75905
1998	11609	Wrigley	1/15/1998	.	15472	.	2	10	1	12907494	22273021
1998	14269	HERBALIFE INTERNAT	4/17/1998	11490	85694	.	2	1	0	5813958	11627924

In order to identify what happened to the firms that exited the sample in 1998, I first checked the CRSP delisting codes. A single-class recapitalization can be confirmed from the delisting code signifying exchange in the case when the superior class that is being eliminated is traded publicly, which usually means both classes are traded (a minority of dual-class firms). It is also possible to rule out a recap based on the delisting codes of the traded class when there was a merger or bankruptcy. I have assumed that if such an event happens within a year of the proxy record date, there could not have been a recapitalization, unless it was a part of the merger agreement or bankruptcy restructuring. I am not interested in the recaps in these cases, only when the company recapitalizes and continues to exist as a single-class company.

Unless there is a delisting code for exchange, or for merger/bankruptcy within a year of the record date, it is unclear what happened. If a merger/bankruptcy happens later, there could be a recapitalization before that. When a firm exits the dual-class sample but there is no delisting code at that time, it may mean there are no available proxy statements. This could happen in the case of a firm that is publicly traded but insiders control more than 51% of the votes so they may not always solicit proxies for the annual shareholder meeting, or in the case of financial distress when all filings become irregular. A variety of other delisting codes may exist around the time of exit, but they are not sufficient to rule out a recapitalization. For example, the company could be delisted for reasons such as “company request” or “insufficient float,” so that it stops being traded publicly. However, even if the firm is not traded publicly but there are more than 300 shareholders, the firm is still required to file with the SEC. It is impossible to identify a recap only when the company reduces its shareholders to under 300 and deregisters with the SEC (“goes dark”).

Thus in most cases CRSP delisting codes were not sufficient, and I had to hand-collect data by going into the SEC filings of the firms, which are available online via EDGAR. A special form regarding a recapitalization is not required. If the articles of incorporation require shareholder approval, an announcement is made in the proxy statement including the date of recapitalization (single-class date). However, if the majority of votes are held by insiders, the decision can be made without eliciting proxies. The most common case is when the articles provide that the classes are exchangeable. The company can redeem all of the outstanding superior class shares in exchange for shares of the inferior class. This is particularly easy if the superior class is held only among a small group of individuals. In these cases there would be an explanation with the date of recapitalization in either the proxy statement or in the footnotes of the financial statements such as 10-K, 10-Q, 10-KSB, etc.

To build the survival-time dataset, it is necessary to identify when the firms became dual-class. For the companies that underwent recapitalization in the period 1994-2003, I also checked the date the firm first became dual-class (dual-class date). When both classes are traded, that date is given by the listing date of the second class in CRSP. When only one class is traded, I searched the SEC filings, the firm's history online or calling investors' relations on a few occasions. I only did that for the firms that recapitalized and only until 2003 because that was the initial sample for the duration model. However, including the censored observations and extending the sample is important for the analysis. Identifying the dual-class date for all observations would be prohibitively time-consuming but fortunately the dual-class date can be assumed to be the same as the first listing date with a small error as explained in the text.

The final year in the dual-class data is 2002. Therefore it is not possible to identify which companies exit the dual-class sample. I reviewed the SEC filings of all 2002 firms until

12/31/2008 (last year of Compustat financial data that I have) except those that I could eliminate based on delisting codes for exchange, or for merger/bankruptcy within one year of the last available proxy record.

A summary of the variables I added to the original dual-class data are given in Table A3.

Table A3: New variables	
CIK: Central Index Key used by the SEC to identify companies	
SIC: Standard Industry Classification Code	
EXIT: Equals one if a firm leaves the dual-class sample, zero otherwise.	
Equals "missing" for 2002 firms because the dual-class sample ends in 2002.	
RECAP: Equals one if a dual-class firm recapitalizes into single-class, zero otherwise.	
If EXIT=0, RECAP=0	
If EXIT=1 and there is a code equal to 331 or 341, RECAP=1. If CLASSNUM>2, SEC filings were checked.	
If EXIT=1 and there is a code equal to 231, 232, 233, 234, 241, 261, 450, 573, 574 within 1 year, RECAP=0.	
If EXIT=1 and there is any other code, or a code at another time, or no delisting code, SEC filings checked.	
If EXIT="missing" (in the case of 2002 dual-class firms) and there is a code equal to 331 or 341, RECAP=1.	
If EXIT="missing" and there is code 231, 232, 233, 234, 241, 261, 450, 573, 574 within 1 year, RECAP=0.	
All other 2002 dual-class firms that could not be eliminated were checked in the SEC filings.	
STILLDC: Equals one if a 2002 firm is still dual-class as of the last SEC filing in 2009, zero otherwise.	
SOURCE: Equals "SEC" if SEC filings were reviewed (DEF 14A, 10-K, 10-Q, 10-KSB, etc).	
Equals "CRSP" if outcome based on CRSP delisting codes.	
DCDT: Date of dual-class capitalization	
In the case of two traded classes, equals the listing date of the second class in CRSP.	
For 1994-2003 recaps of firms with one traded class, SEC filings were reviewed for the dual-class date.	
In all other cases the date of dual-class capitalization is missing.	
SCDT: Date of single-class recapitalization	
Equals delisting date if delisting codes equal to 331 or 341.	
Otherwise, equals the recap date according to the SEC filings; if not available, the date of filing.	
LSTDTA-LSTDTC: First CRSP listing date of classes A-C	
DLSTCDA-DLSTCDC: Delisting codes of classes A-C from CRSP	
DLSTDTA-DLSTDTC: Delisting date of classes A-C from CRSP	
Delisting codes from CRSP available until end of 2008	
231	Merger, shareholders receive common stock
232	Same as 231, but not maintained in CRSP
233	Merger, shareholders receive cash payments
234	Merger, shareholders receive preferred stock
241	Merger, shareholders receive common stock/cash
261	Merger, shareholders receive cash
331	Exchange, other class of common stock
341	Exchange, common stock
550	Delisted, insufficient number of market makers
551	Delisted, insufficient number of shareholders
552	Delisted, price fall below acceptable level
560	Delisted, insufficient capital
561	Delisted, insufficient float or assets
570	Delisted, company request
573	Delisted, deregistration (gone private)
574	Delisted, bankruptcy
580	Delisted, filing delinquent
582	Delisted, not meet requirements of exchange
584	Delisted, not meet guidelines
585	Delisted, protection of investors

Table A4 gives a part of the spreadsheet I used to collect the new data. Note that of the 5 firms that are not in the 1998 sample, three exited due to recapitalizing into single-class. One of them was identified from the CRSP delisting code 341, which indicates that the two traded classes were exchanged for one class of common stock. The single-class date is the date of exchange and the dual-class date is the listing date of the second class. The other two recapitalizations were checked in the SEC filings. There is a delisting code for gvkey 61595 but it happens later and in fact there was a recapitalization before that. I have confirmed that these companies became dual-class at IPO so the dual-class date is the same as the first CRSP listing date. Of the firms that exited the sample, but did not recapitalize, gvkey 5275 was eliminated based on a delisting code signifying merger within a year of the proxy record date. The other one had to be checked in the SEC filings.

Table A4: Example of new data

YEAR	GVKEY	EXIT	RECAP	SOURCE	DCDT	SCDT	LSTDTA	LSTDTB	LSTDTC	DLSTCDA	DLSTDTA	DLSTCDB	DLSTDTB
1997	1034	0	0					2/29/1984					
1997	5275	1	0	CRSP				12/30/1983				233	10/31/1997
1997	11609	0	0					1/30/1926					
1997	12270	1	1	CRSP	3/31/1992	4/30/1997	6/30/1986	3/31/1992	4/30/1997	341	4/30/1997	341	4/30/1997
1997	21158	1	0	SEC			3/30/1990			584	1/29/1999		
1997	61595	1	1	SEC	11/30/1995	6/1/1997	11/30/1995			573	12/31/1999		
1997	63643	1	1	SEC	9/30/1996	5/19/1998		9/30/1996					
1998	1034	0	0					2/29/1984					
1998	5600	0	0					4/30/1997					
1998	6116	0	0					11/29/1996					
1998	11609	0	0					1/30/1926					
1998	14269	0	0				7/31/1987	12/31/1997					
2002 firms													
YEAR	GVKEY	EXIT	RECAP	SOURCE	DCDT	SCDT	LSTDTA	LSTDTB	LSTDTC	DLSTCDA	DLSTDTA	DLSTCDB	DLSTDTB
2002	1034		1	SEC		12/28/2006		2/29/1984					
2002	1038		0	SEC				8/31/1983					
2002	1076		0	SEC	11/30/1992		11/30/1982	11/30/1992					
2002	1239		1	CRSP	4/30/1986	11/3/2005	6/30/1965	4/30/1986				331	11/28/2003
2002	2007		0	CRSP				4/29/1977				233	12/31/2002

The last year is 2002 so I could not identify candidate recapitalizations in the sample of firms and hence EXIT is missing. In the examples I give in Table A4, there are two recapitalizations (RECAP=1) and two firms that were found to be still dual-class in 2009 (RECAP=0). There is also a firm that does not exist after 2002 because of merger (RECAP=0). The observations with RECAP=0 are considered censored. One of the recaps (gvkey 1239) is

based on the CRSP delisting code for exchange. The other was found in the SEC filings but since this recap happens after 2003, I do not have the exact dual-class date. For the censored observations I also do not have the dual-class date unless two classes are traded as in the case of gvkey 1076.

As a result of the data collection, I have a sample of 176 recapitalizations of dual-class firms and 563 censored observations in the period 1994-2009, of which 219 were still dual-class as of 12/31/2008. The data collection process was not automatic and navigating the SEC filings was often complicated by the fact that each case was individual. Therefore it is possible that I have missed a recapitalization or eliminated a dual-class firm, so I do not claim that my final sample is comprehensive. However, the recapitalizations and censored observations included in the sample are confirmed as such and can be used for estimation. Thus I do not expect any biases due to the data collection process.

It is possible to estimate a single-record survival-time model on the basis of the last observation per gvkey where either RECAP=1 or RECAP=0 depending on whether the single-class recapitalization occurs. Each record starts with the dual-class date, or if not available, the first listing date. The records end with the single-class date of the event occurs. Otherwise the observations are censored either as of the last date of observation 12/31/2008 if still dual-class, or as of the delisting date if merger/bankruptcy, or as of the last available proxy date if no delisting code.

To examine recapitalization decisions, it is important to include time-varying financial variables. Thus the final step in the data construction is to split the episode of the single records into episodes according to fiscal years and thus allow financial covariates to vary from the time of becoming dual-class and to the time of single-class recap or censoring. The variables of

interest here are total assets (AT) and sales (SALE), both available from Compustat. The total period for each firm is split into episodes. The beginning of each episode is defined by fiscal year ends as available, except in the first episode when it is the date the firm became dual-class. The episode ends are defined by subsequent fiscal year ends, except in the last episode. In the last episode, either the recap event happens (RECAP=1) so that the episode ends with the single-class date, or it is censored (RECAP=0). If the single-class date or last observation date is before the next available fiscal year end, AT and SALE are a weighted average of the adjacent fiscal year ends. If there is no more financial data after the event date which happens in cases like bankruptcy, then AT and SALE are assumed equal to the previous fiscal year. Table A5 gives an example of the dataset after episode splitting and adding the Compustat variables AT and SALE.

Inside ownership variables from the dual-class dataset are also added in the period 1994-2002, matching proxy record date to previous fiscal year end. For the years prior to 1994, I assume ownership variables fixed as of the first available proxy record, and for the years after 2002, as of the last available proxy in the dual-class data. Hand-collecting data for more than 700 firms for the years not available in the dual-class data would be prohibitively time-consuming. However, ownership structure tends to be more stable over time, so the assumption of fixed inside ownership variables be reasonable. Table A5 also shows some of the inside ownership variables.

Table A5: Survival-time dataset with multiple records

GVKEY	EPBEG	EPEND	RECAP	FYEND	FYEAR	AT	SALE	PROXYDATE	VOTEA	VOTEB	INSIDERA	INSIDERB	NAME
1076	11/30/1992	3/31/1993	0	3/31/1993	1992	108.217	157.632	6/9/1995	1	0	2522840	1082283	1
1076	3/31/1993	3/31/1994	0	3/31/1994	1993	144.917	185.184	6/9/1995	1	0	2522840	1082283	1
1076	3/31/1994	3/31/1995	0	3/31/1995	1994	157.527	228.892	6/9/1995	1	0	2522840	1082283	1
1076	3/31/1995	12/31/1995	0	12/31/1995	1995	158.645	178.224	3/8/1996	1	0	2486523	1048535	1
1076	12/31/1995	12/31/1996	0	12/31/1996	1996	198.103	274.245	3/7/1997	1	0	2471623	4299494	1
1076	12/31/1996	12/31/1997	0	12/31/1997	1997	239.382	310.751	3/27/1998	1	0	2477376	4315295	1
1076	12/31/1997	12/31/1998	0	12/31/1998	1998	272.174	379.659	3/12/1999	1	0	2483081	3682228	1
1076	12/31/1998	12/31/1999	0	12/31/1999	1999	318.408	437.359	3/10/2000	1	0	2482576	3837626	1
1076	12/31/1999	12/31/2000	0	12/31/2000	2000	380.379	502.92	3/9/2001	1	0	2479086	3530051	1
1076	12/31/2000	12/31/2001	0	12/31/2001	2001	397.196	546.681	3/15/2002	1	0	2479593	2658204	1
1076	12/31/2001	12/31/2002	0	12/31/2002	2002	483.648	640.688	3/15/2002	1	0	2479593	2658204	1
1076	12/31/2002	12/31/2003	0	12/31/2003	2003	555.292	766.797	3/15/2002	1	0	2479593	2658204	1
1076	12/31/2003	12/31/2004	0	12/31/2004	2004	700.288	940.98	3/15/2002	1	0	2479593	2658204	1
1076	12/31/2004	12/31/2005	0	12/31/2005	2005	858.515	1124.006	3/15/2002	1	0	2479593	2658204	1
1076	12/31/2005	12/31/2006	0	12/31/2006	2006	979.606	1319.392	3/15/2002	1	0	2479593	2658204	1
1076	12/31/2006	12/31/2007	0	12/31/2007	2007	1113.176	1490.033	3/15/2002	1	0	2479593	2658204	1
1076	12/31/2007	12/31/2008	0	12/31/2008	2008	1235.802	1587.361	3/15/2002	1	0	2479593	2658204	1
1239	4/30/1986	9/30/1986	0	9/30/1986	1986	227.742	435.355	11/28/1994	1	0.1	8844700	1929202	1
1239	9/30/1986	9/30/1987	0	9/30/1987	1987	271.825	514.491	11/28/1994	1	0.1	8844700	1929202	1
1239	9/30/1987	9/30/1988	0	9/30/1988	1988	303.348	604.708	11/28/1994	1	0.1	8844700	1929202	1
1239	9/30/1988	9/30/1989	0	9/30/1989	1989	362.612	717.438	11/28/1994	1	0.1	8844700	1929202	1
1239	9/30/1989	9/30/1990	0	9/30/1990	1990	436.547	795.825	11/28/1994	1	0.1	8844700	1929202	1
1239	9/30/1990	9/30/1991	0	9/30/1991	1991	574.413	873.719	11/28/1994	1	0.1	8844700	1929202	1
1239	9/30/1991	9/30/1992	0	9/30/1992	1992	610.4	1091.286	11/28/1994	1	0.1	8844700	1929202	1
1239	9/30/1992	9/30/1993	0	9/30/1993	1993	593.046	1147.99	11/28/1994	1	0.1	8844700	1929202	1
1239	9/30/1993	9/30/1994	0	9/30/1994	1994	610.208	1216.119	11/28/1994	1	0.1	8844700	1929202	1
1239	9/30/1994	9/30/1995	0	9/30/1995	1995	815.086	1358.219	11/27/1995	1	0.1	8816522	2011110	1
1239	9/30/1995	9/30/1996	0	9/30/1996	1996	909.266	1590.409	11/25/1996	1	0.1	8816522	1709359	1
1239	9/30/1996	9/30/1997	0	9/30/1997	1997	1000.059	1775.258	11/25/1997	1	0.1	17485218	3577661	1
1239	9/30/1997	9/30/1998	0	9/30/1998	1998	1068.184	1834.711	12/1/1998	1	0.1	16563942	3241222	1
1239	9/30/1998	9/30/1999	0	9/30/1999	1999	1184.534	1975.928	12/14/1999	1	0.1	16290834	3518688	1
1239	9/30/1999	9/30/2000	0	9/30/2000	2000	1389.819	2247.163	12/1/2000	1	0.1	17087719	4155253	1
1239	9/30/2000	9/30/2001	0	9/30/2001	2001	1516.501	2494.18	11/30/2001	1	0.1	16466111	4152123	1
1239	9/30/2001	9/30/2002	0	9/30/2002	2002	1729.491	2650.976	11/29/2002	1	0.1	9711596	4408033	1
1239	9/30/2002	9/30/2003	0	9/30/2003	2003	1945.609	2891.417	11/29/2002	1	0.1	9711596	4408033	1
1239	9/30/2003	9/30/2004	0	9/30/2004	2004	2058.78	3257.996	11/29/2002	1	0.1	9711596	4408033	1
1239	9/30/2004	9/30/2005	0	9/30/2005	2005	2302.123	3531.231	11/29/2002	1	0.1	9711596	4408033	1
1239	9/30/2005	11/3/2005	1	9/30/2006	2006	2326.656	3769.419	11/29/2002	1	0.1	9711596	4408033	1
2007	4/29/1977	12/31/1978	0	12/31/1978	1978	351.973	26.424	7/20/1995	1	1	37500000	393250	0
2007	12/31/1978	12/31/1979	0	12/31/1979	1979	425.022	37.17	7/20/1995	1	1	37500000	393250	0
2007	12/31/1979	12/31/1980	0	12/31/1980	1980	668.02	56.644	7/20/1995	1	1	37500000	393250	0
2007	12/31/1980	12/31/1981	0	12/31/1981	1981	1055.676	126.926	7/20/1995	1	1	37500000	393250	0
2007	12/31/1981	12/31/1982	0	12/31/1982	1982	1564.604	179.631	7/20/1995	1	1	37500000	393250	0
2007	12/31/1982	12/31/1983	0	12/31/1983	1983	2123.245	210.735	7/20/1995	1	1	37500000	393250	0
2007	12/31/1983	12/31/1984	0	12/31/1984	1984	2032.721	230.919	7/20/1995	1	1	37500000	393250	0
2007	12/31/1984	12/31/1985	0	12/31/1985	1985	1757.992	197.249	7/20/1995	1	1	37500000	393250	0
2007	12/31/1985	12/31/1986	0	12/31/1986	1986	1028.522	149.335	7/20/1995	1	1	37500000	393250	0
2007	12/31/1986	12/31/1987	0	12/31/1987	1987	715.06	70.607	7/20/1995	1	1	37500000	393250	0
2007	12/31/1987	12/31/1988	0	12/31/1988	1988	831.39	63.24	7/20/1995	1	1	37500000	393250	0
2007	12/31/1988	12/31/1989	0	12/31/1989	1989	406.337	71.483	7/20/1995	1	1	37500000	393250	0
2007	12/31/1989	12/31/1990	0	12/31/1990	1990	234.944	30.568	7/20/1995	1	1	37500000	393250	0
2007	12/31/1990	12/31/1991	0	12/31/1991	1991	266.123	24.363	7/20/1995	1	1	37500000	393250	0
2007	12/31/1991	12/31/1992	0	12/31/1992	1992	322.769	27.364	7/20/1995	1	1	37500000	393250	0
2007	12/31/1992	12/31/1993	0	12/31/1993	1993	368.608	25.034	7/20/1995	1	1	37500000	393250	0
2007	12/31/1993	12/31/1994	0	12/31/1994	1994	331.79	18.138	7/20/1995	1	1	37500000	393250	0
2007	12/31/1994	12/31/1995	0	12/31/1995	1995	296.583	22.301	9/30/1996	1	1	2500000	26379	0
2007	12/31/1995	12/31/1996	0	12/31/1996	1996	375.182	23.294	3/18/1997	1	1	2500000	26379	0
2007	12/31/1996	12/31/1997	0	12/31/1997	1997	451.256	31.447	6/1/1998	1	1	2500000	2002276	0
2007	12/31/1997	12/31/1998	0	12/31/1998	1998	719.997	58.802	3/18/1999	1	1	2500000	2239319	0
2007	12/31/1998	12/31/1999	0	12/31/1999	1999	920.707	74.55	11/30/2000	1	1	2500000	5766388	0
2007	12/31/1999	12/31/2000	0	12/31/2000	2000	2741.379	189.325	4/27/2001	1	1	2500000	8777088	0
2007	12/31/2000	12/31/2001	0	12/31/2001	2001	3060.988	235.487	11/1/2002	1	1	2500000	9579871	0
2007	12/31/2001	12/31/2002	0			3060.988	235.487	11/1/2002	1	1	2500000	9579871	0

For the estimation, the variables are defined after adjusting the financial data for inflation. A summary of all relevant variables is given in Table A6.

Table A6: Summary of variables

Variable	Description	Source
RECAP	Equals one if a dual-class firm recapitalizes into single-class in an episode, zero if the firm is still dual-class or there is a merger/bankruptcy (censored).	
EPBEG	In the first episode, the date of dual-class capitalization, or if unknown, the first listing date; in subsequent episodes, defined by available fiscal year ends.	
EPEND	In the last episode, the date of single-class recapitalization if RECAP=1 or 12/31/08 if RECAP=0 and still dual-class, or the delisting date/last proxy date if RECAP=0 and merger/bankruptcy; in previous episodes, fiscal year end	
INSIDERCF	Percentage of cash-flow rights owned by officers and directors Calculated as INSIDERCF= (DIVAMTA*INSIDERA+…+DIVAMTD*INSIDERD)/(DIVAMTA*SHROUTA+…+DIVAMTD*SHROUTD)	
DIVAMTA-DIVAMTD	Equals dividends per share for classes A-D if dividends are paid; Equals the proportion of dividends across classes if paid in the future; Equals one if dividends not paid	Dual-class data
SHROUTA-SHROUTD	Shares outstanding of classes A-D	Dual-class data
INSIDERA-INSIDERD	Number of shares of classes A-D owned by insiders	Dual-class data
INSIDERV	Percentage of votes held by officers and directors Calculated as INSIDERV= (VOTEA*INSIDERA+…+VOTED*INSIDERD)/(VOTEA*SHROUTA+…+VOTED*SHROUTD)	
VOTEA-VOTED	Number of votes attached to classes A-D	Dual-class data
WEDGE	Equals the difference between INSIDERV and INSIDERCF	
AT	Book value of assets	Compustat Item 6
SIZE	Equals *ln*(AT) adjusted for inflation*	
SALE	Sales during fiscal year	Compustat Item 12
SALESGROWTH	Growth of sales adjusted for inflation* Equals the rate of change from EPBEG to EPEND	
NAME	Equals one if name of person appears in company name at IPO, zero otherwise	Dual-class data
AFTERSOX	Equals one if EPEND is after SOX signed into law on July 30, zero otherwise	
CPI	Annual avg CPI—All Urban Consumers, US city avg, All items, base:1982-84=100	BLS

*AT and SALE are adjusted for inflation by using annual CPI matched to fiscal year.

Exhibit 1
Excerpts from proxy statement of CTE announcing single-class recapitalization

Dear Shareholder:

At the Annual Meeting of Shareholders to be held on September 3, 2003, you will be asked to consider and vote upon a proposal to amend the Company's Articles of Incorporation to reclassify and convert each outstanding share of Class B Common Stock into 1.09 shares of Common Stock. The reclassification and related amendments to the Company's Articles of Incorporation are collectively referred to as the Charter Amendment…

…Level 3 currently owns 1,017,061 shares of CTE Class B Common Stock, or approximately 50.2% of the outstanding CTE Class B Common Stock, and owns no shares of CTE Common Stock. Level 3's shares of CTE Class B Common Stock represent 4.3% of the economic interest of the Company and 29.3% of the total voting power of the Company…If the Charter Amendment is implemented, Level 3 will own 1,108,596 shares of CTE Common Stock, or approximately 4.6% of the economic interest and voting power of the Company…All amounts in this section are based on the number of shares of CTE Class B Common Stock and CTE Common Stock outstanding on the record date of the Proxy Statement, June 20, 2003...

REASONS FOR THE CHARTER AMENDMENT

…Level 3 previously has stated publicly that it would consider monetizing certain of its non-core assets, including its holdings in public companies such as CTE…

…The Charter Amendment will better align the stockholders' voting rights with their economic interests and will establish a simplified one share/one vote capital structure…

…The Charter Amendment will eliminate the opportunity for an investor to gain voting influence that is disproportionate to its economic interest…

…Although the Charter Amendment will provide a fair premium to the holders of CTE Class B Common Stock, the dilution to the holders of CTE Common Stock from an economic perspective will be minimal…

…The Charter Amendment will provide the holders of CTE Class B Common Stock with a substantially more liquid security. The shares of CTE Class B Common Stock are thinly traded on the Nasdaq SmallCap Market, with an average trading volume of 1,197 shares per day during the twelve months ended July 10, 2003. By contrast, the CTE Common Stock trades on the Nasdaq National Market and had an average volume of 115,703 shares per day for the twelve months ended July 10, 2003…

…The Charter Amendment will increase the market float of the CTE Common Stock, which should, in turn, increase its liquidity, trading volume and trading efficiencies. Heightened liquidity may also increase the number of institutional holders that invest in the Company's equity…

…The Charter Amendment will simplify the Company's capital structure, which may generate increased investor interest and a larger investor base and would likely provide greater flexibility and efficiency in pursuing strategic acquisitions and equity financings if, when and to the extent desired or needed by the Company…

…The Charter Amendment would conform the Company's capital structure to that of most other publicly traded corporations and eliminate any perceived negative impact on the market price of the CTE Common Stock and CTE Class B Common Stock that results from having two classes of traded stock with generally similar characteristics except voting rights…

…If a change of control transaction involving the Company were to occur in the future, the Charter Amendment would increase the likelihood of a full takeover premium accruing to all shareholders equally...

…The Special Committee also considered the following factors relating to the Exchange Ratio: the current and historical trading prices and volumes of the CTE Class B Common Stock compared to CTE Common Stock; the trading price differentials between two classes of stock of other similarly situated companies in relation to the Exchange Ratio; and the premiums received in comparable dual class recapitalization transactions…

Exhibit 2
Excerpts from proxy statement of DiamondCluster announcing single-class recapitalization because of SOX

…The annual meeting of stockholders of DiamondCluster International, Inc. will be held on Tuesday, September 23, 2003 for the following purposes:
1. To elect three directors for a term of three years to replace directors whose terms are expiring and two directors for a term of two years to fill vacancies created by the retirement of a director and the resignation of a director.
2. To approve a plan of recapitalization to convert all Class A and Class B Common Stock into a new class of capital stock called Common Stock and adopt related amendments to our Restated Certificate of Incorporation.
3. To ratify the recommendation of the Audit Committee that KPMG LLP be appointed independent auditors for the Company for fiscal year 2004.

PROPOSED PLAN OF RECAPITALIZATION

…During the past year, our Board of Directors engaged in a detailed review of our system of corporate governance. The Board considered the effect of the Sarbanes-Oxley Act and the corporate governance standards being adopted for listed companies by the Nasdaq National Market. After a thorough review, the Board concluded that amendments to the Corporation's Restated Certificate of Incorporation and internal governance agreements should be made to assure continued independence of a majority of the Board and the proper functioning of the Audit and other committees of the Board…

…As a result of this process, the Board of Directors has approved, subject to the approval of our stockholders, a Plan of Recapitalization attached to this proxy statement as Annex B… Pursuant to this Plan of Recapitalization, we propose to convert all Class A and Class B Common Stock into shares of a new class of capital stock called Common Stock, and in connection with the conversion, amend and restate our Restated Certificate of Incorporation to (i) eliminate the Class A and Class B Common Stock and (ii) create a new class of capital stock called Common Stock . All issued shares of Class A Common Stock and all issued shares of Class B Common Stock will be converted into shares of Common Stock on a one-to-one basis…

…Our Class B Common Stock is identical to our Class A Common Stock except in two respects. Our Class A Common Stock is entitled to one vote per share on all matters submitted to a vote of holders of Common Stock whereas our Class B Common Stock is entitled to five votes per share. Additionally, our Class B Common Stock may be owned beneficially or of record only by permitted holders. The new class of Common Stock will be the same as the Class A Common Stock. It will be entitled to one vote per share on all matters submitted to a vote of holders of Common Stock and there were be no restrictions on the ownership or transferability of shares of Common Stock…

…Because each outstanding share of Class B Common Stock has five votes, the holders of the Class B Common Stock may significantly influence all matters brought to a vote of our stockholders. In addition, all of the holders of our Class B Common Stock have granted proxies to our Chief Executive Officer to vote their shares. Accordingly, the Chief Executive Officer currently has the right to vote 100% of the outstanding shares of Class B Common Stock…

…Our Board of Directors has determined that it is no longer in the Company's best interest for our Chief Executive Officer to have such voting control. The Board of Directors has approved this Plan of Recapitalization in order to bring our system of corporate governance more in line with the governance of other public companies. Implementation of this Plan will eliminate the voting control currently held by our Chief Executive Officer…

AUDIT COMMITTEE CHARTER

…The primary function of the DiamondCluster International, Inc. Audit Committee of the Board of Directors is to assist the Board in fulfilling its oversight responsibilities by reviewing and monitoring the integrity of the financial information which will be provided to the shareholders and others, the systems of internal controls which management and the Board of Directors have established, and the audit process…

…The membership of the Audit Committee shall consist of at least three independent members of the Board of Directors who shall serve at the pleasure of the Board of Directors. The members of the Audit Committee shall meet the independence and experience requirements of Nasdaq and the rules and regulations of the Securities and Exchange Commission "(SEC"). At least one member of the Audit Committee shall be a financial expert as defined by the SEC. Audit Committee members and the committee chairman shall be designated by the full Board of Directors…

Tables and Figures

Table 1
Summary statistics

Panel A: Variables

Time	Mean	Std. Dev.	Min	Max	Obs
Duration as dual-class firm before recap	7.87	7.91	0.84	44.95	176
Duration if censored	14.00	12.83	0.38	82.92	563
Ownership and financial variables as of last episode					
InsiderCF	0.39	0.24	0.00	0.99	726
InsiderV	0.57	0.29	0.00	1.00	726
Wedge	0.18	0.19	-0.34	0.97	726
Name	0.23	0.42	0.00	1.00	702
Book value assets (in $mil)*	1,914	9,452	0	150,355	720
Sales growth*	0.11	1.02	-2.69	14.34	704

Panel B: Differences by recap

	Recap	Remain dual-class	Difference (t-stat)
InsiderCF	0.40	0.36	0.04** (2.20)
InsiderV	0.60	0.48	0.12** (4.77)
Wedge	0.19	0.12	0.07** (4.33)
Name	0.25	0.17	0.08** (2.25)
Book value assets (in $mil)*	1,740	2,452	718 (-0.87)
Sales growth*	0.12	0.08	0.04 (0.21)

*Adjusted for inflation using annual avg CPI--All Urban Consumers, US city avg All items
Annual avg by calendar year matched to fiscal year; base: 1982-84=100
** Indicates significance level of 5%.

Table 2
Dual-class firms and recapitalizations into single-class

Panel A: Number of firms in the dual-class sample by year of proxy record date

Year	1994	1995	1996	1997	1998	1999	2000	2001	2002
Number of firms	109	398	434	485	504	484	489	432	379
Firms with two traded classes	33	85	84	81	78	81	74	68	62
Number of exits	6	21	33	47	57	62	72	62	
CRSP delisting code for exchange	0	3	1	1	1	6	5	4	7
Eliminated based on other delisting codes	2	5	17	23	32	22	23	17	17
Checked in SEC filings	4	13	15	23	24	34	44	41	355
Number of recapitalizations	2	9	10	13	17	20	25	18	62

Total number of dual-class firms in sample: 739
Total number of recapitalizations: 176
Number of firms that did not recap: 563
Number of remaining dual-class firms in 2009: 219

Panel B: Firms where exact date of becoming dual-class is available

	Two traded classes		One traded class	
Number of observations	112	100%	111	100%
Dual-class date same as first listing date	18	16%	103	93%
Dual-class date later than IPO	94	84%	8	7%

Panel C: Recapitalizations and censored observations in 1994-2009

	Before SOX (07/30/02)	After SOX
Recaps	116	60
Censored	304	259

Table 3
Choosing the parametric model

	1		2		3	
	PH Weibull		AFT Lognormal		PH Cox	
	Expected sign	Estimated Coefficient	Expected sign	Estimated Coefficient	Expected sign	Estimated Coefficient
INSIDERV	+	-0.486	-	0.406	+	-0.633*
		[0.335]		[0.415]		[0.331]
WEDGE	-	-2.534***	+	3.531***	-	-2.376***
		[0.586]		[0.676]		[0.589]
NAME	-	-0.597***	+	0.753***	-	-0.526**
		[0.206]		[0.246]		[0.205]
SIZE	+	0.026	-	-0.084**	+	0.03
		[0.040]		[0.042]		[0.040]
SALESGROWTH	-	0.099*	+	-0.072	-	0.115*
		[0.051]		[0.065]		[0.052]
AFTERSOX		0.986***		-1.522***		1.101***
		[0.177]		[0.359]		[0.180]
Constant		-7.595***		9.357***		
		[0.549]		[0.348]		
ln(p)		-0.229***				
		[0.071]				
p		0.796				
ln(sigma)				0.574***		
				[0.067]		
sigma				1.776		
AIC		1049.289		1012.444		1972.615
Observations		8961		8961		8961

Coefficients are not exponentiated and represent marginal effects, not hazard/time ratios.
p and sigma are the shape parameters in the Weibull and lognormal functions, respectively.
Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 4
The effect of SOX in the AFT Lognormal model

	1 Marginal effects	2 Time ratios
INSIDERV	0.406	1.501
	[0.415]	[0.623]
WEDGE	3.531***	10.219***
	[0.676]	[4.094]
NAME	0.753***	2.124***
	[0.246]	[0.522]
SIZE	-0.084**	0.919**
	[0.042]	[0.046]
SALESGROWTH	-0.072	0.931
	[0.065]	[0.061]
AFTERSOX	-1.522***	0.218***
	[0.359]	[0.078]
Constant	9.357***	
	[0.348]	
ln_sig	0.574***	0.574***
	[0.067]	[0.067]

Standard errors in brackets
*** p<0.01, ** p<0.05, *
p<0.1

Figure 1
Histograms of duration as dual-class

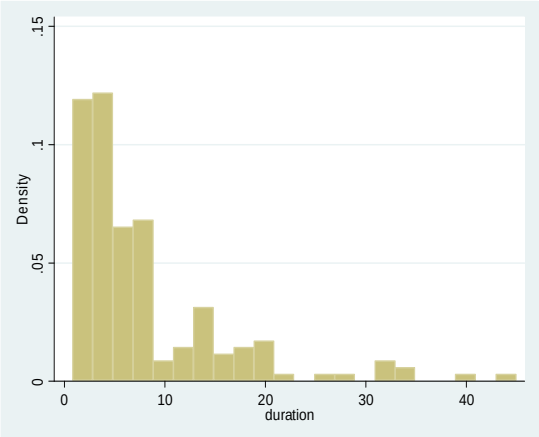


Fig. 2a: Recaps

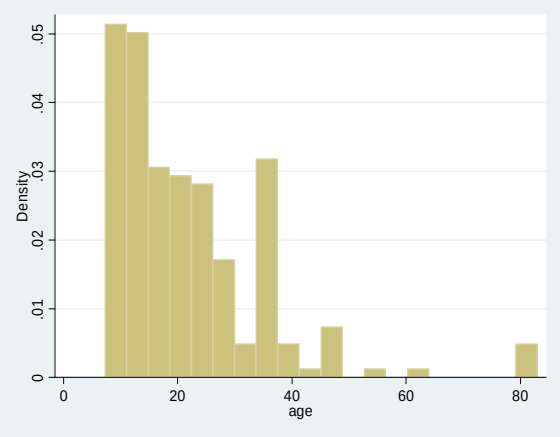


Fig. 2b: Censored

Figure 2
Right-censoring

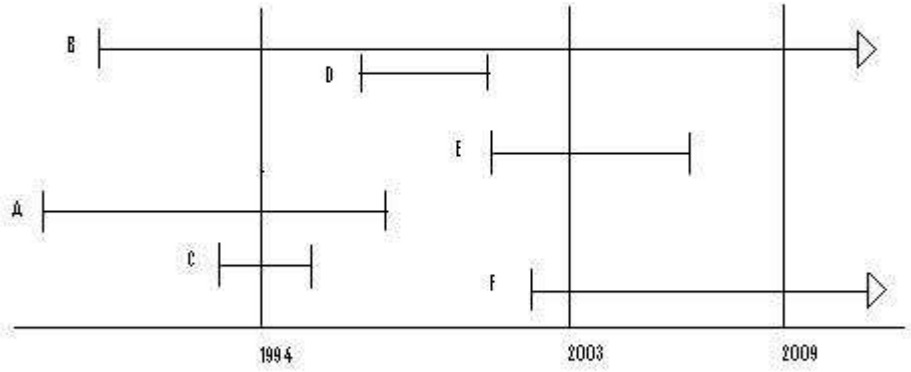


Figure 3
Hazard and survival functions of models in Table 3

Fig. 3a. Model 1: PH Weibull

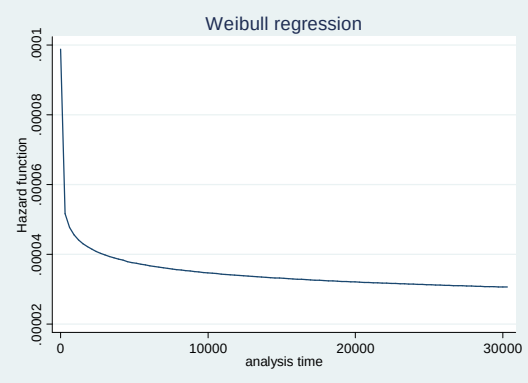


Fig. 3b. Model 1: PH Weibull

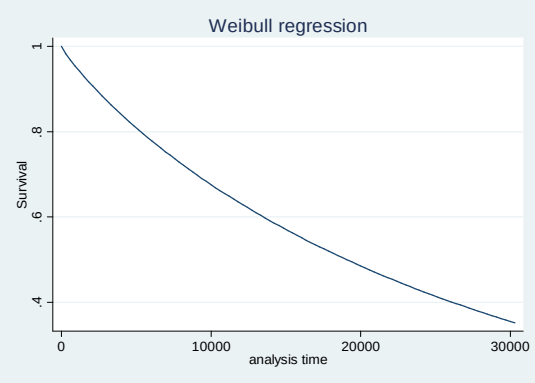


Fig. 3c. Model 2: AFT Lognormal

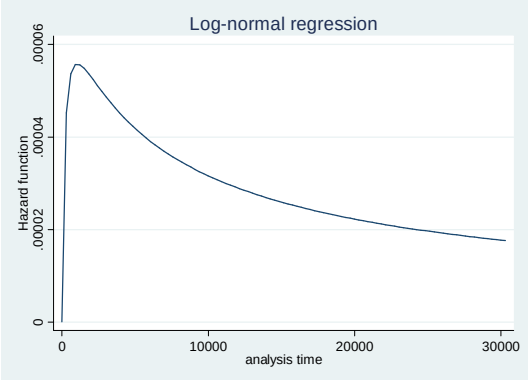


Fig. 3d. Model 2: AFT Lognormal

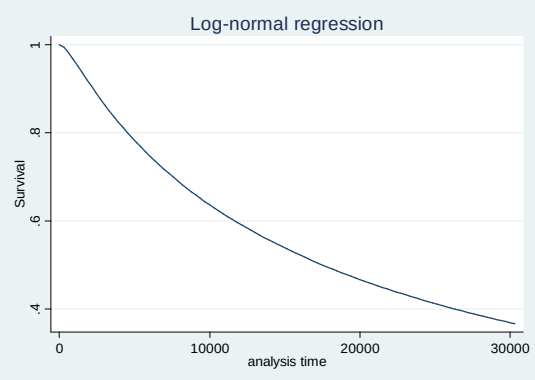


Fig. 3e. Model 3: PH Cox

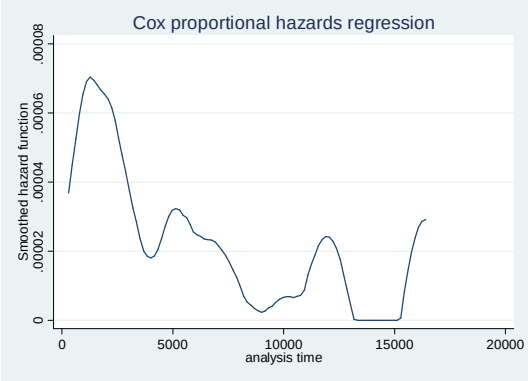


Fig. 3f. Model 3: PH Cox

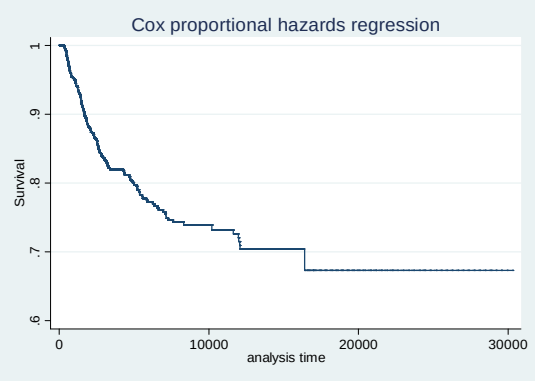


Fig. 3g. Model 3: PH Cox

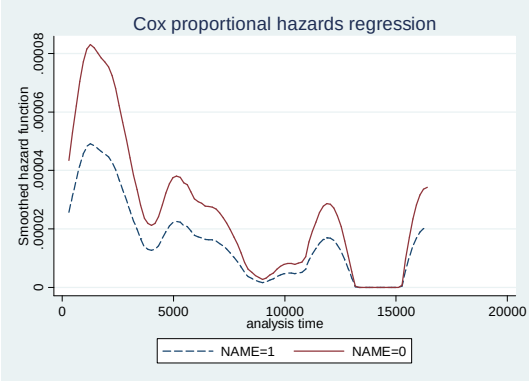


Fig. 3h. Model 3: PH Cox

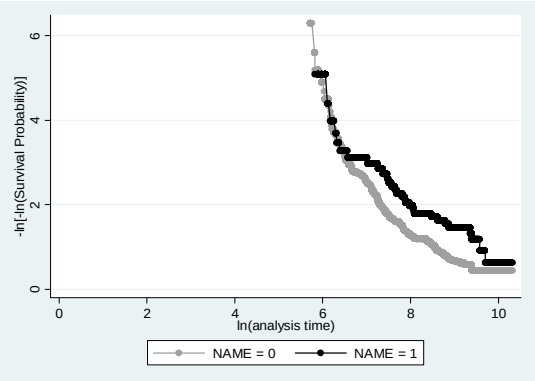


Figure 4
The effect of SOX on hazard and survival functions

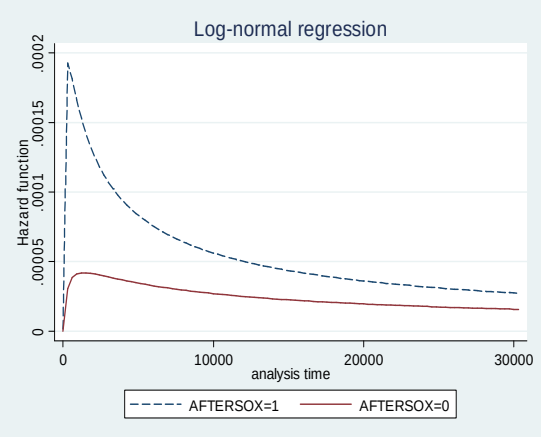


Fig. 4a

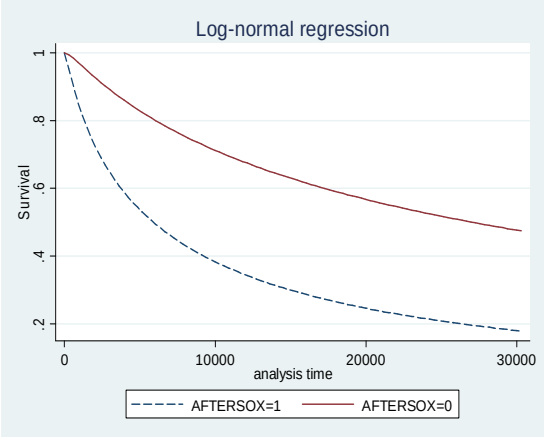


Fig. 4b

REFERENCES

Agrawal, A., Knoeber, C.R., 1996. Firm performance and mechanisms to control agency problems. The Journal of Financial and Quantitative Analysis 31, 377-397

Allison, P., 1984. Event history analysis. Sage University Paper 46

Amit, R., Villalonga, B., 2006. How do family ownership, control and management affect firm value? Journal of Financial Economics 80, 385-417

Amoako-Adu, B., Smith, B., 2001. Dual class firms: capitalization, ownership structure and recapitalization back into single class. Journal of Banking and Finance 25, 1083-1111

Basu, A., Terza, J.V., Rathouz, P.J., 2008. Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. Journal of Health Economics 27, 531-543

Baum, C.F., Schaffer, M.E., Stillman, S., 2003. Instrumental variables and GMM: estimation and testing. The Stata Journal 3, 1-31

Baum, C.F., Schaffer, M.E., Stillman, S., 2007. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. The Stata Journal 7, 465-506

Bebchuk, L.A., Cohen, A., 2005. The costs of entrenched boards. Journal of Financial Economics 78, 409-433

Block, S.B., 2004. The latest movement to going private: an empirical study. Journal of Applied Finance 14, 36-44

Bound, J., Jaeger, D., Baker, R., 1995. Problems with IV estimation when the correlation between the instruments and the endogenous explanatory variables is weak. Journal of Americal Statistical Association 90, 443-450

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

Burkart, M., Gromb, D., Panunzi, F.,1997. Large shareholders, monitoring, and the value of the firm. The Quarterly Journal of Economics, August, 693-728.

Burkart, M., Lee, S., 2008. One share-one vote: the theory. Review of Finance 12, 1-49

Cable, J., Holland, K., 2000. Robust vs. OLS estimation of the market model: implications for event studies. Economics Letters 69, 385-391

Chhaochharia, V., Grinstein, Y., 2007. Corporate governance and firm value: the impact of the 2002 governance rules. The Journal of Finance 62, 1789-1825

Claessens, S., Djankov, S., Fan, J.P.H., Lang, L.H.P., 2002. Disentangling the incentive and entrenchment effects of large shareholdings. The Journal of Finance 57, 2741-2771

Coates, J.C., IV, 2007. The goals and promise of the Sarbanes-Oxley Act. Journal of Economic Perspectives 21, 91-116

Coles, J.L., Lemmon, M.L., Meschke, J.F., 2007. Structural models and endogeneity in corporate finance: the link between managerial ownership and corporate performance.

Cornett, M.M., Vetsuypens, M.R., 1989. Voting rights and shareholder wealth: the issuance of limited voting common stock. Managerial and Decision Economics 10, 175-188

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

Cronqvist, H., Nilsson, M., 2005. The choice between rights offerings and private equity placements. Journal of Financial Economics 78, 375-407

Dahya, J., Dimitrov, O., McConnell, J.J., 2008. Dominant shareholders, corporate boards, and corporate value: a cross-country analysis. Journal of Financial Economics 87, 73-100

DeAngelo, H., DeAngelo, L., 1985. Managerial ownership of voting rights: public corporations with dual classes of common stock. Journal of Financial Economics 14, 33-99

Demsetz, H., Lehn, K., 1985. The structure of corporate ownership: causes and consequences. The Journal of Political Economy 93, 1155-1177

Demsetz, H., Villalonga, B., 2001. Ownership structure and corporate performance. Journal of Corporate Finance 7, 209-233

Dimitrov, V., Jain, P.C., 2006. Recapitalization of one class of common stock into dual-class: growth and long-run stock returns. Journal of Corporate Finance 12, 342-366

Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. Journal of Financial Economics 7, 197-226

Fahlenbrach, R., 2008. Shareholder Rights, Boards, and CEO Compensation. The Ohio State University, Fisher College of Business

Fama, E., French, K., 1993. Common risk factors in the returns in stocks and bonds. Journal of Financial Economics 33, 3-56

Fama, E.F., 1980. Agency problems and the theory of the firm. The Journal of Political Economy 88, 288-307

Fama, E.F., Jensen, M.C., 1983. Separation of ownership and control. Journal of Law and Economics 26, 301-325

Garson, D., 2009. Lecture notes. North Carolina State University

Gomes, A., 2000. Going public without governance: managerial reputation effects. The Journal of Finance 55, 615-646

Gompers, P.A., Ishii, J., Metrick, A., 2008. Extreme governance: an analysis of dual-class firms in the United States.

Grossman, S., Hart, O., 1988. One share-one vote and the market for corporate control. Journal of Financial Economics 20, 175-202

Harris, M., Raviv, A., 1988. Corporate governance: voting rights and majority rules. Journal of Financial Economics 20, 203-235

Hauser, S., Lauterbach, B., 2004. The value of voting rights to majority shareholders: evidence from dual-class stock unifications. Review of Financial Studies 17, 1168-1184

Heinkel, R., Kraus, A., 1988. Measuring event impacts in thinly traded stocks. The Journal of Financial and Quantitative Analysis 23, 71-88

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

Hochberg, Y.V., Sapienza, P., Vissing-Jorgensen, A., 2007. A lobbying approach to evaluating the Sarbanes-Oxley Act of 2002. In: NBER Working Paper No 12952

Jain, P.K., Rezaee, Z., 2006. The Sarbanes-Oxley Act of 2002 and security market behavior: early evidence. Contemporary Accounting Research 23, 629-654

Jarrell, G., Poulsen, A., 1988. Dual-class recapitalizations as antitakeover mechanisms: the recent evidence. Journal of Financial Economics 20, 129-152

Jenkins, S., 2005. Lecture notes. University of Essex, UK.

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., 1986. Agency costs of free cash flow, corporate finance, and takeovers. The American Economic Review 76, 323-329

Kalay, A., Pant, S., 2008. One share-one vote is unenforceable and sub-optimal.

Kamar, E., Karaca-Mandic, P., Talley, E., 2008. Going-private decisions and the Sarbanes-Oxley Act of 2002. In: RAND Institute for Civil Justice working paper series

Kaplan, S.N., Zingales, L., 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? The Quarterly Journal of Economics 112, 169-215

Keifer, N., 1988. Economic duration data and hazard functions. Journal of Economic Literature 26, 646-679

Kleibergen, F., Paap, R., 2006. Generalized reduced rank tests using the singular value decomposition. Journal of Econometrics 127, 97-126

Kole, S., 1995. Managerial ownership and firm performance: incentives or rewards. Advances in Financial Economics 2, 119-149

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 2002. Investor protection and corporate valuation. The Journal of Finance 57, 1147-1170

Lease, R.C., McConnell, J.J., Mikkelson, W.H., 1983. The market value of control in publicly-traded corporations. Journal of Financial Economics 11, 439-471

Lease, R.C., McConnell, J.J., Mikkelson, W.H., 1984. The market value of differential voting rights in closely held corporations. The Journal of Business 57, 443-467

Lehn, K., Netter, J., Poulsen, A., 1990. Consolidating corporate control: dual-class recapitalizations versus leveraged buyouts. Journal of Financial Economics 27, 557-580

Leuz, C., 2007. Was the Sarbanes-Oxley Act of 2002 really this costly? Evidence from event returns and going-private decisions. Journal of Accounting and Economics 44, 146-165

Li, H., Pincus, M., Rego, S.O., 2008. Market reaction to events surrounding the Sarbanes-Oxley Act of 2002 and earnings management. Journal of Law and Economics 51

Lins, K.V., 2003. Equity ownership and firm value in emerging markets. Journal of Financial and Quantitative Analysis 38, 159-184

Litvak, K., 2007. The effect of the Sarbanes-Oxley Act on non-US companies cross-listed in the US. Journal of Corporate Finance 13, 195-228

Masulis, R.W., Wang, C., Xie, F., 2006. Agency problems at dual-class companies.

Maynes, E., Rumsey, J., 1993. Conducting event studies with thinly traded stocks. Journal of Banking and Finance 17, 145-157

McConnell, J.J., Servaes, H., 1990. Additional evidence on equity ownership and corporate value. Journal of Financial Economics 27, 595-612

Mitchell, L.E., 2003. The Sarbanes-Oxley Act and the reinvention of corporate governance. Villanova Law Review 48, 1189

Morck, R., Shleifer, A., Vishny, R., 1988. Management ownership and market valuation: an Empirical Analysis. Journal of Financial Economics 20, 293-315

Moyer, R.C., Rao, R., Sisneros, P.M., 1992. Substitutes for voting rights: evidence from dual-class recapitalizations. Financial Management 21, 35-47

Nenova, T., 2003. The value of corporate voting rights and control: a cross-country analysis. Journal of Financial Economics 68, 325-351

Petersen, M.A., 2006. Estimating standard errors in finance panel data sets: comparing approaches.

Petersen, T., 1986. Fitting parametric survival models with time dependent covariates. Journal of Royal Statistical Society Series C 35, 281-288

Romano, R., 2005. The Sarbanes-Oxley Act and the making of quack corporate governance. Yale Law Journal 114, 1521

Scholes, M., Williams, J., 1977. Estimating betas from nonsynchronous data. Journal of Financial Economics 5, 309-327

Schwert, G.W., 1981. Using financial data to measure effects of regulation. Journal of Law and Economics 24, 121-158

Shleifer, A., Vishny, R., 1986. Large shareholders and corporate control. The Journal of Political Economy 94, 461-488.

Smith, G.P., 2007. A look at the impact of Sarbanes-Oxley on cross-listed firms. Arizona State University

Stein, J.C., 1998. Takeover threats and managerial myopia. The Journal of Political Economy 96, 61-80

Taylor, S., Whittred, G., 1998. Security design and the allocation of voting rights: evidence from the Australian IPO market. Journal of Corporate Finance 4, 107-131

Thompson, S.B., 2006. Simple formulas for standard errors that cluster by both firm and time.

Venkatesh, P.C., 1992. Empirical evidence on the impact of the bid-ask spread on the characteristics of CRSP daily returns. 15, 113-125

Wintoki, M.B., 2007. Corporate boards and regulation: the effect of the Sarbanes-Oxley Act and exchange listing requirements on firm value. Journal of Corporate Finance 13, 229-250

Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data. MIT.

Zhang, I., 2007. Economic consequences of the Sarbanes-Oxley Act of 2002. Journal of Accounting and Economics 44, 74-115

Zingales, L., 1995. What determines the value of corporate votes. The Quarterly Journal of Economics 110, 1047-1073