Information and Control in Financial Markets

Samuel Lee

EFI THE ECONOMIC RESEARCH INSTITUTE

Dissertation for the Degree of Doctor of Philosophy, Ph.D.
Stockholm School of Economics

KEYWORDS: Market liquidity, active investors, investment strategies, noise trading, information acquisition, information markets, takeovers, tender offers, one share - one vote, asymmetric information, signaling, free-rider problem, means of payment, restricted bids, two-dimensional types, private equity, venture capital, fund performance, learning in financial markets, takeover waves, delegated monitoring, social capital, social esteem, mutual altruism, loyalty, social networks, corporate governance, crony capitalism, family firms, boards of directors, organizational culture.

ISBN 978-91-7258-799-1

PRINTED BY:
Elanders, Stockholm 2009

DISTRIBUTED BY:
EFI, The Economic Research Institute
Stockholm School of Economics
P O Box 6501, SE-113 83 Stockholm
www.hhs.se/efi

The flipside of the downside… that would be the upside—na-na-na-na-na

Contents

Introduction

This thesis contains five research papers in finance. The papers were written between the summer of 2005 and the spring of 2009 while I was a PhD student in the finance department of the Stockholm School of Economics. Parts of the first paper were written in New York, during visits to the Columbia Business School in the fall of 2008 and the Stern School of Business in the fall of 2009. Similarly, parts of the fourth paper and the last paper as a whole were completed during the visit to Stern.

My field of interest is applied microeconomics, in particular contract theory and information economics. The present collection of papers uses microeconomic models to explore specific questions related to financial markets and corporate governance. The overarching theme is to understand how the distribution of control or information may influence transactions in financial markets or decisions within an organization.

In the first paper, I study how investors choose among investment strategies—which are based on different types of information—when their strategy choices affect the liquidity of the financial market. The common view is that, for a given informed investor, the presence of other informed investors is detrimental because it reduces the extent to which she can exploit her own information to profit from trading in the market. Therefore, the more informed investors there are, the less profitable it is to become informed. In the paper, the relationship between informed investors is more complex because investors who pursue different strategies do not necessarily harm each other. On the contrary, such investors may supply each other with liquidity. Depending on the circumstances, different investment strategies can therefore be substitutes or complements. A strategy is a substitute for another strategy when a decrease in its cost reduces the demand for the other strategy. Conversely, a strategy is a complement to another strategy when a decrease in its cost raises the demand for the other strategy.

The key result of the paper is that a strategy substitutes for other strategies when it is relatively expensive, but complements them when it is relatively cheap. Thus, as an expensive strategy becomes cheaper, it first crowds other strategies out of the market but later invites them back in. This has implications for various financial market phenomena, such as herding behavior among investors, how prices of different assets move together, and how informative prices are about the true value of the traded assets. It also offers a new perspective on the link between financial markets and markets for information. On the one hand, a monopolistic information seller may sell information cheaper to investors to crowd out other information sources, thereby inducing herding behavior. On the other hand, competition makes particular types of information widely available, thereby creating liquidity for investment strategies based on other types of information. Thus, competition fosters diversity. Finally, financial institutions may sell information to improve liquidity in the market, so that they can better trade on more exclusive information which they do not share with other investors.

In the second paper, Christian At, Mike Burkart and I study whether a potential acquirer, who wants to make a tender offer to acquire a target company, is more or less likely to be successful when the target company has a dual-class share structure. In a dual-class share structure, only a fraction of the outstanding shares is endowed with voting rights. So, the acquirer need not buy as many shares to gain control because she need not extend the tender offer to the holders of the non-voting shares. The question how a dual-class share structure affects tender offers has already been analyzed by Grossman and Hart (1988). In a setting with competing acquirers, they show that a single-class share structure (one share - one vote) is the socially optimal structure which ensures that control eventually goes to the party that maximizes the value of the target company. We revisit the question in a setting without competition. Instead, we assume that the acquirer is better informed than the dispersed shareholders of the target company about what the company would be worth after the takeover. Due to this information asymmetry, some value-improving takeovers may not take place. The main result of the paper is that a dual-class share structure can mitigate this asymmetric information problem. They key to this result is that, while the shareholders of the target company are concerned about receiving (what they perceive to be) a "fair" price for the *dividend* rights that they give up, they do not care about the *voting* rights per se. When they suspect that the price is "unfair", they are only unwilling to sell their dividend rights but—by construction—cannot transfer their voting rights either. A dual-class share structure resolves this problem as it reduces the amount of dividend rights that are attached to the voting rights, and hence the extent to which disagreement about the value of the divident rights impedes the transfer of the voting rights. Thus, unlike in Grossman and Hart (1988), one share - one vote need in general not be the socially optimal security-voting structure.

In the third paper, Mike Burkart and I address a question that arose while we wrote the second paper, namely whether the acquirer can use other means to reveal (signal) what she knows about the impact of the takeover on the value of the dividend rights. To this end, we allow for a richer setting than in the previous paper, analyzing more generally the possibility of signaling in models of tender offers. It turns out that the bidder's signaling ability depends on how the value gains from a takeover are divided into a dividend value (security benefits), which the acquirer cannot withhold from other shareholders, and a control value (private benefits), which the bidder can exclusively keep for herself if she wants to. Central to this result is the so-called free-rider problem: if the price offered per share is smaller than what the share would be worth after the takeover, each (small) shareholder prefers to keep her shares, hoping that the other shareholders will sell their shares so that she can benefit from the *full* increase in the share value. To buy any shares, the acquirer must therefore pay what the shares are worth after the takeover, in which case all the gains in security benefits go to the shareholders. This conflicts with a principle in contract theory, which says that signaling superior information must in general be *costly* to the signaler (if she otherwise would lie). Since the acquirer must forgo all the security benefits, she is not able to signal information by forgoing security benefits. Thus, private benefits are a prerequisite for signaling. Still, they only serve a signaling purpose if the acquirer can *commit* to relinquish private benefits in a manner that allows the shareholders to *infer* the value of the security benefits. This is never satisfied when she has *additional* private information about the value of the private benefits. However, we show that she can design the tender offer in a way that relates forgone private benefits to the security benefits in a predictable manner. For example, dilution, leverage, and toeholds are viable signals, while restricted bids and cash-equity offers in general are not.

The above signaling solutions suffer from the problem that they rely on the acquirer's private benefits. When the benefits are small, relinquishing them may fail to convince the shareholders to accept what actually is a fair price. Hence, we explore an alternative solution which includes offering shareholders derivatives for their shares. Using derivatives as a means of payment de facto allows the acquirer to buy the target shares for cash, strip them of their votes, and restructure the dividend rights into claims that she wants to keep and claims that she wants to return to the shareholders. The first steps give her control, while she can use the last step to self-impose penalties for "lying" about the security benefits. In particular, call options enable the shareholders to seek "damages" from the acquirer if the security benefits later turn out to be worth more than what they were paid in cash. This solution does not rely on private benefits and, more importantly, fully resolves the asymmetric information problem.

In the fourth paper, Daniel Sunesson and I seek to explain a number of empirical facts about the private equity market. A private equity firm is a specialized investment firm which buys, reorganizes, and sells companies. To this end, it raises a private equity fund. The fund looks for suitable target companies and, when a company is found, bargains with its shareholders over the price at which it can buy (a control stake in) the company. Once in control, the fund takes measures to improve the firm in order to profitably re-sell it in the future. Because private equity funds have a finite lifetime, fund activity can fluctuate considerably over time depending on whether existing funds are succeeded by follow-on funds and on whether new funds enter the market.

We develop a simple model which captures the above features. A fixed number of companies becomes improvable, and the improvement can (only) be implemented by private equity firms. The true gains from reorganizing the companies are unknown but they can be inferred from the outcomes of completed reorganizations. Finally, there are many private equity firms which differ in their ability to manage a fund. Over time, each firm must repeatedly decide whether to run a fund or not. The predictions of the model are consistent with empirical evidence. **(1)** Overall fund activity follows wave patterns, whereby periods of little activity are occasionally interrupted by a period of growth (slow boom) followed by a crash (sudden bust). After events that generate high reorganization gains, the few private equity funds that are initially active earn promising returns, which attracts more funds to the market (learning). At the same time, as more funds enter, the pool of target companies is depleted faster (attrition), which ultimately leads from the boom to the bust. **(2)** Fund activity and the valuation levels of target companies move together. When the market becomes more bullish about the reorganization gains, potential target companies increase in value, which in turn affects the negotiations between funds and those companies. Thus, a rise in market confidence not only attracts more private equity funds but also raises the price that these funds must pay to buy companies. **(3)** A period in which the funds as a whole performed well—on the one hand—increases entry in the next period and—on the other hand—decreases subsequent performance by raising prices. That it, good industry performance precedes high entry, which in turn precedes low industry performance. **(4)** Because private equity firms differ in ability, there are persistent performance differences among funds. More interestingly, entry and exit by funds follow a last-in-first-out pattern: the least capable firms are the latest to enter and, by the same token, the earliest to exit. At any point in time, the first-time funds (the latest entrants) are thus the worst performers. If the boom continues, their follow-on funds *relatively* improve as even less capable firms enter the market. The least capable firms enter after highly profitable periods, when valuation levels are high, and during what later turn out to "peak" periods. Firms that raise their first fund during such periods are unlikely to raise a follow-on fund. **(5)** At a given point in time, variation in the size of funds reflects variation in the ability of the private equity firms that run them. Fund size and fund profitability are therefore positively correlated across contemporaneous funds. In contrast, across time, variation in the size of a fund reflects variation in market expectations. When markets become more bullish, a fund increases in size but its profitability drops because of higher prices and the higher costs of running a larger fund. Thus, fund size and fund profitability are negatively correlated across consecutive funds run by the same firm. **(6)** When markets are bullish, prices rise not only because of higher valuation levels but also because of increased fund competition. When market expectations overshoot, it can thus happen that "too much capital chases too few deals."

In the fifth and last paper, Petra Persson and I study a classic governance issue, the problem of delegated monitoring: one party (the principal) hires a second party (the supervisor) to monitor the actions of a third party (the agent). For instance, shareholders appoint board members to check on management; the public appoints a regulator to supervise banks; and banks hire risk managers to keep an eye on the loan officers. While the delegated monitoring problem has been studied before, notably by Diamond (1984) and Holmstrom and Tirole (1997), we revisit the problem to analyze how the supervisor and the agent would behave if they were friends, and what implications this has for whether or not the principal would want them to be friends. We believe that this is important because many situations involving issues of control and governance take place in a social context. Corporate boards are entrenched in social networks, employees in most companies have social relationships with each other, and politicians and corporate elites often convene in the same social circles. In fact, probably most of us mix business with pleasure as our professional and personal networks overlap. However, we know very little about whether—or in what contexts—such social connections are economically desirable; empirical evidence suggests that a given social connection can be beneficial in some contexts but harmful in others.

In the model, the supervisor (she) is hired because there is a conflict of interest between the agent (he) and the principal (she); unless the supervisor monitors the agent, he does not choose the action that the principal wants him to implement. We embed this problem in the following social structure: The supervisor and the agent are friends. Furthermore, the supervisor cares about whether the principal views her as a person of high integrity, i.e., as having fulfilled her monitoring duty. Our analysis yields three main results. **(1)** Stronger friendship between the supervisor and the agent has an ambiguous effect on the quality of the implemented action. Consistent with common sense, friendship makes the supervisor more reluctant to perform her duty as a monitor and to intervene against the agent's interest, so that fewer bad decisions are detected and corrected. We refer to this effect as *capture*. However, friendship also affects the agent's behavior. In case of a bad outcome, the principal puts some blame on the supervisor and hence thinks less highly of her. When the agent is a friend of the supervisor, she partly internalizes the "shame" felt by the supervisor, which makes the agent reluctant to misbehave in the first place. We refer to this effect as *loyalty*. **(2)** The principal always prefers a *full loyalty* situation in which the supervisor never monitors but the agent, on his own accord, acts in the principal's interest. Since this requires a sufficiently high level of friendship, it is not always feasible. In that case, the principal often prefers a *full monitoring* situation in which the supervisor monitors the agent intensively, which requires a low level of friendship. Thus, the principal has a preference for extremes. **(3)** The full loyalty situation requires a lower level of friendship when the supervisor cares a lot about her image. A given social connection between an agent and a supervisor can therefore have very different implications for the principal, depending on the importance of social esteem. This suggests that overlapping personal and professional ties can enhance delegated governance in cultures or contexts where image concerns are important. In the paper, we apply this insight to issues related to crony capitalism, corporate governance, and organizational culture.

This thesis would not have been possible without my advisor and co-author, Mike Burkart, who taught me as much in academia as in the pub, from minor matters such as writing a research paper to more important matters such as how to properly pronounce "Cagliari", how to carry a refrigerator, and when to fill the dishwasher. That which I did not learn from Mike, such as humming for no reason, wearing a bike helmet, and using female pronouns, I was taught by you-know-who-you-are with whom my adventure shall continue. Then there are my third partner-in-crime, Daniel Sunesson, who can walk a marathon and still watch a four-hour opera standing up, and the rest of the gang, Linus Kaisajuntti, Reimo Juks, and Linus Siming whose NY escapades we shall one day see on a movie screen. I also want to mention everyone at SITE, ECGTN, and Erik Berglöf who brought me to Stockholm. Last but not least, a bow and kisses to Umma, Abba, and Johipohi for their love and support from afar. I plan to see you more often now, hopefully in New York. Sunshine, here I come.

Papers

PAPER 1

Market Liquidity, Active Investment, and Markets for Information

Abstract. This paper studies a financial market in which active investors choose among investment strategies that exploit information about different fundamentals. The presence of other active investors generates illiquidity. Though active investors who pursue different strategies can serve as quasi-noise traders for each other, thereby also supplying each other with liquidity. Different strategies can therefore be substitutes or complements. Such externalities in information acquisition have implications for herding behavior, price comovement, liquidity commonality, trade volume, and the informational role of prices. They also affect the role of markets for information. Information market competition fosters investor diversity, while a monopolistic vendor expands sales only to induce investor herding. Furthermore, in order to benefit from quasi-noise trading, a financial institution may engage in both proprietary trading and information sales.

1. Introduction

Many financial market trades are motivated by the desire to profit from superior information about the value of a traded asset, and a key role of asset prices is to reflect the information that underlies these trades. Since there is a plethora of information that is relevant for the value of an asset, active investment managers specialize in a variety of investment strategies.[1] The present paper addresses a number of questions which arise when investment approaches vary along such lines: How do investors select among alternative strategies? How do investors who have chosen different strategies interact,

Thanks to Philippe Aghion, Ulf Axelson, Sabrina Buti, Douglas Diamond, Thierry Foucault, Larry Glosten, Ludovic Phallipou, Petra Persson, Tarun Ramadorai, Per Strömberg, Matti Suominen, Stijn van Nieuwerburgh, Laura Veldkamp, and seminar participants at the European Econometric Society Winter Meetings (Cambridge), European Finance Association Meetings (Athens), Nordic Finance Network Workshop (Helsinki), NYU (Stern), MIT (Sloan), Chicago (Booth), Northwestern (Kellogg), UPenn (Wharton), UC Berkeley (Haas), OSU (Fisher), and the University of Amsterdam for helpful comments. Financial support from the European Corporate Governance Training Network and the Jan Wallander Foundation is gratefully acknowledged.

[1] Differences in fund managers' expertise may in part be reflected by so-called *investment styles* (Chan et al., 2002; Barberis and Shleifer, 2003; Goetzmann and Brown, 2003).

and how do they influence market prices? What is the relation between investment diversity and markets for financial information?

The common view is that different active investors compete with each other when they trade in the same market: by revealing information to the market, the active investors dilute each other's information advantage vis-à-vis uninformed market participants (Grossman and Stiglitz, 1980). Consequently, every active investor's expected profit decreases in the overall number of active investors. A key assumption for this result is that differences in investors' information only arise from noise, so that no group of investors possesses *distinct* information which can by no means be elicited from other investors. The departing point of the present paper is that, once this assumption is abandoned, active investors do not necessarily impede each other.

On the contrary, in a world where investors choose from various information strategies, diversely informed investors can potentially benefit from each other. To be specific, investors who acquire one type of information can sometimes gain from an increase in the population of investors who trade on another type of information. The two investment strategies can therefore be substitutes or complements. As further discussed in the paper, such externalities in information acquisition have implications for various financial market phenomena, such as herding behavior, price comovement, liquidity commonalities, and liquidity spirals. They also provide a novel rationale for the sale of financial information, and for why information market *competition* may be an important determinant of financial market efficiency.

The paper studies a model in which the channel for these externalities is market liquidity. The baseline model involves a static Kyle (1985) market for a single asset. The fundamental value of the asset is determined by two factors. Investors choose among trading strategies that are each based on one of the factors. For the sake of illustration, one can picture the factors as macro-economic and company-specific events, and the investors as specializing either in "macro-trading" or in "stock-picking". Uninformed investors therefore participate in a market with two classes of active investors, each of which possesses a distinct information advantage. In the course of trading, this information asymmetry generates illiquidity, meaning that the investors' orders have an impact on the price. Although the two active investor classes may contribute differently to the illiquidity in the market, all investors—informed and uninformed alike—are exposed to the same illiquidity.

The illiquidity in the market reflects the market-maker's attempt to infer (any type of) value-relevant information that investors impound into the order flow. Trading activity based on either investment strategy improves the information content of the order flow, thereby making the market less liquid than it would be in the absence of the strategy. In this sense, the two strategies deprive each other of liquidity. Nevertheless, provided that both strategies are employed, investors in one class can benefit from more active trading in the other class: Investors who trade on different fundamentals submit orders that are relatively uncorrelated with each other's information. That is, each investor class generates something akin to "noise trading" with respect to the other class. As a consequence, investors in one class can conceal their information better when the trade volume generated by the other class becomes large. Thus, active investors with different strategies can provide each other with liquidity.

When the investors choose what information to acquire prior to trading, the information choice inherits the strategic interdependencies that characterize the subsequent trading game (Hellwig and Veldkamp, 2008). Hence, the aforementioned liquidity linkages sometimes render the two investment strategies substitutes and sometimes complements. When evaluating a strategy, each investor weighs the *cost* of acquiring the pertinent information against the *scarcity* of the information relative to what other investors have chosen. All else equal, every investor prefers a cheap strategy. But as more investors opt to compete in the cheap strategy, the more expensive strategy gradually gains in appeal. Nevertheless, due to the illiquidity created by the investors who employ the cheap strategy, the expensive strategy may prove unprofitable even though it would be profitable in the absence of the cheap strategy. By contrast, if many investors can afford the cheap strategy, their combined trading activity may provide enough liquidity for the expensive strategy to become profitable (again). Thus, a decrease in the cost of a strategy can dampen or stimulate the demand for the other strategy.[2]

These externalities in information acquisition generate interesting predictions about the relation between information costs and several financial market characteristics:

Diversity of investment strategies. Herding (defined as an increased correlation across trades) can occur when information costs decrease or when they increase. The former case is accompanied by an increase in the number of traders and more likely occurs in illiquid markets, whereas the latter case is accompanied by a decrease in the number of traders and more likely occurs in liquid markets. Furthermore, expensive

[2] In most financial market models with endogenously informed traders, information choices exhibit strategic substitutability. Barlevy and Veronesi (2000) and Li and Yang (2008) are two notable exceptions where strategic complementarity arises with respect to the acquisition of the same type of information. Similarly, in Hellwig and Veldkamp (2008), complementarity means that investors want to know what others know, whereas substitutability means that investors want to know what others do not know. In contrast, the present paper analyzes externalities *across* different types of information: one type of information is a substitute (complement) to another type of information, if a decrease in the cost of the former decreases (increases) the demand for the latter.

investment strategies tend to emerge when there already are many active investors (following other investment strategies).

Trading volume and liquidity spirals. A general and continuous decrease in the costs of information propagates existing strategies and spawns new strategies. As a result, total trade volume tends to increase, average trade volume tends to decrease, and—along with the new strategies—new large volume investors tend to emerge. Improvements in the information environment of financial markets may thus explain the trends in stock trading volume documented by Chordia et al. (2008b). Furthermore, in markets that accomodate complementary strategies, adverse cost shocks to one investor class can trigger liquidity spirals: As the class reduces its trade volume, it withdraws liquidity from other investor classes, forcing them to reduce their trade volumes as well. This in turn withdraws liquidity from the first class leading to another round of contractions in trade volume. The spiral not only amplifies the impact on the directly affected strategy but also transmits the shock to other strategies.

Synchronicity in price and in liquidity. When there are several assets with values that are driven by a common factor and asset-specific factors, the investors can specialize on common information or on asset-specific information. The cost of each information type affects to what degree the prices and the liquidities of the different assets comove. Consistent with evidence in Chan and Hameed (2006), assets that receive more attention may comove more with the market unless that attention coincides with a greater dispersion of beliefs. The model also suggests a potential reason for why the liquidity of larger stocks may be more sensitive to, but less explained by, common variation than that of smaller stocks (Chordia et al., 2000; Hasbrouck and Seppi, 2001).

Market efficiency and price informativeness. A decrease in the cost of one information type can make the price *less* informative about the asset's value even though market efficiency improves. While more investors choose to become informed, they opt for more similar strategies. Thus, more of the information acquired by the investors is reflected in the price, but the investors as a whole may acquire less information.

Externalities in information acquisition also yield a novel rationale for selling information. Instead of producing information privately at some cost, investors may purchase information from commercial information vendors, who *can* offer the information at a lower price because of scale economies and the negligible marginal cost of duplicating information. However, the key question is whether a vendor is *willing* to offer a lower price. In a similar framework with a single information type, Admati and Pfleiderer (1988) show that a monopolist would cannibalize on her own revenues if she were to increase the competition among her customers. She therefore sets the highest possible price, so that the financial market is de facto the same as without the information sales.[3] This result naturally carries over to the present setting if the monopolist sells information about both factors.

When an information vendor cannot provide investors with all types of valuation-relevant information, she has an incentive to sell information. A monopolist who produces and sells information about only one of the factors reduces her sales price *as long as* the investment strategy based on the sold information is a substitute for alternative investment strategies. Her motivation for expanding sales is to crowd out the acquisition of alternative information, which allows her to absorb more of the total demand for information and to reduce the illiquidity that other investors impose on her customers. Thus, the information sales promote investment activity at the expense of investment diversity. This makes the asset market more efficient but asset prices potentially less informative (see point *d.* above).

Competition by a vendor who sells another type of information promotes diversity. More interestingly, this is also true for competition by a vendor who sells the *same* type of information. When the vendors offer perfect substitutes, they are more concerned with undercutting each other than with crowding out alternative investment strategies. The price battle propagates investment based on the sold information, and the increase in trade volume increases liquidity for alternative information sources. That is, information market competition generates financial market liquidity, which in turn fosters investor diversity. This supports the view that competitive business media play an important role in financial market development. It also provides a possible explanation for cross-country differences and time trends in the levels of stock price comovement (Morck et al., 2000; Campbell et al., 2001).

When information sales generate liquidity, a financial institution may want to engage in both proprietary trading *and* information sales. Consider a situation in which one investor (she) has exclusive information and another investor (he) has more common information. As it turns out, the exclusive investor is willing to compensate the common investor for giving her information away *for free*. This creates a "herd" of investors who camouflage the exclusive investor's trades to the extent that he gains more than the common investor loses. For a financial institution, a benefit of supplying many investors with mundane information may thus be increased liquidity, which in turn facilitates trading on proprietary information, i.e., information that the institution does not share with outside investors.

[3] Strictly speaking, as the authors show, the no-sale result requires risk-neutrality.

The present paper belongs to the literature on information acquisition in financial markets as well as to the related literature on information sales in financial markets.[4] It revisits both of these themes in a setting where investors choose among different types of information and differently informed investors impose liquidity externalities on each other. The model extends Subrahmanyam and Titman (1999) by allowing investors to choose among the different types of information and by introducing a market for information. The information distribution is thus completely endogenous. Thematically, the paper is most closely related to Fishman and Hagerty (1995) and to Veldkamp (2006a,b). Fishman and Hagerty (1995) first show that sales of financial information can be a means to capture a larger share of overall trading profits. But due to a single-information structure, they cannot address positive externalities among active investors, the diversity of investment strategies, or the crowding out of information. Veldkamp (2006a,b) considers different types of information. In her model, economies of scale in a competitive information market cause information types in higher demand to be supplied at lower prices. Thus, investors who acquire the same information enjoy a positive information market externality. This can lead to herding when the investor population is finite. By contrast, in the present paper, herding behavior results from negative asset market externalities and can occur even though the investor population is infinite.

The remainder of the paper is organized as follows. Section 2 describes the model and the equilibrium in the absence of information sales. Section 3 discusses the implications for financial markets. Section 4 introduces information sales. Section 5 contains concluding remarks. All mathematical proofs are presented in the Appendix.

2. Model

2.1. Active Investment. A single asset with uncertain liquidation value $\tilde{V} \sim N(0, 2\sigma^2)$ is traded. The liquidation value is determined by a pair of fundamental factors, $\Theta = \{A, B\}$. For simplicity, I assume that the factors are independent and equally important. That is, $\tilde{V} = \sum_\Theta \tilde{V}_\theta$ with $\tilde{V}_\theta \overset{iid}{\sim} N(0, \sigma^2)$ for $\theta \in \Theta$.[5]

Investors belong to one of two classes. Each class is informed about a different fundamental. The size of class $\theta \in \Theta$ is n_θ. The ith investor in class θ receives data about V_θ and interprets it with some idiosyncratic bias $\tilde{\epsilon}_{i\theta} \overset{iid}{\sim} N(0, \sigma_\epsilon^2)$. Her information

[4] The large literature on endogenously informed speculators includes *inter alia* Grossman and Stiglitz (1980), Verrecchia (1982), Hellwig and Veldkamp (2008), and Van Nieuwerburgh and Veldkamp (2008a). Examples of the somewhat smaller literature on markets for financial information are Admati and Pfleiderer (1986, 1988, 1990), Allen (1990), Garcia and Sangiorgi (2007), and Cespa (2008).

[5] Such a setting is used in Subrahmanyam and Titman (1996) and Bernhardt and Taub (2006).

is thus a signal $\tilde{s}_{\theta i} \equiv \tilde{V}_\theta + \tilde{\epsilon}_{\theta i}$. In short, investors are sorted into classes with different types of expertise, and individual biases induce some heterogeneity within each class.

Noise traders form a third investor category. Their motives for trading are exogenous to the model, and their total demand for (or supply of) the asset is $\tilde{y} \sim N(0, \sigma_y^2)$. The probability distributions and the class sizes are commonly known. All investors are risk-neutral and there is no discounting.

Trading proceeds as in Kyle (1985). All investors submit quantity orders to a competitive market maker. Order submission is non-cooperative, simultaneous and anonymous. After observing the aggregate order flow, the market-maker sets a uniform price at which she meets the orders. Finally, V becomes public. Yet, V_A and V_B are not observed individually. So, they cannot be traded separately.

It is standard to solve the game for the Bayes-Nash equilibria in linear and symmetric strategies. In such equilibria, each informed investor's strategy $x_{i\theta} = \alpha_{i\theta} s_{i\theta}$ is linear in her signal; and the market-maker's pricing rule $p = \lambda z$ is linear in the net imbalance of the order flow $z \equiv \sum_\Theta \sum_1^{n_\theta} x_\theta(s_{i\theta}) + y$. Moreover, investors in the same class follow the same strategy, $\alpha_{i\theta} = \alpha_\theta$ for $\theta \in \Theta$. A strategy profile is thus a triple $(\alpha_A, \alpha_B, \lambda)$.

The "trading intensity" coefficient α_θ captures how much the order flow of an investor in class θ varies with her information. The "price impact" coefficient λ in turn gauges how sensitively the price reacts to any variation in the order flow. The inverse $1/\lambda$ is a measure of "market liquidity".

LEMMA 1 (Subrahmanyam and Titman, 1999). *There exists a unique Bayes-Nash equilibrium of the trading game in linear and symmetric strategies:*

$$(2.1) \qquad \alpha_\theta^* = \frac{\sigma_y}{\lambda^* \left[(n_\theta + 1)\sigma^2 + 2\sigma_\epsilon^2\right]} \qquad \textit{and} \qquad \lambda^* = \frac{\sigma^2}{\sigma_y} \left[\sum\nolimits_\Theta T(n_\theta)\right]^{\frac{1}{2}}$$

where

$$(2.2) \qquad T(n_\theta) \equiv \frac{n_\theta \left(\sigma^2 + \sigma_\epsilon^2\right)}{\left[(n_\theta + 1)\sigma^2 + 2\sigma_\epsilon^2\right]^2}, \qquad \theta \in \Theta.$$

Individual investors in the same class engage in a Cournot-type competition. Since they trade on similar information, they reinforce each other's impact on the price. To mitigate the cumulative impact, each of them cuts back her individual order. Thus, when the classes differ in size, individual investors in the larger class trade less intensively ($n_{\theta'}^* > n_{\theta''}^* \Rightarrow \alpha_{\theta'}^* < \alpha_{\theta''}^*$).

Investors from different classes do not compete in the above sense, since their trades are uncorrelated ex ante. They nonetheless affect each other. Since the market maker expects informed orders from either class, both classes contribute to the illiquidity in the market [via the subfunctions $T(\cdot)$]. Market liquidity is thus a channel for inter-class externalities. These externalities arise because the investors, even if they possess unrelated information, must trade in the same market.

As in single-class models, the relation between the number of informed traders and market liquidity is ambiguous. The cumulative order flow of more investors conveys more (precise) information, as the idiosyncratic biases tend to offset each other. At the same time, the higher number of traders simply leads to a larger, more volatile order flow. These two effects have countervailing consequences for market liquidity. Since the information effect gradually vanishes, $T(\cdot)$ is strictly quasi-concave and has an interior maximum in $\mathbb{R}^+$.

COROLLARY 1. *λ^* is unimodal in $n_\theta \geq 0$ for all $\theta \in \Theta$.*

The corollary states that the externality that a given investor class imposes on all other investors can increase or decrease in the size of the class.[6] (It also implies that changes in n_A and n_B can have opposite effects on market liquidity.) The non-monotonicity has interesting consequences for the formation of investor classes, to which we turn in the next section.

The assumption that two investor classes receive independent information is not crucial. What matters is that competition is more intense within a class than across classes, and that a class as a whole affects liquidity in a non-monotonic way. By isolating these effects, the independence assumption merely accentuates the results. Neither are the externalities unique to the above market microstructure. For example, Bernhardt and Taub (2006) show that, if the factors are independent, the interaction *between* investor classes is the same for quantity orders and price-quantity schedules.[7]

2.2. Investment Specialization. Instead of being endowed with information, investors must now actively acquire information. Formally, the model is extended to include stage 0, in which investors either choose to remain uninformed or to conduct a fundamental analysis of the asset. Each active investor can only specialize in the analysis of one fundamental factor, and a truthful exchange of private signals among investors is not enforceable.[8] These assumptions are discussed in section **2.3**.

[6] A similar non-monotonicity arises when investors are risk-averse (Subrahmanyam, 1991).

[7] As regards the interaction *within* a class, investors who submit price-quantity schedules would compete more fiercely investor who submit pure quantity orders. This suggests that the (information and magnitude) effects of a growing class size on liquidity would be accelerated. For competitive rational expectations models, I conjecture that investors in one class can benefit from the growth of a sufficiently different class, because they learn more from the price and face less residual uncertainty.

[8] Hong et al. (2007) also study a model in which asset values depend on several factors but investors update their beliefs over the class of single-factor models. In their words, the investors make "simple forecasts".

Fundamental analysis is costly. To produce information about a factor $\theta \in \Theta$, an investor must incur some fixed cost $c_\theta > 0$ to gather and interpret data. For instance, one can think of c_A as the cost of macro-economic analysis and c_B as the cost of company-specific analysis. Accordingly, active investors can be seen as either "macro-traders" or "stock-pickers."

Investors choose their specialization to maximize expected profit

$$\pi_\theta\left(n_\theta, n_{\theta'}\right) = \rho_\theta(n_\theta, n_{\theta'}) - c_\theta$$

where $\rho_\theta(n_\theta, n_{\theta'})$ denotes investor θi's expected *trading* profit (gross of information costs). A pure-strategy subgame perfect equilibrium is defined by a pair of investor classes (n_A, n_B) in conjunction with Lemma 1 such that (i) no active investor prefers to switch class or to be uninformed, and (ii) no uninformed investor prefers to become informed. There is an infinite population of investors who can become informed. Normalizing their outside option to 0, this implies that the expected profit of any investor in equilibrium (if one exists) must be 0. That is, $\pi_\theta\left(n_\theta, n_{\theta'}\right) = 0$ for all $\theta \in \Theta$ in equilibrium.[9]

Informed noise trading. Because investors in the same class compete with each other, their (expected) trading profits decrease in the size of their own class. This competition effect is illustrated by the downward sloping curve in Figure **1**. The effect of class size on trading profits *across* investor classes is ambiguous, as illustrated by the U-shaped curve in Figure **1**.

At first glance, the intuition behind the U-shape comes from Corollary 1. If the growth of an investor class primarily makes its order flow more informative, the price impact increases. This in turn induces investors in the other class to trade less intensively, thus reducing their trading profits. By contrast, if the order flow primarily gains in magnitude, the price impact decreases and investors in the other class gain from trading more intensively.

However, the price impact is the *same* for all investors. As such, it cannot explain why expected profits respond in *different* ways to changes in class size. Indeed, the main point is not that a growth of class A affects the price impact but that it intensifies competition more in class A than in class B. The key difference between the two classes is therefore that they trade on different types of information, so that the trading activity of one class is orthogonal to (part of) the information exploited by the other class. In this sense (alone), it is as if the two investor classes represent *noise* traders to each other. A growth in the magnitude of A- trading can thus provide more camouflage for B-investors: when A-trades swamp the market, they impair the market-maker's ability

9 To simplify matters, I ignore integer problems and treat n_θ as a continuous variable.

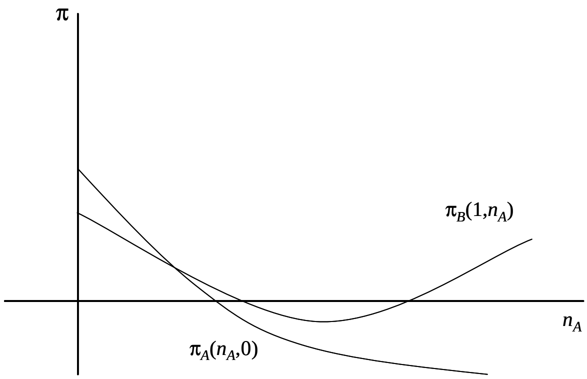

FIGURE 1. Expected profits as a function of the size of class A.

to extract the information contained in B-trades. In response, B-investors trade more intensively and accordingly make higher expected profits. For this reason, π_B begins to increase when n_A becomes sufficiently large.

Crowding out. I restrict attention to cases where each trading strategy is *per se* profitable, i.e., $\pi_\theta(1,0) > 0$ for all $\theta \in \Theta$. Without loss of generality, let $c_A \leq c_B$. Thus, in terms of the previous example, a sole investor prefers macro-economic information over company-specific information, and either of these over no information. In Figure **1**, her preferences are captured by the diverging but positive intercepts.

A sole trader chooses macro-trading (A) over stock-picking (B). Her presence not only decreases the profitability of macro-trading but also the profitability of stock-picking (for other investors), as reflected in Figure **1** by the initial decline in both curves. A second active investor then faces a trade-off: while macro-economic information is cheaper, macro-trading is more competitive. She too prefers macro-trading only if the cost difference $c_B - c_A$ exceeds the difference in trading profits $\rho_B(1,1) - \rho_A(2,0)$. In this fashion, every investor weighs the *cost* of a trading strategy against its *scarcity* (or competitiveness).

As macro-trading propagates, the macro-traders compete each others' profits away, while stock-picking becomes relatively more attractive, as reflected in Figure **1** by the falling margin $\pi_A(n_A,0) - \pi_B(1,n_A)$. At some point, the "marginal" investor either prefers to remain out of the market or to become a stock-picker. In Figure **1**, stock-picking is chosen only if $\pi_A(n_A,0)$ and $\pi_B(1,n_A)$ intersect before $\pi_A(n_A,0)$ hits zero. This intuition is captured in the following result.

Proposition 1. *Let* $\mathcal{R}_A \equiv \{n_A : \pi_A (n_A, 0) \leq 0\}$ *and* $\mathcal{R}_B \equiv \{n_A : \pi_B (1, n_A) \leq 0\}$. *In equilibrium,* $(n_A^*, n_B^*) = (n_A^0, 0)$ *if and only if* $\mathcal{R}_B \neq \varnothing$ *and* $\min \mathcal{R}_A \in \mathcal{R}_B$. *Otherwise, there exists a unique equilibrium* (n_A^*, n_B^*) *with* $n_A^* \geq n_B^* > 0$.

The cheaper trading strategy is always more prevalent. Less obvious is that under certain circumstances the other strategy, despite generating profits for a sole trader, is not pursued in equilibrium. This occurs precisely when the illiquidity created by macro-trading renders stock-picking unprofitable, even though the underlying information is unrelated.

Substitutes or complements. Lemma 1 states that stock-picking is "crowded out" in equilibrium whenever the (unique) root of $\pi_A (n_A, 0)$ falls into a region $\mathcal{R}_B$ where $\pi_B (1, n_A)$ is negative. Now suppose that $\mathcal{R}_B$ is non-empty. As the root of $\pi_A (n_A, 0)$ can be shifted by varying c_A, it follows that there exists a cost range $[\underline{c}_A, \overline{c}_A]$ such that crowding out occurs whenever $c_A \in [\underline{c}_A, \overline{c}_A]$. Intuitively, one can choose the cost of macro-trading such that stock-picking becomes unprofitable. Striking is also the impact of changing the cost of macro-trading outside of this range.

Proposition 2. *A reduction in the cost of A-trading decreases the prevalence of B-trading above some threshold* $\overline{c}_A$ *but increases it below some threshold* $\underline{c}_A$.

The proposition implies that macro-trading is a strategic *substitute* for stock-picking when $c_A \geq \overline{c}_A$ but a strategic *complement* to stock-picking when $c_A \leq \underline{c}_A$ (Figure **2**). When it is difficult to obtain macro-economic information, the volume of macro-trading is small and makes the market very illiquid. This discourages trading on even more inaccessible company-specific information. But when macro-economic information is easily accessible, the macro-trading volume camouflages stock-picking more effectively, thereby raising the profitability of the latter trading strategy.

The subsequent analysis focuses on the case $\mathcal{R}_B \neq \varnothing$. This is less restrictive than it seems, as the analysis can be extended to more than two trading strategies. The assumption $\mathcal{R}_B \neq \varnothing$ simply states that some strategies are crowded out under certain cost schedules.

2.3. Robustness Issues. This section discusses several issues related to the robustness of the above information equilibrium. They are not crucial for understanding the main analysis which continues in Section 3 with a discussion of the economic implications of Propositions 1 and 2.

Cognition. In the above analysis, it is assumed that investors are boundedly rational in the sense that they can base their trading strategy on one fundamental factor only. Such "limits in processing (receiving, storing, retrieving, transmitting) information" (Williamson, 1981, p.553) which induce economic agents to optimally neglect

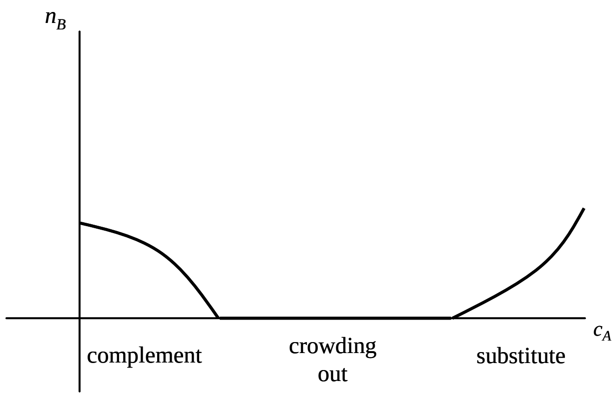

FIGURE 2. Substitutes or complements.

information are commonly referred to as rational inattention. Rational inattention has been documented in financial markets by e.g., Huberman (2001), Huberman and Regev (2001), Massa and Simonov (2005) and Hong et al. (2007).[10]

Notwithstanding, suppose instead that investors can process information about both factors without additional difficulties. That is, they can produce a signal $s_{ABi} = \sum_{\Theta}(V_\theta + \epsilon_{\theta i})$ at cost $c_{AB} = c_A + c_B$. Though there may now exist other equilibria, it is straightforward to see that pure specialization (as assumed in Proposition 1) remains an equilibrium in this setting. In such an equilibrium, no uninformed investor finds it worthwhile to enter with any type of information. By the same token, no active investor finds it worthwhile to incur the extra cost of acquiring additional information. Moreover, it seems reasonable that the marginal cost of processing information is increasing, so that $c_{AB} > c_A + c_B$. In this case, generalism becomes more expensive, and pure specialization becomes a more likely equilibrium outcome.

Communication. The analysis also assumes that investors cannot credibly communicate with each other. Since neither the individual factor realizations V_θ nor the individual error terms ϵ_i are revealed, a misreported signal is never detected. This makes it impossible for investors to commit to truthfully share information, as they have an incentive to shirk effort or to communicate false information to trade privately.

The infeasibility of communication is not as restrictive as it may seem at first glance. In a similar setting, Colla and Mele (2004) show that information sharing,

[10] Theoretical studies of rational inattention in financial markets include Moscarini (2004), Peng (2005) and Peng and Xiong (2006). For more general treatments of rational inattention, see Simon (1957), Kahneman (1973), Sims (2003, 2006) and Gabaix et al. (2006).

because it dilutes each investor's "monopoly" power, arises only if the initial correlation between the signals is sufficiently high. In the present model, information about different fundamental factors is uncorrelated, so that communication across investor classes is unattractive. Thus, the equilibrium in Lemma 1 is robust to communication.

Moreover, communication becomes less attractive when it is costly. Like information acquisition, successful communication typically requires effort. The receiver must exert effort to understand the sender's message, and the sender must exert effort to make her message intelligible to the receiver (Dewatripont and Tirole, 2005). Clearly, such communication costs favor equilibria in which boundedly rational investors specialize in different trading strategies.

Information. One could also consider a setting where investors can mix information about different factors, while choosing the precision of each type of information. Because of the competition effect, two investors would prefer to be as different as possible, and therefore specialize in distinct factors. With more investors, the incentives to avoid competition induce investors to choose different combinations of information about both factors. Nonetheless, liquidity externalities continue to exist, and each investor's decision criterion remains to weigh the cost of a particular type of information against its scarcity. Thus, changes in the cost of one type of information, and the corresponding changes in the demand for this information, should continue to exert positive or negative externalities on the demand for the other type of information.

3. Information Costs and Financial Markets

This section discusses some implications for financial markets and sheds light on a number of empirically observed phenomena. The model is developed further in section 4, where the supply of information is endogenized.

Participants in financial markets not only experience information shocks, which affect the distribution of asset prices, but also shocks to the *access* to information, which affect the composition of active investment. By Proposition 2, even if a shock affects only one investor class, liquidity externalities propagate the shock to other investor classes, and potentially across assets. Spillover effects of this kind induce interdependencies in trading activity and commonalities in prices and liquidity.

In the model, the arrival of new information is reflected in the investors' signals, whereas shocks to the access to information are best viewed as changes in the cost of information (c_θ).[11] In practice, such shocks to access can be both permanent (e.g.,

[11] Alternatively, one can model these shocks as changes in the precision of the signals. For given costs, a decrease (increase) in the precision corresponds to a decrease (increase) in the access to information. For the most part, the predictions of this alternative approach are similar.

the advent of new information technologies) or temporary (e.g., time variation in the media coverage of economic events).

For a single asset, it is immaterial whether the cost changes affect, say, macroeconomic information or asset-specific information. What matters is whether the affected investment strategy acts as a substitute for, or as a complement to, the other investment strategy. In the case of multiple assets, however, it is important whether the changes primarily affect investment strategies based on common factors or those based on asset-specific factors. In this case, the discussion below primarily focuses on common factors.

3.1. Investor Diversity and Herding Episodes. A continuous decrease in the cost of an investment strategy can move the market from the substitute region via the crowding out region to the complement region (Figure 2). During this process, the number of active investors increases, but the number of actively used investment strategies first decreases and then increases.

COROLLARY 2. *As c_A gradually decreases from a level above $\overline{c}_A$ to a level below $\underline{c}_A$, the number of investment strategies employed in the market goes from two to one and then back to two.*

In the substitute region, different investment strategies compete for liquidity in the market. As a result, expansions of one investment strategy come at the expense of the other strategy, and active investors become less heterogeneous. By contrast, in the complement region, the expanding investment strategy *supplies* liquidity to the market, thereby encouraging investors to enter with new investment strategies. In fact, when both strategies are widely used, expansions become mutually reinforcing.

One implication of Corollary 2 is that more sophisticated investment strategies, such as those of some hedge funds, may require a sufficient level of "conventional" informed trading to provide sufficient liquidity. This is at odds with the view that active investors elbow each other over profits but consistent with the view that improved market liquidity attracts more informed trading (Chordia et al., 2008a). What stands out in this setting is that the incentives for costly investment strategies feed on liquidity provided by *informed* trading, as opposed to pure noise trading, by other investor classes.

Another implication pertains to episodes of shifting confidence or, conversely, uncertainty in financial markets. Suppose that exogenous events suddenly make reliable macro-economic forecasts either more or less "difficult" (higher or lower c_A). Such shifts in the information environment can induce more correlated trading, i.e., herding behavior. Interestingly, this is true for both increases and decreases in c_A.

COROLLARY 3. *Shifting* c_A *from a level below* $\underline{c}_A$ *to a level in* $[\underline{c}_A, \overline{c}_A]$ *decreases the number of strategies employed in the market but increases the number of active investors. By contrast, shifting* c_A *from a level above* $\overline{c}_A$ *to a level in* $[\underline{c}_A, \overline{c}_A]$ *decreases both the number of strategies employed in the market and the number of active investors.*

A surge in the supply of macro-economic information can move a (rather inactive) market from the substitute region into the crowding out region. The number of active investors *rises* while the investment strategies become more homogeneous, in what resembles a herding *frenzy*. By contrast, starting from the complement region, such herding occurs only if the supply of macro-economic information drops moving the market into the crowding out region. In that case, the number of active investors *falls* while the investment strategies become more homogeneous, in what resembles a herding *panic*. Moreover, either case, the frenzy or the panic, may lead to more correlated trading not only within one asset market but also *across* different asset markets (see Section 3.3).

The diversity of active investment can vary significantly. Consider three equally effective strategies, $\Theta = \{A, B, C\}$, and suppose that $c_A \leq c_B = c_C$. Clearly, strategy A either crowds out both or none of the other two strategies. Once it becomes sufficiently widespread, both other strategies become equally viable, so that the number of employed investment strategies jumps from one to three. New investment strategies emerge even faster when the information environment improves generally, i.e., when c_A, c_B and c_C decrease. In that case, strategies B and C are not only made (more) attractive by the increased liquidity provided by A-trading, but also by the general decrease in the cost of active investment.

3.2. Trading Volume and Liquidity Spirals. The composition of active investment also influences the distribution of trading volume across investors.[12] For $c_A \geq \overline{c}_A$, a decrease in c_A raises the total trading volume of A-investors. By contrast, it reduces the total trading volume of B-investors. Thus, the impact on overall trading volume appears ambiguous. However, for $c_A \leq \overline{c}_A$, a decrease in c_A increases trading activity either in strategy A (crowding out region) or in both strategies (complement region), meaning that total trading volume eventually increases as the availability of information improves. The evolution of the order size distribution is similarly ambiguous.

[12] Because of differences in the trading intensity across investment strategies, the order flow composition (i.e., the distribution of order sizes) may convey more information than the mere net order imbalance. The information content of the composition would itself depend crucially on the (assumed) composition of pure noise trading. Analyzing such a model is beyond the scope of this paper. A conjecture is that, to the extent that order sizes reveal information, investors cannot fully exploit the benefits of a scarcer strategy because they have to "mimic" smaller order sizes.

Corollary 4. *A decrease in c_A raises α_B^* for $c_A \geq \overline{c}_A$ but reduces α_B^* for $c_A \leq \underline{c}_A$. In contrast, a decrease in c_A always reduces α_A^*.*

For a given signal precision, an active investor's expected (or "average") order size is determined by her trading intensity. Corollary 4 says that, in the substitute region, a decrease in c_A raises the average trade volume of B-investors (as their number shrinks) but reduces the average trade volume of A-investors (as their number grows). Thus, the order size distribution becomes more skewed. In the crowding out region, there are only A-investors, and their average trade volume continues to fall as c_A decreases. When the market enters the complement region, B-investors reappear with initially large average trade volumes. But as c_A decreases further, the average trade volumes of both A-investors and B-investors fall. Similar patterns arise when, starting from the substitute region, both types of information become cheaper.

Chordia et al. (2008b) document that the total stock trading volume at the New York Stock Exchange has steadily grown over the past decades. At the same time, the average order size has decreased, despite persistent activity in large-volume trades. They also report that the growth in trade volume has coincided with an increasing production of private information. These trends are broadly consistent with the above-mentioned patterns. Improvements in the information environment could have led to more (competitive) active investment, decreasing order sizes for conventional investment strategies, and the continuous emergence of new investment strategies traded in large(r) volumes.

Corollary 5. *When $T'(n_A^*) < 0$ and $T'(n_B^*) < 0$, the total trading volume of either investor class decreases in c_A and in c_B.*

The trading volumes in the different strategies, while evolving somewhat differently, are interdependent. As Corollary 5 highlights, this is most apparent in the complement region where a shock to the trading volume of one class can be contagious. If the two strategies are so widespread that they are *mutually* complementary, a sudden increase in, say, c_A can trigger a *liquidity spiral*: Initially, the adverse shock only reduces A-trading. However, this withdraws liquidity from the B-investors, so that B-trading decreases as well. This in turn withdraws liquidity from the A-investors, leading to another round of contractions in trading volume. Eventually, the market reaches a new equilibrium in which each investor classes trades a smaller volume. Importantly, the liquidity spiral not only *amplifies* the negative impact on the directly affected group of investors but also *transmits* it to another group of investors.

3.3. Price Comovement and Liquidity Commonality. The degree to which different stock prices comove is often taken as a (inverse) measure of the amount of company-specific information that is impounded into stock prices (Roll, 1988). It has been documented that the level of comovement is lower in more developed economies and has decreased in the US over the 20th century (Morck et al., 2000; Campbell et al., 2001). Some evidence further suggests that variation in comovement patterns may be related to differences in the information environment (Fox et al., 2003; Bushman et al., 2004; Hameed et al., 2005).

Suppose that several assets share (some) common fundamentals but are also driven by asset-specific fundamentals. Since the availability of the different types of information determines what strategies investors choose, it also affects how closely the prices move together. For example, consider three fundamental factors, $\Theta = \{A, B, C\}$, and two separately traded assets, $\mathcal{M} = \{1, 2\}$. The liquidation values are given by

$$\begin{aligned}\tilde{V}_1 &= \tilde{V}_A + \tilde{V}_B \\ \tilde{V}_2 &= \tilde{V}_A + \tilde{V}_C\end{aligned}$$

where the factors are i.i.d. and have uniform variances. Suppose that the three investment strategies are equally effective and that $c_A \leq c_B = c_C$. Noise traders buy and sell the market portfolio, so that their trades induce market-wide movements (De Long et al., 1990; Morck et al., 2000). There is a different market-maker for each asset. Trade occurs simultaneously in all assets, and market-makers see only their own order flow.[13] As before, let $[\underline{c}_A, \overline{c}_A]$ denote the crowding out region, in this case for both assets.

Further, define the *market index* as $p_{\mathcal{M}} \equiv p_1 + p_2$ and the correlation coefficient between the price of asset $a \in \mathcal{M}$ and the market index as $\rho_{a\mathcal{M}}$. The degree of price comovement can then be measured by the average (absolute) correlation coefficient $\overline{\rho}_{a\mathcal{M}} \equiv \frac{1}{2}|\rho_{1\mathcal{M}}| + \frac{1}{2}|\rho_{2\mathcal{M}}|$. The coefficient $\overline{\rho}_{a\mathcal{M}} \in [0, 1]$ indicates how much the variation of individual prices is explained by market-wide variations and is typically a good approximation of the R^2 in a regression.

Corollary 6. *Prices comove perfectly if and only if* $c_A \in [\underline{c}_A, \overline{c}_B]$.

For $c_A > \overline{c}_A$, active investors are few but pursue diverse strategies (either A, B, or C). However, when c_A falls to intermediate levels, the number of active investors grows, and macro-trading propagates at the expense of asset-specific strategies. As a result, the prices of the two assets comove more. For $c_A < \underline{c}_A$, the asset-specific strategies regain appeal, so that prices incorporate idiosyncratic information again. Corollary 6 can potentially explain why price comovement is different in developed and undeveloped

[13] The assumption that noise traders induce market-wide movements is not important. It only "neutralizes" their influence on the level of price comovement, an influence which is orthogonal to changes in c_A. Similarly, information "spill-overs" between the asset markets would simply introduce an additional channel for price comovements, from which I want to abstract here.

countries, and why it has decreased in the US over time. The observed cross-country variation may indicate that information is more widely available in developed countries than in developing countries. The decline in price comovement in the US may have been driven by improvements in information technology and the development of competitive (business and financial) information markets. Section **4.2** discusses how the level of price comovement can be related to the degree of competition in information markets.[14]

In a sample of emerging markets, Chan and Hameed (2006) find price comovement to be higher for stocks that are followed by more analysts. This relationship is weaker, however, when the higher number of analysts coincides with a higher forecast dispersion. Corollary 6 seems generally consistent with this observation: in inactive markets, an increase in information acquisition goes together with a decrease in information diversity. But when a stock is more actively traded, the increase in analyst following may increase diversity, causing the stock to comove less with the market. To flesh out a cross-sectional prediction, suppose that a more idiosyncratic stock is added to the above model, $\tilde{V}_3 = \tilde{V}_D + \tilde{V}_E$. If information about the common factor A is more accessible, this third asset would not only covary less with the market, but it would also attract fewer active investors than the other assets. As information about A becomes sufficiently widespread, however, investors in stocks 1 and 3 may acquire information about B and C (again). At this point, the inflow of investors interacts with a rise in diversity, so that the positive cross-sectional relation between the number of active investors in a stock and the stock's comovement with the market becomes weaker.[15]

In the model, shocks to c_A also induce common variation in the liquidity of the assets. For example, a decrease in c_A improves liquidity in assets 1 and 2 if both are in the complement region. This is consistent with evidence that the liquidity of individual stocks covaries with market-wide liquidity, as empirically documented (Chordia et al., 2000; Hasbrouck and Seppi, 2001; Huberman and Halka, 2001).

Corollary 7. *A decrease in* c_A *decreases the liquidity of assets* 1 *and* 2 *when* $c_A \geq \overline{c}_A$*, whereas it increases the liquidity of both assets when* $c \leq \underline{c}_A$.

[14] Veldkamp (2006b) also explains price comovement by means of investors' information choices. In her theory, comovement is the result of positive externalities in the information market. Due to scale economies in information production, information in higher demand is supplied at a lower price. As a result, investors benefit from acquiring the same information. By contrast, price comovement in the present model results from negative externalities in the asset market.

[15] It should be pointed out that the dispersion of analyst recommendations need not necessarily be a sign of information diversity but can also indicate less precise information. If so, more dispersion indicates less, not more, information (e.g., Jin and Myers, 2006). The finding by Chan and Hameed pertains, however, not to the level of dispersion *per se* but to its interaction with the number of analysts. Thus, my interpretation presumes that an increase in dispersion indicates an increase in information if it is, at the same time, associated with an increase in analyst coverage.

Illiquidity in this model arises from asymmetric information across investors (including the market-makers) about the value of the stocks. When the availability of macro-economic information changes, so does the information asymmetry between the active investors and the less informed market participants. In response, the (less informed) market-makers' willingness to provide liquidity changes.

When there is heterogeneity with respect to other determinants of liquidity, the sensitivity of liquidity to changes in c_A may vary across assets both in direction and in strength. Suppose that asset 1 is larger and therefore attracts more pure noise trading ($\sigma^2_{\epsilon 1} > \sigma^2_{\epsilon 2}$). As a result, it might also attract more macro-traders than stock 2;[16] making it possible that asset 1 is in the complement region while asset 2 is in the substitute region. In that case, market liquidity in the two assets moves in opposite directions when c_A changes marginally. Such differential responses complicate the influence of common determinants of liquidity across stocks.

Alternatively, suppose that asset 2 is in the crowding out region, while asset 1 is in the region where strategies A and C are mutually complementary. In asset 2, an increase in c_A has an ambiguous effect on liquidity. By contrast, in asset 1, it causes liquidity to spiral downward as trading contracts in both strategies (see previous section). Despite having higher *levels* of liquidity, larger stocks may therefore be more sensitive to common variation in liquidity (Chordia et al., 2000). At the same time, as they attract more diverse strategies, their liquidity may exhibit more idiosyncratic variation so that common factors explain a smaller fraction of their overall liquidity variation (Hasbrouck and Seppi, 2001).

3.4. Market Efficiency and Price Informativeness. A central role of prices is to aggregate dispersed information (Hayek, 1945). The economic literature offers three concepts to describe how well prices fulfill this role: *Market efficiency* measures the extent to which prices reveal the information held by investors (Fama, 1970). *Price informativeness* measures the precision with which a price predicts the true asset value. *Allocative efficiency* measures the degree to which prices aid decision-makers in allocating resources optimally (Tobin, 1982).

Since market efficiency measures the information content of prices relative to the amount of information held by active investors, it can be proxied by the total loss of the (uninformed) noise traders. For example, when information becomes cheaper, more active investors enter the market and compete more intensively. As a result, a larger

[16] This now assumes that the interpretation of macro-economic information is a cognitive task which is specific for each stock. That is, investors not only specialize in a subset of information but also in a subset of assets. This may be due to informational "familiarity" with certain assets (Van Nieuwerburgh and Veldkamp, 2008b).

portion of their information is revealed by the order flow, and the market-maker sets the price closer to the active investors' (average) forecast. This in turn implies that the uninformed traders lose less when trading.

In contrast, price informativeness measures the total amount of information—rather than the proportion of information acquired by investors—that is impounded into prices (Chen et al., 2007). This is captured by the *residual* uncertainty about the true asset value or, more precisely, the conditional variance $\mathrm{Var}(\tilde{V}\,|P) = \sigma^2\left(1 - \rho^2_{V,z}\right)$, where

$$\rho^2_{V,z} = \frac{1}{2}\sum_{\theta=A,B}\frac{n_\theta\sigma^2}{(n_\theta+1)\sigma^2 + 2\sigma^2_\epsilon} \tag{3.1}$$

is derived in the proof to the next corollary. Since residual uncertainty decreases in $\rho^2_{V,z}$, I use the latter as a measure of price informativeness.

In a single-factor framework, the three forms of information efficiency typically move in tandem. For instance, when the number of active investors increases, both market efficiency and price informativeness improve, and so does allocative efficiency to the extent that decision-makers benefit from more information about the asset. In a multi-factor model, this need no longer be the case.

Corollary 8. *For $c_A \geq \overline{c}_A$, a decrease in c_A improves market efficiency but may reduce price informativeness.*

The following comparative statics provide some intuition: $\partial\rho^2_{V,z}/\partial n_\theta > 0$ and $\partial^2\rho^2_{V,z}/\partial n^2_\theta < 0$. That is, price informativeness increases with the prevalence of either type of information. However, the marginal increase is decreasing in the prevalence of a given type of information (because the average error asymptotically converges to 0). For a fixed investor population, this implies that the price is the most informative under a balanced information structure ($n_A = n_B$). All else equal, skewing the information distribution hence reduces price informativeness.[17]

This explains why price informativeness need not necessarily increase in the substitute region, where the expansion of strategy A comes at the expense of strategy B. Intuitively, while the number of active investors increases, they acquire less diverse information. As a result, the market price—while reflecting more of the investors' acquired information—may reflect less total information. It seems counterintuitive that cheaper information can lead to less informative prices. However, the result reflects the trade-off between the level and the diversity of active investment when there are multiple investment strategies.

[17] The optimality of the balanced structure is particular to the assumption that both types of information are equally important. The argument, however, is more general.

Finally, it should be noted that allocative efficiency need not improve even if market efficiency and price informativeness do. In a multi-factor setting, some types of information may be more relevant for allocative decisions than others. For instance, Holmström and Tirole (1993) argue that stock-based compensation schemes can enhance managerial incentives because active investors collect information about managerial effort. Suppose that B relates to managerial effort, whereas A relates to macro-economic events outside of managerial control. A stock-based compensation scheme is less effective if strategy B is crowded out by strategy A—irrespective of the effect on price informativeness. The macro-economic information reflected in the stock price represents "luck" and confounds the role of the price as a signal about effort.[18]

4. Markets for Financial Information

The preceding analysis has focused on exogenous changes in the access to information. These may result from changes in the uncertainty of the economic environment, the transparency of economic policy, corporate disclosure rules, or the quality of accounting standards. Alternatively, one can analyze the endogenous supply of information, e.g., by commercial vendors. Sections 4.1 and 4.2 discuss the pricing incentives in a monopoly and under competition. Section 4.3 considers the situation of two investors, each with a distinct expertise, who can agree to sell part of their information.

4.1. News Monopoly, Predatory Pricing, and Investor Herding. The model is extended as follows. In stage -1, news vendors offer investors subscriptions for A-data. In stage 0, each investor decides whether to remain uninformed, to purchase a subscription, or to produce data privately at cost c_A. A vendor who sells a subscription communicates in stage 1 the promised data to the subscriber. The marginal cost of communicating data is negligible. News vendors add neither bias nor noise to the data.

This section considers a monopolistic vendor. The setting is by no means implausible. While characterized by low marginal costs, the large-scale and timely production and dissemination of information can impose high fixed costs. Due to such fixed costs,

[18] Dow and Gorton (1997) also argue that market efficiency (which they term price efficiency) need not imply allocative efficiency (which they term economic efficiency), though for a different reason. With respect to allocative efficiency, they further distinguish between a retrospective role (evaluating past actions) and a prospective role (reflecting the value of investment opportunities) of prices. Several papers formalize the idea that firm managers may extract information from stock prices to improve their investment decisions (e.g., Subrahmanyam and Titman, 1999, 2001; Dow et al., 2006; Goldstein and Gümbel, 2006). Empirical evidence supports this view (Wurgler, 2000; Baker et al., 2003; Durnev et al., 2004; Chen et al., 2007).

media industries are highly concentrated.[19] In many countries, entry regulations and political capture impose further restrictions on competition (Besley and Prat, 2006).

In the same setting with a single type of information, Admati and Pfleiderer (1988) show that a (risk-neutral) monopolist has no incentive to increase supply, i.e., to lower the price. Intuitively, since she is selling information to investors who compete over trading profits, new subscribers' profits come at the expense of the profit of existing subscribers. In fact, due to the intensified competition, any revenue gain from the new subscribers is always smaller than the revenue loss on the existing subscribers. Hence, she cannibalizes on her own profit if she expands the investor population. This result carries over to the present setting if the monopolist sells both types of information: she would set $p_A = c_A$ and $p_B = c_B$ so that the market would de facto be the same as without the information vendor.

In the present setting, one can alternatively consider the case in which information vendors only provide a subset of the asset-relevant information. Certain types of information may be difficult to communicate or impossible to communicate credibly (Allen, 1990; Michaely and Womack, 1999). A direct sale may therefore be either unprofitable or infeasible.[20] Hence, consider a *partial* information market in which only information about factor A represents marketable news.

First, I establish that a direct, unrestricted sale with a uniform price is optimal.[21]

LEMMA 2. *A direct sale of unlimited subscriptions at a single price is optimal.*

Intuitively, consider any price-quantity schedule posted by a monopolist. Since investors are symmetric within each investor class, the highest subscription price paid in equilibrium must equal the expected trading profit of a subscriber. At all lower prices, subscribers pay less than their reservation price, and the monopolist would earn more by raising the price. Binding quantity restrictions are therefore suboptimal from the vendor's point of view, while slack restrictions are unnecessary.

[19] Before Reuters and Thompson merged in 2007, Bloomberg, Reuters and Thompson accounted for a combined market share of about two-thirds of the US financial information services industry. Notwithstanding, compared to other countries, the US financial information services industry is arguably the most competitive.

[20] For instance, Goetzmann et al. (2004) find that movie scripts that are less certifiable by hard information are sold at a discount. That is, the scriptwriter incurs a higher cost of selling 'soft' information. (If possible, she might prefer to instead produce the movie herself.) As regards speculative information, Roll (1988, p.564), in his pioneering study on stock price comovement, concludes that "the financial press misses a great deal of relevant information generated privately."

[21] In an indirect sale, the seller sets up an investment fund, and investors can "subscribe" to the seller's information by purchasing fund shares. I do not consider subscription fees that are contingent on realized trading profits. In the context of direct sales, this question has not been addressed in the literature, perhaps because, in practice, the subscriber's use of the data, including any ensuing profit, is difficult to monitor.

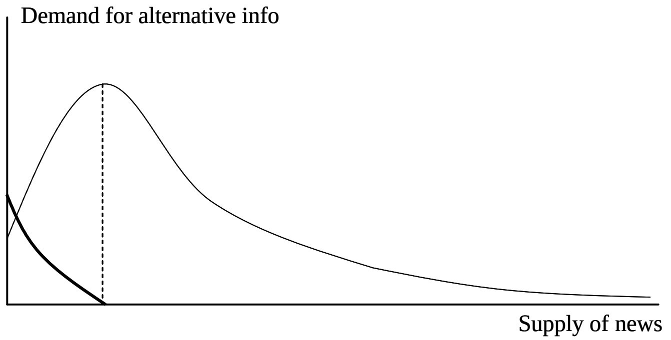

FIGURE 3. Predatory pricing: The thin (thick) line represents the news vendor's profit (demand for alternative information).

Indirect sales can serve as a means to control the number of active investors, and hence to curb competition (Admati and Pfleiderer, 1990). Here, this benefit does not arise because investors can resort to alternative sources of information. Consider an investor who sets up a fund. The fund's optimal investment strategy is that of a single investor. Once sufficiently many participate in the fund, others prefer to compete with self-collected A-information. The fund's expected profit is therefore $\rho_A(n_A^*, n_B^*) - c_A$. If the fund manager instead could instead sell the data at $p_A = c_A$, she would extract the expected trading profits of all A-investors and earn $n_A^*\rho_A(n_A^*, n_B^*) - c_A$.

Turning to the main analysis, let $p_A \in [0, c_A]$ denote the subscription price. The number of subscriptions, $n_A^*(p_A)$, is endogenously determined as a function of the price. The monopolist chooses p_A to maximize her total profit

$$\Pi(p_A) = \begin{cases} n_A^*(p_A)p_A - c_A & \text{if } p_A \leq c_A \\ 0 & \text{if } p_A > c_A \end{cases} . \tag{4.1}$$

For $p_A \leqslant c_A$, free entry of active investors ensures that subscriptions are sold until the subscribers' expected profits are driven to zero. As a result, A-investors' expected trading profits are fully extracted by the monopolist. Hence,

$$\Pi(p_A) = n_A^*(p_A)\rho(n_A^*(p_A), n_B^*(p_A)) - c_A. \tag{4.2}$$

This expression highlights not only that p_A jointly determines n_A^* and n_B^* but also that maximizing the monopolist's profit is tantamount to maximizing the total expected trading profits from strategy A.

PROPOSITION 3. *The news monopolist sets p_A such that* $\min \mathcal{R}_A = \min \mathcal{R}_B$.

Unlike in Admati and Pfleiderer (1988), the news vendor is not a pure monopolist as her product "competes" against information about B. This affects her pricing incentives for two reasons. First, the presence of B-investors reduces market liquidity and thereby the trading profits of the monopolist's subscribers. Second, the resources which B-investors expend to acquire information do not translate into revenues for the monopolist. With a lower price, the monopolist can crowd out B-investors and attract more subscriptions or sway B-investors to become subscribers. As a result, she reduces the price just enough to render strategy B unattractive (Figure **3**).

Relative to the outcome in the absence of a news market, the monopolist increases trading volume and the number of active investors, which improves market efficiency. At the same time, she homogenizes investor expectations, which induces herding behavior and potentially decreases price informativeness (Corollary 8). Thus, in the present model, the impact of information sales on financial market quality is not necessarily benign. Indeed, in the absence of competitors, a news vendor increases supply *only to* reduce investors' incentives to seek alternative sources of information.

An anecdote about *Reuters* in its early days illustrates this crowding-out rationale.[22] In the 1850s, *Reuters* controlled telegraph lines as well as the right to circulate news received from ships of the *Austrian Lloyd's.* To develop its business, *Reuters* offered the main London newspapers subscriptions to its international news dispatches at £30 per month, significantly less than a newspaper's cost of running its own correspondent network. At this price, *Reuters* had to attract a critical number of daily newspapers to make the service profitable.

The Times initially resented any dependence on *Reuters* and preferred its own network. However, the value of its network deteriorated, as rival newspapers gained access to foreign news: "Good though its own network was, it needed to know each evening what telegrams from Reuter were likely to appear in the columns of its competitors next morning, even though it did not necessarily want to print the telegrams itself." Eventually, *The Times* took out a *Reuters* subscription. By 1861, *Reuters* supplied almost all London newspapers with identical foreign dispatches. Importantly, these short dispatches differed in nature from the reports of newspaper correspondents: "*The Times* kept a correspondent at the front, the famous W. H. Russell, who wrote long mailed dispatches. These made a great impression by revealing military shortcomings, but they were not intended to give the latest news."

The aim of *Reuters'* pricing strategy was to replace the newspapers' own networks. While this gave each newspaper cheaper access to foreign news, it caused foreign news coverage to become more homogenous. Moreover, the telegraph dispatches may at the

[22] The historical account is taken from Read (1999, p.24f).

time have crowded out the more complex information typically provided by overseas correspondents.

4.2. News Competition, Trading Volume, and Investor Diversity. Introducing a competing news vendor who sells information about factor B naturally promotes diversity.[23] Surprisingly, so does competition by a vendor who also sells information about factor A. I model competition in news markets, which are characterized by high fixed and entry costs, as a contested monopoly in which the incumbent vendor is threatened by the entry of a competitor. To emphasize the impact of competition on the diversity of investment strategies, I revisit the example with two stocks, $\mathcal{M} = \{1, 2\}$, and three fundamental factors, $\Theta = \{A, B, C\}$, from Section 3.3.

To model contestability, let the incumbent incur a non-negligible fixed cost $K_I \in [K_L, K_H]$ of running the sales operation. Unlike c_A, this cost is private information, though its distribution is common knowledge. To enter the market, the rival must incur an up-front (sunk) cost S. For simplicity, the rival's operative costs are $K_E < K_L$ and commonly known.[24]

The timing of the entry game (which takes place in stage -2) is as follows. First, the incumbent precommits to a subscription price p_A^I, which is renegotiable in stage -1. Second, the rival observes p_A^I and then decides whether or not to enter the market at cost S. Third, if the rival enters, the two vendors engage in price competition (in stage -1). Otherwise, the game proceeds as in the monopoly case except that the incumbent cannot raise the price over its pre-commitment level p_A^I.

Since the incumbent is bound to be undercut by the rival if the latter enters, the incumbent either loses the market or preempts entry. The rival enters when she expects post-entry profits to be at least as large as her entry cost S. If she enters, the equilibrium price will equal the incumbent's break-even price, which implies a revenue of $E\left(K_I \left| p_A^I\right.\right) + c_A$. Hence, the rival enters only if

$$\mathrm{E}\left(K_I \left| p_A^I\right.\right) + c_A \geq S + K_E + c_A.$$

[23] Though, even if we introduce a B-vendor into the monopoly setting, predatory pricing by the A-vendor remains a possible equilibrium. For example, p_A may be set such that $\min R_A = \max R_B$, and n_A is just small enough for the market not to reach the complement region. At this point, a single investor finds B-trading unprofitable. Selling information about B to more than one investor is more profitable, if crowding out A-investors improves the liquidity for B-investors. Yet, at $\min R_A = \max R_B$, the A-strategy is a complement to the B-strategy, so that expanding the supply of B-information to crowd out A-investors reduces liquidity for B-investors. Thus, a B-vendor may not fare better than a single B-investor, and stays out of the market.

[24] Introducing uncertainty or private information about S or K_E, or allowing for $K_E \geq K_L$, makes the extension more realistic, but the mechanics and basic intuition behind the results remain the same. Note also that this setup includes Bertrand competition as a special case ($S = 0$ and $K_E = K_I$.)

To preempt entry, the incumbent must therefore choose p_A^I to signal that her operative costs to do not exceed K_I^+, as defined by

$$K_I^+ = S + K_E.$$

LEMMA 3. *The following is a Pareto-dominant Perfect Bayesian Equilibrium: For $\mathrm{E}(K_I) \leq S + K_E$, the incumbent is uncontested and sets the monopoly price. For $\mathrm{E}(K_I) > S + K_E$, the incumbent is contested and deters entry if $K_I \leq K_I^*$ but loses the market if $K_I > K_I^*$ where $K_I^* < K_H$. When the incumbent is contested, the equilibrium price is below the monopoly price and (weakly) increases in S and K_E.*

If entry is not worthwhile when facing an incumbent of average efficiency, entry is deterred without further ado. The incumbent is a de facto monopolist who need not be concerned with signaling high efficiency. But if average efficiency is too low to deter entry, a sufficiently efficient incumbent wants to signal a higher efficiency, because a higher efficiency implies a lower post-entry profit for the rival. This can be achieved by precommitting to a lower price which a less efficient incumbent could not afford to mimic. Thus, a sufficiently efficient incumbent ($K_I \leq K_I^*$) reduces the price and successfully defends the market. The deterrence price and the cut-off type K_I^* increase in S and K_E. Intuitively, when entry is cheaper and the rival is more efficient, the incumbent must precommit to a lower price to deter entry. If the incumbent is too inefficient ($K_I > K_I^*$), she cannot credibly deter entry and eventually loses the market to the rival.

A reasonable interpretation of S and K_E is that these costs capture regulatory barriers to entry into information markets and the progress in information technology, respectively. For instance, the entry cost S might consist of two components, technological expenditures (S_1) and the cost of overcoming regulatory "red tape" (S_2).[25] $(S_1 K_E)^{-1}$ can then be interpreted as a measure of technological efficiency, and S_2^{-1} as a measure of the openness of the market. With this interpretation, Lemma 3 can be invoked to explain cross-country variation in price comovement as driven by variation in $(S_1 K_E)^{-1}$ and S_2^{-1}.

PROPOSITION 4. *Progress in information technology and less entry regulation increase information market competition, which promotes investor diversity and reduces price comovement.*

Rival news vendors supplying information about A compete in perfect substitutes, while information about B or C represents an imperfect substitute. Competition is

[25] Djankov et al. (2002) document that regulatory entry barriers tend to be higher in countries with less democratic governments (cf. http://www.doingbusiness.org/). Djankov et al. (2003) show that the media sector in such countries is often concentrated and government-controlled.

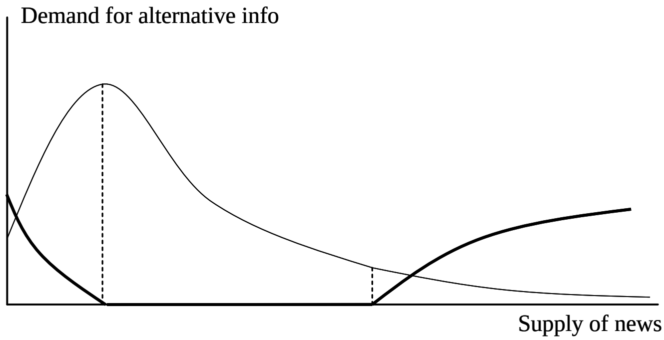

FIGURE 4. Competition promotes diversity: The thin (thick) line represents the news vendors' profit (demand for alternative information).

thus "hierarchical" in that the news vendors first and foremost compete with each other, while their impact on strategies B and C plays a secondary role for their pricing incentives. The degree of competition increases in the threat of entry, and hence in $(S_1 K_E)^{-1}$ and S_2^{-1}. Importantly, the competition among the news vendors exerts a positive externality on the demand for alternative information: the price battle propagates investment based on the sold information, and the rise in trade volume creates liquidity for alternative investment strategies (Figure 4). As a result, prices incorporate more asset-specific information, and comove less.

The main insight is that the competition among the news vendors is not only beneficial because it disseminates news more effectively but also because it encourages investors to tap alternative sources of information. This suggests that competition policy in news markets may have a significant impact on the diversity of active investment strategies and thereby on the quality of financial markets.

4.3. Proprietary Trading and Newsletters. Some financial institutions engage in proprietary trading activities and, at the same time, provide information to other investors. At first glance, it seems puzzling that they distribute information and trade actively, in particular if sharing information creates more competitors. In the present model, such behavior can be rational when the institution owns different types of information, some of which is also known to others while some is (more) exclusive.

Consider two investors (1 and 2) endowed with information about A and B, respectively. Information about A is also held by $n_A - 1$ other investors, whereas information about B is exclusive to investor 2. Suppose that investor 2 can make a take-it-or-leave-it (cash) offer P to investor 1 such that, if accepted, investor 1 has to give away her information to (infinitely) many other investors. The question is whether there exists a price P such that both investors can benefit from such an agreement.

PROPOSITION 5. *There exists a non-empty interval $\mathcal{N}_A$ such that, when $n_A \in \mathcal{N}_A$, the two investors would benefit from the following agreement: investor 2 engages in a B-strategy and pays $P > 0$ to investor 1, who in return shares her data with infinitely many other investors.*

The intuition behind this result is straightforward. The presence of $n_A - 1$ other investors pursuing the A-strategy reduces the expected profit of both investor 1 and investor 2. While giving away information about A for free eliminates investor 1's profit, it also eliminates any negative externality that the A-investors exert on investor 2's profit $[T(\infty) \to T(0)]$. As long as $\pi_B(1,0) > \pi_B(1,n_A) + \pi_A(n_A,1)$, there are gains from trade that the two investors can share. Intuitively, by flooding the market with information about A, they create a "herd" of A-investors among which investor 2 can "hide". In analogy, for a financial institution, a benefit of supplying many investors with mundane information may thus be increased liquidity, which in turn facilitates trading on proprietary information, i.e., information that the institution does not share with outside investors. It is worth emphasizing that the shared information is not a means of manipulation; the investors who receive, or buy, the mundane information are not deceived in any way.

5. Concluding Remarks

This paper studies a financial market in which fundamentals are driven by several factors, and active investors choose which factor to base their trading strategies on. The central questions are how active investors who pursue different strategies interact when trading in the same market, and how this interaction affects their strategy choice when entering the market. Contrary to common wisdom, active investors in this setting can benefit from increased trading by other active investors, as long as they pursue a different strategy. On the one hand, trading in any strategy is a source of information asymmetry, and hence reduces market liquidity. On the other hand, active investors who trade on different information can serve as quasi-noise traders for each other. Different investment strategies can therefore be substitutes or complements: if a strategy becomes cheaper, it may decrease or increase the demand for other strategies.

Such externalities in information acquisition can have interesting consequences for financial markets. The paper discusses a number of implications for herding behavior, trade volumes, price comovement, liquidity commonality, and the informational role of prices. However, the discussion only scratches the surface of these issues. A fully-fledged analysis warrants a more general framework with more factors, more assets, dynamics, and a generalized variance-covariance structure. For example, it would be interesting to analyze how shocks to the informational environment lead to contagion effects across different markets, taking into account both prices, liquidity, and trading volume. In addition, it would be interesting to relax the assumption that (all) investors have to specialize on a single factor. This can lead to hierachical information structures, in which some investors know what other know *and more*. Moreover, one could distinguish between "generalists" (broadly researching many types of information) and "specialists" (deeply researching one type of information). Such questions are left for future research.

The paper also examines how the externalities in information acquisition affect the sale of information. A monopolist, who provides information to a particular investor class, expands supply only to crowd out other types of information. Such predatory pricing increases the number of active investors but decreases the diversity of their information. Accordingly, prices may become less informative although the market becomes more efficient. Competition—even among vendors who sell the same information—promotes diversity. As prices fall, the investment strategies based on sold information propagate. The resulting increase in trade volume creates liquidity that makes alternative investment strategies more profitable. Thus, information market competition improves financial market liquidity, stimulates private information acquisition, and fosters investor diversity. Finally, the framework can explain why a financial institution may engage in proprietary trading while also selling information to other investors. By supplying many investors with mundane information, the institution can improve the market liquidity. This in turn allows it to trade more profitably on proprietary information that it does not share with other investors. An interesting question, also left for future research, is how a financial institution should be organized to optimize its (different) managers' incentives to acquire and to coordinate different types of information.

Appendix

Proof of Proposition 1. The main proof makes use of the following auxiliary result.

LEMMA 4. *For $n_\theta, n_{\theta'} \geqslant 0$ and $n \geqslant n_\theta$,*

(a): $\pi_\theta(n_\theta, \cdot)$ *decreases in* n_θ,
(b): $\pi_\theta(\cdot, n_{\theta'})$ *has a unique minimum in* $\mathbb{R}^+$,
(c): $\pi_\theta(n_\theta, n - n_\theta)$ *decreases in* n_θ.

Proof of Lemma 4. Given that c_θ is a constant, I need only consider the behavior of $\rho_\theta(n_\theta, n_{\theta'})$ or, more precisely,

$$\begin{aligned}
\rho_\theta(n_\theta, n_{\theta'}) &= \mathrm{E}[(\tilde{V} - \tilde{p})\tilde{x}_{i\theta}] = \mathrm{E}[(\tilde{V} - \lambda\tilde{z})\alpha_\theta\tilde{s}_{i\theta}] \\
&= \alpha_\theta \mathrm{E}(\tilde{V}\tilde{s}_{i\theta}) - \lambda\mathrm{E}\left[\left(\sum_{i=1}^{n_\theta}\alpha_\theta\tilde{s}_{i\theta} + \sum_{l=1}^{n_{\theta'}}\alpha_{\theta'}\tilde{s}_{l\theta'} + \tilde{y}\right)\alpha_\theta\tilde{s}_{i\theta}\right] \\
&= \alpha_\theta\left[\mathrm{E}\left(\tilde{V}_\theta\tilde{s}_{i\theta}\right) + \mathrm{E}\left(\tilde{V}_{\theta'}\tilde{s}_{i\theta}\right)\right] - \lambda\left[\sum_{i=1}^{n_\theta}\alpha_\theta^2\mathrm{E}(\tilde{s}_{\hat{i}\theta}\tilde{s}_{i\theta}) + \sum_{l=1}^{n_{\theta'}}\alpha_{\theta'}\alpha_\theta\mathrm{E}(\tilde{s}_{l\theta'}\tilde{s}_{i\theta}) + \alpha_\theta\mathrm{E}(\tilde{y}\tilde{s}_{i\theta})\right].
\end{aligned}$$

Since $\mathrm{E}\left(\tilde{V}_\theta\tilde{s}_{i\theta}\right) = \sigma^2$, $\mathrm{E}\left(\tilde{V}_{\theta'}\tilde{s}_{i\theta}\right) = 0$, $\mathrm{E}\left(\tilde{s}_{i\theta}^2\right) = \sigma^2 + \sigma_\epsilon^2$, $\mathrm{E}(\tilde{s}_{\hat{i}\theta}\tilde{s}_{i\theta}) = \sigma^2$, $\mathrm{E}\left(\tilde{s}_{l\theta'}\tilde{s}_{i\theta}\right) = 0$, and $\mathrm{E}\left(\tilde{y}\tilde{s}_{i\theta}\right) = 0$ for all θ, $\theta' \neq \theta$, i, $l \neq i$, this becomes

$$\rho_\theta(n_\theta, n_{\theta'}) = \alpha_\theta\sigma^2\left(1 - \lambda n_\theta\alpha_\theta\right) = \alpha_\theta\sigma^2\left[\frac{\sigma^2 + 2\sigma_\epsilon^2}{(n_\theta+1)\,\sigma^2 + 2\sigma_\epsilon^2}\right] \tag{A1}$$

where the last equality follows from Lemma 1. Substituting for α_θ and λ in (A1) gives

$$\rho_\theta(n_\theta, n_{\theta'}) = \left[\frac{\sigma_y\sigma^2\left(\sigma^2 + 2\sigma_\epsilon^2\right)}{\left[\sigma^2\left(n_\theta+1\right) + 2\sigma_\epsilon^2\right]^2}\right]\left[\frac{n_\theta\left(\sigma^2 + \sigma_\epsilon^2\right)}{\left[(n_\theta+1)\sigma^2 + 2\sigma_\epsilon^2\right]^2} + T_{\theta'}\right]^{-1/2}.$$

Now define $A \equiv [\rho_\theta(n_\theta, n_{\theta'})]^{-2}$. This gives

$$A = \frac{n_\theta\left(\sigma^2 + \sigma_\epsilon^2\right)\left[\sigma^2\left(n_\theta+1\right) + 2\sigma_\epsilon^2\right]^2}{\sigma_y^2\sigma^4\left(\sigma^2 + 2\sigma_\epsilon^2\right)^2} + \frac{\left[\sigma^2\left(n_\theta+1\right) + 2\sigma_\epsilon^2\right]^4}{\sigma_y^2\sigma^4\left(\sigma^2 + 2\sigma_\epsilon^2\right)^2}T_{\theta'}.$$

A strictly increases in n_θ, which proves (a). Recall that $T_{\theta'}$ has a unique maximum. For a given n_θ, the maximum corresponds to a unique minimum of A, which proves (b). The proof of (c) proceeds in steps (i)-(iii).

(i) First define the functions

$$\lambda^n(n_\theta) \equiv \frac{\sigma^2}{\sigma_y}\left[\frac{n_\theta\left(\sigma^2 + \sigma_\epsilon^2\right)}{\left[(n_\theta+1)\sigma^2 + 2\sigma_\epsilon^2\right]^2} + \frac{\left(n - n_\theta\right)\left(\sigma^2 + \sigma_\epsilon^2\right)}{\left[(n - n_\theta + 1)\sigma^2 + 2\sigma_\epsilon^2\right]^2}\right]^{1/2}$$

and

$$\alpha_\theta^n(n_\theta) \equiv \frac{\sigma^2\left[\lambda^n(n_\theta)\right]^{-1}}{\sigma^2(n_\theta+1) + 2\sigma_\epsilon^2}.$$

Note that $\lambda^n(n_\theta)$ and hence $\alpha_\theta^n(n_\theta)$ are continuously differentiable for $n_\theta \in [0, n]$. Using the first equality in (A1), write

$$\rho_\theta(n_\theta, n - n_\theta) \equiv \rho_\theta^n(n_\theta) = \alpha_\theta^n(n_\theta)\,\sigma^2\left[1 - n_\theta\alpha_\theta^n(n_\theta)\,\lambda^n(n_\theta)\right].$$

Since this is an algebraic combination of continuously differentiable functions for $n_\theta \in [0, n]$, it is also continuously differentiable for $n_\theta \in [0, n]$.

(ii) For a fixed n, it follows from Lemma 1 that $n_\theta < n_{\theta'} \Rightarrow \alpha_\theta > \alpha_{\theta'} \Rightarrow \rho_\theta^n(n_\theta) > \rho_{\theta'}^n(n_{\theta'})$ (*). Simple inspection of the formulae in Lemma 1 further shows that, for a fixed population, a trader from group θ with $n_\theta = x$ faces exactly the same decision problem as would a trader from group θ' with $n_{\theta'} = x$. That is, $\rho_\theta^n(x) = \rho_{\theta'}^n(x)$. Together with (*), this "symmetry" implies that

$$\rho_\theta^n(n_\theta) > \rho_\theta^n(n - n_\theta) \quad \text{for } n_\theta \in (0, n/2),$$

which in turn implies that $\rho_\theta^n(n_\theta)$ is *decreasing* over some range in $[0, n]$.

(iii) Now suppose that $\rho_\theta^n(n_\theta)$ is also *increasing* over some range in $[0, n]$. Since $\rho_\theta^n(n_\theta)$ is continuously differentiable, this requires the existence of some $x^* \in (0, n)$ such that $\partial\rho_\theta^n/\partial n_\theta\big|_{x^*} = 0$. Using Lemma 1 and $n_{\theta'} = n - n_\theta$ to eliminate α_θ, λ and $n_{\theta'}$ from $\rho_\theta^n(n_\theta)$, one can verify (e.g., in *Maple*) that $\partial\rho_\theta^n/\partial n_\theta = 0$ has no real solution. Hence, $\rho_\theta^n(n_\theta)$ cannot be increasing in $[0, n]$. This proves part (c). ■

Preliminary. I show that $\pi_A(n_A, 0) = \pi_B(1, n_A)$ has a unique solution. That they cross at least once follows from parts (a) and (b) of Lemma 4. To show that they cross at most once, it suffices to show that $\rho_A(n_A, 0) = \rho_B(1, n_A)$

has a unique solution. After substituting for α_θ and λ in (A1), the respective functions are given by

$$\rho_A(n_A,0) = \frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left[(n_A+1)\,\sigma^2+2\sigma_\epsilon^2\right]\sqrt{n_A\left(\sigma^2+\sigma_\epsilon^2\right)}}$$

$$\rho_B(1,n_A) = \frac{\sigma^2\sigma_y\left(\sigma^2+2\sigma_\epsilon^2\right)\left((n_A+1)\sigma^2+2\sigma_\epsilon^2\right)}{2\left(\sigma^2+\sigma_\epsilon^2\right)^{3/2}\sqrt{\left((n_A+1)\sigma^2+2\sigma_\epsilon^2\right)^2+4n_A\left(\sigma^2+\sigma_\epsilon^2\right)^2}}$$

Equating these expressions and rearranging yields

$$1=\frac{\left((n_A+1)\,\sigma^2+2\sigma_\epsilon^2\right)^2\sqrt{n_A}}{2\left(\sigma^2+\sigma_\epsilon^2\right)\sqrt{\left((n_A+1)\sigma^2+2\sigma_\epsilon^2\right)^2+4n_A\left(\sigma^2+\sigma_\epsilon^2\right)^2}}$$

On both sides, I square, take the inverse and further rearrange to get

$$4\left(\sigma^2+\sigma_\epsilon^2\right)^2\left(\frac{1}{\left((n_A+1)\,\sigma^2+2\sigma_\epsilon^2\right)^2 n_A}+\frac{1}{\left((n_A+1)\,\sigma^2+2\sigma_\epsilon^2\right)^4}\right)=1.$$

The left-hand side goes to infinity for $n_A\to 0$ and strictly decreases in n_A. Thus, $\pi_A\left(n_A,0\right)=\pi_B\left(1,n_A\right)$ has exactly one solution, to the left of which $\pi_A\left(n_A,0\right)>\pi_B\left(1,n_A\right)$ and to the right of which $\pi_A\left(n_A,0\right)<\pi_B\left(1,n_A\right)$.

Main proof. For use below, I define $\mathcal{R}_B=\left[\underline{n}_A,\overline{n}_A\right]$ and $\min\mathcal{R}_A=n_A^0$.

First part: $\min\mathcal{R}_A\in\mathcal{R}_B\neq\varnothing$. I start with the *sufficient condition*, $\underline{n}_A\leqslant n_A^0\leqslant\overline{n}_A\Rightarrow\left(n_A^*,n_B^*\right)=\left(n_A^0,0\right)$. It is straightforward to verify that the set of inequalities $\underline{n}_A\leqslant n_A^0\leqslant\overline{n}_A$ is equivalent to the condition $\pi_B\left(1,n_A^0\right)\leqslant\pi_A\left(n_A^0,0\right)=0$. Suppose that this condition holds. I now check different candidate equilibria. **(i)** Note that $\left(n_A^0,0\right)$ trivially satisfies the free-entry condition. **(ii)** Conjecture an equilibrium with $n_A>n_A^0$ and $n_B\geqslant 0$. For all $n_A>n_A^0$, $\pi_A\left(n_A,n_B\right)<\pi_A\left(n_A^0,n_B\right)<\pi_A\left(n_A^0,0\right)=0$, that is, A-traders would incur a loss. Hence this cannot be an equilibrium. **(iii)** Conjecture an equilibrium with $n_A<n_A^0$ and $n_B\geqslant 0$ and distinguish the cases (a) $n_A+n_B=n\leqslant n_A^0$ and (b) $n_A+n_B>n_A^0$. **(iiia)** First, recall that $\pi_A\left(n_A^0,0\right)\geqslant\pi_B\left(1,n_A^0\right)$ and that $\pi_A\left(n_A,0\right)$ and $\pi_B\left(1,n_A\right)$ cross only once. This implies that $\pi_A\left(n,0\right)>\pi_B\left(1,n-1\right)$ because $n<n_A^0$ (+). Further, by Lemma 4(c), $\pi_A\left(1,n-1\right)>\pi_A\left(2,n-2\right)>\cdots>\pi_A\left(n,0\right)$ and $\pi_B\left(1,n-1\right)>\pi_B\left(2,n-2\right)>\cdots>\pi_B\left(n-1,1\right)$. Together with (+), these imply that $\pi_A\left(n-x,x\right)>\pi_B\left(x+\varepsilon,n-x-\varepsilon\right)$ for all $\varepsilon\in(0,n-x]$. That is, B-investors would on the margin switch to the A-strategy. Hence, this cannot be an equilibrium. **(iiib)** Denote $\hat{n}_B=n_A^0-n_A$ so that $\hat{n}_B>n_B$ (because $n_A<n_A^0$). This provided, note that $\pi_A\left(n_A^0,0\right)=0>\pi_B\left(1,n_A^0-1\right)\geqslant\pi_B\left(\hat{n}_B,n_A\right)>\pi_B\left(n_B,n_A\right)$. The first two inequalities follow from Lemma 4(c), and the last inequality follows from Lemma 4(a). Together, they imply that this cannot be an equilibrium because B-investors would expect to make a loss.
Finally, the *necessary condition*, $\left(n_A^*,n_B^*\right)=\left(n_A^0,0\right)\Rightarrow\pi_B\left(1,n_A^0\right)\leqslant\pi_A\left(n_A^0,0\right)$, holds because, if the latter inequality were violated, A-investors would on the margin switch to the B-strategy, and $\left(n_A^0,0\right)$ would not be an equilibrium.

Second part: $\min\mathcal{R}_A\notin\mathcal{R}_B$ or $\mathcal{R}_B=\varnothing$. I must show that, when $n_A^0\notin\left(\underline{n}_A,\overline{n}_A\right)$, there exists a unique pair $\left(n_A^*,n_B^*\right)$ that satisfies $n_A^*>n_B^*>0$ and $\pi_\theta(n_\theta^*,n_{\theta'}^*)=0$ for $\theta=A,B$. I proceed in steps (i)-(vi).
(i) Note that neither $(0,0)$, $(0,n_B)$ nor $(n_A,0)$ can be an equilibrium. This follows respectively from $\pi_\theta\left(1,0\right)>0$ for $\theta=A,B$, $\pi_A\left(1,n_B-1\right)>\pi_B\left(n_B,0\right)$ (Lemma 4 (c) and $c_A\leqslant c_B$), and the first part of proof.
(ii) Note that there exists an information structure where positive profits are equally shared. Consider a point n where $\pi_B\left(1,n\right)>\pi_A\left(n,0\right)>0$, the existence of which follows from $n_A^0\notin\left(\underline{n}_A,\overline{n}_A\right)$. In this case, it is on the margin profitable for an A-investor to switch to the B-strategy. In doing so, she marginally lowers the expected profit in class B but marginally raises it in class A (Lemma 4 (c)). Still, it might be profitable for the next marginal A-investor to switch. Yet, since $\pi_B\left(n-1,1\right)<\pi_A\left(1,n-1\right)$ and the profit functions are continuous, there exists a unique x^* such that $\pi_B\left(n-x^*,x^*\right)=\pi_A\left(x^*,n-x^*\right)$. Given $\pi_A\left(n,0\right)>0$, these expected profits must be positive for both classes.
(iii) It follows from the monotonicity in Lemma 4 (c) that such an indifference point exists for any n (though not always with positive profits).
(iv) If—starting from $\pi_B\left(n_B,n_A\right)=\pi_A\left(n_A,n_B\right)>0$—the total population is changed, both n_A and n_B have to move in the same direction to maintain the indifference. To see this, note that after substituting for α_θ and λ in (A1), $\pi_A\left(n_A,n_B\right)=\pi_B\left(n_A,n_B\right)$ can be written out and rearranged to

$$\left(\frac{1}{\left((n_A+1)\,\sigma^2+2\sigma_\epsilon^2\right)^2}-\frac{1}{\left((n_B+1)\,\sigma^2+2\sigma_\epsilon^2\right)^2}\right)=\frac{c_A-c_B}{\sigma^4\left(\sigma^2+2\sigma_\epsilon^2\right)}\lambda$$

Suppose that this holds for a given n_A and n_B. Now suppose that the change in population lowers λ. In order for the equation to still hold, the term in the parentheses must become smaller. It cannot be that only one class grows because, if so, the class that does not grow in size would end up with positive profits (lower price impact, same or lower number of investors). Similarly, it cannot be that only one class shrinks when λ increases. To maintain the equality, both classes have to grow (shrink) when λ falls (rises). Now consider a population change without any change in λ. For a fixed λ, the expected profit in class θ is only determined by n_θ. In fact, it is strictly decreasing in n_θ. Thus, if the population change raises n_A (and hence lowers π_A), it must also raise n_B (to equivalently lower π_B). Again, n_A and

n_B must move in the same direction.

(v) As both classes grow, the investors' total expected profits eventually decrease (and even become negative). To see this, write them as

$$n_A \frac{\sigma^2 \sigma_y \left(\sigma^2 + 2\sigma_\epsilon^2\right)}{\left(\left(n_A+1\right)\sigma^2 + 2\sigma_\epsilon^2\right)^2} \left(\frac{n_A\left(\sigma^2+\sigma_\epsilon^2\right)}{\left[(n_A+1)\sigma^2+2\sigma_\epsilon^2\right]^2} + \frac{n_B\left(\sigma^2+\sigma_\epsilon^2\right)}{\left[(n_B+1)\sigma^2+2\sigma_\epsilon^2\right]^2}\right)^{-1/2} - n_A c_A$$
$$+ n_B \frac{\sigma^2 \sigma_y \left(\sigma^2 + 2\sigma_\epsilon^2\right)}{\left(\left(n_B+1\right)\sigma^2 + 2\sigma_\epsilon^2\right)^2} \left(\frac{n_A\left(\sigma^2+\sigma_\epsilon^2\right)}{\left[(n_A+1)\sigma^2+2\sigma_\epsilon^2\right]^2} + \frac{n_B\left(\sigma^2+\sigma_\epsilon^2\right)}{\left[(n_B+1)\sigma^2+2\sigma_\epsilon^2\right]^2}\right)^{-1/2} - n_B c_B =$$
$$\sigma^2 \sigma_y \left(\frac{n_A\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(\left(n_A+1\right)\sigma^2+2\sigma_\epsilon^2\right)^2} + \frac{n_B\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(\left(n_B+1\right)\sigma^2+2\sigma_\epsilon^2\right)^2}\right)^{1/2} - n_A c_A - n_B c_B$$

The total trading profits (first term) are concave in the sense that, if both n_A and n_B increase the marginal trading gain decreases. Moreover, total trading profits converge to zero as both n_A and n_B approach ∞. By contrast, the total information costs (last two terms) increase linearly. Thus, once total expected profits have fallen to zero, they must continue to decrease.

(vi) Since the above functions are continuous, the preceding arguments imply that—starting from $\pi_B(n_B, n_A) = \pi_A(n_A, n_B) > 0$—there exist a unique population size $n^* > n_A + n_B$ such that, at the respective indifference point, total trading profits and hence average trading profits are zero. This point identifies the unique equilibrium information structure. ■

Proof of Proposition 2. It suffices to show that, in equilibrium, n_A is monotonically decreasing in c_A. Starting from an equilibrium ($\pi_A^* = \pi_B^* = 0$), consider an exogenous increase in c_A. Keeping n_B^* and n_A^* fixed, the expected profit in class A drops and becomes negative, whereas expected profits in class B remain unaffected. Keeping the total number of investors (n) fixed, A-investors therefore start switching to the B-strategy. By Lemma 4 (c), this gradually increases π_A and decreases π_B until the expected profits in the two classes are equal again. Starting from $\pi_B^* = 0$, this adjustment process ends up at $\pi_A^{**} = \pi_B^{**} < 0$. That is, there are still too many investors in the market. By step (iv) of the second part of the previous proof, a decrease in the total investor population implies that both classes, A and B, must shrink. Thus, as a result of an increase in c_A, class A shrinks because A-investors either switch to the B-strategy or exit the market. By symmetry, analogous arguments apply in the case of a decrease in c_B. ■

Proof of Corollaries 2 and 3. Corollary 2 and parts of Corollary 3 follow directly from Proposition 1. It remains to prove that n is strictly decreasing in c_A. That this must hold in the crowding out region is trivial to see. In the complement region, a decrease in c_A not only increases the number of A-investors but—because of liquidity-provision—also the number of B-investors. Hence, n must be decreasing in c_A. Finally, consider an equilibrium in the substitute region ($\pi_A^* = \pi_B^* = 0$). Keeping n_B^* and n_A^* fixed, an increase in c_A causes the expected profit in class A to become negative. Keeping n fixed, A-investors therefore start switching to the B-strategy. By Lemma 4 (c), this gradually increases π_A and decreases π_B until the expected profits in the two classes are equal again. Starting from $\pi_B^* = 0$, this adjustment process ends up at $\pi_A^{**} = \pi_B^{**} < 0$. To ensure break-even, the population must shrink further, i.e., n must decrease. ■

Proof of Corollary 6. As mentioned in the text, I measure price comovement by the average (absolute) correlation coefficient between individual asset prices and the market index:

$$\bar{\rho}_{\mathcal{M}} = |\mathcal{M}|^{-1} \sum_{a \in \mathcal{M}} |\rho_{a\mathcal{M}}|$$

where

$$\rho_{a\mathcal{M}} \equiv \frac{\mathrm{Cov}(p_a, p_{\mathcal{M}})}{\sqrt{\mathrm{Var}(p_a)\mathrm{Var}(p_{\mathcal{M}})}} \quad \text{and} \quad p_{\mathcal{M}} = \sum_{a \in \mathcal{M}} p_a.$$

The average correlation coefficient is typically a good approximation of the R^2 in a regression.

Each A-investor trades in both assets. Let α_{aA} denote A-investors' trading intensity when trading in asset a. By definition of the market makers' pricing functions,

$$p_1 = \lambda_1 \left(\sum_{i=1}^{n_A} x_{1A}(\tilde{s}_{iA}) + \sum_{l=1}^{n_B} x_B(\tilde{s}_{lB}) + \tilde{y}\right) \quad \text{and} \quad p_2 = \lambda_2 \left(\sum_{i=1}^{n_A} x_{2A}(\tilde{s}_{iA}) + \sum_{l=1}^{n_C} x_C(\tilde{s}_{lC}) + \tilde{y}\right)$$

which implies the following 'market index'

$$p_{\mathcal{M}} = \lambda_1 \left(\sum_{i=1}^{n_A} x_{1A}(\tilde{s}_{iA}) + \sum_{l=1}^{n_B} x_B(\tilde{s}_{lB}) + \tilde{y}\right) + \lambda_2 \left(\sum_{i=1}^{n_A} x_{2A}(\tilde{s}_{iA}) + \sum_{l=1}^{n_C} x_C(\tilde{s}_{lC}) + \tilde{y}\right).$$

Price variances are given by

$$\begin{aligned}
\mathrm{Var}\,(p_1) &= (\lambda_1)^2\,\mathrm{Var}\left(\sum_{i=1}^{n_A} x_{1A}(\tilde{s}_{iA}) + \sum_{l=1}^{n_B} x_B(\tilde{s}_{lB}) + \tilde{y}\right) = (\lambda_1)^2\,\mathrm{Var}\left(\sum_{i=1}^{n_A} \alpha_{1A}\tilde{s}_{iA} + \sum_{l=1}^{n_B} \alpha_B \tilde{s}_{lB} + \tilde{y}\right) \\
&= (\lambda_1)^2 \left[\mathrm{Var}\left(\sum_{i=1}^{n_A} \alpha_{1A}\left(\tilde{V}_A + \tilde{\epsilon}_i\right)\right) + \mathrm{Var}\left(\alpha_B \sum_{l=1}^{n_B}\left(\tilde{V}_B + \tilde{\epsilon}_l\right)\right) + \sigma_y^2\right] \\
&= (\lambda_1)^2 \left[(n_A\alpha_{1A})^2\,\sigma^2 + (\alpha_{1A})^2\,n_A\sigma_\epsilon^2 + (n_B\alpha_B)^2\,\sigma^2 + (\alpha_B)^2\,n_B\sigma_\epsilon^2 + \sigma_y^2\right]
\end{aligned}$$

and, analogously,

$$\mathrm{Var}\,(p_2) = (\lambda_2)^2 \left[(n_A\alpha_{2A})^2\,\sigma^2 + (\alpha_{2A})^2\,n_A\sigma_\epsilon^2 + (n_C\alpha_C)^2\,\sigma^2 + (\alpha_C)^2\,n_C\sigma_\epsilon^2 + \sigma_y^2\right].$$

The variance of the market index is

$$\begin{aligned}
\mathrm{Var}\,(p_{\mathcal{M}}) &= \mathrm{Var}\left(\lambda_1 \sum_{i=1}^{n_A} x_{1A}(\tilde{s}_{iA}) + \lambda_2 \sum_{i=1}^{n_A} x_{2A}(\tilde{s}_{iA}) + (\lambda_1+\lambda_2)\,\tilde{y} + \lambda_1 \sum_{l=1}^{n_B} x_B(\tilde{s}_{lB}) + \lambda_2 \sum_{l=1}^{n_C} x_C(\tilde{s}_{lC})\right) \\
&= \mathrm{Var}\left(\lambda_1 \sum_{i=1}^{n_A} x_{1A}(\tilde{s}_{iA}) + \lambda_2 \sum_{i=1}^{n_A} x_{2A}(\tilde{s}_{iA})\right) + \mathrm{Var}\left[(\lambda_1+\lambda_2)\,\tilde{y}_2\right] \\
&\quad + (\lambda_1)^2 \left[(n_B\alpha_B)^2\,\sigma^2 + (\alpha_B)^2\,n_B\sigma_\epsilon^2\right] + (\lambda_2)^2 \left[(n_C\alpha_C)^2\,\sigma^2 + (\alpha_C)^2\,n_C\sigma_\epsilon^2\right] \\
&= (\lambda_1\alpha_{1A} + \lambda_2\alpha_{2A})^2 \left[(n_A)^2\,\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_1+\lambda_2)^2\,\sigma_y^2 \\
&\quad + (\lambda_1)^2 \left[(n_B\alpha_B)^2\,\sigma^2 + (\alpha_B)^2\,n_B\sigma_\epsilon^2\right] + (\lambda_2)^2 \left[(n_C\alpha_C)^2\,\sigma^2 + (\alpha_C)^2\,n_C\sigma_\epsilon^2\right]
\end{aligned}$$

The covariances between individual asset prices and the market index are given by

$$\begin{aligned}
\mathrm{Cov}\,(p_1, p_{\mathcal{M}}) &= \mathrm{Cov}\left(\lambda_1\left(\sum_{i=1}^{n_A} x_{1A}(\tilde{s}_{iA}) + \sum_{l=1}^{n_B} x_B(\tilde{s}_{lB}) + \tilde{y}\right), \lambda_1\left(\sum_{i=1}^{n_A} x_{1A}(\tilde{s}_{iA}) + \sum_{l=1}^{n_B} x_B(\tilde{s}_{lB}) + \tilde{y}\right) + \lambda_2\left(\sum_{i=1}^{n_A} x_{2A}(\tilde{s}_{iA}) + \tilde{y}\right)\right) \\
&= \mathrm{Cov}\left(\lambda_1 \sum_{i=1}^{n_A} \alpha_{1A}\tilde{s}_{iA} + \lambda_1 \sum_{l=1}^{n_B} \alpha_B\tilde{s}_{lB} + \lambda_1\tilde{y}, (\lambda_1\alpha_{1A} + \lambda_2\alpha_{2A}) \sum_{i=1}^{n_A} \tilde{s}_{iA} + \lambda_1 \sum_{l=1}^{n_B} \alpha_B\tilde{s}_{lB} + (\lambda_1+\lambda_2)\,\tilde{y}\right) \\
&= \mathrm{Cov}\left(\lambda_1 \sum_{i=1}^{n_A} \alpha_{1A}\tilde{s}_{iA} + \lambda_1 \sum_{l=1}^{n_B} \alpha_B\tilde{s}_{lB} + \lambda_1\tilde{y}, (\lambda_1\alpha_{1A} + \lambda_2\alpha_{2A}) \sum_{i=1}^{n_A} \tilde{s}_{iA}\right) + \\
&\quad \mathrm{Cov}\left(\lambda_1 \sum_{i=1}^{n_A} \alpha_{1A}\tilde{s}_{iA} + \lambda_1 \sum_{l=1}^{n_B} \alpha_B\tilde{s}_{lB} + \lambda_1\tilde{y}, \lambda_1 \sum_{l=1}^{n_B} \alpha_B\tilde{s}_{lB}\right) + \\
&\quad \mathrm{Cov}\left(\lambda_1 \sum_{i=1}^{n_A} \alpha_{1A}\tilde{s}_{iA} + \lambda_1 \sum_{l=1}^{n_B} \alpha_B\tilde{s}_{lB} + \lambda_1\tilde{y}, (\lambda_1+\lambda_2)\,\tilde{y}\right) \\
&= (\lambda_1\alpha_{1A})\,(\lambda_1\alpha_{1A} + \lambda_2\alpha_{2A}) \left[(n_A)^2\,\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_1\alpha_B)^2 \left[(n_B)^2\,\sigma^2 + n_B\sigma_\epsilon^2\right] + (\lambda_1)\,(\lambda_1+\lambda_2)\,\sigma_y^2
\end{aligned}$$

and, similarly,

$$\mathrm{Cov}\,(p_2, p_{\mathcal{M}}) = (\lambda_2\alpha_{2A})\,(\lambda_1\alpha_{1A} + \lambda_2\alpha_{2A}) \left[(n_A)^2\,\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_2\alpha_C)^2 \left[(n_C)^2\,\sigma^2 + n_C\sigma_\epsilon^2\right] + (\lambda_2)\,(\lambda_1+\lambda_2)\,\sigma_y^2.$$

The correlation coefficients are thus

$$\rho_{1\mathcal{M}} = \frac{1}{\sqrt{(\lambda_1)^2 \left[(n_A\alpha_{1A})^2\,\sigma^2 + (\alpha_{1A})^2\,n_A\sigma_\epsilon^2 + (n_B\alpha_B)^2\,\sigma^2 + (\alpha_B)^2\,n_B\sigma_\epsilon^2 + \sigma_y^2\right]}}$$

$$\frac{(\lambda_1\alpha_{1A})\,(\lambda_1\alpha_{1A} + \lambda_2\alpha_{2A}) \left[(n_A)^2\,\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_1\alpha_B)^2 \left[(n_B)^2\,\sigma^2 + n_B\sigma_\epsilon^2\right] + (\lambda_1)\,(\lambda_1+\lambda_2)\,\sigma_y^2}{\sqrt{Z}}$$

and

$$\rho_{2\mathcal{M}} = \frac{1}{\sqrt{(\lambda_1)^2\left[(n_A\alpha_{1A})^2\sigma^2 + (\alpha_{1A})^2 n_A\sigma_\epsilon^2 + (n_B\alpha_B)^2\sigma^2 + (\alpha_B)^2 n_B\sigma_\epsilon^2 + \sigma_y^2\right]}}$$

$$\frac{(\lambda_2\alpha_{2A})(\lambda_1\alpha_{1A} + \lambda_2\alpha_{2A})\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_2\alpha_C)^2\left[(n_C)^2\sigma^2 + n_C\sigma_\epsilon^2\right] + (\lambda_2)(\lambda_1+\lambda_2)\sigma_y^2}{\sqrt{Z}}.$$

where

$$\begin{aligned} Z \equiv{}& (\lambda_1\alpha_{1A} + \lambda_2\alpha_{2A})^2\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_1+\lambda_2)^2\sigma_y^2 \\ &+ (\lambda_1)^2\left[(n_B\alpha_B)^2\sigma^2 + (\alpha_B)^2 n_B\sigma_\epsilon^2\right] + (\lambda_2)^2\left[(n_C\alpha_C)^2\sigma^2 + (\alpha_C)^2 n_C\sigma_\epsilon^2\right] \end{aligned}$$

Simple inspection shows that these correlation coefficients must be smaller than 1 as long as $(n_B, n_C) \neq (0,0)$. This is intuitive. For instance, if $n_c > 0$, the price of asset 1 is independent of the C-factor, whereas the market index is not.

By contrast, suppose that the investment strategies are crowded out ($n_B = n_C = 0$). In this case, A-investors face no rival investor class in either asset market. Using $\alpha_{1A} = \alpha_{2A} = \alpha_A$, the correlation coefficient then becomes

$$\begin{aligned} \rho_{1\mathcal{M}} &= \frac{(\lambda_1\alpha_A)(\lambda_1\alpha_A + \lambda_2\alpha_A)\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_1)(\lambda_1+\lambda_2)\sigma_y^2}{\sqrt{\left\{(\lambda_1\alpha_A)^2\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_1)^2\sigma_y^2\right\}\left\{(\lambda_1\alpha_A + \lambda_2\alpha_A)^2\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_1+\lambda_2)^2\sigma_y^2\right\}}} \\ &= \frac{(\alpha_A)^2(\lambda_1)(\lambda_1+\lambda_2)\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right] + (\lambda_1)(\lambda_1+\lambda_2)\sigma_y^2}{(\lambda_1)(\lambda_1+\lambda_2)\sqrt{(\alpha_A)^4\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right]^2 + 2(\alpha_A)^2\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right]\sigma_y^2 + \sigma_y^4}} \\ &= \frac{(\alpha_A)^2\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right] + \sigma_y^2}{\sqrt{(\alpha_A)^4\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right]^2 + 2(\alpha_A)^2\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right]\sigma_y^2 + \sigma_y^4}} \\ &= \frac{(\alpha_A)^2\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right] + \sigma_y^2}{\sqrt{\left((\alpha_A)^2\left[(n_A)^2\sigma^2 + n_A\sigma_\epsilon^2\right] + \sigma_y^2\right)^2}} \\ &= 1. \end{aligned}$$

Intuitively, price movements in this case result exclusively from the (market-wide) trades of A-investors and noise traders. ■

Proof of Corollary 7. Given the assumption that $c_A \leq c_B = c_C$ and that the two asset markets are otherwise isomorphic, investment strategy A either crowds out both or none of the other strategies. When $c_A \geq \overline{c}_A$, either market is in the substitute region. Here, a decrease in c_A crowds out the asset-specific strategies. Crowding out immediately implies that λ^* (which is the only channel through which the strategies affect each other) must increase in equilibrium; otherwise, the profitability of the asset-specific strategies would (weakly) increase, as would then the number of active investors in those strategies. Similarly, when $c_A \geq \overline{c}_A$, a decrease in c_A must increase equilibrium liquidity; otherwise, the number of investors in the asset-specific strategies would not increase. ■

Proof of Corollary 8. Market efficiency. I need to show that total trading profits are lower under $(\underline{n}_A, 0)$ than under (n_A, n_B). First, note that $\pi_A(n_A + n_B, 0) < \pi_A(n_A, n_B) = \pi_B(n_B, n_A)$ where the inequality follows from Lemma 4 (c), and the equality follows from the fact that (n_A, n_B) denotes an equilibrium outcome. These relations imply that

$$(n_A + n_B)\,\pi_A(n_A + n_B, 0) < n_A\pi_A(n_A, n_B) + n_B\pi_B(n_B, n_A).$$

This inequality can be rearranged to

$$\begin{aligned} (n_A + n_B)\left[\rho_A(n_A + n_B, 0) - c_A\right] &< n_A\left[\rho_A(n_A, n_B) - c_A\right] + n_B\left[\rho_B(n_B, n_A) - c_B\right] \\ (n_A + n_B)\,\rho_A(n_A + n_B, 0) - (n_A + n_B)\,c_A &< n_A\rho_A(n_A, n_B) + n_B\rho_B(n_B, n_A) - n_A c_A - n_B c_B \\ (n_A + n_B)\,\rho_A(n_A + n_B, 0) &< n_A\rho_A(n_A, n_B) + n_B\rho_B(n_B, n_A) - n_B\,(c_B - c_A) \end{aligned}$$

which – due to $c_B > c_A$ – implies

$$(n_A + n_B)\,\rho_A(n_A + n_B, 0) < n_A\rho_A(n_A, n_B) + n_B\rho_B(n_B, n_A)$$

where the left-hand side and the right-hand side represent total trading profits for $(n_A + n_B, 0)$ and (n_A, n_B) respectively. Finally, it is well-known that $n\rho_A(n,0) < (n')\,\rho_A(n',0)$ if $n' > n$ (see, e.g., Admati and Pfleiderer, 1988). Since $n' > n_A + n_B$, it follows that

$$\underline{n}_A\rho_A(\underline{n}_A,0) < (n_A+n_B)\,\rho_A(n_A+n_B,0) < n_A\rho_A(n_A,n_B) + n_B\rho_B(n_B,n_A)$$

which proves the proposition.

Price informativeness. First, note that $\mathrm{Var}(\tilde{V}\,|p) = \mathrm{Var}(\tilde{V}\,|z)$ as the price is based exclusively on order flow information. This conditional variance is given by $\mathrm{Var}(\tilde{V}\,|z) = \mathrm{Var}(\tilde{V})(1-\rho^2_{V,z})$ where

$$\begin{aligned}
\rho^2_{V,z} &= \frac{[\mathrm{Cov}(\tilde{V},\tilde{z})]^2}{\mathrm{Var}(\tilde{V})\mathrm{Var}(\tilde{z})} = \frac{\mathrm{Cov}(\tilde{V},\tilde{z})}{\mathrm{Var}(\tilde{V})}\lambda = \frac{(n_A\alpha_A + n_B\alpha_B)\,\sigma^2}{2\sigma^2}\lambda \\
&= \frac{1}{2}\left(n_A\frac{\sigma^2\lambda^{-1}}{\sigma^2(n_A+1)+2\sigma^2_\epsilon} + n_B\frac{\sigma^2\lambda^{-1}}{\sigma^2(n_B+1)+2\sigma^2_\epsilon}\right)\lambda \\
&= \frac{1}{2}\sum_{\theta=A,B}\frac{n_\theta\sigma^2}{\sigma^2(n_\theta+1)+2\sigma^2_\epsilon}
\end{aligned}$$

Note that

$$\frac{\partial\rho^2_{V,z}}{\partial n_\theta} = \frac{\sigma^2\left(\sigma^2+2\sigma^2_\epsilon\right)}{[\sigma^2(n_\theta+1)+2\sigma^2_\epsilon]^2} > 0 \quad\text{and}\quad \frac{\partial\rho^2_{V,z}}{\partial n_\theta} = -\frac{2\sigma^4\left(\sigma^2+2\sigma^2_\epsilon\right)}{[\sigma^2(n_\theta+1)+2\sigma^2_\epsilon]^3} < 0.$$

I now show that price informativeness can decrease when B-information is crowded out. Define $\mathrm{I}\,(n_A,n_B) = 2\rho^2_{V,z}$ as a measure of price informativeness. Suppose that $n^0_A(c_A) < \underline{n}_A$, and denote the equilibrium information structure by (n_A,n_B). Price informativeness is then given by

$$\mathrm{I}\,(n_A,n_B) = \sum_{\theta=A,B}\frac{n_\theta\sigma^2}{\sigma^2(n_\theta+1)+2\sigma^2_\epsilon}.$$

Comparing this to the information structure $(\underline{n}_A,0)$ with price informativeness

$$\mathrm{I}\,(\underline{n}_A,0) = \frac{\underline{n}_A\sigma^2}{\sigma^2(\underline{n}_A+1)+2\sigma^2_\epsilon},$$

it remains to show that $\mathrm{I}\,(\underline{n}_A,0) < \mathrm{I}\,(n_A,n_B)$ is feasible.

To this end, consider a numeric example where, for a given c_A, $n_B = 1$. This is true in equilibrium if $\rho_B(1,n_A) = c_B$, i.e. when the participation constraint of a single B-investor is binding

$$\text{(A2)}\qquad \frac{\sigma^2\sigma_y\left(\sigma^2+2\sigma^2_\epsilon\right)}{\left(2\sigma^2+2\sigma^2_\epsilon\right)^2}\left[\frac{1}{4\left(\sigma^2+\sigma^2_\epsilon\right)} + \frac{n_A\left(\sigma^2+\sigma^2_\epsilon\right)}{[(n_A+1)\sigma^2+2\sigma^2_\epsilon]^2}\right]^{-1/2} = c_B,$$

When the participation constraint of an A-trader is also binding,

$$\text{(A3)}\qquad \frac{\sigma^2\sigma_y\left(\sigma^2+2\sigma^2_\epsilon\right)}{\left((n_A+1)\,\sigma^2+2\sigma^2_\epsilon\right)^2}\left(\frac{n_A\left(\sigma^2+\sigma^2_\epsilon\right)}{[(n_A+1)\sigma^2+2\sigma^2_\epsilon]^2} + \frac{1}{4\left(\sigma^2+\sigma^2_\epsilon\right)}\right)^{-1/2} = c_A,$$

the structure (n_A,n_B) is an equilibrium.

Suppose these conditions hold for c_A, and consider a change in cost to $c'_A < c_A$. This will increase n_A,.and may in turn crowd out the B-investor. When this happens, the equilibrium number of A-investors is determined by

$$\text{(A4)}\qquad \frac{\sigma^2\sigma_y\left(\sigma^2+2\sigma^2_\epsilon\right)}{\left((n_A+1)\,\sigma^2+2\sigma^2_\epsilon\right)^2}\left(\frac{n_A\left(\sigma^2+\sigma^2_\epsilon\right)}{[(n_A+1)\sigma^2+2\sigma^2_\epsilon]^2}\right)^{-1/2} = c'_A.$$

I now compute the equilibrium information structure(s) for $\sigma^2 = 1$, $\sigma_y = 10$, $\sigma^2_\epsilon = 10$, $c_A = 1/2$ and $c'_A = 1/(2.1)$. (I use c_B as my degree of freedom to ensure that $n_B = 1$.) Using condition A3, I first determine the equilibrium number of A-investors for c_A,

$$\frac{210}{(n_A+21)^2}\left(\frac{11n_A}{(n_A+21)^2} + \frac{1}{44}\right)^{-1/2} = \frac{1}{2},$$

which yields $n_A = 12.236$. Similarly, I compute the number of A-investors that enter once the B-investor has been crowded out via condition A4,

$$\frac{210}{\frac{1}{2.1}\sqrt{11}} = (n_A+21)\sqrt{n_A},$$

as $\underline{n}_A = 14.238$. The B-investor will indeed drop out because the maximum of $T(n_A)$ is at $n = 21$ in this case. That is, the increase in n_A occurs in a range where the liquidity externality on B-investors is negative.

Price informativeness in this example drops from

$$\mathrm{I}\left(n_A,1\right)=\frac{12.236}{(12.236+1)+20}+\frac{1}{(1+1)+20}=0.413\,61$$

to

$$\mathrm{I}\left(\underline{n}_A,0\right)=\frac{14.238}{(14.238+1)+20}=0.404\,05.$$

That is, the price becomes less informative. ■

Proof of Lemma 2. First, note that price discrimination without rationing at each price is equivalent to selling unlimited subscriptions at the lowest offered price.

Second, consider price-quantity schedules of the following form: The seller determines a set of prices $p_A^n > p_A^{n-1} > \cdots > p_A^1$ and the maximum number of subscriptions $y_A^1, y_A^2, \ldots, y_A^n$ she is willing to sell at each price. Suppose that a quantity restriction is binding in the sense that more traders would like to purchase a subscription at that price, say p_A^i. There are two possible cases. (i) If it is unprofitable for any additional trader to purchase data at p_A^{i+1}, the information seller fares better by increasing y_A^i and hence the number of subscriptions sold at p_A^i. (ii) If there are traders who purchase data at p_A^{i+1}, the information seller fares better by setting $y_A^i=0$. To see this, note that all A-traders make the same trading profit, irrespective of the individual price paid for the data. Thus, if some traders do not incur a loss when buying data at p_A^{i+1}, the y_A^i traders that buy data at p_A^i make a (total) profit of at least $y_A^i(p_A^i - p_A^{i+1})$. When $y_A^i=0$, these traders would be willing to buy subscriptions at p_A^{i+1}. Thus, having a binding quantity restriction is not optimal. But quantity restrictions that are not binding are unnecessary.

Finally, consider the indirect sale of information, e.g., through a fund. In this case, the seller trades on behalf of its subscribers, and a contract prespecifies a fixed subscription fee and a profit-sharing rule. If only a fixed fee is paid, the seller's trading strategy and hence the fund's expected profit is equivalent to that of a single trader. That is, the seller is better off directly selling the data (to many traders). Now suppose that the sharing rule induces the seller to trade as aggressively as n_A traders. That is, the fund's expected profit will be equal to the combined expected profit of n_A individual traders with *exactly* the same signal. Competition is more intensive for any number of traders with 'photocopied' errors than for the same number of traders with 'personalized' errors (Admati and Pfleiderer, 1986). This is because, in the first case, the traders commonly know that all of them will submit identical orders, perfectly reinforcing each other's impact on the price. As a result, selling data to n_A traders that interpret it differently generates a higher expected profit than trading as aggressively as n_A traders with the same signal. Since the seller extracts the entire trading profits of her subscribers, she is better of selling her signal directly. ■

Proof of Proposition 3. Suppose that $n_A^0(c_A) < \underline{n}_A$ and define $\overline{p}_A$ by $n_A^0(\overline{p}_A) = \underline{n}_A$. Now consider any price $p_A \in (\overline{p}_A, c_A]$. For any such price, $n_A^0(p_A) < \underline{n}_A$, both types of data will be acquired in equilibrium, and the monopolist's gross profit (which is equal to A-traders' total trading profits) is given by

$$\begin{aligned}\Pi_A^g\left(n_A,n_B\right) &= n_A\frac{\sigma^2\sigma_y\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(\left(n_A+1\right)\sigma^2+2\sigma_\epsilon^2\right)^2}\left(\frac{n_A\left(\sigma^2+\sigma_\epsilon^2\right)}{\left[(n_A+1)\sigma^2+2\sigma_\epsilon^2\right]^2}+\frac{n_B\left(\sigma^2+\sigma_\epsilon^2\right)}{\left[(n_B+1)\sigma^2+2\sigma_\epsilon^2\right]^2}\right)^{-1/2}\\ &= \frac{\sigma^2\sigma_y\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(\sigma^2+\sigma_\epsilon^2\right)}\left(\frac{\left((n_A+1)\sigma^2+2\sigma_\epsilon^2\right)^2}{n_A\left(\sigma^2+\sigma_\epsilon^2\right)}+\frac{n_B}{n_A^2\left(\sigma^2+\sigma_\epsilon^2\right)}\frac{\left((n_A+1)\sigma^2+2\sigma_\epsilon^2\right)^4}{\left((n_B+1)\sigma^2+2\sigma_\epsilon^2\right)^2}\right)^{-1/2}.\end{aligned}$$

Now suppose that the monopolist lowers the price to $\overline{p}_A$, thereby crowding out all B-traders. Her gross profit in this case is given by

$$\begin{aligned}\Pi_A^g\left(\underline{n}_A,0\right) &= \underline{n}_A\frac{\sigma^2\sigma_y\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(\left(\underline{n}_A+1\right)\sigma^2+2\sigma_\epsilon^2\right)^2}\left(\frac{\underline{n}_A\left(\sigma^2+\sigma_\epsilon^2\right)}{\left((\underline{n}_A+1)\sigma^2+2\sigma_\epsilon^2\right)^2}\right)^{-1/2}\\ &= \frac{\sigma^2\sigma_y\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(\sigma^2+\sigma_\epsilon^2\right)}\left(\frac{\underline{n}_A\left(\sigma^2+\sigma_\epsilon^2\right)}{\left((\underline{n}_A+1)\sigma^2+2\sigma_\epsilon^2\right)^2}\right)^{1/2}\\ &= \frac{\sigma^2\sigma_y\left(\sigma^2+2\sigma_\epsilon^2\right)\sqrt{\underline{n}_A\left(\sigma^2+\sigma_\epsilon^2\right)}}{\left(\sigma^2+\sigma_\epsilon^2\right)\left(\left(\underline{n}_A+1\right)\sigma^2+2\sigma_\epsilon^2\right)}\end{aligned}$$

It is straightforward but tedious to show that $\underline{n}_A > n_A$. I therefore omit the proof, which rests on the logic that, unless their number increases in response to a fall in p_A, A-traders would earn a positive profit (which cannot be an equilibrium).

I now need to show that $\Pi_A^g(\underline{n}_A, 0) > \Pi_A^g(n_A, n_B)$ or, equivalently, that

$$\frac{\Pi_A^g(\underline{n}_A,0)}{\Pi_A^g(n_A,n_B)} = \frac{\frac{\sigma^2\sigma_y(\sigma^2+2\sigma_\epsilon^2)\sqrt{\underline{n}_A(\sigma^2+\sigma_\epsilon^2)}}{(\sigma^2+\sigma_\epsilon^2)((\underline{n}_A+1)\sigma^2+2\sigma_\epsilon^2)}}{\frac{\sigma^2\sigma_y(\sigma^2+2\sigma_\epsilon^2)}{(\sigma^2+\sigma_\epsilon^2)}\left(\frac{((n_A+1)\sigma^2+2\sigma_\epsilon^2)^2}{n_A(\sigma^2+\sigma_\epsilon^2)}+\frac{n_B}{n_A^2(\sigma^2+\sigma_\epsilon^2)}\frac{((n_A+1)\sigma^2+2\sigma_\epsilon^2)^4}{((n_B+1)\sigma^2+2\sigma_\epsilon^2)^2}\right)^{-1/2}}$$
$$= \frac{\sqrt{\underline{n}_A(\sigma^2+\sigma_\epsilon^2)}}{((\underline{n}_A+1)\sigma^2+2\sigma_\epsilon^2)}\left(\frac{((n_A+1)\sigma^2+2\sigma_\epsilon^2)^2}{n_A(\sigma^2+\sigma_\epsilon^2)}+\frac{n_B}{n_A^2(\sigma^2+\sigma_\epsilon^2)}\frac{((n_A+1)\sigma^2+2\sigma_\epsilon^2)^4}{((n_B+1)\sigma^2+2\sigma_\epsilon^2)^2}\right)^{1/2}.$$

is greater than 1. This condition can be written as

$$1+\frac{\frac{n_B}{((n_B+1)\sigma^2+2\sigma_\epsilon^2)^2}}{\frac{n_A}{((n_A+1)\sigma^2+2\sigma_\epsilon^2)^2}} > \frac{\frac{n_A}{((n_A+1)\sigma^2+2\sigma_\epsilon^2)^2}}{\frac{\underline{n}_A}{((\underline{n}_A+1)\sigma^2+2\sigma_\epsilon^2)^2}}.$$

The value of the left-hand side is greater than 1. The value of the right-hand side is smaller than 1 if $\frac{\underline{n}_A}{((\underline{n}_A+1)\sigma^2+2\sigma_\epsilon^2)^2} > \frac{n_A}{((n_A+1)\sigma^2+2\sigma_\epsilon^2)^2}$, or equivalently

$$T(\underline{n}_A) > T(n_A).$$

This is true because, by the definition of $\underline{n}_A$, $T(n_A)$ is increasing in n_A for $n_A < \underline{n}_A$. Thus, the monopolist is better off crowding out B-information. ■

Proof of Lemma 3. If $\mathrm{E}(K_I) \leq S + K_E$, any uninformative precommitment price preempts entry. That provided, it is clearly a Perfect Bayesian Equilibrium for all incumbent types to choose the monopoly price. It is also straightforward to see that, from the incumbent's perspective, this is a Pareto-dominant equilibrium.

However, if $\mathrm{E}(K_I) > S + K_E$, a pooling price does not preempt entry. Therefore, some (of the more efficient) incumbent types have an incentive to reveal their type in order to deter the challenger. I first conjecture a Perfect Bayesian equilibrium such that all types below a cut-off type K_I^c preempt entry by setting a uniform price p_A^c and all types above K_I^c surrender the market. Incentive-compatibility requires that

$$n_A^*(p_A^c)\rho(n_A^*(p_A^c), n_B^*(p_A^c)) - c_A - K_I < 0 \quad \text{for all } K_I > K_I^c$$

and that

$$n_A^*(p_A^c)\rho(n_A^*(p_A^c), n_B^*(p_A^c)) - c_A - K_I \geq 0 \quad \text{for all } K_I \leq K_I^c.$$

This trivially implies that p_A^* must satisfy

$$n_A^*(p_A^c)\rho(n_A^*(p_A^c), n_B^*(p_A^c)) - c_A - K_I^c = 0.$$

To deter entry, the cut-off value must further satisfy

$$\mathrm{E}(K_I \leq K_I^c) \leq K_I^+.$$

Since $\mathrm{E}(K_I) > S + K_E$ implies $K_I^+ < \mathrm{E}(K_I) < K_H$ and K_I is continuously distributed, there exists a unique $K_I^* \in [K_L, K_H)$ such that all $K_I^c < K_I^*$ satisfy this condition. Since p_A^c increases in the cut-off value K_I^c, the Pareto-dominant Perfect Bayesian equilibrium (from the incumbent's perspective) is to set the cut-off value as high as possible, that is, to $K_I^c = K_I^*$. To establish Pareto-dominance formally, it is easy to verify that all types above K_I^* earn zero profits in any Perfect Bayesian equilibrium, and that all types below K_I^* prefer a higher cut-off value not only because it preempts entry for more incumbent types but also because it increases the precommitment price and hence the profit of any incumbent type. If there is entry, price is set to incumbent's break-even price.

Lower S or K_E make it less likely that the information market is uncontestable ($\mathrm{E}(K_I) \leq S + K_E$). When the market is contestable, lower S or K_E decrease K_1^* and thereby also the equilibrium price. To see this, first note that K_I^+ increases in S and K_E (by definition: $K_I^+ = S + K_E$), that K_I^* increases in K_I^+ (by definition: $\mathrm{E}\left(K_I \leq K_I^*\right) = K_I^+$), and that the equilibrium price p_A^* increases in K_I^* (by definition: $n_A^*(p_A^*)\rho(n_A^*(p_A^*), n_B^*(p_A^*)) = K_I^* + c_A$). By implication, K_1^* and p_A^* increase in S and K_E. ■

Proof of Proposition 4. By Proposition 3, the monopolist expands the population of A-investors just until it reaches the crowding out region. Any further decrease in p_A can only increase the number of B-investors. That p_A tends to be lower for lower S_1, S_2, or K_E follows from Lemma 3 and the fact that K_I^+ increases in S_1, S_2, or K_E. ■

Proof of Proposition 5. It suffices to show that there are aggregate gains from trade that the two investors can achieve by entering into the contract. Without the contract, their joint expected trading profit is given by the sum

$$\begin{aligned}\rho_B(1,n_A)+\rho_A(n_A,1) &= \left[\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(2\sigma^2+2\sigma_\epsilon^2\right)^2}\right][T(1)+T(n_A)]^{-1/2}+\left[\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left[\sigma^2\left(n_A+1\right)+2\sigma_\epsilon^2\right]^2}\right][T(1)+T(n_A)]^{-1/2} \\ &= \left[\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(2\sigma^2+2\sigma_\epsilon^2\right)^2}+\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left[\left(n_A+1\right)\sigma^2+2\sigma_\epsilon^2\right]^2}\right][T(1)+T(n_A)]^{-1/2}.\end{aligned}$$

If investor 1 gives away her information for free, then there will be infinitely many A-investors and $T(\infty)\approx 0$. In that case, their joint trading profit is equal to investor 2's profit as a sole active investor:

$$\rho_B(1,\infty)=\left[\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(2\sigma^2+2\sigma_\epsilon^2\right)^2}\right][T(1)+0]^{-1/2}.$$

I want to know whether, for some initial $n_A>1$, the inequality $\rho_B(1,\infty)>\rho_B(1,n_A)+\rho_A(n_A,1)$ is satisfied. Substituting the above expressions into the inequality gives

$$\begin{aligned}\left[\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(2\sigma^2+2\sigma_\epsilon^2\right)^2}\right][T(1)]^{-1/2} &> \left[\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(2\sigma^2+2\sigma_\epsilon^2\right)^2}+\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left[\left(n_A+1\right)\sigma^2+2\sigma_\epsilon^2\right]^2}\right][T(1)+T(n_A)]^{-1/2} \\ \left[\frac{T(1)+T(n_A)}{T(1)}\right]^{1/2} &> \left[\frac{\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(2\sigma^2+2\sigma_\epsilon^2\right)^2}+\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left[\left(n_A+1\right)\sigma^2+2\sigma_\epsilon^2\right]^2}}{\frac{\sigma_y\sigma^2\left(\sigma^2+2\sigma_\epsilon^2\right)}{\left(2\sigma^2+2\sigma_\epsilon^2\right)^2}}\right] \\ \left[1+\frac{T(n_A)}{T(1)}\right]^{1/2} &> 1+\frac{4\left(\sigma^2+\sigma_\epsilon^2\right)^2}{\left(\sigma^2+2\sigma_\epsilon^2+\sigma^2 n_A\right)^2} \\ 1+\frac{T(n_A)}{T(1)} &> 1+\frac{8\left(\sigma^2+\sigma_\epsilon^2\right)^2}{\left(\sigma^2+2\sigma_\epsilon^2+\sigma^2 n_A\right)^2}+\frac{16\left(\sigma^2+\sigma_\epsilon^2\right)^4}{\left(\sigma^2+2\sigma_\epsilon^2+\sigma^2 n_A\right)^4}\end{aligned}$$

Finally, substituting for $T(\cdot)$ and rearranging yields

$$4n_A-8>\frac{16\left(\sigma^2+\sigma_\epsilon^2\right)^2}{\left(\sigma^2+2\sigma_\epsilon^2+\sigma^2 n_A\right)^2}.$$

The left-hand side is increasing in n_A, whereas the right-hand side is decreasing in n_A. Hence, this inequality is satisfied for sufficiently high n_A. ■

References

Admati, Anat R. and Paul Pfleiderer, 1986, "A Monopolistic Market for Information," Journal of Economic Theory, 39, 400-38.

Admati, Anat R. and Paul Pfleiderer, 1988, "Selling and Trading on Information in Financial Markets", American Economic Review, 78, 96-103.

Admati, Anat R. and Paul Pfleiderer, 1990, "Direct and Indirect Sale of Information," Econometrica, 58, 901-928.

Allen, Franklin, 1990, "The Market for Information and the Origin of Financial Intermediation," Journal of Financial Intermediation, 1, 3-30.

Barberis, Nicholas and Andrei Shleifer, 2003, "Style Investing," Journal of Financial Economics, 68, 161–199.

Baker, Malcolm, Jeremy Stein, and Jeffrey Wurgler, 2003, "When Does the Market Matter? Stock Prices and the Investment of Equity-Dependent Firms," Quarterly Journal of Economics, 118, 969-1006.

Barlevy, Gadi and Pietro Veronesi, 2000, "Information Acquisition in Financial Markets," Review of Economic Studies, 67, 79–90.

Besley, Timothy and Andrea Prat, 2006, "Handcuffs for the Grabbing Hand?: Media Capture and Government Accountability," American Economic Review 96, 720–736.

Bushman, Robert, Joseph Piotroski and Abbie Smith, 2004, "What Determines Corporate Transparency?" Journal of Accounting Research, 42, 207-252.

Campbell, John Y., Martin Lettau, Burton G. Malkiel and Yexiao Yu, 2001, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," Journal of Finance, 56, 1-43.

Cespa, Giovanni, 2008, "Information Sales and Insider Trading with Long-lived Information," Journal of Finance, 63, 639-672.

Chan, Louis K. C., Hsiu-Lang Chen and Josef Lakonishok, 2002, "On Mutual Fund Investment Styles," Review of Financial Studies, 15, 1407-1437.

Chan, Kalok and Allaudeen Hameed, 2006, "Stock Price Synchronicity and Analyst Coverage in Emerging Markets," Journal of Financial Economics, 80, 115–147.

Chen, Qi, Itay Goldstein and Wei Jiang, 2007, "Price Informativeness and Investment Sensitivity to Stock Price," Review of Financial Studies, 20, 619-650.

Chordia, Tarun, Richard Roll and Avanidhar Subrahmanyam, 2000, "Commonality in Liquidity," Journal of Financial Economics 56, 3-28.

Chordia, Tarun, Richard Roll and Avanidhar Subrahmanyam, 2008a, "Liquidity and Market Efficiency," Journal of Financial Economics 87, 249–268.

Chordia, Tarun, Richard Roll and Avanidhar Subrahmanyam, 2008b, "Why Has Trading Volume Increased?," Emory University and UCLA.

Colla, Paolo and Antonio Mele, 2008, "Information Linkages and Correlated Trading," Review of Financial Studies, forthcoming.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers and Robert J. Waldmann, 1990, "Noise Trader Risk in Financial Markets," Journal of Political Economy, 98, 703-738.

Dewatripont, Mathias and Jean Tirole, 2005, "Modes of Communication," Journal of Political Economy, 113(6), 1217-1238.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer, 2002, "The Regulation of Entry," Quarterly Journal of Economics, 117, 1-37.

Djankov, Simeon, Caralee McLiesh, Tatiana Nenova and Andrei Shleifer, 2003, "Who Owns the Media?" Journal of Law and Economics, 46, 341-381.

Dow, James, Itay Goldstein and Alexander Gümbel, 2006, "Incentives for Information Production in Markets when Prices Affect Real Investments," London Business School, The Wharton School, Said Business School.

Dow, James and Gary Gorton, 1997, "Stock Market Efficiency and Economic Efficiency: Is There a Connection?" Journal of Finance, 52, 1087-1129.

Durnev, Artyom, Randall Morck, Bernard Yeung and Paul Zarowin, 2003, "Does Greater Firm-specific Return Variation Mean More or Less Informed Stock Pricing?" Journal of Accounting Research, 41, 797-836.

Durnev, Artyom, Randall Morck and Bernard Yeung, 2004, "Value-Enhancing Capital Budgeting and Firm-specific Stock Return Variation," Journal of Finance, 59, 65-105.

Fama, Eugene F., 1970, "Efficient Capital Markets: A Review of Theory and Empirical Work," Journal of Finance, 25, 383-417.

Figlewski, Stephen, 1982, "Information Diversity and Market Behavior," Journal of Finance, 37, 87-102.

Fishman, Michael and Kathleen Hagerty, 1995, "The Incentive to Sell Financial Market Information," Journal of Financial Intermediation, 4, 95-115.

Fox, Merritt, Artyom Durnev, Randall Morck and Bernard Yeung, 2003, "Law, Share Price Accuracy and Economic Performance: The New Evidence," Michigan Law Review, 102, 331-386.

Gabaix, Xavier, David Laibson, Guillermo Moloche, Stephen E. Weinberg, 2006, "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model," American Economic Review, 96, 1043–1068.

Garcia, Diego and Francesco Sangiorgi, 2007, "Information Sales and Strategic Trading," University of North Carolina at Chapel Hill and Stockholm School of Economics.

Goetzmann, William N. and Stephen J. Brown, 2003, "Hedge Funds with Style," Journal of Portfolio Management, 29, 101-112.

Goetzmann, William N., Vicente Pons-Sanz and S. Abraham Ravid, 2004, "Soft Information, Hard Sell: The Role of Soft Information in the Pricing of Intellectual Property," Yale ICF Working Paper 04-16.

Goldstein, Itay and Alexander Gümbel, 2006, "Manipulation and the Allocational Role of Stock Prices," Review of Economic Studies, forthcoming.

Grossman, Sanford and Joseph Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets," American Economic Review, 70, 81-101.

Hameed, Allaudeen, Randall Morck and Bernard Yeung, 2005, "Information Markets, Analysts, and Comovement in Stock Returns," New York University.

Hasbrouck, Joel and Duane Seppi, 2001, "Common factors in prices, order flows, and liquidity," Journal of Financial Economics 59, 383-411.

Hayek, Friedrich A., 1945, "The Use of Knowledge in Society," American Economic Review, 35, 519-530.

Hellwig, Christian and Laura Veldkamp, 2008, "Knowing What Others Know: Coordination Motives in Information Acquisition," Review of Economic Studies, forthcoming.

Holmström, Bengt and Jean Tirole, 1993, "Market Liquidity and Performance Monitoring," Journal of Political Economy, 101, 678-709.

Hong, Harrison, Jeremy Stein and Jialun Yu, 2007, "Simple Forecasts and Paradigm Shifts," Journal of Finance, 62(3), 1207-1242.

Hong, Harrison, Walter Torous and Rossen Valkanov, 2007, "Do Industries Lead Stock Markets?" Journal of Financial Economics, 83, 367–396.

Horwitz, Robert B., 2004, "On Media Concentration and the Diversity Question," Information Society, fothcoming.

Huberman, Gur, 2001, "Familiarity Breeds Investment," Review of Financial Studies, 14, 659–680.

Huberman, Gur and Tomer Regev, 2001, "Contagious Speculation and a Cure for Cancer: A Non-event That Made Stock Prices Soar," Journal of Finance, 56, 387–396.

Huberman, Gur and Dominika Halka, 2001, "Systematic liquidity," Journal of Financial Research, 24, 161-178.

Jin, Li and Stewart C. Myers, 2007, "R^2 Around the World: New Theory and New Tests," Journal of Financial Economics, 79, 257–292.

Kahneman, Daniel, 1973, Attention and Effort, Prentice-Hall, Englewood Cliffs, New Jersey.

Kyle, Albert S., 1985, "Continuous Auctions and Insider Trading," Econometrica, 53, 1315-1335.

Li, Yan and Liyan Yang, 2008, "Complementarities in Information Acquisition with Heterogeneous Investment Opportunities," Cornell University.

Read, Donald, 1992, "The Power of News: The History of Reuters 1849-1989," Oxford University Press.

Massa, Massimo and Andrei Simonov, 2005, "Hedging, Familiarity and Portfolio Choice," Review of Financial Studies, 19, 633-685.

Michaely, Ron and Kent Womack, 1999, "Conflicts of Interest and the Credibility of Underwriter Analyst Recommendations," Review of Financial Studies, 12, 653-686.

Morck, Randall, Bernard Yeung and Wayne Yu, 2000, "The Information Content of Stock Markets: Why Do Emerging Markets Have Synchronous Stock Price Movements?" Journal of Financial Economics, 58, 215-260.

Moscarini, Guiseppe, 2004, "Limited Information Capacity as a Source of Inertia," Journal of Economic Dynamics and Control, 28, 2003 – 2035.

New York Times, "Combined Thomson-Reuters Would Challenge Bloomberg," 9 May 2007.

Peng, Lin, 2005, "Learning with Information Capacity Constraints," Journal of Financial and Quantitative Analysis, 50(2), 307-329.

Peng, Lin and Wei Xiong, 2006, "Investor Attention, Overconfidence and Category Learning," Journal of Financial Economics, 80, 563–602.

Piotroski, Joseph D. and Darren T. Roulstone, 2004, "The Influence of Analysts, Institutional Investors, and Insiders on the Incorporation of Market, Industry, and Firm-Specific Information into Stock Prices," Accounting Review, 79, 1119-1151.

Read, Donald, 1999, The Power of News: The History of Reuters, Oxford University Press, Oxford.

Roll, Richard, 1988, "R^2," Journal of Finance, 43, 541-566.

Simon, Herbert, 1957, "A Behavioral Model of Rational Choice," in: Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting, Wiley, New York.

Sims, Christopher A., 2003, "Implications of Rational Inattention," Journal of Monetary Economics, 50, 665–690.

Sims, Christopher A., 2006, "Rational Inattention: Beyond the Linear-Quadratic Case," American Economic Review, 96, 158-163.

Subrahmanyam, Avanidhar and Sheridan Titman, 1999, "The Going-Public Decision and the Development of Financial Markets," Journal of Finance 54, 1045-1082.

Subrahmanyam, Avanidhar, 1991, "Risk Aversion, Market Liquidity, and Price Efficiency," Review of Financial Studies, 4, 417-441.

Subrahmanyam, Avanidhar and Sheridan Titman, 2001, "Feedback from Stock Prices to Cash Flows," Journal of Finance, 56, 2389-2413.

Tobin, James, 1982, "On the Efficiency of the Financial System," Lloyd's Banking Review, 153, 1-15.

Van Nieuwerburgh, Stijn and Laura Veldkamp, 2008a, "Information Immobility and the Home Bias Puzzle," Journal of Finance, forthcoming.

Van Nieuwerburgh, Stijn and Laura Veldkamp, 2008b, "Information Acquisition and Under-Diversification," New York University.

Veldkamp, Laura, 2006a, "Media Frenzies in Markets for Financial Information," American Economic Review, 96, 577-601.

Veldkamp, Laura, 2006b, "Information Markets and the Comovement of Asset Prices," Review of Economic Studies, 73, 823-845.

Verrecchia, Robert E., 1982, "Information Acquisition in a Noisy Rational Expectations Economy," Econometrica, 50, 1415-30.

Williamson, Oliver, 1981, "The Economies of Organization: The Transaction Cost Approach," American Journal of Sociology, 87, 548-577.

Wurgler, Jeffrey, 2000, "Financial Markets and the Allocation of Capital," Journal of Financial Economics, 58, 187-214.

PAPER 2

Security-Voting Structure and Bidder Screening

with Christian At and Mike Burkart

ABSTRACT. This paper shows that non-voting shares can promote takeovers. When the bidder has private information, shareholders may refuse to tender because they suspect to sell at an ex-post unfavourable price. The ensuing friction in the sale of cash flow rights can prevent an efficient sale of control. Separating cash flow and voting rights mitigates this externality, thereby facilitating takeovers. In fact, the fraction of non-voting shares can be used to discriminate between efficient and inefficient bidders. The optimal fraction decreases with managerial ability, implying an inverse relationship between firm value and non-voting shares. As non-voting shares increase control contestability, share reunification programs entrench managers of widely held firms, whereas dual-class recapitalizations can increase shareholder wealth.

1. Introduction

Dual-class shares in publicly traded firms continue to be controversial. The New York Stock Exchange used to deny listings to firms with multiple share classes but abandoned this requirement in 1986. Similarly, the European Commission recently withdrew a proposal to mandate the one share - one vote principle, which would have banned shares with differential voting rights and voting restrictions. If these provisions had been adopted, they would have affected a large number of firms: According to a survey commissioned by the Association of British Insurers, 29 percent of the top 300 European companies in 2005 had dual-class share structures. In the US dual-class shares are less frequent but still common. They are used in about 6 percent of all publicly-traded firms (Gompers et al., 2007).

Proponents of the one share - one vote rule argue that it is most conducive to an efficient allocation of corporate control. The theoretical foundation of this view

Thanks to Tore Ellingsen, Mara Faccio, Denis Gromb, Oliver Hart, Ulrich Hege, Henry Hu, Linus Kaisajuntti, Pedro Matos, Ernst Maug, Anete Pajuste, Salvatore Piccolo and seminar participants at the FIRS Conference (Anchorage), Amsterdan/Vanderbilt conference on Shareholder Voting, the ESSFM (Gerzensee), HEC (Paris), the LBS/ECGI/ESRC Conference on Corporate Governance, London School of Economics, SIFR/SITE (Stockholm), and the Universities of Frankfurt and Zurich for helpful comments and discussions. Financial support from the European Corporate Governance Training Network (ECGTN) and the Jan Wallander Foundation is gratefully acknowledged.

is the analysis of Grossman and Hart (1988) and Harris and Raviv (1988). In their framework, the security benefits and private benefits vary across bidders who compete for a dispersedly held firm. Since bidders compete for voting shares, one share - one vote prevents that bidders' willingness-to-pay for control and their ability to create value diverge, thereby ensuring an efficient control allocation. By contrast, deviations from one share - one vote bear the risk that an inefficient bidder with large private benefits outbids more efficient bidders. At the same time, a dual-class share structure may be in the shareholders' interest as it allows to extract more surplus from the winning bidder.

It has to be noted that the security-voting structure matters for the control allocation in this framework only if the bidders' ranking according to security benefits differs from their ranking according to private benefits. If the most efficient bidder also has the largest private benefits, she wins the bidding contest irrespective of the security-voting structure. Moreover, the security-voting structure is immaterial for bid price and shareholder wealth in the absence of (effective) competitors. Due to the target shareholders' free-rider behavior (Grossman and Hart, 1980), the bid price must under full information match the winning bidder's security benefits. Nonetheless, one share - one vote is optimal in the sense that no other security-voting structure leads to a more efficient control allocation in this framework.

The present paper shows that asymmetric information undermines the dominance of one share - one vote. Asymmetric information can lead to disagreement about what constitutes an acceptable price which in turn may prevent a control transfer. The root of this failure is that cash flow and control rights have to be jointly traded. Separating cash flow and voting rights mitigates the impact that the disagreement about the value of the cash flow rights has on the trade of votes. Thus, contrary to the prevailing view, one share - one vote does not ensure an efficient control allocation and is typically inferior to a dual-class share structure.

We develop this idea in a tender offer model with atomistic target shareholders and absent (effective) competition. Instead, it assumes a single bidder who has private information about her ability to create value. As a result, the bid price is determined by the shareholders' free-rider behavior and must at least equal the expected post-takeover share value. Costly bids are feasible because the bidder can extract part of the value generated by her as private benefits. Within this framework, we demonstrate that one share - one vote maximizes the severity of the asymmetric information problem, thereby deterring too many value-increasing takeovers.

In our model, all types of bidders who make a bid in equilibrium offer the same price. While the equilibrium price is equal to the average post-takeover share value, some (overvalued) bidder types pay more and some (undervalued) types pay less than their true post-takeover share value. In addition, there is a cut-off value, and types who generate less value are deterred as they would make a loss when offering the equilibrium price. Hence, the presence of asymmetric information has two effects. First, it causes redistribution among all bidder types who actually make a bid. Second, it exacerbates the free-rider problem as ceteris paribus more bids fail than under symmetric information.

Separating cash flow and voting rights affects the takeover outcome by altering how shareholders up-date their expectations. (More) non-voting shares reduce the fraction of return rights that bidders purchase and therefore render a bid ceteris paribus more profitable for overvalued types. Hence, some formerly frustrated types can earn a profit and now make a bid. In response, shareholders revise their beliefs about the post-takeover share value downward. This in turn lowers the bid price at which shareholders are willing to tender and makes the takeover profitable for further types.

The monotone relationship between the fraction of voting shares and the cut-off value implies that the security-voting structure can be used to discriminate among desirable and undesirable bids. Unless takeover costs are either too large or too small relative to the bidder's private benefits, the socially optimal structure implements the first-best outcome: all and only bids with value improvements in excess of the takeover cost succeed. Since target shareholders abstract from the takeover cost, they only bear part of the social cost of deterring a bid. Therefore, they choose a lower cut-off value, or equivalently, a higher fraction of non-voting shares. Our theory provides a rationale for dual-class recapitalization. It increases the likelihood of a subsequent takeover which in turn translates into a higher share price.

The optimal security-voting structure varies with the incumbent management's quality. More non-voting shares increase the probability that the incumbent manager is replaced, which is warranted for less able managers. Conversely, it is optimal to protect very good management from the takeover threat with the one share - one vote structure. Thus, our model consents with the common perception that the merit of the one share - one vote structure is to prevent value-decreasing bids (Grossman and Hart, 1988). But it also implies that share class unification entrenches managers of widely held firms which is neither socially desirable nor in the shareholders' interest.

As a firm's current market value improves with the quality of its management, our model predicts that firms with (more) non-voting shares have lower market values. While this prediction is consistent with the empirical evidence, the underlying intuition runs counter to the usual explanation that dual-class shares destroy value because they enable corporate insiders to extract more private benefits (Bebchuk et al., 2000). In our model, the use of non-voting shares is an optimal response to low firm value under the incumbent management, as it increases the likelihood of a value-increasing takeover.

Another important determinant of the optimal security-voting structure is the bidder's ability to extract private benefits. Weak legal shareholder protection (higher extraction rates), like non-voting shares, promotes takeovers. Thus, the optimal fraction of voting shares increases with private benefits, suggesting that the case for one share - one vote is strongest in countries with weak investor protection.

Finally, shareholders prefer to promote takeovers by lowering the fraction of voting shares rather than allowing bidders to extract more private benefits. Extraction transfers wealth from shareholders to bidders, whereas the security-voting structure merely affects the extent of redistribution among bidder types. This result stands out against the common view which advocates private benefits (higher extraction rates) as a means to mitigate the free-rider problem and the one share - one vote structure as a means to deter value-decreasing bidders (Grossman and Hart, 1980, 1988).

As noted earlier, the existing literature on the security-voting structure considers a competitive setting to derive the optimality of one share - one vote. While a successful bid must exceed any (potential) rival offer, it must also win shareholder approval. We intentionally presuppose that shareholder approval is the binding constraint and consider a single bidder who makes one offer. The single-bidder assumption does not literally rule out other parties interested in controlling the firm. It merely presumes that no competitor can create nearly as much value as the bidder under consideration. In fact, the optimal security-voting structure in our framework is such that the incumbent manager would not want to match the equilibrium offer, as she does not create enough value.

Takeover studies document that the single-bidder setting is empirically relevant. For instance, Betton and Eckbo (2000) report that 62 percent of all US tender offer contests (1,353) between 1971 to 1990 involve only one bid. This does not imply that shareholder approval is the binding constraint in all these cases. The single bid may instead have been set above the target shareholders' reservation price to deter potential rivals. However, this hardly holds for the 22 percent of single bids which failed. Further support for shareholder approval being the binding constraint comes from multiple-bid takeovers. In this subsample, all bids are made by the same bidder in 41 percent of the cases. In addition, these bid revisions can only in very few cases be attributed to rumored competition.

Besides the single-bidder setting, our analysis builds on two further assumptions. First, we assume that the bid price must satisfy the free-rider condition. The key consequence of the free-rider behaviour in our model is that the bidder has private information about the target shareholders' reservation price. Thus, our main result that the separation of cash flow and voting rights reduces inefficiencies in control transactions applies more generally to settings where the buyer knows more about the sellers' outside option. As regards tender offers, the results extend to any ownership structure where the majority of voting rights are dispersedly held. Such ownership patterns are by no means unusual.[1]

Second, we assume a constant extraction rate. This is meant to reflect circumstances where the bidder's ability to expropriate shareholders is self-evident (e.g., from her corporate charter) or is determined by industry and institutional characteristics. The constant extraction rate implies a positive correlation between security and private benefits. As a result, shareholders overvalue in equilibrium bidder types with small private benefits and undervalue those with large private benefits. Increasing the fraction of voting shares therefore discourages types with a low propensity to bid. When security benefits and private benefits are inversely related, low private benefit types are undervalued, and redistribution among types encourages takeovers. Thus, some of our results - like others in this literature - are sensitive to the relationship between security benefits and private benefits. Nonetheless, the conclusion remains that one share - one vote need not be socially optimal. Section 4.3 discusses these issues in more detail.

Several papers analyze tender offer games with a single bidder who has private information. Grossman and Hart (1981) establish that takeovers require an information advantage about the value improvement brought about by the bidder. That is, if takeovers were solely motivated by the bidder's knowledge that the target is undervalued, rational shareholders would not tender. Shleifer and Vishny (1986) show that the acquisition of a stake prior to the tender offer provides a partial solution to the free-rider problem. Their (partial) pooling equilibrium anticipates the equilibrium outcome in our benchmark case with a single share class. The difference is that the source of the bidder's gains is private benefit extraction rather than toeholds. In a model with two bidder types, Marquez and Yilmaz (2005) demonstrate that uncertainty about the post-takeover security benefits may make bids profitable for the type with high security benefits and low private benefits. Hirshleifer and Titman (1990) and Chowdry

[1] In the sample of Gompers et al. (2007), which comprises all dual-class firms in the US between 1995 and 2002, corporate insiders do not have the majority of votes in about a third of the observations. In the sample of Pajuste (2005), which covers 493 dual-class firms from seven European countries (Denmark, Finland, Germany, Italy, Norway, Sweden and Switzerland) during 1996 to 2002, the two largest shareholders own together less than 20 percent of the votes in about a quarter of the firms. In the subsample of all firms (63) that were taken over, the majority of Swedish targets (16 out of 25) had widely held dual-class shares.

and Jegadeesh (1994) analyze models in which takeover outcomes are probabilistic and equilibrium offers fully reveal the bidder's type.[2] None of these papers examine the role of the security-voting structure.

As discussed above, Grossman and Hart (1988) and Harris and Raviv (1988) show that forcing a would-be acquirer to purchase all return rights ensures an efficient control allocation in a bidding competition but may not maximize shareholder wealth. Bergström et al. (1997) and Cornelli and Felli (2000) revisit these effects in the context of the mandatory bid rule and the sale of a bankrupt firm. In Burkart et al. (1998), deviations from one share - one vote can be socially optimal, though there is no comprehensive analysis of the optimal security-voting structure. Moreover, the mechanisms through which non-voting shares affect the takeover outcome differ. In their model, the fraction of voting shares determines the bidder's private benefits as opposed to the shareholders' expectations about the post-takeover security benefits. Gromb (1992) shows in a framework with a finite number of shareholders that non-voting shares mitigate the free-rider problem. Reducing the number of voting shares makes each voting shareholder more likely to be pivotal and increases their tendering probability.[3] His model, contrary to ours, predicts that voting shares trade at a discount.

The paper is organized as follows. Section 2 outlines the model and derives the pooling equilibrium in the simple case with a single share class and value-increasing bidders. Section 3 solves the model for a dual-class target and demonstrates that deviations from one share - one vote mitigate the asymmetric information problem. Section 4 introduces value-decreasing bidders and shows that the security-voting structure can be used to screen bidder types. We derive the socially and the shareholders' optimal security-voting structure and examine the comparative static properties of these structures. Concluding remarks are in Section 5, and the mathematical proofs are in the Appendix.

2. Model

Consider a widely held firm that faces a single potential acquirer, henceforth the bidder B. If the bidder gains control, she can generate revenues V. While the bidder learns her type prior to making the tender offer, target shareholders merely know that the revenues V are distributed on $\left[\underline{V}, \overline{V}\right]$ according to the continuously differentiable density function $g(V)$. The cumulative density function is denoted by $G(V)$.

[2] Separating cash flow and voting rights also promotes takeovers in variants of the tender offer game that allow for separating equilibrium outcomes (Burkart and Lee, 2008).

[3] For the same reason, super-majority rules increase the takeover probability (Holmström and Nalebuff, 1992).

In addition, the bidder can divert part of the revenues as private benefits. The non-contractible diversion decision is modelled as the bidder's choice of $\phi \in [0, \bar{\phi}]$, such that security benefits (dividends) are $X = (1-\phi)V$ and her private benefits are $\Phi = \phi V$. The upper bound $\bar{\phi} \in (0,1)$ is commonly known and identical for all bidder types. The latter assumption will be relaxed in Section 4.3 where we allow bidder types to differ in their extraction abilities.

Tender offers are the only admissible mode of takeover, and a successful offer requires that the bidder attracts at least 50 percent of the firm's voting rights. To illustrate the workings of the model, we first consider the one share - one vote structure and defer the analysis of dual-class shares to subsequent sections. If the takeover succeeds, the bidder incurs a fixed cost K of administrating the takeover which is independent of her type and common knowledge.

If the takeover does not materialize, the incumbent manager remains in control. The incumbent can generate revenues V^I which are known to all shareholders. Like the bidder, she can extract a fraction $\phi \in [0, \bar{\phi}]$ of the revenues. Hence, shareholders obtain $X^I = (1-\phi)V^I$ in the absence of a takeover. Initially, we restrict attention to value-increasing bids and set $V^I = \underline{V}$. The sequence of events unfolds as follows.

In stage 1, the bidder learns her type V and decides whether to make a take-it-or-leave-it, conditional, unrestricted tender offer. If she does not make a bid, the game moves directly to stage 3. If she makes a bid, she offers to purchase all shares for a total price P, provided that at least 50 percent of the shares (voting rights) are tendered. Moreover, the offer must be for cash, which precludes that its terms depend on the future observation of V. We discuss this last assumption on the bid form at the end of this section.

In stage 2, the target shareholders non-cooperatively decide whether to tender their shares. Shareholders are homogeneous and atomistic and do not perceive themselves as pivotal for the tender offer outcome.

In stage 3, if at least 50 percent of the shares are tendered, the bidder gains control and pays the price P and the cost K. Otherwise, the incumbent manager remains in control. In either case, the controlling party chooses which fraction ϕ of the revenues to divert as private benefits, subject to the constraint $\phi \leq \bar{\phi}$.

Given that private benefit extraction entails no deadweight loss, the stage 3 diversion decision is straightforward. Setting $\phi = \bar{\phi}$ is a successful bidder's (weakly) dominant strategy, and the post-takeover security benefits are independent of the size of the bidder's final stake. If the bid fails or does not materialize, the incumbent manager chooses likewise the maximum extraction rate $\bar{\phi}$ as she owns no equity.

Since shareholders are atomistic, each of them tenders at stage 2 only if the offered price at least matches the expected security benefits. Shareholders condition their expectations on the offered price P, the known takeover cost K and the anticipated extraction decision $\phi = \bar{\phi}$. Hence, a successful tender offer must satisfy the free-rider condition

$$P \geq \mathrm{E}\left(X|P, K, \phi\right) = (1-\bar{\phi})\mathrm{E}\left(V|P, K\right).$$

For simplicity, we assume that shareholders tender unless retaining is weakly dominant. This ensures the existence of a unique equilibrium outcome: When the free-rider condition is violated, the bid fails. Otherwise, success is the unique equilibrium outcome, and the bidder acquires all shares.[4]

At stage 1, the bidder is willing to offer at most $V-K$ as a successful offer attracts all shares. Thus, the bidder's participation constraint is simply $V-K \geq P$.

To avoid trivial outcomes, we impose a joint restriction on takeover cost, maximum extraction rate and the support of bidder types.

Assumption 1. *$\bar{\phi}\underline{V} < K < \bar{\phi}\overline{V}$.*

These restrictions ensure that some but not all bidder types can make a profitable bid when paying a price equal to their respective post-takeover security benefits. This in turn excludes outcomes where either all or no bidder types make an offer.

In any Perfect Bayesian Equilibrium, the bidder must have correct expectations about which bid prices are acceptable and prefer the smallest successful offer. Given shareholders by assumption tender when the free-rider condition is satisfied, this immediately rules out equilibria in which offers succeed at different prices. As there can only be a single equilibrium price P^*, shareholders infer from observing a bid that it may come from any type who makes a non-negative profit at that price. Thus, the shareholders' conditional expectations about the post-takeover security benefits are

$$\mathrm{E}\left(X|P^*, K\right) = (1-\bar{\phi})\mathrm{E}\left(V|V \geq P^* + K\right).$$

Given the distribution of V, a bid is therefore made and succeeds in equilibrium if the bidder's participation constraint

$$V - P^* \geq K \tag{2.1}$$

[4] Given a bid is conditional, a shareholder who believes the bid to fail is indifferent between tendering and retaining. Imposing this belief on all shareholders and breaking the indifference in favour of retaining supports failure as an equilibrium, irrespective of the offered price (Burkart et al., 2006). To avoid co-existence of success and failure as equilibrium outcomes, it is typically assumed that shareholders tender when they are indifferent. Contrary to our assumption, this precludes failure as the equilibrium outcome for a conditional bid, and hence the existence of an equilibrium when the free-rider condition is violated.

and the free-rider condition

$$P^* \geq \left(1 - \bar{\phi}\right) \int_{P^*+K}^{\overline{V}} \frac{g(V)}{1 - G[P^* + K]} V dV \tag{2.2}$$

hold.

There exists a continuum of prices that satisfy these conditions and so constitute Perfect Bayesian Equilibria of the tender offer game. Following Shleifer and Vishny (1986), we select the minimum bid equilibrium which is the unique equilibrium satisfying the credible beliefs criterion of Grossman and Perry (1986). All other equilibria require shareholders to believe that bidders generate, on average, security benefits that are smaller than the offered equilibrium price. (Details of the equilibrium selection are provided in the Appendix.)

PROPOSITION 1. *Given the one share - one vote structure, only types $V \in [V^c, \overline{V}]$ make a bid and offer the same price P^* where $V^c = P^* + K$ and*

$$P^* = \min\left\{P : P = (1 - \bar{\phi})\mathrm{E}\left(V | V \geq P + K\right)\right\}.$$

Since a bidder can appropriate part of the revenues as private benefits, some value-increasing bids succeed in equilibrium despite the target shareholders' free-rider behavior. Nonetheless, all bidder types below the cut-off value V^c are frustrated.[5] Asymmetric information aggravates the free-rider problem which becomes most apparent when considering the bidder's participation constraint $\bar{\phi}V \geq [(P^* - X) + K]$. In a full information setting, free-riding would imply $P^* = X$, and all bidder types with $\bar{\phi}V \geq K$ would make a successful bid. Under asymmetric information, $P^* = X$ holds on average but not for each individual bidder type. Instead, some types pay more and others less than their respective post-takeover security benefits. The mispricing deters some types whose private benefits are sufficient to cover the takeover cost. That is, the cut-off value V^c under asymmetric information exceeds $K/\bar{\phi}$ (the cut-off value under full information).[6] Moreover, bidder types $V \in [K/\bar{\phi}, V^c)$ cannot succeed with a lower offer because all types $V \geq V^c$ would then make the same offer, and target shareholders would on average be offered less than the post-takeover security benefits. Hence, asymmetric information exacerbates the free-rider problem and prevents some bids, even though they would be value-increasing.

[5] In an extension with private benefit extraction, Chowdry and Jeegadesh (1994) derive an equilibrium in which a subset of types also offer an uninformative bid price.

[6] To see this, rewrite the cut-off type's participation constraint, $V^c - P^* \geq K$, as $\bar{\phi}V^c - K \geq P^* - X^c$. If $P^* > X^c$, this inequality requires that $\bar{\phi}V^c > K$, or equivalently $V^c > K/\bar{\phi}$. Finally, note that indeed $P^* = \mathrm{E}(X | X \geq X^c) > X^c$ as $X^c < \overline{X}$ by Assumption 1.

COROLLARY 1. *The takeover probability decreases with the takeover cost and increases with private benefits.*

The ex ante probability of a takeover corresponds to the probability that a bidder type exceeds the cut-off value V^c. Accordingly, the corollary follows from the fact that the cut-off value increases in K but decreases in $\bar{\phi}$.

When the takeover cost increases, any bidder who can still break even must on average generate higher revenues. As a bid signals higher post-takeover security benefits, target shareholders only tender at a higher price. This increases the cut-off value, thereby decreasing the takeover probability.

By contrast, larger private benefits ($\bar{\phi}$-values) not only enable bidders to recoup the takeover cost more easily but also lower the post-takeover share value. Both effects induce target shareholders to revise their expectations about the post-takeover share value downward. This lowers the equilibrium bid price and cut-off value.

In our framework with cash offers, the equilibrium price only reveals that the bidder's valuation V is above the cut-off value V^c. As a consequence, the free-rider problem is aggravated, and more bidder types are deterred relative to the full information setting. Relaxing the restrictions on the bid form can help to overcome the asymmetric information problem. Indeed, an all-security exchange offer replicates the full information outcome, although it does not reveal the bidder's type: Shareholders accept a bid that exchanges each share against a new share, thereby preserving their fraction of the cash flow rights. If the offer were to exchange shares at less than a one-to-one ratio, each shareholder would reject it. Moreover, all bidder types whose private benefits are sufficient to cover the takeover cost are willing to make such a one-to-one security exchange offer.[7]

Although the all-security exchange offer resolves the asymmetric information problem, it is unconvincing for two reasons. First, it leaves all cash flow rights with the shareholders, making it equivalent to a simple replacement of management. This begs the question why a takeover is needed in the first place. Second, the bidder gains control only if she offers non-voting equity, or at least separates the majority of the votes from the cash flow rights. Thus, the takeover implements a new security-voting structure. More generally, once bidders are allowed to freely recombine votes and cash flow rights, the existing security-voting structure becomes irrelevant for the takeover outcome (Hart, 1995). Like previous models in this literature (e.g., Grossman and

[7] Contrary to the bilateral bargaining models (e.g., Hansen, 1987; Eckbo et al., 1990), Burkart and Lee (2008) show that neither restricted bids nor the means of payment (mix of a cash and equity) can serve as a signal in tender offer games.

Hart, 1988), we treat the security-voting structure as a constraint rather than a choice variable of the bidder.

3. Non-Voting Shares and Takeover Activity

We now explore the impact of dual-class shares on the takeover outcome. More specifically, the target firm has a fraction $s \in (0,1]$ of voting shares entitled to the same (pro-rata) cash flow rights as the $1-s$ non-voting shares. Here we treat the fraction s as a parameter and analyze its optimal choice in the next section.

The takeover bid and the decision to tender proceed under the same premises as before. In addition, the tender offer may discriminate between share classes but not within the same class. Thus, the bidder may quote different prices for voting and non-voting shares. However, if she submits a price for a certain share class, she has to buy all tendered shares from that class, conditional upon a control transfer.[8]

To derive the equilibrium, we initially assume that the bidder only offers to buy voting shares. As we will show below, this is (part of) the optimal bidding strategy. Since either all or none of the voting shareholders tender in equilibrium, a bidder has to pay sP to gain control. Hence, the bidder's participation constraint is

$$\bar{\phi}V - s\,(P-X) \geq K. \tag{3.1}$$

Upon observing a bid, shareholders infer that the bidder can make a profit when buying all s voting shares at that price. Consequently, their expectations are

$$\mathrm{E}\,[X\,|V \geq V^c(s,P)] = \left(1-\bar{\phi}\right)\int_{V^c(s,P)}^{\overline{V}} \frac{g(V)}{1-G[V^c(s,P)]} V dV \tag{3.2}$$

where

$$V^c(s,P) \equiv \frac{sP+K}{\bar{\phi}+s\left(1-\bar{\phi}\right)}.$$

In equilibrium, the bid price must at least match these expectations. As before, we impose the credible beliefs criterion to select the minimum bid equilibrium.

Lemma 1. *Given $s \in [0,1]$, only types $V \in [V^c(s),\bar{V}]$ make a bid and offer the same price $sP^*(s)$ for the voting shares where $V^c(s) = V^c(s,P^*(s))$ and*

$$P^*(s) = \min\left\{P : P = (1-\bar{\phi})\mathrm{E}\left(V|V \geq V^c(s,P)\right)\right\}.$$

Lemma 1 replicates Proposition 1 for a target firm with dual-class shares. All types who make a bid offer the same price which is equal to their average post-takeover

[8] The assumption that a bid has to be unrestricted for a given class is not crucial. Indeed, one can easily replicate the analysis of intra-class restricted bids by redefining s. For example, restricted offers for half of the voting shares are equivalent to unrestricted offers for all $s' = s/2$ voting shares.

security benefits. Hence, a given type purchases the s voting shares either at a premium ($P^*(s) > X$) or at a discount ($P^*(s) < X$), but the gains of the undervalued types are exactly offset by the losses of the overvalued types. Furthermore, all types $V < V^c(s)$ abstain from bidding because the cost of purchasing overpriced (voting) shares exceeds their private benefits net of the takeover cost.

Since non-voting - like voting - shareholders only tender if the bidder offers at least the post-takeover security benefits, a bidder needs to offer the same price to purchase the non-voting shares. Accordingly, only undervalued types have an incentive to extend the offer P^* to non-voting shares. As shareholders are aware of this, they would reject bids for all shares. Thus, acquiring only voting shares is optimal for all types who make a bid in equilibrium: Overvalued types limit the loss on the shares purchased in the offer, while undervalued types avoid to reveal that they purchase the voting shares at a discount.

Even though non-voting shareholders are excluded from the offer, both classes of shareholders realize the same expected payoff. Conditional on a takeover, voting shareholders receive the bid price in cash, whereas non-voting shareholders retain shares of uncertain value. In equilibrium, the mispricing cancels out on average such that the expected post-takeover share value equals the cash amount paid to the voting shareholders.

Equal expected payoffs in an uncontested takeover translate into a zero voting premium only if a bidding contest is from an ex-ante perspective a zero probability event. Otherwise, voting shares trade at a premium that reflects the odds that the market puts on a future bidding contest. Voting premia do not arise in our model precisely because it analyses the takeover outcome when a bidding contest does not materialize.[9]

The comparative-static properties of the minimum bid equilibrium are key to our subsequent analysis of the optimal security-voting structure. Lemma 1 shows that each security-voting structure s maps into a unique minimum bid equilibrium. (Being the minimum of a closed subset of $\mathbb{R}$, $P^*(s)$ is always unique.) Moreover, we show in the Appendix that the equilibrium price $P^*(s)$ and cut-off value $V^c(s)$ are continuously increasing in the fraction s of voting shares. This has a straightforward implication for the takeover probability.

Proposition 2. *Non-voting shares promote takeovers.*

[9] Furthermore, the result of equal expected returns to both classes of shareholders is specific to the minimum bid equilibrium. In any other Perfect Bayesian Equilibrium, tendering (voting) shareholders receive on average more than the expected post-takeover security benefits, and voting shares trade accordingly at a premium.

In equilibrium, the marginal type who makes a bid ($V = V^c(s)$) purchases the voting shares at a loss that is exactly offset by her private benefits net of the takeover cost. A lower fraction of voting shares enables her to earn a positive profit as she has to purchase fewer overvalued shares. In addition, fewer voting shares render a bid feasible for some previously deterred types whose participation constraint (3.1) is now satisfied. Hence, a higher fraction of non-voting shares induces more types to bid, even if the price were to remain unchanged.

Shareholders infer that less exposure to mispricing extends the pool of types making a bid. They revise their expectations accordingly and are willing to tender at a lower price. This in turn further relaxes the participation constraint (3.1) and induces additional types to bid, thereby reinforcing the reduction in the minimum acceptable bid price.

4. Optimal Security-Voting Structure

The quality of a security-voting structure is determined by the extent to which it frustrates value-decreasing bids but encourages value-increasing bids. To examine both dimensions, we introduce value-decreasing bidder types and let $V^I \in \left[\underline{V}, \overline{V}\right]$.

The presence of value-decreasing bidder types does not affect our preceding analysis and results. Indeed, the share value under the incumbent management does not matter for the takeover outcome. Each shareholder tenders if the offered bid price matches the conditional expected post-takeover security benefits. Similarly, the decision to make a tender offer depends for a given price solely on the bidder's type. Hence, Lemma 1 continues to hold for any $V^I \in \left[\underline{V}, \overline{V}\right]$, and success of a value-decreasing bid is an equilibrium outcome in our setting whenever $V^I > V \geq V^c(s)$.

It has to be noted that failure of a conditional tender offer - whether value-decreasing or increasing - can in general be supported as an equilibrium outcome (see fn. 4). However, our assumption that shareholders tender unless retaining is weakly dominant eliminates failure as an equilibrium if the bid satisfies the free-rider condition. That is, when success and failure of a given bid can be supported as equilibrium outcomes, we select success even if the bidder is inferior to the incumbent manager.

Alternatively, one may assume that shareholders reject all bids below the current share value. This amounts to shareholders playing weakly dominated strategies whenever a bid is lower than the current share value but higher than the post-takeover security benefits. More broadly, this selection criterion abstracts from coordination problems among dispersed shareholders, such as the pressure-to-tender problem, and resulting undesirable takeover outcomes. But these are precisely the major issues in the literature on takeover regulation (e.g., Bebchuk and Hart, 2001). Addressing these concerns, we select success as the equilibrium outcome and analyze how the security-voting structure can help to overcome coordination problems.

Given our selection criterion, V^I affects neither takeover probability nor takeover outcome. Nonetheless, being the revenues when the takeover fails, V^I matters for the choice of the security-voting structure. To analyze this choice, we assume that the social planner decides on the fraction $s \in (0, 1]$ of voting shares, knowing the current share value, the takeover cost K, the upper bound $\bar{\phi}$ and the distribution of bidder types. Later (Section 4.2), we also derive the shareholders' preferred security-voting structure under the same informational assumptions.

4.1. Social Planner's Choice. From a social perspective, the takeover cost is a deadweight loss, while it is immaterial how the revenues are shared between shareholders and bidder or incumbent manager. Hence, the expected social welfare is

$$W = (1 - \Pr(V \geq V^c))V^I + \Pr(V \geq V^c)\left(\mathrm{E}\left[V \middle| V \geq V^c\right] - K\right).$$

Takeovers are socially desirable if they increase revenues by more than the takeover cost. That is, the socially optimal cut-off value is equal to $V^I + K$. Indeed, inserting the takeover probability converts the social welfare into

$$W = V^I + [1 - G(V^c)] \int_{V^c}^{\overline{V}} \frac{g(V)}{1 - G(V^c)} \left(V - V^I - K\right) dV,$$

and the first-order condition with respect to V^c yields

$$V^c_{soc} = V^I + K. \tag{4.1}$$

Since $\partial^2 W / \partial \left(V^c\right)^2\big|_{V^c_{soc}} = -g(V^c_{soc}) < 0$, the first-order condition identifies the unique optimum. Implementing the optimal cut-off value is straightforward in view of the inverse relationship between s and V^c (Proposition 2).

PROPOSITION 3. *Each firm has a unique socially optimal fraction of non-voting shares which decreases with the revenues generated by the incumbent manager.*

Due to the monotone relationship between s and V^c, there is in turn a unique fraction of voting shares that implements the cut-off value $V^I + K$ (or the closest achievable value). Thus, each firm as defined by its V^I has a unique socially optimal security-voting structure which increases in V^I. As long as the optimal security-voting structure includes both voting and non-voting shares ($0 < s^* < 1$), it achieves the first-best control allocation: It frustrates all and only value-decreasing bids. This does not hold for the two corner solutions $s^* = 1$ and $s^* = 0$. If V^I is sufficiently high, the one share - one vote structure is constrained optimal in the sense that not all, though as many as possible, value-decreasing bids are frustrated. Similarly, for sufficiently low V^I complete separation of cash flow and voting rights does not ensure that all value-increasing bids succeed.

Variations in the optimal fraction of non-voting shares across firms translate into different degrees of control contestability. Low values of V^I go together with high fractions of non-voting shares. Such stark deviations from one share - one vote are necessary to elicit bids from the many bidder types that can generate higher revenues (net of the takeover cost) than the incumbent manager. By contrast, one share - one vote is optimal if the firm is run by a sufficiently competent manager. Since most bidder types are in this case less competent, the optimal takeover barrier is set high. In all other cases, one share - one vote offers incumbent managers too much protection. Thus, we find that deviations from one share - one vote are in many cases socially optimal. At the same time, our theory concurs with the argument that one share - one vote is effective in deterring value-decreasing bids (Grossman and Hart, 1988).

In recent years, many dual-class firms have unified their shares into a single class. In some cases the abolition of dual-class shares has been a voluntary decision, in others it has been a response to an anticipated new regulation. Proposition 3 pertains to stock reunifications when undertaken by firms with dispersedly held votes.

COROLLARY 2. *Share-class reunifications are always in the managers' interest but need not be socially efficient.*

The one share - one vote structure minimizes the threat (probability) of a takeover. While this is always in the interest of the incumbent manager, it is socially desirable only if the incumbent manager is of high quality. The abolition of dual-class shares can have a sizeable entrenchment effect even for an average manager. To illustrate this point, we solve the model for a general uniform distribution (see Appendix), and apply the parametric solutions to the following numeric example. Suppose V^I is \$5M and bidder types are uniformly distributed between \$4M and \$6M. In addition, let K be \$.1M and $\bar{\phi} = .05$. The socially optimal security-voting structure in this case, $s^* \approx .36$, leads to a value-increasing takeover with a probability of 45 percent. Once the firm unifies its shares, the takeover probability is reduced to 19 percent. Or put differently, one share - one vote frustrates about 58 percent of all potential value-increasing takeovers.

Thus, our theory implies that announcements of share-class reunifications may lead to both negative or positive stock price reactions. By contrast, empirical studies on the abolition of dual-class share structures report significant positive announcement returns (e.g., Hauser and Lauterbach, 2004; Smart et al., 2008). These studies do, however, not distinguish between sample firms with dispersed votes and (the majority of sample) firms with a controlling shareholder. A proper test of our prediction would require to analyze the subsample of dispersedly held firms separately.[10]

Deviations from one share - one vote promote takeovers only if the bidder can discriminate between voting and non-voting shares. The requirement to make voting and non-voting shareholders the same offer replicates the one share - one vote structure, which may deter some value-increasing bids.[11] Due to the deterrence effect, this so-called "coattail" provision need not be socially optimal nor in the target shareholders interest, even though they lead to higher takeover premia.

The same argument pertains to restricted bids for single-class targets. Like non-voting shares, restricted bids reduce the fraction of cash-flow rights the bidder has to purchase to gain control. Consequently, a straightforward implication of Proposition 3 is that the mandatory bid rule can deter too many takeovers. While restricted bids are functionally similar to non-voting shares, they are by no means equivalent. First, partial bids must be for at least 50 percent of the cash-flow rights to ensure a voting majority, whereas the fraction of cash-flow rights attached to voting shares can be lower. Hence, a dual-class structure is in principle a more powerful instrument to screen bidders. Second, the security-voting structure is set by the target firm, while the fraction of shares to which the bid is restricted is in the bidder's discretion.

Besides the incumbent manager's quality, the takeover cost K and the maximum extraction rate $\bar{\phi}$ are further determinants of the socially optimal security-voting structure. Hence, variations in these parameters also alter the optimal fraction of voting shares.

The size of the takeover cost has two opposite effects on the optimal security-voting structure. On the one hand, higher costs raise the socially optimal cut-off value, as the revenues generated by the bidder must exceed current revenues by a larger margin. On the other hand, higher costs require larger private benefits to break even. This deterrence effect is reinforced by the adjustment of the shareholders' expectations about the post-takeover security benefits. Therefore, the latter effect dominates and the optimal fraction of non-voting shares increases with the takeover cost.

Higher extraction rates enable bidders to recoup the takeover costs more easily and lower the shareholders' expectations about post-takeover share value. Hence, higher

[10] In the sample of Pajuste (2005), six out of seven Swedish firms that unified their shares had dispersed control. Interestingly, the author suggests that the unification was motivated by the threat of a takeover. The provided information does not allow to discriminate whether the unification served to prevent value-decreasing bids or to entrench management, either of which being compatible with our theory.

[11] Since 1987 the Toronto Stock Exchange requires any firm which newly issues shares with superior voting rights to include a provision that obliges would-be acquirers to extend the offer at the same terms to all classes of shares (Allaire, 2006).

extraction rates and non-voting shares are substitutes: Both promote takeovers. As a result, more voting shares are required to implement a given cut-off value, when other governance mechanisms put weaker constraints on private benefit extraction.

PROPOSITION 4. *The optimal fraction of non-voting shares increases in the quality of shareholder protection.*

The result suggests that the rationale for one share - one vote is strongest in countries with weak shareholder protection whereas shares with differential voting rights may be desirable in environments where extraction is limited by strong institutions. It furthermore weakens the case for regulatory harmonization across diverse governance systems. Incidentally, a recent European study (European Commission, 2007) reports that non-voting preference shares and multiple voting shares appear to be most frequently used in the UK and Sweden, both commonly considered to be countries with strong shareholder protection (Nenova, 2003).[12]

4.2. Shareholders' Choice. As the equilibrium bid price always equals the expected post-takeover share value, voting and non-voting shareholders have homogeneous preferences. Hence, we can describe their collective and individual preferences by the aggregate wealth function

$$\Pi = (1-\phi)\, V^I + [1 - G(V^c)] \int_{V^c}^{\overline{V}} \frac{g(V)}{1-G(V^c)} \,(1-\phi)\left(V - V^I\right) dV. \tag{4.2}$$

In contrast to the social planner, shareholders are only concerned about the security benefits. Simplifying and deriving the first-order condition with respect to V^c yields

$$V^c_{sh} = V^I.$$

Target shareholders benefit from a takeover whenever the bidder can generate more revenues than the incumbent manager, irrespective of the takeover cost. Thus, they prefer a lower cut-off value than socially optimal and choose an accordingly lower fraction of voting shares. The privately and socially optimal security-voting structure only coincide when both are corner solutions, i.e., when either complete separation is socially optimal or one share - one vote is privately optimal. In all other cases, shareholders prefer too many takeovers.

Unless the incumbent manager is of high quality, shareholders benefit from a dual-class structure as it increases expected takeover gains. Accordingly, shares under the

[12] The study commissioned by the European Commission analyses the use of control-enhancing mechanisms in 320 randomly selected large companies across 16 EU countries. As each country sample contains only 20 firms, the cross-country comparisons are tentative.

dual-class structure command a higher price. This in turn translates into higher proceeds for a shareholder who wants to exit.

PROPOSITION 5. *A dual-class share structure may increase the proceeds from selling out in the stock market.*

Our theory argues that the adoption of dual-class shares - whether before or after going public - can be part of an optimal sale procedure.[13] Consistent with this prediction, several empirical studies report positive abnormal returns following the announcement of dual-class recapitalizations (Adams and Ferreira, 2008). In particular, Bauguess et al. (2007) report that most dual-class recapitalizations in their sample go together with sell-outs by dominant shareholders.

Like the social planner, shareholders have an interest to protect competent managers. Their preferred level of control contestability, and hence the optimal fraction of non-voting shares, decreases in the incumbent's ability.

PROPOSITION 6. *Under both the privately and the socially optimal security-voting structure, firms with (more) non-voting shares have lower market values.*

A firm's market value increases with the incumbent manager's ability, even though the probability of a value-increasing takeover decreases. To see why this is the case, compare two firms, 1 and 2, with $V_1^I < V_2^I$. Under the privately optimal structure, all bids that increase shareholder value succeed, and firm 1 is more likely to be taken over. The difference in the takeover probabilities is $1 - G(V_1^I) - \left[1 - G(V_2^I)\right] = G(V_2^I) - G(V_1^I)$, which is the probability that firm 1 is taken over by a bidder with valuation $V \in (V_1^I, V_2^I]$, thereby (partially) catching up with firm 2's current value. That is, firm 1's higher takeover probability stems only from the potential value improvements that firm 2 has already realized. Thus, firm 2's shares must have a higher market value under the privately optimal security structure. Clearly this reasoning also applies to the firms' socially optimal structures that takes the takeover costs into account.

Proposition 6 is consistent with empirical studies reporting a valuation discount for dual-class firms (Villalonga and Amit, 2007; Gompers et al., 2007). However, our result differs from the standard explanation, typically raised with respect to controlling shareholders, that the use of dual-class shares induces corporate insiders to extract more private benefits at the expense of shareholder value (e.g., Bebchuk et al., 2000; Masulis et al., 2009). In our model with dispersed control, the causality runs in the

[13] Bebchuk and Zingales (2000) offer a similar rationale for dual-class shares. In their model, a dual-class structure enables the firm founder to extract more rents from a future acquirer. Here, it encourages takeover bids once ownership has been dispersed.

opposite direction: It is the low firm value under the incumbent manager that induces shareholders to choose (more) non-voting shares. In doing so, they increase the likelihood that a better management team will acquire the firm.

Given that non-voting shares and extraction rates are substitutes, the question arises which combination of s and $\bar{\phi}$ target shareholders prefer. To this end, we compare two regimes implementing the same takeover probability. More precisely, consider the alternatives $\left\{\bar{\phi}', s'\right\}$ and $\left\{\bar{\phi}'', s''\right\}$ where $\bar{\phi}' < \bar{\phi}''$, $s' < s''$ and $V^c|_{s',\bar{\phi}'} = V^c|_{s'',\bar{\phi}''}$.

PROPOSITION 7. *For a given takeover probability, shareholder wealth is higher in the regime with less extraction and more non-voting shares.*

This result reverses the role commonly attributed to the security-voting structure and private benefit extraction (e.g., Grossman and Hart; 1980, 1988). In our setting, shareholders do not choose a high $\bar{\phi}$ to promote takeovers and a high s to frustrate inefficient bids. Instead, a low $\bar{\phi}$ is used to deter undesirable bidders and a low s is used to encourage the others. The security-voting structure affects the redistribution among bidder types. More specifically, reducing the fraction of voting shares promotes takeovers by reducing the gains that high bidder types earn from mimicking low bidder types. By contrast, the extraction rate affects how the takeover surplus is split between shareholders and bidders. When using $\bar{\phi}$ to encourage bids, shareholders bribe bidders out of their own pockets. From a social perspective, the regimes are equivalent as they both implement the same takeover probability and hence the same control allocation.

4.3. Heterogenous Extraction Abilities. So far, we assumed that the security and private benefits are positively correlated. In our view, this is a plausible assumption because it reflects circumstances in which private benefits are primarily determined by firm characteristics and the institutional environment rather than the bidder's identity. At the same time we agree that theoretical reasoning alone does not preclude alternative correlations. As in other models, our optimal security-voting structure and its comparative-static properties are sensitive to the posited relationship between security and private benefits.

When security benefits and private benefits are assumed to be inversely related, non-voting shares have the opposite effect and discourage takeovers. Suppose that all bidder types create the same revenues V but differ in their extraction abilities $\bar{\phi}$, which are uniformly distributed on the unit interval. When the target has a single share class, all types bid $P^* = \mathrm{E}\left[1 - \bar{\phi}\right] V = V/2$ in equilibrium, provided that the takeover cost is not too large ($V > 2K$). Since all shareholders tender in a successful bid, *every* type enjoys the same surplus $V - P^* - K$. In this equilibrium, *no* bid is frustrated while all types $\bar{\phi} < K/V$ would fail under symmetric information.

In the setting with negative correlation, the free-rider problem is mitigated rather than exacerbated by asymmetric information. The pooling price allows low extraction types to earn information rent that enables them to make a profit even though their private benefits do not cover the takeover cost.[14] More non-voting shares decrease the information rents and discourage these types. Complete separation eliminates all information rents, and - as under full information - only types whose private benefits exceed the takeover cost make a bid. Hence, either one share - one vote or complete separation are the socially optimal structure, depending on the current revenues V^I. For $V > V^I + K$ one share - one vote is optimal. Otherwise, complete separation is optimal without preventing value-decreasing bids by all types $\bar{\phi} \in (K/V, 1]$.

In the general case where bidder types differ both in their ability to generate revenues and extract private benefits, non-voting shares will exhibit both effects: they will encourage some bidder types but at the same time deter others. Which effect dominates and, in particular, which security-voting structure is socially optimal will depend on the extent of the assumed uncertainty about V and $\bar{\phi}$. An optimal security-voting structure will in general not perfectly discriminate between value-increasing and value-decreasing types. For any optimal cut-off value there will be higher V-types who are deterred because their extraction rates are too low. Conversely, there will be lower V-types who bid because they can extract more private benefits. While a full analysis of the general case is beyond the scope of this paper, the identified mechanism would carry over: The fraction of voting shares affects the bidders' participation constraint and hence shareholders' expectations about the post-takeover security benefits. The preceding analysis offers little reason to believe that one share - one vote would consistently outperform other structures in the generalized (two-dimensional) setting.

5. Concluding Remarks

This paper identifies a new mechanism through which the security-voting structure influences the tender offer outcome. When the bidder has private information about the post-takeover security benefits, she and the target shareholders may not agree on a mutually acceptable price, and the takeover fails. Non-voting shares mitigate this problem because the bidder can acquire control while buying fewer cash flow rights. Conversely, one share - one vote maximizes the risk that disparate information about

[14] This result suggests that more information about the bidder type may not always be desirable, as pointed out by Marquez and Yilmaz (2005). They briefly analyse supermajority rules which force bidders to acquire a larger fraction of cash flow rights when restricted bids are permitted. In our setting, the security-voting structure determines this fraction and allows that fraction to be both larger and smaller than 50 percent. Furthermore, Marquez and Yilmaz only consider a two-type setting where asymmetric information promotes takeovers and more takeover activity is by assumption beneficial. Therefore, they do not identify the benefits of separating cash flow and voting rights.

the security benefits prevents a takeover. Therefore, one share - one vote is in general not optimal.

While developed in a takeover model with atomistic shareholders, the insight that separating cash flow and voting rights can help to bring about efficient control transactions is not confined to tender offers. The essence of the free-rider behaviour is to create a link between the bidder's private information and each target shareholder's outside option which in turn can lead to disagreement about the purchase price. Such disagreement is by no means limited to settings with an infinite number of uninformed shareholders. For instance, it can also arise when current owners suspect a potential buyer of wanting to purchase their firm because it is (currently) undervalued. To invalidate the suspicion the buyer may have to offer a price that renders the acquisition unprofitable despite its value improvements. Separating the trade of cash flow and control rights can overcome the deadlock.

In the recurrent debate about the optimality of one share - one vote, dual-class shares are often criticized because they allow owners to lock-in control without a corresponding majority stake. By contrast, the current paper shows that widely held dual-class shares increase control contestability, thereby promoting value-increasing takeovers. For that reason, share reunification programs in dispersedly held firms entrench professional managers and may not be in the dispersed shareholders' best interest. Conversely, dual-class recapitalization may increase share value and be a means for dominant shareholders to improve the terms at which they sell out in the market. Furthermore, the optimal security-voting structure depends on the quality of both the incumbent manager and the governance mechanisms limiting managerial self-dealing. Hence, our analysis casts doubt on the merits of mandating a uniform security-voting structure across firms or countries.

Appendix

Proof of Proposition 1. Define the function $f(P) \equiv (1-\bar{\phi})\mathrm{E}\left(V|V \geq P+K\right)$ for $P \in \left[\underline{V}-K, \overline{V}-K\right]$. This function has the following properties.

(a): $f(P)$ is continuous.

(b): $f(\underline{V}-K) = (1-\bar{\phi})\mathrm{E}\left(V\right) > \underline{V}-K$.

(c): $f(\overline{V}-K) = (1-\bar{\phi})\overline{V} < \overline{V}-K$.

While property (a) follows from the continuity of the density function $g(\cdot)$, (b) and (c) follow from Assumption 1. Indeed, $\bar{\phi}\underline{V} < K$ is equivalent to $(1-\bar{\phi})\underline{V} > \underline{V}-K$ which implies $(1-\bar{\phi})\mathrm{E}(V) > \underline{V}-K$. Similarly, $(1-\bar{\phi})\overline{V} < \overline{V}-K$ follows from $\bar{\phi}\overline{V} > K$.

Properties (a) to (c) imply that there exists at least one fixed point of $f(P)$. Denote the smallest fixed point by P^*. From properties (a) to (c), it follows that (2.2) is satisfied for some $P \geq P^*$, whereas it is violated for all $P < P^*$. Denoting $\mathcal{S}_f \equiv \{P : P \geq f(P)\}$, it follows that $P^* = \min \mathcal{S}_f$.

Any element in $\mathcal{S}_f$ can be supported as a Perfect Bayesian Equilibrium by imposing appropriate out-of-equilibrium beliefs, e.g., $\mathrm{E}[V|\,P] = \overline{V}$ for all $P \neq P^+$ where P^+ is some element in $\mathcal{S}_f$. The credible beliefs criterion imposes that target shareholders believe a deviating (out-of-equilibrium) bid to come only from types that would want the bid to succeed. For any Perfect Bayesian Equilibrium with $P^+ > P^*$, denote the set of types whose participation constraint is satisfied for P^* by $D \equiv \left\{V \in \left[\underline{V}, \overline{V}\right] : V \geq P^* + K\right\}$ and its complement by D^C. No bidder type in D^C would want to bid P^* and succeed, whereas bidder types in D would want to bid and succeed. Consequently, the credible beliefs criterion imposes that shareholders believe $\Pr\left[V \in D|\,P = P^*\right] = 1$. Given such beliefs and sequential rationality, shareholders would accept the deviation bid P^*. Hence, no Perfect Bayesian Equilibrium with $P^+ > P^*$ survives the credible beliefs refinement. If $P^+ = P^*$, there exists no bid price to which any bidder would like to deviate as any lower price is rejected. Hence, P^* is the unique price (Perfect Sequential Equilibrium) that satisfies the credible beliefs criterion. ■

Proof of Corollary 1. In the minimum bid equilibrium, $V^c = P^* + K$ and $P^* = \mathrm{E}(X|\,V \geq V^c) = \mathrm{E}(X|\,V \geq P^* + K)$. From the proof of Proposition 1, we know that $P < \mathrm{E}(X|\,V \geq P + K)$ for any $P < P^*$. Now consider the effect of an increase in the takeover cost from K to $\hat{K}$ where $\hat{K} > K$. All else equal, the cut-off value increases. Thus, a necessary condition for $\hat{V}^c \leq V^c$ is that $\hat{P}^* < P^*$. However, this would violate the free-rider condition. To see this, note that $P < \mathrm{E}(X|\,V \geq P + K)$ implies $P < \mathrm{E}(X|\,V \geq P + \hat{K})$, with the former condition being satisfied for any $P < P^*$. Hence, it must be that $\hat{P}^* > P^*$. This in turn implies that $\hat{V}^c > V^c$. The positive relation between takeover probability and extraction rate $\bar{\phi}$ follows directly from the proof of Proposition 1 and the fact that an increase in $\bar{\phi}$ reduces $f(P) = (1-\bar{\phi})\mathrm{E}(V|V \geq P+K)$ for any given P. ■

Proof of Lemma 1. The free-rider condition is now given by

$$P \geq (1-\bar{\phi})\mathrm{E}\left[V|V \geq V^c(s,P)\right] = (1-\bar{\phi})\int_{V^c(s,P)}^{\overline{V}} g(V)V dV/\left(1-G\left[V^c(s,P)\right]\right)$$

Define $h(s,P) \equiv (1-\bar{\phi})\mathrm{E}\left[V|V \geq V^c(s,P)\right]$ for $P \in \left[\underline{V}-K, \overline{V}-K\right]$. As $V^c(s,P)$ is continuous in P, so is $h(s,P)$. (Note that $h(s,P) = f(P)$ for $s=1$.) Like $f(P)$ in the proof of Proposition 1, $h(s,P)$ satisfies property (b)

$$g(s,\underline{V}-K) = (1-\bar{\phi})\mathrm{E}\left[V \middle| V \geq \frac{s\underline{V}+(1-s)K}{\bar{\phi}+s\left(1-\bar{\phi}\right)}\right] \geq (1-\bar{\phi})\mathrm{E}\left(V\right) > \underline{V}-K$$

and property (c)

$$g(s,\overline{V}-K) = (1-\bar{\phi})\mathrm{E}\left[V \middle| V \geq \frac{s\overline{V}+(1-s)K}{\bar{\phi}+s\left(1-\bar{\phi}\right)}\right] \leq (1-\bar{\phi})\overline{V} < \overline{V}-K.$$

Hence, existence and uniqueness of the minimum bid equilibrium follow from the proof of Proposition 1. ■

Proof of Proposition 2. Since $h(s,P) \equiv (1-\bar{\phi})\mathrm{E}\left[V|V \geq V^c\right]$ is an increasing function of V^c, we know that $h(s,P)$ is increasing in s if and only if

$$\frac{\partial V^c(s,P)}{\partial s} > 0. \tag{5.1}$$

Partially differentiating $V^c(s,P)$ and imposing (5.1) gives

$$P > \left(1-\bar{\phi}\right)K/\bar{\phi}. \tag{5.2}$$

This condition is satisfied as the free-rider condition

$$P \geq (1-\bar{\phi})\mathrm{E}\left[V \middle| V \geq \frac{sP+K}{\bar{\phi}+s\left(1-\bar{\phi}\right)}\right]$$

implies

$$P > (1-\bar{\phi})\frac{sP+K}{\bar{\phi}+s\left(1-\bar{\phi}\right)}$$

which is equivalent to (5.2). Given condition (5.2) holds, $h(s,P)$ is increasing in s for all potential solutions to $P \geq h(s,P)$, including the minimum bid equilibrium $P^*(s)$. ■

Example: Share Class Unification. The bidder's participation constraint is $V \geq V^c(s,P)$. Given the uniform distribution, shareholders' expectations about the post-takeover share value conditional upon observing a bid P are

$$h\left(s,P\right)=\int_{V^c(s,P)}^{\overline{V}}\frac{\left(1-\bar{\phi}\right)V}{\overline{V}-V^c(s,P)}=\frac{1-\bar{\phi}}{2}\left(\overline{V}+\frac{sP+K}{\bar{\phi}+s\left(1-\bar{\phi}\right)}\right)$$

which is linear in P. The unique fixed point of $h\left(s,P\right)$ is given by

$$P^{*}=\frac{\left(\bar{\phi}+s\left(1-\bar{\phi}\right)\right)\overline{V}+K}{\left(2\bar{\phi}+s\left(1-\bar{\phi}\right)\right)/\left(1-\bar{\phi}\right)},$$

and the equilibrium cut-off value $V^c(s)$ is

$$\left.\frac{sP+K}{s\left(1-\bar{\phi}\right)+\bar{\phi}}\right|_{P=P^{*}}=\frac{s\left(1-\bar{\phi}\right)\overline{V}+2K}{2\bar{\phi}+s\left(1-\bar{\phi}\right)}.$$

The socially optimal s^* satisfies $V^c(s^*)=V^I+K$. Using the numbers, we have

$$\frac{s^{*}\left(1-0.05\right)6+0.2}{0.1+s^{*}\left(0.95\right)}=5+0.1\Leftrightarrow s^{*}=0.362\,57,$$

with takeover probability $\Pr\left[V\geq V^c(s^*)\right]\frac{\overline{V}-\left(V^I+K\right)}{\overline{V}-\underline{V}}=\frac{6-(5+0.1)}{6-4}=0.45$.

Under one share - one vote, we have

$$V^c(1)=\frac{\left(1-0.05\right)6+0.2}{1+0.05}=5.619,$$

with takeover probability $\Pr\left[V\geq V^c(1)\right]=\frac{6-5.619}{6-4}=0.1905$. Hence, only a fraction $\frac{0.1905}{0.45}=0.423$ of the value-increasing takeovers would succeed under one share - one vote.

Proof of Proposition 4. The threshold V^c_{soc} is independent of both $\bar{\phi}$ and s. For a given $\bar{\phi}$, the optimal s^* must be such that $V^c(s^*,\bar{\phi})=V^c_{soc}$. Consider the extraction rates $\bar{\phi}<\bar{\phi}'$ and the corresponding optimal security-voting structures s^* and s'^*. The first step is to establish that $V^c(s^*,\bar{\phi})>V^c(s^*,\bar{\phi}')$. To see this, insert the explicit expressions for $V^c(s^*,\bar{\phi})$ and $V^c(s^*,\bar{\phi}')$ to obtain

$$\frac{s^*P^*(s^*,\bar{\phi})+K}{\bar{\phi}\left(1-s^*\right)+s^*}>\frac{s^*P^*(s^*,\bar{\phi}')+K}{\bar{\phi}'\left(1-s^*\right)+s^*}. \tag{5.3}$$

The inequality holds because (a) the denominator on the right-hand side is larger and (b) because $P^*(s^*,\bar{\phi}')<P^*(s^*,\bar{\phi})$ follows from $\partial V^c(s,P)/\partial\bar{\phi}<0$ for any given s and P and from the arguments made to show that $h(s,P)$ hence decreases with $\bar{\phi}$ (similarly to the proof of Proposition 2). Given $V^c(s^*,\bar{\phi})>V^c(s^*,\bar{\phi}')$ and $\partial V^c(s,P^*\left(s\right))/\partial s>0$ (Proposition 2), $s'^*>s^*$ must hold to implement $V^c(s'^*,\bar{\phi}')=V^c_{soc}$. ■

Proof of Proposition 6. We prove this proposition only for the case of privately optimal structures. The proof for socially optimal structures is analogous. Under the privately optimal structure, every bidder with $V^B \geq V^I$ (or $X^B \geq X^I$) succeeds. The probability that the bidder is of such a type is given by

$$\Pr(V \geq V^I) = 1 - G(V^I).$$

The expected security benefits conditional on a successful takeover are

$$(1-\bar{\phi})\mathrm{E}(V \left| V \geq V^I \right.) = (1-\bar{\phi}) \int_{V^I}^{\overline{V}} g(V) V dV / \left(1 - G(V^I)\right).$$

The unconditional expected gain from a takeover is therefore

$$\Pr(V \geq V^I)(1-\bar{\phi})\mathrm{E}(V \left| V \geq V^I \right.) = (1-\bar{\phi}) \int_{V^I}^{\overline{V}} g(V) V dV,$$

and the current market value, which also takes into account the security benefits in the absence of a takeover, is

$$(1-\bar{\phi}) \left[G(V^I)V^I + \int_{V^I}^{\overline{V}} g(V) V dV \right].$$

Taking the partial derivative of the term in the brackets with respect to V^I gives

$$g(V^I)V^I + G(V^I) - g(V^I)V^I = G(V^I) > 0.$$

Since higher V^I also imply a smaller optimal fraction of non-voting shares, the proposition follows. ■

Proof of Proposition 7. Let $\phi' < \phi''$ and choose s' and s'' such that $V^c|_{s',\bar{\phi}'} = V^c|_{s'',\bar{\phi}''} = v$, where $v \in \left[\underline{V}, \overline{V}\right]$. Following the proof of Proposition 4, we know that $\phi' < \phi''$ implies $s' < s''$. Comparing shareholder wealth (4.2) across the two regimes and noting that the cut-off value is identical,

$$(1-\phi') \left[V^I + \int_v^{\overline{V}} g(V) \left(V - V^I\right) dV \right] > (1-\phi'') \left[V^I + \int_v^{\overline{V}} g(V) \left(V - V^I\right) dV \right],$$

proves the result. ■

References

Adam, Renée and Daniel Ferreira, 2008, "One Share, One Vote: The Empirical Evidence," *Review of Finance*, 12(1), 51-91.

Allaire, Yvan, 2006, "Dual-Class Share Structures in Canada: Review and Recommendations," The Institute for Governance of Public and Private Organizations, Policy Paper No. 1.

Association of British Insurers, 2005, "Application of the One Share - One Vote Principle in Europe," March.

Bauguess, Scott, Myron Slovin and Marie Sushka, 2007, "Recontracting Shareholder Rights at Closely Held Firms," Arizona State University.

Bebchuk, Lucian A. and Oliver Hart, 2001, "Takeover Bids vs. Proxy Fights in Contests for Corporate Control," Harvard Law School, John M. Olin Discussion Paper No. 336.

Bebchuk, Lucian A., Reiner R. Kraakman and George Triantis, 2000, "Stock Pyramids, Cross-Ownership, and the Dual Class Equity: The Creation and Agency Costs of Separating Control from Cash Flow Rights," in Morck, Randall K. (ed.) *Concentrated Corporate Ownership*, 295-315, University of Chicago Press, Chicago and London.

Bebchuk, Lucian A. and Luigi Zingales, 2000, "Ownership Structures and the Decision to Go Public," in Morck, Randall K. (ed.), *Concentrated Corporate Ownership,* 55-75, University of Chicago Press, Chicago and London.

Bergström, Clas, Peter Högfeldt and Johan Molin, 1997, "The Optimality of the Mandatory Bid Rule," *Journal of Law, Economics, and Organization*, 13, 433-51.

Betton, Sandra and B. Espen Eckbo, 2000, "Toeholds, Bid Jumps, and Expected Payoffs in Takeovers," *Review of Financial Studies*, 13(4), 841-82.

Burkart, Mike, Denis Gromb and Fausto Panunzi, 1998, "Why Takeover Premia Protect Minority Shareholders," *Journal of Political Economy*, 106(1), 172-204.

Burkart, Mike, Denis Gromb and Fausto Panunzi, 2006, "Minority Blocks and Takeover Premia," *Journal of Institutional and Theoretical Economics*, 162(1), 32-49.

Burkart, Mike and Samuel Lee, 2008, "Signaling in Tender Offer Games," Stockholm School of Economics.

Chowdhry, Bhagwan and Narasimhan Jegadeesh, 1994, "Pre-Tender Offer Share Acquisition Strategy in Takeovers," *Journal of Financial and Quantitative Analysis,* 29(1), 117-29.

Cornelli, Francesca and Leonardo Felli, 2000, "How to Sell a (Bankrupt) Company," CESifo Working Paper Series No. 292.

Eckbo, B. Espen, Ronald M. Giammarino and Robert L. Heinkel, 1990, "Asymmetric Information and the Medium of Exchange in Takeover: Theory and Tests," *Review of Financial Studies*, 3(4), 651-75.

European Commission, 2007, *Report on the Proportionality Principle in the European Union.*

Faccio, Mara and Larry H. P. Lang, 2002, "The Ultimate Ownership of Western European Countries," *Journal of Financial Economics*, 65, 365-95.

Gompers, Paul A., Joy Ishii and Andrew Metrick, 2007, "Extreme Governance: An Analysis of Dual-Class Firms in the United States," Harvard Business School.

Gromb, Denis, 1992, "Is One Share/One Vote Optimal?," Ecole Polytechnique, Discussion Paper No. 378.

Grossman, Sanford J. and Oliver D. Hart, 1980, "Takeover Bids, the Free-rider Problem, and the Theory of the Corporation," *Bell Journal of Economics*, 11(1), 42-64.

Grossman, Sanford J. and Oliver D. Hart, 1981, "The Allocational Role of Takeover Bids in Situations of Asymmetric Information," *Journal of Finance*, 36(2), 253-270.

Grossman, Sanford J. and Oliver D. Hart, 1988, "One Share-One Vote and the Market for Corporate Control," *Journal of Financial Economics*, 20, 175-202.

Grossman, Sanford J. and Motty Perry, 1986, "Perfect Sequential Equilibria," *Journal of Economic Theory*, 39, 97-119.

Hansen, Robert G., 1987, "A Theory for the Choice of Exchange Medium in Mergers and Acquisitions," *Journal of Business*, 60(1), 75-95.

Harris, Milton and Artur Raviv, 1988, "Corporate Governance, Voting Rights and Majority Rules," *Journal of Financial Economics*, 20, 203-35.

Hart, Oliver, 1995, *Firms, Contracts and Financial Structure*, Oxford University Press, Oxford.

Hauser, Shmuel and Beni Lauterbach, 2004, "The Value of Voting Rights to Majority Shareholders: Evidence from Dual-Class Stock Unifications," *Review of Financial Studies*, 17(4), 1167-84.

Hirshleifer, David and Sheridan Titman, 1990, "Share Tendering Strategies and the Success of Hostile Takeover Bids," *Journal of Political Economy*, 98(2), 295-324.

Holmström, Bengt and Barry Nalebuff, 1992, "To the Raider Goes the Surplus? A Reexamination of the Free Rider Problem," *Journal of Economics and Management Strategy*, 1, 37-62.

Marquez, Robert and Bilge Yilmaz, 2005, "Takeover Bidding and Shareholder Information," mimeo, University of Pennsylvania.

Masulis, Ronald W., Cong Wang and Fei Xie, 2009, "Agency Costs at Dual-Class Companies," forthcoming *Journal of Finance.*

Nenova, Tatiana, 2003, "The Value of Corporate Voting Rights and Control: A Cross-Country Analysis," *Journal of Financial Economics*, 68, 325-351.

Pajuste, Anete, 2005, "Determinants and Consequences of the Unification of Dual-Class Shares," ECB Working Paper No. 465/March 2005.

Shleifer, Andrei and Robert W. Vishny, 1986, "Large Shareholders and Corporate Control," *Journal of Political Economy*, 94(3, Part 1), 461-88.

Smart, Scott, Ramabhadran Thirumalai and Chad Zutter, "What's in a Vote? The Short- and Long-Run Impact of Dual-Class Equity on IPO Firm Value", *Journal of Accounting and Economics*, 45, 94-115.

Villalonga, Belen and Raphael H. Amit, 2007, "How are U.S. Family Firms Controlled?" forthcoming *Review of Financial Studies.*

PAPER 3

Signaling in Tender Offer Games

with Mike Burkart

Abstract. We examine whether a bidder can use the terms of the tender offer to signal the post-takeover security benefits. As atomistic shareholders extract all the gains in security benefits, signaling equilibria are subject to a constraint that is absent from bilateral trade models. The buyer (bidder) must enjoy gains from trade that are excluded from bargaining (private benefits), but can nonetheless be relinquished and enable shareholders to draw inference about the security benefits. Restricted bids and cash-equity offers do not satisfy these requirements. Dilution, debt financing, probabilistic takeover outcomes and toeholds are all viable signals because they make bidder gains depend on the security benefits in a predictable manner. In all the signaling equilibria, lower-valued types must forgo a larger fraction of their private benefits and these signaling costs prevent some takeovers. When there is additional private information about the private benefits as in the case of two-dimensional bidder types, fully revealing equilibria cease to exist. This does not hold once bidders can offer not only cash or equity but also (more) elaborate contingent claims. Offers which include options avoid inefficiencies and implement the symmetric information outcome.

1. Introduction

Presumably the best-known friction plaguing the market for corporate control is the free-rider problem (Grossman and Hart, 1980; Bradley, 1980): Small shareholders perceive their individual decision as being negligible for the tender offer outcome, and hence do not tender unless the offer price matches at least the post-takeover share value. As a result, they extract all the gains in share value, which in turn may deter potential bidders.

Another friction which has received less attention in the takeover literature is asymmetric information. In principle, both the bidder and the target shareholders can possess relevant private information. Contrary to merger negotiations between two management teams, the information advantage in tender offers is likely to be one-sided.

Thanks to Patrick Bolton and seminar participants at the London School of Economics, NHH Bergen, Stockholm School of Economics, and the ECGTN Conference in Barcelona for helpful comments. Financial support from the European Corporate Governance Training Network and the Jan Wallander Foundation is gratefully acknowledged.

Dispersed target shareholders do not actively monitor the firm and seldom possess information which is not already impounded in the stock price. By contrast, the bidder has typically spent resources to identify the target and to devise post-takeover (restructuring) plans. To succeed, the bidder therefore has to credibly communicate that the offer price adequately compensates target shareholders. Otherwise, the offer will be rejected even though it may entail a takeover premium.

The potential interaction of information asymmetries and coordination failure makes tender offers distinct from the standard bilateral trade setting. One-sided asymmetric information does not affect the bilateral trade outcome when the informed party makes a take-it-or-leave-it offer. This is not true in tender offer games. The first-mover advantage fails to endow the bidder with the ability to appropriate (part of the) gains in share value. Due to their free-riding behavior, uninformed shareholders have the bargaining power, even though the informed party (bidder) moves first.

This paper explores how and when a bidder can signal her private information about the post-takeover security benefits to dispersed shareholders. To the best of our knowledge, this questions has not yet been systematically analyzed. Though it has been shown that separating equilibria can be constructed in tender offer games (Hirshleifer and Titman, 1990; Chowdhry and Jegadeesh, 1994). The equilibria in those papers require dispersed shareholders to randomize their tendering decisions in a "coordinated" way to produce specific takeover probabilities. The present paper derives the general principle lying beneath the existence of signaling equilibria, thereby encompassing both previous results and novel signaling devices.

Our analysis begins by considering a tender offer game with one-dimensional asymmetric information. Only the bidder knows the post-takeover security benefits. In addition, a successful bidder enjoys private benefits and cannot commit not to extract these benefits. In this setting, an impossibility result obtains: as shareholders extract all the gains in security benefits, the bidder cannot reveal her type through the offer terms. Thus, neither restricted bids nor cash-equity offers are viable signals, which stands out against findings from bilateral merger models (e.g., Eckbo et al., 1990). The result exposes a fundamental conflict between incentive-compatibility and free-riding. The incentive-compatibility constraints require that high-valued bidders who have an incentive to mimic low-valued types earn information rents. However, the shareholders' free-rider behavior precludes that these rents stem from gains in security benefits, as they are fully appropriated by the target shareholders. Thus, private benefits are a prerequisite for incentive-compatible revealing offers.[1]

[1] The result can also be cast in terms of signaling costs. A trustworthy signal must be costly for the bidder. For example, the bidder might *voluntarily* forgo gains in security benefits. Yet, the free-rider problem precludes this possibility by *forcing* the bidder to forgo all these gains.

But even if the takeover is associated with private benefits, they can only serve a signaling purpose if the bidder is able to commit to relinquish part of the private benefits. In addition, the bidder must abandon the private benefits in a manner that enables shareholders to infer the post-takeover security benefits. Hirshleifer and Titman (1990)'s probabilistic separating equilibrium showcases this principle. In their equilibrium, target shareholders randomize their tendering decision such that bids at lower prices fail with higher probability. The higher failure probability deters high-valued bidders from mimicking low-valued types. Crucially, the deterrence operates exclusively through the risk of forgoing private benefits. If the bidder lacks such benefits, or high-valued bidders have substantially lower private benefits, the separating equilibrium breaks down. That is, bidders can credibly signal low security benefits only if lower bids forgo relatively more of their (expected) private benefits, and if the corresponding loss is larger for high-valued bidders.

A principal contribution of this paper is to identify this mechanism as a broad principle for the viability of signaling in tender offer games. Accordingly, signaling equilibria can be implemented through dilution, toeholds and debt-financing, even when tendering decisions are deterministic. These devices all allow the bidder to choose how much of the proceeds to divert or withhold from target shareholders, and the shareholders to infer the post-takeover security benefits. For instance, the bidder can signal low security benefits by committing to dilute minority shareholders less. (In the other cases, a low-valued bidder chooses a smaller toehold or less debt-financing to signal her type.) While this reduces private benefits, it allows her to succeed at a lower price. Dilution as a signal has unexpected implications. Firm quality and firm-level minority shareholder protection are inversely related as higher-valued bidders choose to divert more. Moreover, since higher-valued bidders pay higher prices, acquirers with weaker minority protection pay higher bid premia. Yet, the high bid premia neither reflect overpaying nor wasteful empire-building but merely the fact that the bidder creates more value.

In these signaling equilibria, lower-valued bidders must forgo more private benefits, either through failure or through reducing their level of private benefit extraction. As a result, the equilibrium outcomes typically exhibit inefficiencies at the "bottom": low-valued bidders are more prone to fail or do not even submit a bid. In particular, when takeover outcomes are deterministic, only bidders above a cut-off type make a bid in equilibrium, and a lower cut-off value implies more value-improving takeover activity.

Furthermore, signaling equilibria are not robust to the introduction of additional private information. In a setting with two-dimensional bidder types, signaling is no longer feasible. Essentially, relinquishing a given fraction of private benefits is no longer a viable signal when target shareholders are now unable to infer how costly such an offer is for the bidder.

Another finding is that bid restrictions, though either insufficient or redundant as a signal, promote takeover activity. This is because smaller transaction sizes mitigate the asymmetric information problem: With fewer traded shares, a bidder gains less (in total) from paying a price below the post-takeover share value. This reduces the incentives to mimic low-valued bidders, so that these types do not need to sacrifice as much private benefits to credibly reveal low security benefits. Thus, more restricted bids translate into smaller signaling costs.

The positive impact of bid restrictions on takeover activity could be taken further if control did not require a majority stake. This insight leads us to reformulate the asymmetric information problem in tender offer games: control transfers are impaired because control must be transferred along with *misvalued* cash flow rights. The appropriate solution is therefore to separate votes from cash flow rights. Indeed, we show that the use of non-voting shares or financial derivatives can generate signaling equilibria that completely eliminate the impact of asymmetric information. These financial instruments allow the bidder to buy the target shares against cash, strip the shares of their votes, repartition the cash flow rights and reissue only those cash flow rights that she wants to shed. While the first two steps give the bidder control, the last two steps can be used to self-impose penalties for "lying" about the post-takeover security benefits. In particular, call options enable target shareholders to seek "damages" from the bidder ex post if the security benefits turn out to be higher than ex ante professed. This makes the bid price de facto contingent on the post-takeover security benefits, thereby overcoming the information asymmetry. When the value improvement is deterministic, the options are never exercised so that the bidder essentially succeeds with a simple cash offer.

The use of derivatives allows to implement the symmetric information outcome because of a crucial difference between the tender offer game and most other signaling models in corporate finance. In tender offer games, the gains from trade are typically realized upon the transfer of control, as opposed to the transfer of cash flow rights. Thus, in the market for corporate control, company shares represent a bundle of two "goods", cash flow and voting rights. Separating these goods is beneficial when frictions in the trade of one impose a negative externality on the trade of the other.

Grossman and Hart (1981) and Shleifer and Vishny (1986) offer the first analyses of asymmetric information in tender offer games. Both papers focus exclusively on pooling equilibria. At et al. (2008) revisit the pooling equilibrium and show that dual-class share structures mitigate the asymmetric information problem. Hirshleifer and Titman (1990) and Chowdry and Jegadeesh (1994) study tender offer games in which takeover outcomes are probabilistic. As we demonstrate, their separating equilibria are applications of a general principle which does not rely on the probabilistic tender offer outcome.[2]

Several papers show that the choice of payment method can overcome asymmetric information problems in mergers (Hansen, 1987; Fishman, 1989; Eckbo et al., 1990; Berkovitch and Naranayan, 1990). Importantly, all of these papers consider bilateral merger negotiations and hence abstract from the free-riding problem. With the exception of Berkovitch and Naranayan, these papers consider two-sided asymmetric information settings in which target shareholders know more either about the share value under the incumbent manager or the takeover synergies. Thus, the settings differ from ours in precisely those aspects that are characteristic of tender offers. The same holds true for Brusco et al. (2007) and Ferreira et al. (2007) who study cash-equity offers in a mechanism design framework. The problem they explore becomes rather simple under our informational assumptions, and pure cash offers would always implement the full information outcome.

The paper proceeds as follows. The next section presents the basic model with non-transferable private benefits. Section 3 shows that this model has no signaling equilibria and explains the importance of the free-rider problem for this result. In addition, we demonstrate that neither bid restrictions nor cash-equity offers are viable signals in this setting. Section 4 shows why and how signaling equilibria can be implemented. We first develop the principle in an abstract setting and then apply it to familiar variants of the tender offer game. We also show that additional private information eradicates fully revealing equilibria. Section 5 demonstrates how the symmetric information outcome can be implemented through the use of derivatives. Concluding remarks are in Section 6, and mathematical proofs are in the Appendix.

2. Model

Our basic setting closely follows existing tender offer models with asymmetric information (Shleifer and Vishny, 1986; Hirshleifer and Titman, 1990), while remaining agnostic about the specific source of bidder gains. There is a widely held firm that faces a single potential acquirer, henceforth the bidder. If the bidder gains control, she can generate security benefits X. The bidder learns her type prior to making the tender

[2] Signaling can also serve the purpose of deterring potential rivals as in Fishman (1988) and Liu (2008), rather than reducing the information gap between bidder and target shareholders.

offer, whereas target shareholders merely know that X is distributed on $\mathcal{X} = [0, \overline{X}]$ according to the continuously differentiable density function $g(X)$. The cumulative distribution function is denoted by $G(X)$. If the takeover does not materialize, the incumbent manager remains in control. The incumbent generates security benefits which are known to all shareholders and normalized to zero. Thus, we restrict attention to the case of value-improving bids.

In addition, control confers exogenous private benefits $\Phi \geq 0$ on the bidder. The private benefits are only known to the bidder and for simplicity a deterministic function of her type.[3] Furthermore, the bidder cannot commit not to extract the private benefits once she is in control. As the private benefits accrue exclusively to the bidder, they are de facto non-transferable. Our specification of private benefits can accommodate various sources of bidder gains, such as dilution (Grossman and Hart, 1980) and toeholds (Shleifer and Vishny, 1986). Though, for the sake of notational simplicity, we subsequently assume that the bidder has no initial stake.

As the firm has a one share - one vote structure, a successful tender offer must attract at least 50 percent of the firm's shares. The tender offer is conditional, and therefore becomes void if less than 50 percent of the shares are tendered. In addition, the bidder can restrict the offer to a fraction $r \in [0.5, 1]$ of the shares. For simplicity, we assume that there are no takeover costs. Hence, the benchmark (full information) outcome is that all takeovers succeed.[4]

The timing of the model is as follows. In stage 0, the bidder learns her type X. In stage 1, she then decides whether to make a take-it-or-leave-it, conditional, restricted tender offer in cash. (Alternative means of payment will be considered later.) If she does not make a bid, the game moves immediately to stage 3. Otherwise, she offers to purchase a fraction r of the outstanding shares at a price rP.

In stage 2, the target shareholders non-cooperatively decide whether to tender their shares. Shareholders are homogeneous and atomistic. In stage 3, the incumbent manager remains in control if the fraction of tendered shares β is less than 50 percent. Otherwise, the bidder gains control and pays βP unless the offer is oversubscribed, in which case she pays rP, and tendering shareholders are randomly rationed. We henceforth refer to the basic model as the tender offer game with non-transferable private benefits.

[3] We analyse the more general case with two-dimensional bidder types later in the paper (section 4.2).

[4] Like other tender offer models exploring the free-rider problem, we assume that the firm's outstanding shares of mass 1 are dispersed among an infinite number of shareholders whose individual holdings are both equal and indivisible. When either of these assumption is relaxed, the Grossman and Hart (1980) result that all the gains in security benefits go to the target shareholders no longer holds (Holmström and Nalebuff, 1992).

3. Non-Transferable Private Benefits

In this exchange game, the bidder has private information and moves first. If target shareholders could coordinate, they would accept the offer whenever the price at least matches the security benefits under the incumbent manager. Thus, their reservation price would be independent of the bidder's type, and the bidder would succeed and appropriate the entire value improvement from the takeover.

However, as the shareholders are atomistic and decide non-cooperatively, their reservation price depends on the bidder's type. Each of them tenders at stage 2 only if the offered price at least matches the expected security benefits. Since shareholders condition their expectations on the offer terms (r, P), a successful tender offer must satisfy the free-rider condition $P \geq \mathrm{E}(X|r, P)$. We assume that shareholders do not play weakly dominated strategies. This eliminates failure as an equilibrium outcome when the free-rider condition is strictly satisfied.[5]

When the bid price exactly equals the expected post-takeover share value, the target shareholders are strictly indifferent between tendering and retaining their shares. That is, they are indifferent between these actions irrespective of their beliefs about the takeover outcome, so that the weak dominance criterion does not pin down a tendering strategy. The prevalent way of resolving the indeterminacy when $P = \mathrm{E}(X\,|r, P)$ is to assume that each shareholder tenders in this case, and hence the bid succeeds with certainty.[6] Alternatively, one may assume that strictly indifferent shareholders randomize, and that this leads to a probabilistic outcome.[7] Subsequently, we focus on deterministic outcomes and examine the conditions under which fully revealing equilibira exist in which (some) bidders signal their type through the chosen offer terms. The exception is Section 4.1.5 where we consider probabilistic outcomes of the basic model with non-transferable private benefits. Finally, to keep focus on the feasibility of signaling, we abstract from pooling equilibrium outcomes that are extensively studied in Shleifer and Vishny (1986), Marquez and Yilmaz (2005) and At et al. (2008).

---

[5] Given a bid is conditional, a shareholder who believes the bid to fail is indifferent between tendering and retaining. Imposing this belief on all shareholders and breaking the indifference in favour of retaining supports failure as an equilibrium, irrespective of the offered price (Burkart et al., 2006). To avoid co-existence of success and failure as equilibrium outcomes, it is typically assumed that shareholders tender when they are indifferent (e.g., Shleifer and Vishny, 1986). Contrary to our assumption, this precludes failure as the equilibrium outcome for a conditional bid, and hence the existence of an equilibrium when the free-rider condition is violated.

[6] A common motivation for this approach is that the bidder could sway the shareholders by raising the price infinitesimally. Although this argument holds under full information, it does not apply in the asymmetric information setting, as even small price increases affect shareholders' expectations about the post-takeover security benefits.

[7] Judd (1985) shows that a continuum of i.i.d. variables can yield a stochastic aggregate outcome.

3.1. Impossibility of Signaling. Under the assumption that each shareholder tenders in case she is strictly indifferent, all shares ($\beta = 1$) are tendered in a successful takeover. Accordingly, a successful restricted bid is oversubscribed, and the bidder randomly selects the fraction r among all shareholders whose shares she purchases. The remaining $1 - r$ shareholders cannot sell and become minority shareholders.

The bidder's expected profit from a bid (r, P) is

$$\Pi(r, P) = q(r, P)\left[\Phi(X) + r\,(X - P)\right]$$

where $q(r, P)$ denotes the success probability which is equal to 1 for $P \geq \mathrm{E}\left(X\,|r, P\right)$ and 0 otherwise. In a fully revealing equilibrium, the offer terms must be distinct across types that make a (successful) bid. This requires that each equilibrium offer satisfies the free-rider condition, $P(X) \geq X$, and the bidder's incentive-compatibility constraint

$$\Phi(X) + r\,(X)\left[X - P(X)\right] \geq \Phi(X) + r\,(X - P)$$

for all $r \in [0.5, 1]$ and $P \in \mathbb{R}$.

Theorem 1. *In the deterministic tender offer game with non-transferable private benefits, no fully revealing equilibrium exists.*

Given that $P(X) \geq X$, a truthful bidder at best breaks even on the purchased shares, and her expected profit cannot exceed $\Phi(X)$. However, each type offering her actual security benefits cannot be an equilibrium outcome. If a type x would succeed with an offer rx, any type $X > x$ would mimic type x to acquire shares at a price below their true value X. This also holds if each type would choose a different bid restriction $r(\cdot)$. Type X's profits are higher when buying $r(x)$ shares at a discount compared to buying $r(X)$ shares at their fair price whether $r(x)$ is smaller or larger than $r(X)$. These arguments eliminate $P(X) = X$ combined either with a common r or a type-contingent $r(\cdot)$ as possible equilibria. They also rule out outcomes in which some types offer more than their true security benefits but less than the highest type's security benefits. Successful offers with $P(x) \in (x, \overline{X})$ would be mimicked by bidders of type $X > P(x)$. Thus, a bidder can credibly signal her type only by offering a sufficiently large premium such that $P \geq \overline{X}$.

Revealing her type with an offer $P \geq \overline{X}$ is, however, not an attractive option for the bidder.[8] She can instead make a bid $P = \overline{X}$ and restrict it to $r = 0.5$, the minimum

[8] In fact, there exists an incentive-compatible schedule $\{(r(\cdot), P(\cdot))\}$ which entails that lower-valued bidders offer higher (per-share) prices but purchase fewer shares. Higher-valued bidders abstain from mimicking as they would forego a larger more valuable equity stake. Conversely, higher-valued bidders offer larger total amounts, $r(\cdot)P(\cdot)$, which are unattractive for lower-valued bidders who would at the same expense purchase a less valuable equity stake.

fraction required to gain control. The less costly offer $(0.5, \overline{X})$ succeeds as it satisfies the free-rider condition for all types (and any possible shareholder beliefs).

The inexistence result extends to settings where the private benefits are not a deterministic function of the bidder's type, but follow some — possibly type-contingent — density function. Indeed, the constraints in the bidder's maximization problem are not affected by the non-transferable private benefits. They cancel out in the incentive-compatibility constraint and they are not part of the free-rider condition.

Also, note that letting bidders choose the fraction of shares that they acquire does not enable them to signal their type. The sole function of the bid restriction is to limit the fraction of shares the bidder purchases in exchange for cash. This makes restricted bids in this setting equivalent to bids in which target shareholders are in part compensated through equity. Indeed, it is immaterial whether the bidder makes a partial bid for cash only or acquires all shares in exchange for some cash and $1 - r$ shares in the target firm under her control. Moreover, control requires that the partial bid is for at least half the shares or that the equity component does not exceed the cash component in the cash-equity offer. By virtue of this equivalence, any fully revealing equilibrium in cash-equity offers would also have to exist in restricted cash only offers.

PROPOSITION 1. *Allowing cash-equity offers in the deterministic tender offer game with non-transferable private benefits does not make fully revealing equilibria feasible.*

Proposition 1 contrasts sharply with results from bilateral merger models where cash-equity offers can reveal the bidder's type (Hansen, 1987; Berkovitch and Naranayan, 1990; Eckbo et al., 1990).[9] Our basic framework differs in two key respects. First, target shareholders have no private information. Instead, they face a collective action problem, i.e., are unable to coordinate their individual tendering decisions.[10] Second, the takeover is not undertaken to combine assets from two firms but to replace the incumbent managers. How or whether the free-rider problem affects signaling equilibria is the subject matter of the next section, while the role of bidder assets will be explored later in the paper (Section 4.1.3).

[9] Contrary to negotiated mergers, tender offers are usually cash offers. In fact, the mode of acquisition is one of the most important determinants of the payment method (e.g., Martin, 1996). The standard explanation focuses on regulatory delays associated with equity offers, i.e., the greater cost of using equity as a means of payment. Proposition 1 and also subsequent Theorem 3 suggest that means of payment do not help to overcome asymmetric information problems in tender offers.

[10] In merger models, the shareholders' reservation price is typically the stand-alone value of the target firm that although unknown to the bidder does not depend on her type. A notable exception is Berkovitch and Narayanan (1990) where shareholders' outside option is to wait for a competing bid, and the option value depends on the quality of the initial bidder relative to potential competitors.

3.2. Free-Riding and Information Rents. To illustrate the role of the free-rider condition, we digress to a modified setting in which the bidder is able to appropriate part of the security benefits. Abstracting from a specific extensive form, we assume that the bidder has bargaining power $\omega \in (0,1)$ such that shareholders, if fully informed, would tender at a price $P = (1-\omega)\,X$. Accordingly, the bidder would under full information appropriate a value improvement ωX on the purchased shares. Like the private benefits Φ, these gains depend on the bidder type and a successful takeover. But unlike the private benefits, the gains are transferable. That is, the bidder can (commit to) leave part of ωX to the shareholders. A second, purely simplifying modification is the absence of private benefits ($\Phi = 0$).

Given shareholders do not observe the bidder's type, they condition their beliefs on the offer terms and tender only if $P \geq (1-\omega)\,\mathrm{E}\,(X\,|r,P)$. Consequently, some bidders may not succeed or realize less than the full information profit ωX. That is, some bidders may have to offer more than $(1-\omega)\,X$ or set $r < 1$ to signal their type, while others may find such signals too costly.

Proposition 2. *In the tender offer game with bidder bargaining power $\omega \in (0,1]$, a fully revealing equilibrium exists. All types above the cut-off type $X^c(\omega) \in [0, \overline{X})$ make a bid, and higher types buy more shares at a higher price and make a larger profit.*

Incentive-compatibility requires that both the fraction of shares offered to buy $r(\cdot)$ and the price $P(\cdot)$ increase with the bidder type. Higher-valued bidders offer to buy more shares at higher prices. On the one hand, bidders refrain from mimicking lower-valued types because the increase in the profit margin is offset by the smaller fraction of shares that can be bought at lower prices. On the other hand, bidders do not mimic higher-valued types as buying more shares requires paying a higher price.[11] Lower-valued bidders credibly reveal their type not only by bidding for fewer shares but also by purchasing them at lower relative discounts $|P(X) - X|\,/X$, thereby having to concede an increasing fraction of their full information profit ωX to the target shareholders. As a consequence, there exists a cut-off type X^c who just breaks even, offering exactly $P = X^c$. Conversely, the highest-valued bidder reaps her full information gains $\omega\overline{X}$. She can purchase all shares at $P = (1-\omega)\,\overline{X}$ because target shareholders always tender for $P \geq (1-\omega)\,\overline{X}$ irrespective of their beliefs.

[11] Proposition 2 implies that the bidder faces an upward-sloping supply curve: a larger demand for shares reveals a higher valuation which in turn raises target shareholders' ask price. This is akin to the downward-sloping demand curve that a privately informed issuer meets when selling securities (e.g., DeMarzo and Duffie, 1999). Though contrary to the informed seller setting, gains from trade materialize in the tender offer game only if the bidder acquires a control stake. Therefore, trade collapses once the incentive-compatible supply of shares is less than 0.5.

Due to above equivalence result (Proposition 1), Proposition 2 can also be phrased in terms of cash-equity offers. In this interpretation, the equilibrium offer schedule entails that higher-valued bidders use more cash and less equity. This is the same result as in the bilateral merger models of Eckbo et al. (1990) and Berkovitch and Narayanan (1990), though there is a subtle difference. In our setting, the bidder wants to signal low rather than high security benefits. This shifts the emphasis from cash as a high-value signal to equity as a low-value signal. Equity is a credible signal for low-valued bidders because relinquishing equity is costlier for high-valued bidders.

The positive relation between equilibrium profits and bidder types in Proposition 2 is a common feature of adverse selection models: incentive-compatibility requires that the types who have incentives to mimic others earn information rents (e.g., Laffont and Martimort, 2002). In our setting, higher-valued bidders receive these rents, and incentive-compatibility dictates the rate at which equilibrium profits decrease. Given the slope, the profit levels are determined by the boundary condition $\Pi^*(\overline{X}) = \omega\overline{X}$. That is, the highest type's equilibrium profit determines at which type the incentive-compatible profit falls to 0.

Corollary 1. *As the bidder's bargaining power ω approaches 0, the cut-off type $X^c(\omega)$ converges to $\overline{X}$.*

Corollary 1 brings to light the impact that the free-riding behavior has on the feasibility of revealing offers when private benefits are non-transferable. The subset of bidders that can signal their type without incurring a loss (on the purchased shares) decreases in the fraction of the share value improvement that each bidder would be able to appropriate under full information. In other words, as the target shareholders' free-riding behavior becomes more severe, the bidder's ability to signal her type gradually deteriorates. In the limit ($\omega = 0$), no bidder type is able to make a profit on the purchased shares, and the separating equilibrium breaks down.

To be noted is that separation fails even though the bidder's objective function satisfies the single-crossing property.[12] The impact of the free-rider behavior on the (in)existence of signaling equilibria can be interpreted in two ways. From the perspective of lower-valued types, the free-rider condition eliminates the possibility of producing a costly signal. Given that target shareholders extract all the gains in security benefits, the bidder cannot surrender (part of) these gains to signal her type. From the perspective of higher-valued types, the free-rider condition wipes out information rents. A bidder who at best breaks even on truthfully purchased shares always mimics a lower bid price.

[12] For each fixed (r, P), $-\frac{\partial\Pi/\partial r}{\partial\Pi/\partial P}$ is strictly monotone in X.

When bidders have no bargaining power ($\omega = 0$) and no private benefits ($\Phi = 0$), neither a fully revealing nor a pooling equilibrium exists, and trade virtually collapses as only the highest type makes a bid. It is worth comparing this outcome to Milgrom and Stokey (1982)'s classic no-trade theorem. The theorem says that asymmetrically informed but rational parties cannot agree to a transaction unless there are aggregate gains to be shared. In the tender offer game, such gains are present, since the takeover improves the target's value. Still, the tender offer is from the bidder's perspective equivalent to a trade without any aggregate gains, as they are entirely appropriated by the free-riding shareholders. That is, trade breaks down because, albeit the bidder makes a take-it-or-leave-it offer, shareholders have all the bargaining power.

4. Relinquishing Private Benefits

The preceding discussion suggests that key to signaling is the existence of bidder gains that are excluded from bargaining (private benefits) but can nonetheless be relinquished in a manner which allows inference about the security benefits. Subsequently, we show that this principle can indeed be applied to construct fully revealing equilibria. We first derive it in a general setting and then implement it in well-known variants of the tender offer game in which the bidder can choose how much private gains she appropriates.

4.1. Fully revealing equilibria. Suppose that the bidder can (commit to) transfer any fraction of her private benefits Φ to the target shareholders, such that she retains a fraction $\alpha \in [0, 1]$. The tender offer is then a triple (r, α, P); it specifies the fraction of shares offered to acquire, the retention rate of private benefits and the per share cash price. For a given offer, the bidder's payoff from a successful takeover is

$$\Pi(r, \alpha, P; X) = \alpha\Phi(X) + r(X - P). \tag{4.1}$$

If a signaling equilibrium exists, the equilibrium outcome can also be implemented as a direct (truth-telling) mechanism. Let $\hat{X}$ denote a bidder's self-reported type. The bidder's problem can thus be formulated as

$$\max_{\hat{X}} \Pi(\hat{X}; X) = \alpha(\hat{X})\Phi(X) + r(\hat{X})[X - P(\hat{X})]. \tag{4.2}$$

In a fully revealing equilibrium, the solution to this problem must be $\hat{X} = X$ for all $X \in \mathcal{X}$. This ensures *local* optimality. The corresponding first-order condition is

$$r(X)P'(X) = \alpha'(X)\Phi(X) + r'(X)[X - P(X)]. \tag{4.3}$$

Under a fully revealing offer schedule $\{r(\cdot), \alpha(\cdot), P(\cdot)\}$, equation (4.3) holds for all $X \in \mathcal{X}$. If the schedule further ensures quasi-concavity of the objective function in (4.2) for all $X \in \mathcal{X}$, the solution to (4.3) also identifies a *global* optimum, thus supporting the proposed schedule as a fully revealing equilibrium.

THEOREM 2. *In the tender offer game where bidders can commit to relinquish any fraction of their private benefits, a fully revealing equilibrium exists if* $\Phi(\cdot)$ *is a non-decreasing function. All types above the cut-off type* $X^c \in [0, \overline{X})$ *make a bid, and higher types relinquish a smaller fraction of their private benefits. The bid restriction is a redundant signal.*

As in Proposition 2, lower-valued bidders have to forego a larger fraction of their takeover gains. The difference is that $\Phi(X)$ and $\alpha(X)$ respectively play the roles of ωX and $r(X)$. Incentive-compatibility thus requires that both $\alpha(\cdot)$ and $P(\cdot)$ increase with the bidder type. Lower-valued types offer a smaller price but, to avoid being mimicked by higher-valued types, relinquish a larger fraction of their private benefits, and possibly purchase their shares at higher relative premia $[P(X) - X]/X$. Because lower-valued types are further endowed with (weakly) less private benefits, there exists a cut-off type X^c whose retained private benefits $\alpha\Phi(X^c)$ just equal her total takeover premium $r(P - X^c)$. The highest-valued type keeps her entire private benefits $\Phi(\overline{X})$ and purchases her shares at $P = \overline{X}$.

There are many feasible equilibrium schedules, leading to different cut-off types. At the same time, every possible cut-off type can be implemented by many equilibrium schedules that differ in how they weight $\alpha(X)$ and $P(X) - X$ as a means to reduce bidder profits. The bid restrictions—while redundant as a signal—also affect the cut-off type, an effect to which we turn our attention in Section 5.1. Finally, the assumption that $\Phi(\cdot)$ is non-decreasing ensures that the bidder's objective function is quasi-concave under any locally optimal schedule. It is moreover representative of the main examples that follow below.

4.1.1. *Dilution.* In our first example, we endogenize private benefit consumption by letting the successful bidder choose what fraction ϕ of the firm's total post-takeover value $V \in \mathcal{V}$ to divert for private purposes. Dilution does not dissipate value, so that a successful bid generates security benefits $X(V) = (1 - \phi)V$ and private benefits $\Phi(V) = \phi V$.

Diversion as a source of private benefits was first introduced by Grossman and Hart (1980), who assume a given dilution rate ϕ. Burkart et al. (1998) endogenize the ex post dilution decision by assuming that diversion dissipates value, so that ϕ is determined by the bidder's post-takeover equity stake. This implies that the bidder can de facto pre-commit to a dilution rate by choosing what fraction of target shares she acquires.

Here, we take a simpler approach and assume that the bidder can pre-commit not to extract more than $\phi \in [0, \bar{\phi}]$ of V independent of the bid restriction. The upper bound $\bar{\phi} < 1$ is an exogenous limit set by legal shareholder protection or more broadly the corporate governance system.[13] In practice, a takeover activist or private equity fund may include self-imposed constraints in the restructuring plans or in the post-takeover governance structure (e.g., management team or board composition) which are announced as part of the takeover proposal. Similarly, the takeover proposal may contain provisions that restrict certain types of post-takeover transactions that would be detrimental to minority shareholders.

The tender offer terms are now given by the triple (r, ϕ, P). If ϕ were uniform across bidder types, the setting becomes equivalent to the tender offer game with non-transferable private benefits, in which no separating equilibrium exists. However, as ϕ is a choice variable, Theorem 2 can be implemented.

PROPOSITION 3. *In the tender offer game where bidders can commit to divert up to a fraction $\bar{\phi}$ of the firm value, a fully revealing equilibrium exists. All types above the cut-off type $V^c(\bar{\phi}) \in [0, \overline{V})$ make a bid, and higher types dilute a larger fraction as private benefits.*

Compared to the general setup, $\bar{\phi}V$ plays the role of $\Phi(X)$, and lowering ϕ amounts to relinquishing a larger fraction of $\Phi(X)$. The incentive for a bidder to voluntarily dilute less ($\phi < \bar{\phi}$) does not arise in the symmetric information settings of Grossman and Hart (1980) or Burkart et al. (1998). It reflects the signaling role of private benefits. As both $\phi(\cdot)$ and $P(\cdot)$ are increasing in the bidder type, diversion and absolute bid premia are positively related in equilibrium: a bidder who pays a higher bid premium extracts more private benefits. Conversely, a bidder who pays a lower bid premium must "tie her hands", conceding some post-takeover decision flexibility. In a sense, it is as if bidders can buy more dicretion (control) by paying a higher price.

The cut-off type is decreasing in the overall scope for dilution which is determined by the quality of the corporate governance system. Better minority shareholder protection (lower $\bar{\phi}$-values) reduces the possibility for bidders to reveal their type by voluntarily constraining their ability to dilute minority shareholders. In parallel to Corollary 1, $V^c(\bar{\phi})$ converges to $\overline{V}$ as $\bar{\phi}$ approaches 0, and the fully revealing equilibrium collapses in the limit.

[13] Since the bid restriction is a redundant signal, we could in principle construct fully revealing equilibria where dilution depends on the bidder's equity stake. To this end, $r(\cdot)$ is set to pin down an increasing $\phi(\cdot)$. Given $r(\cdot)$ and $\phi(\cdot)$, $P(\cdot)$ is then chosen to satisfy the first-order condition. That is, the validity of the principle carries over to the inefficient diversion model of Burkart et al. (1998).

4.1.2. *Leverage.* Müller and Panunzi (2004) show that a leveraged tender offer, in which the takeover bid is partly financed through debt, can implement a takeover outcome that is quite similar to the outcome implemented by dilution. Initially, the bidder sets up a shell company that issues debt backed by claims on the target's assets. She then makes a tender offer for the target shares and, if the bid is successful, merges the acquired firm with the shell company. The fact that the combined firm is indebted lowers its share value, which in turn lowers the bid price at which target shareholders are willing to tender their shares. Thus, this "bootstrap acquisition" allows the bidder to acquire the target at a lower price, and hence to appropriate part of the takeover gains.

Given that leverage functions as a source of "private benefits", Theorem 2 suggests that it can also serve as a signaling device. To illustrate this possibility in the simplest fashion, consider the modified type space $\mathcal{X}' = [\underline{X}, \overline{X}]$ where $\underline{X} > 0$, and suppose that the bidder can raise up to $D \in [0, \overline{D}]$ where $\overline{D} \ll \underline{X}$ is an exogenously imposed limit. Thus, the tender offer terms are now given by the triple (r, D, P).

Proposition 4. *In the tender offer game where bidders can raise up to $\overline{D}$ in debt to finance the bid, a fully revealing equilibrium exists. All types above the cut-off type $V^c(\overline{D}) \in [0, \overline{V})$ make a bid, and higher types raise more debt.*

Again, Proposition 4 merely showcases the general principle, with $\overline{D}$ playing the role of $\Phi(X)$, and lower D amounting to relinquishing more private benefits. A high-valued bidder does not mimic a lower type because the gains from purchasing the shares at a discount are offset by the reduction in leverage, whereas a lower-valued bidder does not mimic a higher type because the increase in leverage is offset by the premium she must pay for the shares. In equilibrium, both $D(\cdot)$ and $P(\cdot)$ are increasing in the bidder type, so that more expensive takeovers go together with a larger amount of debt finance. In fact, given the simple setting, the leverage ratio $D(\cdot)/P(\cdot)$ is also increasing in the bidder type. Similarly to before, the cut-off type is decreasing in $\overline{D}$, and takeover activity—absent other signaling devices—vanishes as $\overline{D}$ approaches 0, i.e., debt finance "dries" up.

A thorough analysis of leverage as a signaling device in tender offers should, of course, relax the constraint $\overline{D} \ll \underline{X}$ and, more importantly, account for insolvency risks and financial distress costs, as in the more general framework of Müller and Panunzi (2004). Here, we simply point out that bidders can potentially use tender offer leverage as a means to reveal private information about the post-takeover share value to free-riding shareholders.

4.1.3. *Bidder assets.* The canonical tender offer game abstracts from bidder assets other than cash. In an extended setting with bidder assets $A > 0$, the bidder could use claims to such assets to pay target shareholders, and the willingness to do so might reveal her type. Theorem 2 suggests that the viability of such signals requires that the bidder assets appreciate in value as a result of the takeover. In addition, these "synergy" gains must be exclusionary in that they do not accrue to non-tendering target shareholders and, as such, do not affect their reservation price (free-rider condition). In this sense, they are equivalent to private benefits.

Suppose that the bidder owns cash and a firm that generates security benefits $A(\mathbb{I}, X) \equiv \mathbb{I}_{\beta \geq 0.5} Z + \lambda X$. Both the value $Z \geq 0$ and the parameter $\lambda > 0$ are commonly known and the same for all types. The indicator function $\mathbb{I}_{\beta \geq 0.5}$ takes the value 1 if the bid succeeds and 0 otherwise. That is, the value Z only materializes in the event of a succcessful takeover. Consider a takeover in which the bidder combines the two firms in a holding company H. Target shareholders are offered a cash price $C(\beta)$ and $s(\beta)$ shares in the holding company, where β is the fraction of shares tendered. Furthermore, target shareholders are cash-constrained, and the bidder is unwilling to relinquish majority control of the holding company. These assumptions impose restrictions on the set of admissible offers, modeled as a *cash* constraint $C(\beta) \geq 0$ and a *control* constraint $s(\beta) \in [0, 0.5]$.

PROPOSITION 5. *In the tender offer game with exclusinary synergy gains, a fully revealing equilibrium exists. All types above the cut-off type $X^c(Z) \in [0, \overline{X})$ make a merger bid, and lower types pay more in equity of the merged company.*

Here, $1 - s$ and Z respectively play the role of α and $\Phi(\cdot)$. In equilibrium, $s(\cdot)$ is decreasing, whereas $C(\cdot)$ is increasing, in the bidder type. That is, a lower-valued bidder pays less in cash but gives the target shareholders a larger fraction of the post-merger equity. Even though λX is perfectly correlated with X, the mere fact that the bidder assets are informative about the post-takeover share value is neither necessary nor sufficient to obtain a fully revealing equilibrium. The key is that the bidder assets appreciate in value as a result of the takeover, i.e., that the bidder enjoys some exclusionary gains. Indeed, the cut-off type is decreasing in Z, and takeover activity—again, absent other signaling devices—vanishes as Z approaches 0. Crucial to the bidder's signaling ability is hence how the synergy gains are divided between the two firms. (For example, a merger between a soft drink producer and a fast food chain may have separable effects on the sales in each firm.)

The equilibrium schedule in Proposition 5 is similar to those found in the literature on the means of payment in bilaterally negotiated mergers. Though contrary to bilateral merger models, tender offer games do not require two-sided asymmetric information to generate a role for cash-equity offers involving bidder assets. It is enough that the bidder has private information about the post-takeover value improvement in the target. To some extent, we thus extend the finding that cash-equity offers have signaling value to tender offer games. However, we show in Section 4.2 that this result is very sensitive about the informational assumptions about the value of the bidder assets.

4.1.4. *Toehold acquisition.* Another source of bidder gains, first studied by Shleifer and Vishny (1986), are pre-bid stakes (toeholds) that the bidder acquires prior to the announcement of the tender offer. Suppose that the bidder can purchase up to a fraction $\bar{t}$ of the target shares in the open market—for simplicity, at the price of $P = X^I$—before she must make her takeover intentions public. The upper bound $\bar{t}$ represents a mandatory disclosure, or mandatory bid, rule that prevents the bidder from acquiring an even larger pre-bid stake in the target.

PROPOSITION 6. *In the tender offer game where bidders can acquire up to a fraction $\bar{t}$ of the firms' shares prior to the bid, a fully revealing equilibrium exists. All types above the cut-off type $V^c(\bar{t}) \in [0, \overline{V})$ make a bid, and higher types acquire larger toeholds.*

The signaling potential of endogenous toeholds has already been analyzed within a probabilistic tender offer game by Chowdhry and Jegadeesh (1990). (We turn to the probalistic tender offer game in the next subsection.) Our analysis highlights that the choice of toehold size as a signaling strategy is a particular implementation of a general principle, and that the implementation does not require a probabilistic setting.

4.1.5. *Probabilistic outcomes.* We revisit the tender offer game with non-transferable private benefits. Contrary to before, we now assume that strictly indifferent shareholders randomize their tendering decision, and that this results in an uncertain tender offer outcome. A prerequisite for a probabilistic outcomes is that the offered prices exactly match the expected security benefits. Otherwise, shareholders either always or never tender. Given an offer $P = E(X \mid r, P)$, the success probability $q(r, P)$ can lie anywhere in $[0, 1]$, and the expected fraction of shares $\gamma(r, P)$ acquired by a successful bidder can lie anywhere in $[0.5, r]$. The bidder's expected profit from a bid (r, P) is therefore

$$\Pi(r, P) = q(r, P) \left[\Phi(X) + \gamma(r, P) (X - P)\right].$$

Compared with the general setting, $q(r, P)$ and $q(r, P)\gamma(r, P)$ respectively play the role of α and r in (4.1). This suggests that a signaling equilibrium is possible as long as $q(r, P)$ allows the bidder to relinquish her private benefits in a manner that reveals her type.

PROPOSITION 7. *In the probabilistic tender offer game with non-transferable private benefits, a fully revealing equilibrium exists if $\Phi(\cdot)$ is a non-decreasing function. All types make a bid, though the bids of higher types are more likely to succeed.*

In equilibrium, both $q(r,P)$ and P are increasing in the bidder's type. A lower-valued bidder pays a smaller price but her bid is less likely to succeed. The higher failure rate protects her bid from being mimicked by higher-valued types. Importantly, this deterrence effect *exclusively* operates through the risk of losing *private* benefits. In fact, if $\Phi(\cdot)=0$ or even if merely $\Phi(\overline{X})=0$, the signaling equilibrium breaks down.

Proposition 7 is but another application of Theorem 2. Its specific feature is that bidders do not signal their type through conceding private benefits *to the shareholders*, but rather through "burning" private benefits by way of failure. Common to the other applications, the outcome is inefficient: While all types actually make a bid in the equilibrium, bids do not always succeed and some takeovers fail to occur. Furthermore, the bid restriction remains a redundant signal and there are multiple feasible equilibrium schedules. Hirshleifer and Titman (1990).select a schedule where all types restrict their bid as much as possible ($r=0.5$).

4.2. Another Impossibility of Signaling. In our discussion of Theorem 2 and its various applications, we have emphasized how crucial the *existence* of private benefits is for the bidder's ability to signal her type. However, their existence is merely a necessary but not a sufficient condition. The bidder must also be able to relinquish private benefits in a manner which is *informative*. More specifically, the target shareholders must be able to evaluate how costly a given signal is to the bidder, or else the signal may not be sufficiently credible.

To examine the informativeness of private benefits, we abandon the assumption that the bidder's security benefits perfectly predict her private benefits (i.e., that Φ is a deterministic function of X). Instead, we assume that the bidder types are *two-dimensional*, (X,Φ), and continuously distributed on $[\underline{X},\overline{X}]\times[\underline{\Phi},\overline{\Phi}]$. The bidder is informed about both dimensions of her type. In contrast, the target shareholders neither know how much a particular bidder will improve the share value nor how much she values control over the firm. The setting is otherwise the same as in Theorem 2. Though for expositional convenience, we characterize the tender offer terms by the triple (r,α,C) where $C\equiv rP$.

Consider the type $(\overline{X},\underline{\Phi})$ and an arbitrary type $(X,\Phi)\neq(\overline{X},\underline{\Phi})$. In any fully revealing equilibrium, type $(\overline{X},\underline{\Phi})$ cannot be held to a profit lower than $\underline{\Phi}$ because she can always succeed with the bid $(r,1,\overline{X})$. At the same time, she cannot earn more than $\underline{\Phi}$ because of the free-rider condition, which ensures that the target shareholders receive at least $\overline{X}$. In order for type $(\overline{X},\underline{\Phi})$ not to mimic type (X,Φ), the latter type must make an offer (r,α,C) which satisfies $\underline{\Phi}\geq r\overline{X}+\alpha\underline{\Phi}-C$, or equivalently

$$C\geq\underline{C}\equiv r\overline{X}-(1-\alpha)\underline{\Phi}. \tag{4.4}$$

In addition, a truthful offer by (X, Φ) must also yield a higher profit than the "out-of-equilibrium" offer $\left(0.5, 1, 0.5\overline{X}\right)$ which succeeds irrespective of target shareholder beliefs. That is, her offer (r, α, C) must satisfy $rX + \alpha\Phi - C \geq 0.5(X - \overline{X}) + \Phi$, or equivalently

$$C \leq \overline{C} \equiv (r - 0.5)\, X + 0.5\overline{X} - (1 - \alpha)\Phi. \tag{4.5}$$

The constraints (4.4) and (4.5) are simultaneously satisfied if $\overline{C} \geq \underline{C}$ holds. Straightforward manipulations yield $(r - 0.5)\left(X - \overline{X}\right) \geq (1 - \alpha)\left(\Phi - \underline{\Phi}\right)$. This condition is violated, unless all types with $X < \overline{X}$ make the pooling offer $\left(0.5, 1, 0.5\overline{X}\right)$.

THEOREM 3. *In the tender offer game with two-dimensional types,* $(X, \Phi) \in [\underline{X}, \overline{X}] \times [\underline{\Phi}, \overline{\Phi}]$*, no fully revealing equilibrium exists even when bidders can commit to relinquish any fraction of their private benefits.*

Signaling breaks down in the two-dimensionsal case because the private information about Φ undermines the "credibility" of α (i.e., the relinquishing of private benefits) as a signal. Since Φ is not a deterministic function of X, target shareholders cannot infer from the offer (r, α, C) how costly it is for a bidder of (any) type to concede $1 - \alpha$ of her private benefits. The uncertain relation between Φ and X "jams" the signal. Theorem 3 thus weakens the case for fully revealing equilibria that rely on the use of the bidder's private benefits as a means of revealing information about the security benefits.

That Φ is in this sense uninformative about X is a more reasonable assumption in some settings than in others. In the case of dilution, toeholds, or leverage, the relation between Φ and X is relatively well-defined by the "extraction" technology. By comparison, it is plausible that at least some bidder assets are unrelated to the target firm, or the takeover, and that the bidder has superior information about their value. Theorem 3 implies that, if this is the case, the signaling equilibrium of Proposition 5 collapses, reinforcing the conclusion from Corollary 1 that cash-equity offers are a poor signaling device in tender offers.

5. Separating Votes and Cash Flows

In the above signaling equilibria, lower-valued bidders must relinquish more private benefits. As a result, the equilibrium outcomes typically exhibit inefficiencies at the "bottom": low-valued bidders are more likely to fail or do not even submit a bid. In particular, when takeover outcomes are deterministic, only bidders above a cut-off type make a bid in equilibrium, and a lower cut-off value implies more value-improving takeover activity. This raises the question of what determines the level of the cut-off type.

5.1. Efficiency of Restricted Bids. To address this question, we return to the general framework with one-dimensional bidder types, i.e., the setting of Theorem 2. Recall that, unlike the price P and the private benefit retention rate α, the bid restriction r is a redundant signal in that setting. However, it turns out that the chosen bid restrictions affect the cut-off type and hence takeover activity.

THEOREM 4. *In the tender offer game where bidders can commit to relinquish any fraction of their private benefits and* $\Phi(\cdot)$ *is a non-decreasing function, restricting bids (more) promotes takeover activity.*

Like Theorem 2, this result can be applied to the different tender offer games studied in Section 4.1. A common feature of the signaling equilibria analyzed so far is therefore that bid restrictions, though either insufficient or redundant as a signal, promote takeover activity. This is because smaller transaction sizes mitigate the asymmetric information problem: With fewer traded shares, a bidder gains less (in total) from paying a price below the post-takeover share value. This reduces the incentives to mimic low-valued bidders, so that these types do not need to sacrifice as much private benefits to credibly reveal low security benefits. Thus, more restricted bids translate into smaller signaling costs.

In essence, Theorem 4 says that efficiency is *de*creasing in the amount of cash flow rights traded.[14] Thus, contrary to standard bilateral trade models under asymmetric information, quantity rationing is not an inefficient outcome but a means of mitigating the trade inefficiency. This counterintuitive implication obtains because the gains from trade in tender offer models are typically contingent on the transfer of votes, rather than the transfer of cash flow rights *per se* [in contrast to security issuance models such as Myers and Majluf (1984) or DeMarzo and Duffie (1999)].[15]

5.2. Dual-Class Offers. In fact, that a share (trade) represents a (trade of a) bundle of two goods with potentially distinct values is a defining feature of tender offer games. This provides a new angle on the asymmetric information problem in tender offer games: value-improving control transfers are impeded because control must be transferred in conjunction with misvalued cash flow rights. Hence, efficiency would be

[14] At et al. (2008) show that Theorem also holds in pooling equilibrium outcomes.

[15] This is not true if the bidder's incentives to improve the target value vary with her post-takeover stake. Separating cashflow and voting rights can in this case be beneficial even in the absence of asymmetric information (Burkart et al., 1998).

further improved if the bidder could gain majority control without acquiring a majority stake, i.e., if she could separate votes and cashflows to prevent the informational frictions of the cash flow trade from spilling over into the vote trade.[16]

Perhaps the most straightforward way to unbundle control and ownership is to make a dual-class security-exchange offer. In a dual-class offer, the bidder offers to exchange each of the target's voting shares against a non-voting share. Shareholders accept the bid as long as it preserves their fraction of the cashflow rights. If the offer were to exchange shares at less than a one-to-one ratio, each shareholder would reject it. By construction, the bidder pays exactly the post-takeover security benefits to gain control. This replicates the full information outcome without revealing the bidder's type.

Despite resolving the asymmetric information problem, the dual-class offer is problematic because it leaves all cash flow rights with the shareholders. That is, the bidder has no equity interest in the firm after the takeover. On the one hand, this makes the offer equivalent to a simple replacement of management, which begs the question why a takeover is necessary in the first place. On the other hand, it makes the offer prone to abuse by value-decreasing bidders (or "fly-by-night" operators), since it does not require the bidder to put up any cash (Bebchuk and Hart, 2001). Cash payments put at least some (lower) bounds on the bidder's quality.

5.3. Contingent Claims. Dual-class offers or other extreme solutions, which leave the bidder with no equity interest in the firm ex post, are unnecessary. The bidder is merely unwilling to pay for cashflows which she knows do not to exist. The ideal solution is therefore to let the bidder acquire the target in exchange for cash and a set of securities which leave the "non-existing" cashflow to target shareholders. Such a trade can be implemented with call options. It merely requires that every type $X \in \mathcal{X}$ purchases a target share in exchange for cash X and a (cash-settled) call option with an exercise price of X. To see that this is incentive-compatible, consider two arbitrary types X and x with $X > x$. If the high type mimics the lower bid, she ex ante pays a cash price x for shares that are worth X. However, ex post she cannot capitalize on this gain, as the target shareholders will exercise their options once the actual value improvement becomes known to the market. Conversely, the low type does not mimic the high type because she would pay X for shares that are worth x. Thus, the offer schedule is incentive-compatible. Moreover, every bidder succeeds, irrespective of how many shares she acquires or whether she enjoys any private benefits.

[16] For this reason it can be both socially and privately optimal for target shareholders to adopt a dual-class share structure (At et al., 2008).

Proposition 8. *There exists a fully revealing equilibrium in which the bidder purchases each target share for $P(X) = X$ in cash and a call option with the exercise price $S(X) = X$. The equilibrium implements the full information outcome.*

Financial engineering enables the bidder (i) to trade economic ownership void of voting rights and (ii) to issue contingent claims. The first step of the transaction consists of acquiring the target shares and stripping them of their votes. In the second step, the bidder re-issues the cash flow rights, restructured into claims that punish her for "lying" about the security benefits. The call options which are executed when the post-takeover security benefits are higher than professed penalize the pretense of low security benefits—ex post when the true value is observed. This makes the offer equivalent to the simplest solution to the asymmetric information problem: a bid price which is contingent on the post-takeover share value.

The above offer transfers, cash aside, only future claims but no actual future cash flows to target shareholders. This is an artefact of the assumption that the post-takeover security benefits X are deterministic (i.e., perfectly known by the bidder). Yet, the main intuition carries over to a setting with stochastic cash flows. Let $X \in [0, \infty]$ be a random variable. Suppose that there are n bidder types $\theta \in \Theta \equiv \{1, 2, \ldots, n\}$, each knowing the probability density function $f_\theta(X)$ of her post-takeover cash flows. In addition, we assume that the family of densities $\{f_\theta(X)\}_{\theta \in \Theta}$ satisfies the strong monotone likelihood ratio property (SMLRP).

Assumption 1 (SMLRP). *For all $\theta' > \theta$, $f_{\theta'}(X)/f_\theta(X)$ is strictly increasing.*

To construct a fully revealing equilibrium, we allow the bidder to pay the target shareholders with various contingent claims. In the particular case considered, the bidder—apart from paying cash—issues bonds and barrier options.[17] The barrier option used to construct the equilibrium is a (cash-settled) "knock-in" (or "up-and-in") call option. This is a latent call option with an exercise price of S that becomes activated only once the value of the target X exceeds some "trigger" level $T > S$. When they are exercised, these options dilute the value of the firm's equity. Hence, the combination of bonds and knock-in options can alternatively be interpreted as a reduced-form implementation of convertible bonds. To simplify the exposition, we further assume that $\Phi = 0$.

Theorem 5. *In the tender offer game with stochastic post-takeover security benefits, a fully revealing equilibrium exists if Assumption 1 holds. All bidder types make a*

[17] A barrier option is a type of financial option where the option to exercise depends on the (price of the) underlying crossing or reaching a given barrier level.

bid and purchase the target shares for a combination of cash, bonds and knock-in call options.

The equilibrium offers in Theorem 5 represent a security design solution. This becomes clear from a comparison to models of external financing under asymmetric information (e.g., Myers and Majluf, 1984; Duffie and DeMarzo, 1999). In those models, the security issuer wants to signal a high value. To this end, she issues debt-like claims, selling low cashflow realizations and retaining high cashflow realizations. By contrast, in the tender offer game, the bidder primarily wants to signal a low value, which is why she uses knock-in options that are designed to leave high cash flow realizations to the target shareholders.

A recent strand of literature emphasizes the potential adverse effects of using contingent claims to unbundle control and ownership (e.g., Hu and Black, 2006, 2008). This literature contends that re-allocating control void of economic ownership can lead to inefficient decision-making ("empty voting"). By taking the post-takeover target value as given, the present framework abstracts from how the security design affects the bidder's post-takeover incentives to improve the target. Yet, such incentives are likely to depend on the nature of the economic interest that she retains in the firm. For example, the bonds and the knock-in options in the equilibrium of Theorem 5 would have countervailing effects on incentives.

Theorem 5 suggests quite the opposite: bundling—rather than unbundling—control and ownership can be the cause of an inefficient control allocation. The key to this result is that the two counterparties value cash flow rights and voting rights differently, so that asymmetric information about one set of rights can impede the transfer of the other set of rights. Such situations are by no means confined to tender offers. It can arise in any transaction involving bundles of goods, in which the buyer and the seller not only value the goods differently but also have different levels of information about their quality. As shown here, adopting this view for the analysis of votes and cash flow rights can potentially lead to new insights about the role of security design in corporate governance.

6. Concluding Remarks

This paper analyzes tender offers in which a single bidder is better informed about the post-takeover share value than dispersed target shareholders. Two key features of the tender offer process render this situation very different from standard bilateral trade models. First, free-riding shareholders have full bargaining power over the value improvement in the target shares, even though the better informed bidder makes a take-it-or-leave-it offer. Second, the parties in a tender offer bargain both over control (voting rights) and over ownership (cash flow rights) in the target firm. That is, unlike other signaling models in finance, a share (trade) represents a (trade of a) bundle of two goods with potentially distinct values.

We demonstrate that these differences lead to constraints as well as solutions that are absent in bilateral trade models. Because the bidder is forced to concede all gains in share value to the shareholders, she cannot signal her type by *voluntarily* giving up such gains. Neither restricted bids nor cash-equity offers are therefore viable signals in tender offers. Instead, the bidder must enjoy private benefits that are not only excluded from bargaining but can also be forgone in a manner which allows inference about the post-takeover share value. This is never possible when the bidder has additional private information about the private benefits, as in the case of two-dimensional bidder types. In the one-dimensional case, relinquishing private benefits is a viable signal because they depend on the security benefits in a predictable manner. Dilution, debt financing or toeholds can serve this purpose. The underlying principle in all cases is the same: the bidder must forgo (more) private benefits to signal a low(er) type. Unfortunately, some low-value bidders may find it too costly to signal their type even if the takeover would be efficient.

Such inefficiencies can be overcome if the bidder can include derivatives in the tender offer terms. Derivatives allow the bidder to separate cash flow rights from voting rights. This separation prevents that the information problems in the trade of cash flow rights spill over into, and thereby impair, the trade of voting rights. As a result, control can be transferred efficiently irrespective of any disagreement between the bidder and the target shareholders about the value of the post-takeover cash flow rights.

Our analysis has implications for the design of takeover bids. For instance, it suggest that derivatives as a means of payment should play a more prominent role in tender offers than the combination of cash and equity. Furthermore, acquiring firms may signal their quality through self-imposed restrictions on post-takeover decisions or the amount of takeover leverage. The main theoretical contribution of this paper is to study how the interaction of asymmetric information and collective action problems, in a specific market setting, may bear on the optimal design of a trade contract. We believe that there are many situations other than tender offers in which such interactions are potentially important.

Appendix

Proof of Proposition 1. Corollary 1 follows from the equivalence of mixed offers and restricted cash-only offers which the subsequent lemma establishes. Consider a bid for r shares that offers a cash price C and t shares in the post-takeover firm.

LEMMA 1. *Under full information, the restricted mixed offer (r, C, t) and the restricted cash-only offer (r^{co}, C^{co}) with $C^{co} = C$ and $r^{co} = r - t$ are payoff-equivalent.*

PROOF. To succeed, the mixed offer must satisfy the free-rider condition $C + tX \geq rX$, or equivalently

$$C/r + (t/r)\, X \geq X. \tag{6.1}$$

Given the condition is satisfied, all shareholders tender, and the bidder's payoff is

$$\Phi(X) + r\left[X - (C/r + (t/r)\, X)\right]. \tag{6.2}$$

Rearranging the free-rider condition (6.1) to

$$C \geq (r - t)\, X$$

and the bidder's payoff (6.2) to

$$\Phi(X) + (r - t)\, X - C$$

shows that the restricted cash-*only* offer (r^{co}, C^{co}) with $C^{co} = C$ and $r^{co} = r - t$ is payoff-equivalent for any X. □

Hence, if a fully revealing equilibrium in mixed offers were to exist, a fully revealing equilibrium in cash-only offers would also exist. As Proposition 1 rules out the latter, a mix of cash and equity is not a viable signal. ■

Proof of Proposition 2. We first characterize properties of an incentive-compatible r-P-schedule.

LEMMA 2. *In a fully revealing equilibrium, $r(\cdot)$, $P(\cdot)$ and $\Pi(\cdot)$ must be increasing.*

PROOF. Without loss of generality, choose an arbitrary pair of types, X and x, and let $X > x$. A fully revealing schedule $\{(r(\cdot), P(\cdot))\}$ must satisfy the non-mimicking constraints

$$r\,(X)\,[X - P\,(X)] \geq r\,(x)\,[X - P\,(x)] \quad \text{for } (x, X) \in \mathcal{X}^2. \tag{6.3}$$

We show by contradiction that (6.3) requires $r\,(X) > r\,(x)$. The non-mimicking constraints for type X and x are respectively

$$C\,(X) - r\,(X)\,X \leq C\,(x) - r\,(x)\,X \quad \text{and} \quad C\,(x) - r\,(x)\,x \leq C\,(X) - r\,(X)\,x \tag{6.4}$$

where $C(\cdot) \equiv r(\cdot)\,P(\cdot)$. For $r(X) = r(x)$, the inequalities hold jointly only if $C(X) = C(x)$, and hence $P(X) = P(x)$, in which case the two offers would be identical. For $r(X) < r(x)$, rewrite (6.4) as

$$C(x) \geq C(X) + [r(x) - r(X)]\,X \quad \text{and} \quad C(x) \leq C(X) + [r(x) - r(X)]\,x.$$

Since $C(X) + [r(x) - r(X)]\,X > C(X) + [r(x) - r(X)]\,x$, the constraints cannot hold jointly. Thus, the non-mimicking constraints are violated unless $r(\cdot)$ is increasing. Given $r(X) > r(x)$, condition (6.3) implies that the bid price and the bidder's profit must also be increasing in her type. To this end, we rewrite (6.4) as

$$r(X)\,[X - P(X)] \geq r(x)\,[X - P(x)] \quad \text{and} \quad \frac{r(X)}{r(x)}\,[x - P(X)] \leq x - P(x)\,.$$

Given that $r(X)/r(x) > 1$, the second inequality implies $P(X) > P(x)$. Furthermore, as $r(x)\,[X - P(x)] > r(x)\,[x - P(x)]$, the first inequality implies $r(X)\,[X - P(X)] \geq r(x)\,[x - P(x)]$. Thus, higher types must pay higher prices and make higher profits. $\square$

Local Optimality. Lemma 2 states necessary conditions for incentive-compatibility. To derive a particular schedule, we assume that $r(\cdot)$ and $C(\cdot)$ are continuously differentiable functions, and rephrase the bidder's optimization problem as a direct truth-telling mechanism:

$$\max_{\hat{X} \in \mathcal{X}} \left\{ r(\hat{X})X - C(\hat{X}) \right\}. \tag{6.5}$$

In equilibrium, the first-order condition must hold at $\hat{X} = X$, i.e.

$$r'(X)X = C'(X). \tag{6.6}$$

Quasi-concavity. Condition (6.6) is sufficient to ensure incentive-compatibility if the objective in (6.5) is quasi-concave (and out-of-equilibrium beliefs are suitably chosen). Substituting (6.6) into the derivative of the objective function gives

$$\frac{\partial}{\partial \hat{X}} \left[r(\hat{X})X - C(\hat{X}) \right] \bigg|_{C'(\hat{X}) = r'(\hat{X})\hat{X}} = r'(\hat{X})X - r'(\hat{X})\hat{X} = r'(\hat{X})(X - \hat{X}).$$

Given that $r'(\cdot) > 0$ (Lemma 2), it follows that the derivative switches sign for all types $X \in (0, \overline{X})$ once (from positive to negative), and the objective function is strictly quasi-concave.

Cut-off type. Condition (6.6) puts a constraint on how equilibrium profits $\Pi^*(X) = r(X)X - C(X)$ vary across types. By the envelope theorem,

$$\frac{\partial \Pi^*(X)}{\partial X} = \underbrace{r'(X)X - C'(X)}_{=0} + r(X) = r(X). \tag{6.7}$$

That is, the marginal change in profits is given by the bid restriction $r(X)$.

Given that bidders have bargaining power ω, shareholders always tender at the price $P = (1-\omega)\overline{X}$. As type $\overline{X}$ buys shares below their true value, she buys all shares and makes a profit $\Pi^* \left(\overline{X}\right) = \omega\overline{X}$. Since profits decrease at the rate $r(X)$ (condition (6.7)), the cut-off type type X^c, making zero profits, is defined by

$$\int_{X^c}^{\overline{X}} r(u)du = \omega\overline{X}. \tag{6.8}$$

Out-of-equilibrium beliefs. The proposed schedule can be supported as a signaling equilibrium with out-of-equilibrium beliefs that any deviation from the schedule comes from the highest-valued bidder type $\overline{X}$. Under these beliefs, the target shareholders do not tender their shares in response to a deviation bid $(\tilde{r}, \tilde{P})$ unless $\tilde{P} \geq (1-\omega)\,\overline{X} = P(\overline{X})$. Consider two cases. (i) For bidder types $X \in [P(\overline{X}), \overline{X}]$, the deviation bid $(\tilde{r}, P(\overline{X}))$ would yield a positive profit. Yet, it is dominated by the $(r(\overline{X}), P(\overline{X})) = (1, P(\overline{X}))$, the equilibrium bid of the highest type, which we know to be mimicking-proof. Hence, by implication, the deviation is unattractive to these types. (ii) For bidder types $X \in [0, P(\overline{X}))$, the deviation bid would yield a loss and is therefore unattractive to these types. ∎

Proof of Corollary 1. From the definition of the cut-off type (equation (6.8)), it follows that $\partial X^c/\partial\omega < 0$ and $\lim_{\omega\to 0} X^c = \overline{X}$. ∎

Proof of Theorem 2. Given that $\Phi'(\cdot) \geq 0$, there exists a schedule $\{\alpha(\cdot), r(\cdot), P(\cdot)\}$ with $\alpha' > 0$, $r' \geq 0$ and $P' > 0$ that can be supported as a fully revealing equilibrium.

Quasi-concavity. Suppose that the proposed schedule satisfies (4.3) for all $X \in \mathcal{X}$. This schedule then satisfies quasi-concavity of the objective function. Specifically, we show that

$$\partial\Pi/\partial\hat{X} = \alpha'(\hat{X})\Phi(X) + r'(\hat{X})[X - P(\hat{X})] - r(\hat{X})P'(\hat{X}) \tag{6.9}$$

is non-negative when $\hat{X} \leq X$ and non-positive when $X \geq \hat{X}$.

Condition (4.3) implies that

$$r(\hat{X})P'(\hat{X}) = \alpha'(\hat{X})\Phi(\hat{X}) + r'(\hat{X})[\hat{X} - P(\hat{X})].$$

Substituting the right-hand side into (6.9) and rearranging yields

$$\partial\Pi/\partial\hat{X} = \alpha'(\hat{X})\left[\Phi(X) - \Phi(\hat{X})\right] + r'(\hat{X})\left[X - \hat{X}\right]. \tag{6.10}$$

The assumption $\Phi'(\cdot) \geq 0$ implies that $\Phi(X) \geq \Phi(\hat{X})$ when $\hat{X} \leq X$ and that $\Phi(X) \leq \Phi(\hat{X})$ when $\hat{X} \geq X$. Given that $\alpha' > 0$ and $r' \geq 0$, it follows that

$$\partial\Pi/\partial\hat{X} \text{ is } \begin{cases} \text{non-negative} & \text{for } \hat{X} < X \\ 0 & \text{for } \hat{X} = X \\ \text{non-positive} & \text{for } \hat{X} > X \end{cases} .$$

Thus, the proposed schedule makes $\Pi(\hat{X}; X)$ weakly quasi-concave for all $X \in \mathcal{X}$. This also holds for $r'(\hat{X}) = 0$, in which case all bidder types propose the same bid restriction.

Local optimality. Condition (4.3) is a functional equation for $\alpha(\cdot)$, $r(\cdot)$ and $P(\cdot)$ with *two degrees of freedom*. To derive an example of an incentive-compatible schedule, we set $r(\cdot) = 0.5$. Then, condition (4.3) simplifies to

$$\alpha'(X) = \frac{P'(X)}{2\Phi(X)}.$$

Integrating on both sides over $[X, \overline{X}]$ yields

$$\int_X^{\overline{X}} \alpha'(u)du = \int_X^{\overline{X}} \frac{P'(u)}{2\Phi(u)}du \quad \Leftrightarrow \quad \alpha(\overline{X}) - \alpha(X) = \int_X^{\overline{X}} \frac{P'(u)}{2\Phi(u)}du.$$

As the highest-valued type does not have to relinquish any private benefits $[\alpha(\overline{X}) = 1]$,

$$\alpha(X) = 1 - \int_X^{\overline{X}} \frac{P'(u)}{\Phi(u)}du. \tag{6.11}$$

One possible price schedule is $P(X) = X$ in which case $\alpha(X) = 1 - \int_X^{\overline{X}} [\Phi(u)]^{-1}\, du$. As shareholders receive $P(X) + [1 - \alpha(X)]\, \Phi(X)$, the free-rider condition is also satisfied.

Cut-off type. The condition (4.3) puts a constraint on how equilibrium profits vary across types in equilibrium. By the envelope theorem, we have that equilibrium profits must be increasing at the rate

$$\partial\Pi^*/\partial X = \alpha(X)\Phi'(X) + r(X),$$

for *any* schedule that satisfies (4.3).

Given an equilibrium exists, the cut-off type X^c is given by

$$\int_{X^c}^{\overline{X}} \{\alpha(u)\Phi'(u) + r(u)\}\, du = \Phi(\overline{X}). \tag{6.12}$$

Under the proposed equilibrium schedule, bidder types below X^c incur a loss under the proposed schedule. Hence, they prefer *not* making a bid over making the bid prescribed by the proposed schedule. The option of notmaking a bid does not undermine the non-mimicking constraints. Under the proposed schedule, the bidder prefers a loss-making offer to offers made by higher-valued types. A fortiori, she also prefers a zero-profit offer over the latter.

Out-of-equilibrium beliefs. The proposed schedule can be supported as a signaling equilibrium under the out-of-equilibrium beliefs that any deviation from the schedule comes from the highest bidder type, $\overline{X}$. Under these beliefs, the target shareholders do not tender their shares in response to a deviation bid $(\tilde{r}, \tilde{\alpha}, \tilde{P})$ unless $\tilde{P} \geq \overline{X}$. Any such bid, however, is weakly dominated by $\left(0.5, 1, \overline{X}\right)$, which is the equilibrium bid of the highest type. Since $\left(0.5, 1, \overline{X}\right)$ is mimicking-proof, any successful deviation bid is—by implication—unattractive under the proposed out-of-equilibrium beliefs.

Bid restriction. The above proves that the bid restriction r is a redundant signal. This follows because, as we have shown, the schedule

$$\{r^*(X), \alpha^*(X), P^*(X)\} = \begin{cases} \left\{0.5, 1 - \int_X^{\overline{X}} [\Phi(u)]^{-1}\, du, X\right\} & \text{for } X \in [X^c, X] \\ \{0, 0, 0\} & \text{for } X \in [0, X^c) \end{cases}$$

can, among others, be supported as a fully revealing equilibrium.

By contrast, the private benefit retention rate α and the price P are indispensable as signals. First, if α is invariant across types, Theorem 1 applies. Second, a uniform price in a fully revealing equilibrium must satisfy $P = \overline{X}$. But then all bidder types $X < \overline{X}$ prefer the offer $(0.5, 1, \overline{X})$, which always succeeds irrespective of shareholder beliefs, to any other offer with $P = \overline{X}$. Hence, they would pool. ■

Proof of Proposition 3. The main difference between the extraction setting and the general setup is that relinquishing private benefits (lower ϕ) simultaneously increases the security benefits, and—by the free-rider condition—affects the shareholders' reservation price. More precisely, for a given α, the shareholders' reservation price becomes $X + (1 - \alpha)\, \Phi(X)$. Remaining in the general setup, assume now that, for a chosen price P, the bidder must pay $P + (1 - \alpha)\, \Phi(X)$. Incorporating this "mark-up", the bidder's objective function becomes

$$\begin{aligned} \Pi(r, \alpha, P; X) &= \alpha\Phi(X) + r\, [X - P - (1 - \alpha)\, \Phi(X)] \\ &= \alpha\Phi(X) - (1 - \alpha)\, r\Phi(X) + r\, (X - P) \\ &= \underbrace{[\alpha - (1 - \alpha)\, r]}_{\equiv\psi}\Phi(X) + r\, (X - P)\,. \end{aligned}$$

The last expression is isomorphic to (4.1), with the replacement of α with ψ amounting to but a change of notation. To capture the extraction setting, all that remains to do is to define

$$X \equiv \left(1 - \bar{\phi}\right) V \quad \text{and} \quad \Phi(X) \equiv \bar{\phi} V.$$

Note that $\Phi\,(\cdot)$ is increasing in the bidder's type. ■

Proof of Proposition 4. As in the dilution setting, relinquishing private benefits (lower D) simultaneously increases the security benefits. Within the general setup, the relevant objective function is therefore

$$\Pi(r, \alpha, P; X) = \psi\Phi(X) + r\,(X - P)$$

where ψ is defined as in the previous proof$X \equiv V - \overline{D}$ and $\Phi(X) \equiv \overline{D}$. Note that $\Phi(\cdot)$ is constant across bidder types. ■

Proof of Proposition 5. If the bid succeeds ($\beta \geq 0.5$), the holding company is worth $H(\beta, X) = A(1, X) + \beta X$. Under full information, shareholders do not tender unless $C(\beta) + s(\beta)H(\beta, X) \geq X$. To ensure a successful merger ($\beta = 1$), the bidder must choose $s(\beta)$ and $C(\beta)$ such that

$$s(\beta) \geq \frac{X - C(\beta)}{Z + \beta X + \lambda X} \tag{6.13}$$

for all $\beta \in [0.5, 1]$. In this case, all shareholders tender their shares whenever they believe that more than half the shares are tendered, and the bidder must ultimately pay $C(1)$ and $s(1)$. To simplify the exposition, we omit β and express the bidder's offer as a pair (s, C) which must satisfy the free-rider condition for $\beta = 1$, i.e., $s \geq (X - C)/(Z + X + \lambda X)$.[18] Note that condition (6.13) violates neither the cash constraint nor the control constraint if $C(\beta)$ is chosen sufficiently high.

For a given cash price C and equity component s, the bidder's profit from a successful merger is therefore

$$\Pi(X) = (1 - s)H\,(1, X) - C - \lambda X = (1 - s)Z + (1 - s)\left(X - \frac{s\lambda X + C}{1 - s}\right).$$

Now define $\alpha \equiv 1 - s$, $r = \alpha$, $\Phi(X) = Z$, and $P = \frac{s\lambda X + C}{1 - s}$. The bidder can use s to adjust α and r, and she can use C to adjust P. From Theorem 2, it follows that α must be increasing which in turn implies that s must be decreasing. The constraint $\alpha = r$ results from the fact that the bidder merges the firms and pays the target shareholders with holding company shares.

In this setting, the cut-off type is not necessarily determined by the participation constraint ($\Pi \geq 0$). As lower types issue more equity, they may also run either into the control constraint $s(\cdot) \leq 0.5$ or into the cash constraint $C(\cdot) \geq 0$. The latter may occur because the bidder can in principle become a net *issuer*, rather than a net

[18] Even without a contingent offer, there exists a self-fulfilling equilibrium in which the merger succeeds for (C, t) as long as it satisfies the free-rider condition for $\beta = 1$: If each shareholder believes that all other shareholders tender, she also tenders. Hence, once can alternatively focus on non-contingent offers, and select merger success as the equilibrium outcome whenever it is consistent with the free-rider condition.

purchaser, of financial claims. The cash constraint is relevant for bidders for whom A is very large relative to X. Notwithstanding, the participation constraint becomes binding as Z decreases. In particular, $Z = 0$ is equivalent to $\Phi(X) = 0$, and hence causes signaling breaks down. (It is straightforward to verify that using λX (instead of Z) as the synergy gains leads to similar results; in particular, signaling breaks down when $\lambda = 0$.) ■

Proof of Proposition 6. The main difference between the toehold setting and the general setup is that acquiring a toehold of t shares (i) leaves only $1 - t$ shares that can potentially be acquired in the tender offer, and (ii) allows the bidder to gain control by purchasing only $0.5 - t$ additional shares in the takeover. The bidder's profit function is therefore

$$\Pi(r_t, t, P; V) = tV + r_t\,(1-t)\,(V-P)$$

where $r_t \in \left[\frac{0.5-t}{1-t}, 1\right]$ is the fraction of the *remaining* shares that the bidder offers to purchase in the tender offer. Now define

$$t = \alpha\bar{t} \quad \text{and} \quad r \equiv r_t\,(1-t)$$

so that

$$\Pi(r, t, P; V) = \alpha\bar{t}V + r(V-P).$$

This objective function is isomorphic to (4.1). All we need to do in this case is define $X \equiv V$ and $\Phi(X) \equiv \bar{t}V$. Note that $\Phi(\cdot)$ is increasing in the bidder's type. ■

Proof of Proposition 7. The objective function in the probabilistic tender offer game is:

$$\begin{aligned}\Pi(r,P) &= q(r,P)\left[\Phi(X) + \gamma(r,P)\,(X-P)\right] \\ &= \underbrace{q(r,P)}_{\alpha}\Phi(X) + \underbrace{q(r,P)\gamma(r,P)}_{r}\,(X-P)\,.\end{aligned}$$

The last expression is isomorphic to (4.1), except that r can take values below 0.5. Provided that $\Phi(\cdot)$ is a non-decreasing function, Theorem 2 can thus be applied. ■

Proof of Theorem 4. We show that every implementable level of takeover activity can also be implemented by a schedule with $P(X) = X$ and $r'(X) = 0$ (step 1) and that minimizing r promotes takeover activity (step 2).

Step 1. Consider an equilibrium schedule which implements a cut-off type X'. This type satisfies $\Pi(X'; X') = \alpha(X')\,\Phi(X') + r(X')\,[X' - P(X')] = 0$. Keeping the cut-off type constant, we maintain the zero-profit condition by setting $\alpha(X') = 0$ and $P(X') = X'$.

This provided, the *same* cut-off type can also be characterized by

$$
\begin{aligned}
\Phi(\overline{X}) - \alpha(X')\Phi(X') &= \Phi(\overline{X}) \\
\alpha(\overline{X})\Phi(\overline{X}) - \alpha(X')\Phi(X') &= \Phi(\overline{X}) \\
\left[\alpha(u)\Phi(u)\right]_{X'}^{\overline{X}} &= \Phi(\overline{X}) \\
\int_{X'}^{\overline{X}} \left\{\alpha(u)\Phi'(u) + \alpha'(u)\Phi(u)\right\} du &= \Phi(\overline{X}) \qquad (6.14)
\end{aligned}
$$

where we use the fact that $\alpha(\overline{X}) = 1$.

Next, recall that (6.12) characterizes a unique cut-off type X^c for any equilibrium schedule,

$$\int_{X^c}^{\overline{X}} \left\{\alpha(\hat{X})\Phi'(\hat{X}) + r(\hat{X})\right\} d\hat{X} = \Phi(\overline{X}), \qquad (6.15)$$

and consider an equilibrium schedule which specifies $P\,(X) = X$ and $r\,(X) = r\,(X')$. Under this schedule, the first-order condition (4.3) becomes $r(X') = \alpha'(\hat{X})\Phi(\hat{X})$. Substituting $r\,(X') = r(\hat{X}) = \alpha'(\hat{X})\Phi(\hat{X})$, equation (6.15) becomes equation (6.14), and $X^c = X'$.

Step 2. Given step 1, we can restrict our attention to schedules with $P(X) = X$ and $r(X) = r$. Under these schedules, the bidder's equilibrium profit is

$$\Pi^*\,(X) = \alpha\,(X)\,\Phi\,(X)\,.$$

When $\alpha\,(\cdot)$ declines at a lesser rate, bidder profits deteriorate slower, as the type decreases. Using conditon (4.3), $P'(X) = 1$ and $r'(X) = 0$, the slope of $\alpha\,(\cdot)$ is

$$\alpha'(X) = \frac{r}{\Phi(X)}$$

and is increasing in r for any $X \in \mathcal{X}$. That is, bidder profits deteriorate at a lesser rate and the cut-off type is lower when r is smaller. ■

Proof of Theorem 5. We first establish the following two auxiliary results.

LEMMA 3. *For all $\theta' > \theta$, there exists a unique $X_\theta(\theta') \in (0,\infty)$ s.t.*

$$f_\theta(X)\begin{cases} > f_{\theta'}\,(X) & \textit{for all } X < X_\theta(\theta') \\ < f_{\theta'}\,(X) & \textit{for all } X > X_\theta(\theta') \end{cases}.$$

PROOF. By SMLRP, for all $\theta' > \theta$, there is a unique $X_\theta(\theta') \in (0,\infty)$ s.t.

$$f_{\theta'}(X)/f_\theta(X)\begin{cases} < 1 & \text{for } X < X_\theta(\theta') \\ = 1 & \text{for } X \in \mathcal{X}\,(\theta,\theta') \\ > 1 & \text{for } X > X_\theta(\theta') \end{cases}.$$

Otherwise, if $f_{\theta'}(X)/f_\theta(X)$ is either always larger or always smaller than 1, it cannot be that $F_\theta(\infty) = F_{\theta'}(\infty)$. This implies the result. □

LEMMA 4. *For all* $\theta'' > \theta' > \theta$, $X_{\theta'}(\theta'') \geq X_\theta(\theta')$.

PROOF. Suppose to the contrary that

$$(\bigstar) \qquad X_{\theta'}(\theta'') < X_\theta(\theta').$$

By Lemma 3, it then follows that

$$\begin{array}{lll}
(a) & \text{For } X \in (0, X_{\theta'}(\theta'')): & \frac{f_{\theta''}(X)}{f_{\theta'}(X)} < 1 \text{ and } \frac{f_{\theta'}(X)}{f_\theta(X)} < 1 \Rightarrow \frac{f_{\theta''}(X)}{f_\theta(X)} < 1 \\
(b) & \text{For } X = X_{\theta'}(\theta''): & \frac{f_{\theta''}(X)}{f_{\theta'}(X)} = 1 \text{ and } \frac{f_{\theta'}(X)}{f_\theta(X)} < 1 \Rightarrow \frac{f_{\theta''}(X)}{f_\theta(X)} < 1 \\
(c) & \text{For } X \in (X_{\theta'}(\theta''), X_\theta(\theta')): & \frac{f_{\theta''}(X)}{f_{\theta'}(X)} > 1 \text{ and } \frac{f_{\theta'}(X)}{f_\theta(X)} < 1 \Rightarrow \frac{f_{\theta''}(X)}{f_\theta(X)} \gtrless 1 \\
(d) & \text{For } X = X_\theta(\theta'): & \frac{f_{\theta''}(X)}{f_{\theta'}(X)} > 1 \text{ and } \frac{f_{\theta'}(X)}{f_\theta(X)} = 1 \Rightarrow \frac{f_{\theta''}(X)}{f_\theta(X)} > 1 \\
(e) & \text{For } X \in (X_\theta(\theta'), \infty): & \frac{f_{\theta''}(X)}{f_{\theta'}(X)} > 1 \text{ and } \frac{f_{\theta'}(X)}{f_\theta(X)} > 1 \Rightarrow \frac{f_{\theta''}(X)}{f_\theta(X)} > 1
\end{array}$$

Observe that (i) $f_{\theta''}(X) = f_{\theta'}(X)$ for $X = X_{\theta'}(\theta'')$ and (ii) $f_{\theta'}(X) = f_\theta(X)$ for $X = X_\theta(\theta')$. SMLRP implies that $f_{\theta''}(X)/f_\theta(X) \leq 1$ in case (c), and hence that (iii) $f_{\theta''}(X) = f_\theta(X)$ for $X = X_\theta(\theta')$. Points (ii) and (iii) together imply that (iv) $f_{\theta''}(X) = f_{\theta'}(X)$ for $X = X_\theta(\theta')$. Given that $f_{\theta''}(X) > f_{\theta'}(X)$ in case (c), points (iv) and (i) can only be reconciled with SMLRP if $X_{\theta'}(\theta'') = X_\theta(\theta')$. However, this contradicts inequality ($\bigstar$). □

Main proof. The proof proceeds in two steps. In the first step, we compare adjacent types and analyze local incentive-compatibility. In the second step, we show that an offer which is locally mimicking-proof is also globally mimicking-proof.

Local incentive-compatibility. Consider a type θ who, for each target share, offers a cash price P_θ, a debt claim with face value D, and a (cash-settled) knock-in call option with exercise price S_θ and trigger level T_θ.

Absent private benefits, a fully efficient equilibrium requires that the bidder's cash price is weakly lower than the expected value of the cash flow rights that she acquires. At the same time, the free-rider condition requires that the cash price is weakly higher than the expected value of the transferred cash flow rights. Both constraints can only be satisfied simultaneously if they are both binding:

$$P_\theta = \int_D^{T_\theta} (X - D)\, f_\theta(X) dX + \int_{T_\theta}^{\infty} (S_\theta - D)^+ f_\theta(X) dX.$$

Consequently, every truthful offer must yield zero bidder profits.

(i) The next higher type $\theta + 1$ does not mimic θ iff

$$-P_\theta + \int_D^{T_\theta} (X - D)\, f_{\theta+1}(X) dX + \int_{T_\theta}^{\infty} (S_\theta - D)^+ f_{\theta+1}(X) dX \leq 0.$$

Substituting for P_θ, the inequality can be written as

$$(S_\theta - D)^+ \int_{T_\theta}^{\infty} [f_{\theta+1}(X) - f_\theta(X)]\, dX \leq \int_D^{T_\theta} [f_\theta(X) - f_{\theta+1}(X)]\,(X - D)\, dX. \tag{6.16}$$

Set $T_\theta = X_\theta(\theta + 1)$. By Lemma 3, both integrals are then *strictly* positive for any $D < T_\theta = X_\theta(\theta + 1)$, in which case there exists a $S_\theta > 0$ such that (6.16) is satisfied.

(ii) Analogously, the next lower type $\theta - 1$ does not mimic θ iff

$$(S_\theta - D)^+ \int_{T_\theta}^{\infty} [f_{\theta-1}(X) - f_\theta(X)]\, dX \leq \int_D^{T_\theta} [f_\theta(X) - f_{\theta-1}(X)]\,(X - D)\, dX. \tag{6.17}$$

Set $D = X_{\theta-1}(\theta)$. By Lemma 3, the right-hand side is then *strictly* positive. By Lemma 4, $T_\theta = X_\theta(\theta+1) \geq X_{\theta-1}(\theta)$ so that the left-hand side integral is *strictly* negative. So, (6.17) holds.

Global incentive-compatibility. We now consider in turn types higher than $\theta + 1$ and types lower than $\theta - 1$.

(i) Given $T_\theta = X_\theta(\theta+1)$ and $D = X_{\theta-1}(\theta)$, consider now the incentive-compatibility constraint of an arbitrary type $\theta^+ > \theta + 1$ vis-à-vis type θ:

$$[S_\theta - X_{\theta-1}(\theta)]^+ \int_{X_\theta(\theta+1)}^{\infty} [f_{\theta^+}(X) - f_\theta(X)]\, dX \leq \int_{X_{\theta-1}(\theta)}^{X_\theta(\theta+1)} [f_\theta(X) - f_{\theta^+}(X)]\,[X - X_{\theta-1}(\theta)]\, dX.$$

Defining $\eta(X) \equiv f_{\theta+1}(X) - f_{\theta^+}(X)$, write the inequality as

$$[S_\theta - X_{\theta-1}(\theta)]^+ \int_{X_\theta(\theta+1)}^{\infty} [f_{\theta+1}(X) - f_\theta(X) - \eta(X)]\, dX \leq \int_{X_{\theta-1}(\theta)}^{X_\theta(\theta+1)} [f_\theta(X) - f_{\theta+1}(X) + \eta(X)]\,[X - X_{\theta-1}(\theta)]\, dX. \tag{6.18}$$

By Lemma 4, $X_{\theta+1}(\theta^+) \geq X_\theta(\theta + 1)$ so that $\eta(X) > 0$ for all $X < X_\theta(\theta + 1)$. This implies that the right-hand side of (6.18) is larger than the right-hand side of (6.16), and hence *strictly* positive. Turning to the left-hand side, because

$$-\int_{X_\theta(\theta+1)}^{\infty} \eta(X) dX = \int_0^{X_\theta(\theta+1)} \eta(X) dX - \int_0^{\infty} \eta(X) dX = \int_0^{X_\theta(\theta+1)} \eta(X) dX > 0,$$

the integral on the left-hand side of (6.18) is larger than the integral on the left-hand side of (6.16), and hence *strictly* positive. We conclude that—for $T_\theta = X_\theta(\theta + 1)$ and $D = X_{\theta-1}(\theta)$—there exists a *strictly* positive price, $S_\theta > 0$, such that no type $\theta^+ > \theta$ mimics type θ.

(ii) Given $T_\theta = X_\theta(\theta{+}1)$ and $D = X_{\theta-1}(\theta)$, consider now the incentive-compatibility constraint of an arbitrary type $\theta^- < \theta - 1$ vis-à-vis type θ:

$$[S_\theta - X_{\theta-1}(\theta)]^+ \int_{X_\theta(\theta+1)}^{\infty} [f_{\theta^-}(X) - f_\theta(X)]\, dX \leq \int_{X_{\theta-1}(\theta)}^{X_\theta(\theta+1)} [f_\theta(X) - f_{\theta^-}(X)]\,[X - X_{\theta-1}(\theta)]\, dX.$$

Defining $\zeta(X) \equiv f_{\theta-1}(X) - f_{\theta^-}(X)$, write the inequality as

$$[S_\theta - X_{\theta-1}(\theta)]^+ \int_{X_\theta(\theta+1)}^{\infty} [f_{\theta-1}(X) - f_\theta(X) - \zeta(X)]\, dX \leq \int_{X_{\theta-1}(\theta)}^{X_\theta(\theta+1)} [f_\theta(X) - f_{\theta-1}(X) + \zeta(X)]\,[X - X_{\theta-1}(\theta)]\, dX. \tag{6.19}$$

By Lemma 4, $X_{\theta-1}(\theta) \geq X_{\theta^-}(\theta{-}1)$ so that $\zeta(X) > 0$ for all $X > X_{\theta-1}(\theta)$. This implies that the right-hand side of (6.19) is larger than the right-hand side of (6.17), and hence *strictly* positive. Turning to the left-hand side, again by Lemma 4, $X_\theta(\theta + 1) \geq X_{\theta-1}(\theta) \geq X_{\theta^-}(\theta - 1)$ so that $\zeta(X) > 0$ for all $X > X_\theta(\theta + 1)$. This implies that the left-hand side integral of (6.19) is smaller than the left-hand side integral of (6.17), and hence *strictly* negative. So, (6.19) holds. We conclude that—for $T_\theta = X_\theta(\theta + 1)$, $D = X_{\theta-1}(\theta)$, and $S_\theta > 0$—no type $\theta^- < \theta$ mimics type θ. ■

References

At, Christian, Mike Burkart and Samuel Lee, 2007, "Security-Voting Structure and Bidder Screening," FMG Discussion Paper No 575.

Bebchuk, Lucian A. and Oliver Hart, 2001, "Takeover Bids vs. Proxy Fights in Contests for Corporate Control," CEPR Discussion Paper, No. 3073, ECGI - Finance Working Paper No. 4.

Berkovitch, Elazar and M. P. Narayanan, 1990, "Competition and the Medium of Exchange," *Review of Financial Studies*, 3(2), 153-174.

Bradley, Michael, 1980, "Interfirm Tender Offers and the Market for Corporate Control," *Journal of Business*, 53, 345-376.

Brusco, Sandro, Guiseppe Lopomo, David T. Robinson and S. Viswanhatan, 2007, "Efficient Mechanisms for Mergers and Acquisitions," *International Economic Review*, 48 (3), 995-1035.

Burkart, Mike, Denis Gromb and Fausto Panunzi, 2006, "Minority Blocks and Takeover Premia," *Journal of Institutional and Theoretical Economics*, 162(1), 32-49.

Chowdhry, Bhagwan and Narasimhan Jegadeesh, 1994, "Pre-Tender Offer Share Acquisition Strategy in Takeovers," *Journal of Financial and Quantitative Analysis*, 29(1), 117-129.

DeMarzo, Peter and Darrell Duffie, 1999, "A Liquidity-Based Model of Security Design," *Econometrica*, 67(1), 65-99.

Eckbo, B. Espen, Ronald M. Giammarino and Robert L. Heinkel, 1990, "Asymmetric Information and the Medium of Exchange in Takeover: Theory and Tests," *Review of Financial Studies*, 3(4), 651-675.

Ferreira, Daniel, Emanuel Ornelas and John L. Turner, 2007, "Unbundling Ownership and Control," ECGI Working Paper 172/2007.

Fishman, Michale, J., 1988, "A Theory of Pre-emptive Takeover Bidding," *Rand Journal of Economics*, 19(1), 88–101.

Fishman, Michael J., 1989, "Preemptive Bidding and the Role of the Medium of Exchange in Acquisitions," *Journal of Finance*, 44(1), 41-57.

Grossman, Sanford J. and Oliver D. Hart, 1980, "Takeover Bids, the Free-rider Problem, and the Theory of the Corporation," *Bell Journal of Economics*, 11(1), 42-64.

Grossman, Sanford J. and Oliver D. Hart, 1981, "The Allocational Role of Takeover Bids in Situations of Asymmetric Information", *Journal of Finance* 36(2), 253-270.

Grossman, Sanford J. and Motty Perry, 1986, "Perfect Sequential Equilibria," *Journal of Economic Theory*, 39, 97-119.

Hansen, Robert G., 1987, "A Theory for the Choice of Exchange Medium in Mergers and Acquisitions," *Journal of Business*, 60(1), 75-95.

Hirshleifer, David and Sheridan Titman, 1990, "Share Tendering Strategies and the Success of Hostile Takeover Bids," *Journal of Political Economy*, 98(2), 295-324.

Holmström, Bengt and Barry Nalebuff, 1992, "To the Raider Goes the Surplus? A Reexamination of the Free Rider Problem," *Journal of Economics and Management Strategy*, 1(1), 37-62.

Hu, Henry T.C. and Bernard Black, 2006, "The New Vote Buying: Empty Voting and Hidden (Morphable) Ownership," *Southern California Law Review*, 79(4), 811-908.

Hu, Henry T.C. and Bernard Black, 2008, "Equity and Debt Decoupling and Empty Voting II: Importance and Extensions," *University of Pennsylvania Law Review*, 156, 625-739.

Judd, Kenneth L., 1985, "The law of large numbers with a continuum of IID random variables," *Journal of Eonomic Theory*, 35(1), 19-25.

Laffont, Jean-Jacques and David Martimort, 2002, *The Theory of Incentives*, Princeton University Press, Princeton N.J.

Liu, Tingjun, 2008, "Takeover Bidding with Signaling Incentives," Carey School of Business, Arizona State University.

Marquez, Robert and Bilge Yilmaz, 2005, "Takeover Bidding and Shareholder Information," mimeo, University of Pennsylvania.

Martin, Kenneth J., 1996, "The Method of Payment in Corporate Acquisitions, Investment Opportunities, and Management Ownership," *Journal of Finance*, 51(4), 1227-1246.

Milgrom, Paul and Nancy Stokey, 1982, "Information, Trade and Common Knowledge," *Journal of Economic Theory*, 26(1), 17-27.

Müller, Holger and Fausto Panunzi, 2004, "Tender Offers and Leverage," *Quarterly Journal of Economics*, 119(4), 1217-1248.

Myers, Stewart and Nicholas S. Majluf, 1984, "Corporate Financing and Investment Decisions when Firms have Information that Investors do not have," *Journal of Financial Economics*, 13, 187-221.

Shleifer, Andrei and Robert W. Vishny, 1986, "Large Shareholders and Corporate Control," *Journal of Political Economy*, 94(3, Part 1), 461-488.

PAPER 4

Goldrush Dynamics of Private Equity

with Daniel Sunesson

ABSTRACT. We present a simple dynamic model of entry and exit in a private equity market with heterogeneous fund managers, a depletable stock of target companies, and learning about investment profitability. Its predictions match a number of stylized facts: Aggregate fund activity follows waves with endogenous transitions from booms to busts. Supply and demand in the private equity market are inelastic, and the supply comoves with investment valuations. High industry performance precedes high entry, which in turn precedes low industry performance. Differences in fund performance are persistent, first-time funds underperform the industry, and the first-time funds that are raised in boom periods are unlikely to be suceeded by follow-on funds. Fund performance and fund size are positively correlated across private equity firms, but negatively correlated across consecutive funds by the same firm. Finally, boom periods can make "too much capital chase too few deals."

1. Introduction

Capital commitments and investments in the private equity industry are cyclical. This has been documented for the venture capital industry by Gompers and Lerner (2000) and Lerner (2002), and for the buyout industry by Kaplan and Stein (1993) and Kaplan and Strömberg (2009). To give a recent example, the global buyout volume shrunk from about \$527 billion in early 2007 to about \$124 billion by mid-2008. Such boom-bust patterns suggest that the private equity business is transitory, expanding and contracting as the opportunities for profitable control investments emerge and disappear.

We develop a simple model which captures this transient nature. It produces waves which endogenously transition from booms to busts. Furthermore, the dynamics of entry, prices and returns *within* a wave match a wide range of empirical patterns: the inelasticity of private equity supply to private equity demand and vice versa; the

Thanks to Viral Acharya, Ulf Axelson, Patrick Bolton, Guillermo Caruana, Mariassunta Giannetti, Holger Müller, Raffael Repullo, Javier Suarez, Per Strömberg, and seminar participants at the Corporate Governance Lunch Seminar of the Stern School of Business and the ECGTN Conference in Barcelona for helpful comments. Financial support from the European Corporate Governance Training Network, the Jan Wallander Foundation and the Hedelius Foundation is gratefully acknowledged.

procyclicality of capital inflow and investment valuations; persistent performance differences across private equity firms; the underperformance of first-time funds; the positive (negative) relationship between entry and past industry returns (subsequent industry returns); the positive (negative) relationship between fund performance and fund size in the cross-section (in the time-series); and the notion of "too much capital chasing too few deals."[1]

The basic idea behind the model is to liken private equity waves to goldrushes. A goldrush starts with the discovery of gold which attracts gold diggers who settle nearby in the hope of making a fortune. As more gold is extracted over time, more gold diggers migrate to the area until all claims are staked. When the gold reserves dry up, the gold diggers either retire or migrate to the next discovery.[2] Our model essentially draws an analogy between gold discoveries and the emergence of private equity investment opportunities, gold diggers and private equity firms, claims and investments, as well as gold and investment returns.

In the model, a fixed population of companies becomes improvable because of a latent productivity shock. To keep matters simple, the improvement can only be realized by private equity firms, investment firms specialized in acquiring and reorganizing companies. To do so, a private equity firm must raise a private equity fund, find a target company, and negotiate a price at which the company's shareholders are willing to sell the company.[3] There are many private equity firms that repeatedly decide whether to raise a fund to acquire a company, i.e, whether to enter the private equity market. Each firm's entry decision depends on its own management ability, the number of available target companies, and the expected gains from reorganization. Importantly, the true expected gains are unknown but can be partially inferred from completed reorganizations. This learning process creates a link between past and current entry decisions.

The model yields a private equity wave under the plausible assumption that—absent positive experiences—the market's (prior) expectations are low. In that case, only few private equity funds are raised at the outset. When the true shock is low, these early funds earn disappointing returns, and investment activity subsequently subsides.

[1] These empirical patterns are documented by Gompers and Lerner (1999, 2000), Kaplan and Schoar (2005), Acharya et al. (2007), and Hochberg et al. (2008). The reported performance patterns in private equity stand in stark contrast to the evidence in the mutual fund industry (Malkiel, 1995; Berk and Green, 2004) and the investment management industry (Busse et al., 2008).

[2] An example is the Klondike Goldrush. In August 1896, gold was discovered in the Klondike river. By the summer of 1897, the nearby town of Dawson had grown to a population of 3,500. Around that time, steamships unloaded about one and half million dollars worth of Klondike gold in San Francisco and Seattle. Within half a year, the population of Dawson climbed to over 30,000. In the summer of 1899, the goldrush was officially over.

[3] Throughout the paper, we refer to private equity *firms* and target *companies*.

Conversely, when the true shock is high, the early funds earn promising returns, which attracts other private equity firms to the market. As fund activity rises, the magnitude of the shock is revealed at a faster rate, which in turn accelerates entry. This feedback loop between learning and entry fuels the boom. The countereffect is that the influx of new funds depletes the pool of target companies faster. The accelerating attrition ultimately leads to the bust.

The wave pattern reflects the inelasticity of demand and supply in the private equity market. Since the demand for private equity arises from exogenous shocks, it does not respond positively to supply. On the contrary, increases in supply reduce the demand faster. The supply of private equity is inelastic because the private equity firms only learn gradually about the profitability of investing. The *speed of learning* depends on the degree of investment specificity and on the market's prior beliefs. The more idiosyncratic a target company is, the less informative is its reorganization outcome about the prospects of reorganizing other companies. Furthermore, if the market ex ante perceives a high shock as very unlikely, it is more reluctant to interpret successful outcomes as a sign of general profitability. The *speed of entry* depends on the skill distribution among private equity firms. For instance, a skill pyramid with "few at the top, and many at the bottom" produces few entrants when expectations are low but many entrants when expectations are high. The combination of slow learning with a skill pyramid leads to waves with slow starts, explosive booms and sudden crashes.

When the market becomes more confident about the expected reorganization value, potential target companies increase in value, which in turn affects the negotiations between funds and target shareholders. Thus, a rise in market confidence not only attracts more private equity funds to the market but also raises the price that these funds must pay to acquire target companies. In other words, aggregate fund activity and valuation levels are jointly determined by market expectations and hence move together, consistent with the evidence in Kaplan and Stein (1993) and Gompers and Lerner (2000). However, even when expectations increase, the *true* profitability remains unaffected by learning. Thus, higher valuation levels do not imply that investments are more profitable. On the contrary, as valuations increase relative to "fundamentals", average fund profitability declines during a wave. This decline is reinforced by the entry of less skilled private equity firms. Similarly, the model yields a rationale for the positive relationship between entry and past industry performance, and for the negative relationship between entry and subsequent fund performance, documented by Kaplan and Schoar (2005). High industry performance today raises market confidence, which increases fund activity tomorrow and—at the same time—decreases future fund performance via higher prices.

At the fund level, the heterogeneity among private equity firms immediately implies persistent differences in fund performance: a fund that has outperformed the industry is likely to continue to outperform the industry with its follow-on funds. A more interesting prediction of the model is that a private equity firm's time of entry is related to its quality. In any period, only the private equity firms above a certain threshold quality raise a fund, and the threshold is decreasing in the level of market confidence. Dynamically, this means that entry and exit follow a last-in-first-out pattern: As the level of market confidence varies over time, the least skilled private equity firms are always the latest to enter and, by the same token, always the first to exit the market. Thus, at any point in time, the first-time funds (the latest entrants) underperform the industry. However, their follow-on funds—if the boom continues—improve in relative performance as private equity firms of even lower quality will enter the market. The lowest-quality firms enter after highly profitable periods, when valuation levels are high, and during periods in which fund activity will ex post turn out to have peaked. Due to the last-in-first-out pattern, such firms are likely to exit the market soon after. Or putting it differently, funds first raised in boom times are less likely to see follow-on funds. These predictions are consistent with the evidence in Kaplan and Schoar (2005).

Kaplan and Schoar (2005) also study the relation between fund size and fund profitability and report that the relationship is—on the one hand—positive and concave across different funds, and—on the other hand—negative across consecutive funds of the same private equity firm. While our baseline model assumes a uniform and constant fund size, these patterns naturally arise in an extension that allows private equity firms to run larger funds at an increasing marginal cost. The firms' marginal cost functions reflect their management ability. For any given level of market confidence, cross-sectional variation in size is driven by variation in ability: larger funds are managed by better private equity firms, which is the reason why they are more profitable. By contrast, for a given firm (quality), time variation in fund size is driven by time variation in market confidence, i.e., purely by learning. When market confidence is higher, a private equity fund makes more acquisitions. At the same time, the fund pays higher prices (due to increased valuation levels) and operates at a higher average cost (due to its larger size). Thus, as the true profitability of investing is time-invariant, the fund's true expected profit (per investment) is inversely related to its size during a wave.

Finally, we study the effects of fund competition in a simple model extension which incorporates search frictions into the private equity market. In the presence of such frictions, a fund's bargaining power vis-à-vis a target company is weaker when there are more competing funds or fewer target companies in the market. This reinforces the link between market confidence and acquisition prices: when the market becomes more confident, the prices rise not only because a target's total expected reorganization value increases but also because the entry of new funds shifts bargaining power to the targets. Compared to the absence of competition, fund profitability drops faster as a result of fund entry or target attrition, i.e., when "more money chases fewer deals." Such congestion effects slow down entry and precipitate exit so that fund activity both builds up and declines more gradually than in the basic model. Thus, fund competition affects neither the boom-bust pattern nor the last-in-first-out pattern of fund activity, but it "smoothes" the wave.

The phenomenon of waves has previously been analyzed theoretically by Jovanovic and Rousseau (2002), Shleifer and Vishny (2002) and Rhodes-Kropf and Robinson (2008) in the context of mergers and acquisitions; and by Inderst and Müller (2004) and Michelacci and Suarez (2004) in the context of venture capital markets. These papers address neither the role of learning or attrition nor the endogenous *intra*-wave dynamics of investment, prices, and returns.

We are not the first to study the impact of learning on financial decisions. For instance, learning models have been used to explain financial innovations (Persons and Warther, 1997), stock market prices (Timmermann, 1993, 1996; Veronesi, 1999; Pastor and Veronesi, 2008), going public decisions (Pastor et al., 2006; He, 2007), and business cycles (Veldkamp, 2005; Van Nieuwerburgh and Veldkamp, 2006). Contemporaneous work by Hochberg et al. (2008) and Glode and Green (2008) also incorporates learning into models of the private equity market. In both models, fund investors (limited partners) learn about the ability of fund managers (general partners). By contrast, in our model, fund managers learn about market conditions which affect the profitability of private equity investments.

The remainder of the paper is organized as follows. Section 2 presents the basic model. Section 3 derives the competitive Markov equilibrium. Section 4 analyzes the equilibrium dynamics. Section 5 presents the model extensions which incorporate fund size and fund competition. Section 6 concludes the paper.

2. Model

Consider a risk-neutral economy in discrete time, $t \in \mathbb{Z}_0^+$, with a fixed population of $\mathcal{N}$ companies. Initially, each company is run by an incumbent manager, and its discounted dividend value under the incumbent manager is normalized to 0.

In period 0, the economy experiences a latent productivity shock. The shock makes each company—if appropriately reorganized—improvable. A company's value after reorganization, V, is gamma-distributed with shape parameter $\alpha > 0$ and scale parameter $1/\beta > 0$. The mean of the gamma distribution, $\overline{V} = \alpha/\beta$, reflects the expected reorganization value. We assume that α is commonly known, whereas β is unobserved. Since a lower β translates into a higher expected reorganization value, this implies that the market is uncertain about the magnitude of the shock. The market's initial beliefs about β are also represented by a gamma distribution, with known shape and inverse scale parameters $\tau > 0$ and $1/\gamma > 0$ respectively.[4]

We assume that the incumbent managers cannot generate the value improvement. We also abstract from the possibility that they procure the necessary human capital through consulting services or the labor market. Instead, let there be $\mathcal{M}$ outside management teams who can carry out this task provided that they make a control investment in the company and set up the necessary operations. We henceforth refer to these potential investor-managers as private equity firms.[5]

In every period $t \geq 1$, each private equity *firm* decides whether or not to enter the market for corporate control for the duration of that period. To enter, the firm must raise and operate a fund which imposes a per-period cost (e.g., due to search activities, due diligence, negotiations, legal expenses). The cost is fixed but varies across private equity firms: $C_1 < C_2 < \cdots < C_{\mathcal{M}}$. For later use, we define a continuously increasing function $C(\cdot)$ with $C(i) = C_i$ for all $i \in \{1, 2, \ldots, \mathcal{M}\}$. This function reflects the talent distribution among private equity managers and is commonly known. To ensure interior equilibria, let $C(1) = 0$ and $C(\mathcal{M}) = \infty$.[6]

A private equity *fund* seeks to invest in companies. We assume that (human) resource or time constraints impose a limit on the number of investments that a fund can undertake simultaneously. To keep matters simple, we normalize this limit to one company per period. (Endogenous limits are discussed in Section 5.1.) In every period, each active fund is paired with a potential target (or portfolio) company. Once paired, they negotiate the price at which the fund can purchase (a control stake in) the company. Negotiations are modeled as Nash bargaining with $\omega \in (0, 1)$ denoting

4 The gamma distribution rules out negative value improvements and allows for a tractable Baysian analysis. The basic results should carry over to any stochastic setting with parameter uncertainty where agents update their beliefs about the mean of the underlying probability distribution.

5 Private equity funds often enforce changes in the governance of their portfolio firms (Gertner and Kaplan, 1996; Acharya and Kehoe, 2008; Cornelli and Karakas, 2008). Acharya and Kehoe (2008) report that one-third of CEOs in buyout targets are fired in the first 100 days.

6 The formulation of heterogeneity in terms of cost is not to be taken too literally. Similar results obtain when private equity firms instead differ in their ability to improve their portfolio companies. We choose the cost formulation because it makes the analysis more tractable.

the relative bargaining power of the fund. If a negotiation fails, the involved parties part and neither is paired again in the ongoing period. Otherwise, the fund purchases and reorganizes the company. A reorganized company harbors no further potential for improvement. Thus, there is attrition.

$M_t \leq \mathcal{M}$ and $N_t \leq \mathcal{N}$ respectively denote the number of private equity funds (operated) and potential target companies (available) *in* period t. For $M_t > N_t$, we adopt the convention that the most efficient funds are paired with a company first. Similarly, for $M_t < N_t$, we adopt the convention that those companies which have been in negotiations previously are paired with a fund first.

The timing of the model is as follows. In period 0, everyone in the economy learns about the occurrence of the shock but does not observe its magnitude, i.e. β. In each subsequent period $t \geq 1$, events unfold in the below order:

(1) Everyone enters the period with beliefs $\overline{V}_t = E_t(\overline{V})$.
(2) All private equity firms decide whether to raise a fund for the current period.
(3) Funds are paired with a target company and bargain over the purchase price.
(4) Funds that have successfully negotiated the price acquire their targets.
(5) Acquired companies are reorganized and their new value becomes public.
(6) Everyone updates their beliefs.

3. Competitive Markov Equilibrium

The key decisions in the model are the private equity firms' repeated decisions of whether or not to raise a fund. Let $a_t^i \in \{1, 0\}$ denote firm i's decision in period t, where $a_t^i = 1$ if the firm decides to raise a fund, and $a_t \equiv (a_t^1, \ldots, a_t^{\mathcal{M}})$. We assume competitive behavior and rational expectations. That is, each private equity firm ignores its own impact on *aggregate* variables but has unbiased expectations about (the evolution of) these variables.

In each period t, the history of all previous investment outcomes is commonly known. The history has a *direct* impact on the payoffs from t onward only through its impact on the state variables $\overline{V}_t$ and N_t. Given a state $(\overline{V}_t, N_t)$, firm i chooses a_t^i to maximize the sum of its discounted expected future per-period profits:

$$\Pi^i(a_t, \overline{V}_t, N_t) = \mathrm{E}_t \left[\sum\nolimits_{\tau=t}^{\infty} \delta^{\tau-t} \pi_\tau^i(a_\tau, \overline{V}_\tau, N_\tau) \,\middle|\, \overline{V}_t, N_t \right]$$

where $\pi_t^i(a_t, \overline{V}_t, N_t)$ is i's period-t profit, and $\delta \in [0, 1]$ is the discount factor.

Our analysis focuses on Markov strategies which depend on the history solely through the current state of the world (see Maskin and Tirole, 2001). In a Markov equilibrium, the optimal entry strategies and the equilibrium profits can therefore be written as $a_t^* = a_t(\overline{V}_t, N_t)$ and $\Pi^i(a_t^*, \overline{V}_t, N_t)$. Given optimal future behavior, this allows us to decompose $\Pi^i(a_t, \overline{V}_t, N_t)$ into the profit from the current period and a future "franchise" value:

$$\pi_t^i(a_t, \overline{V}_t, N_t) + \delta \mathrm{E}_t[\Pi^i(a_{t+1}^*, \overline{V}_{t+1}, N_{t+1}) \left| \overline{V}_t, N_t \right.].$$

Importantly, i's decision today affects the future only through its impact on the aggregate state variables $\overline{V}_{t+1}$ and N_{t+1}. Under competitive behavior, each private equity firm ignores this (intertemporal) impact. Consequently, the firm treats the entry decisions in different periods as *independent* options—behaving de facto *as if* it were myopic. Intuitively, the private equity firm perceives the impact of its current investment decision on future market conditions as so small that its sole decision criterion is the immediate payoff.[7]

The dynamics of the competitive Markov equilibrium are the focus of the subsequent analysis. The key driver of these dynamics is a feedback loop between entry decisions and market conditions: entry today depends on how market conditions have evolved, which in turn depends on past entry decisions. Accordingly, we first analyze the entry decisions for given market conditions, and then the market conditions for a given history of entry decisions.

3.1. Entry decisions. To determine entry in period t for a given state $(\overline{V}_t, N_t)$, we must first determine the outcome of the ensuing bargaining stage. Let P_t^i denote the purchase price that the fund (of firm) i and its potential target company bargain over. Furthermore, let O_t^i and O_t^c respectively denote the outside options (threat points) of the fund and the company. The Nash bargaining solution is given by

$$P_t^i = \arg\max(\overline{V}_t - P_t^i - O_t^i)^{\omega}(P_t - O_t^c)^{1-\omega}. \tag{3.1}$$

To derive the bargaining solution, we need to specify the outside options. For the quasi-myopic fund, the outside option is to save the amount—rather than to invest it in the company—for one period at the risk-free rate, which yields P_t/δ. Its current outside option is today's net present value of saving the amount, which is $O_t^i = \delta\,(P_t/\delta) - P_t = 0$. By contrast, the target company's outside option is the expected payoff from returning to the market in the hope of being acquired in the future. Suppose that a company which has been in negotiations previously is certainly paired with a fund in the next period. (This holds in equilibrium: a unilateral deviator would be the only

[7] The assumption of competitive behavior has two principal consequences: On the one hand, firms with negative expected current profits do not take into account the possibility of active *experimentation*. As a result, they become *adaptive* learners (Van Nieuwerburgh and Veldkamp, 2004; Veldkamp, 2004). On the other hand, firms with positive expected current profits neglect the possibility of *procrastinating* entry to learn more from information produced by others.

such company, and would hence be paired with fund 1 in the next period). As in the literature on search markets, a deviator's payoff from a future match is the payoff from a successful deal, i.e., the future "inside" option. Thus, the company's current outside option is $O_t^c = \delta E_t[P_{t+1}]$.

Given these outside options, the Nash bargaining solution is $P_t = (1-\omega)\overline{V}_t + \omega\delta E_t[P_{t+1}]$. To get a closed-form solution, we conjecture an equilibrium outcome in which the price is a *linear* function of $\overline{V}_t$ such that $P_t = \psi\overline{V}_t$. Since $E_t[\overline{V}_{t+1}] = \overline{V}_t$ (by the Law of Iterated Expectations), it then follows that $E_t[P_{t+1}] = E_t[\psi\overline{V}_{t+1}] = \psi\overline{V}_t = P_t$. Thus, *if* P_t is a linear function of $\overline{V}_t$, it is a martingale. Conversely, if P_t is a martingale, the Nash bargaining solution is indeed linear in $\overline{V}_t$: substituting $E_t[P_{t+1}] = P_t$ into the bargaining solution yields

$$P_t = \frac{1-\omega}{1-\omega\delta}\overline{V}_t. \tag{3.2}$$

Thus, $\psi = \frac{1-\omega}{1-\omega\delta)}$ is a rational equilibrium outcome. Consistent with intuition, a more patient firm (lower δ) bargains for a higher price ($\partial\psi/\partial\delta > 0$). Furthermore, since a failure to agree is inefficient, all negotiations lead to a transaction.

Having derived the bargaining solution, we now turn to the entry decision. A quasi-myopic private equity firm raises a fund (only) if the current expected profit from investing is positive. That is, the firm enters the market if $C_i \leq \overline{V}_t - P_t = (1-\psi)\overline{V}_t$ and is sure to be matched with a target company. Since this is true for all private equity firms, there exists a cut-off cost C_{i^*} such that all and only firms with $C_i \leq C_{i^*}$ raise a fund. In fact, i^* is equivalent to M_t, the total number of funds raised in t. It is defined by $C(i^*) = (1-\psi)\overline{V}_t$ as long as $i^* < N_t$; and by $i^* = N_t$ otherwise.

LEMMA 1. *There exists a competitive Markov equilibrium in which all and only private equity firms with $C_i \leq C(M_t) = \min\{(1-\psi)\overline{V}_t, C(N_t)\}$ enter the market with a fund in period t. The total number of funds M_t is increasing in $\overline{V}_t$ but decreasing in N_t, while the acquisition price P_t is increasing in $\overline{V}_t$.*

The equilibrium is intuitive: More skilled private equity firms are more inclined to enter so that, in every period, the relatively "best" firms raise a fund. Furthermore, more private equity funds are raised when the expected reorganization value is higher (or the funds have more bargaining power). The number of funds is also (weakly) increasing in the target stock N_t, i.e., the number of available target companies. Though the target stock only matters when it becomes a binding constraint ($N_t \leq M_t$). In Section 5.2, we discuss possible channels for market congestion, whereby the attrition in the target stock has a more continuous impact on fund activity.

3.2. Market conditions. Lemma 1 characterizes the equilibrium outcome for a given state process $\{\overline{V}_t, N_t\}$. We now turn to the determination of this process. The target stock N_t monotonically decreases as more and more investments are completed. More specifically, if M^t denotes the number of investments consummated prior to t, the target stock at the beginning of period t is $N_t = \mathcal{N} - M^t$.

Past investment also allows market participants to make inference about the true β, i.e., to learn about the magnitude of the shock. In this respect, the revenue generated by each reorganization represents a noisy signal about $\overline{V}$. We assume that reorganization revenues are observable to other market participants. This assumption is not to be taken literally, since private equity firms are in practice known to be secretive about their returns. Rather, it parsimoniously captures the notion that information about superior profitability leaks—at least informally—to other potential targets or to investors who are interested in starting their own private equity funds. The information spillover is central to the dynamics, as it creates a link between past performance and future market entry.

Let v_j denote the revenue generated by investment j. A history of investment outcomes is $\mathcal{H}_t = \{v_j\}_{j=1}^{M^t}$, and the historic average is $\overline{v}^t = \sum_{j=1}^{M^t} (v_j/M^t)$. Given a history $\mathcal{H}_t$, the posterior distribution of $\overline{V}$ is *inverse* gamma with shape and scale parameters $\tau + M^t\alpha$ and $\alpha\,(\gamma + M^t\overline{v}^t)$ respectively. (Details of the Bayesian updating process are provided in the Appendix.) In period t, the market's expectations about the reorganization value are equal to the mean of the inverse gamma distribution, $\overline{V}_t = E(V\,|\mathcal{H}_t)$, or more precisely

$$\overline{V}_t = \frac{\alpha\,(\gamma + M^t\overline{v}^t)}{\tau + M^t\alpha - 1}. \tag{3.3}$$

The conditional expectation (3.3) contains all distributional parameters except β, about which inference is being made. Recall that α is the known shape parameter of the V-distribution, whereas τ and $1/\gamma$ are the parameters of the distribution representing the market's initial (period-0) beliefs about the true β.

LEMMA 2. *$\overline{V}_t$ is ceteris paribus (i) increasing in $\overline{v}^t$, (ii) increasing in M^t if and only if $\overline{v}^t \geq \alpha\gamma/(\tau - 1)$, and (iii) increasing in α and γ but decreasing in τ.*

Current expectations increase with the historic average, because good past outcomes indicate that the reorganization value is high. In addition, if the historic average is high (low) relative to initial expectations, current expectations increase (decrease) in the number of past investments. The reason is that additional observations increase the precision of the estimate (in either direction). Finally, current expectations are higher when the initial expectations $\overline{V}_0 = E(\alpha/\beta\,|\mathcal{H}_0)$ were high, which explains why they are increasing in α and decreasing in $E(\beta) = \tau/\gamma$. In the subsequent analysis, we assume that $\overline{V}_0$ is strictly positive, though very small. This is meant to capture that, absent positive experiences, the market is sceptical about the prospects of reorganization.

4. Investment, Price, and Return Dynamics

We now study entire equilibrium paths to characterize the dynamics of aggregate fund activity, prices and returns. A conceptual difficulty is that, even for a given β, the economy evolves stochastically so that there is no *unique* equilibrium path. To describe "typical" properties of an equilibrium path, we characterize the path that is obtained when every reorganization yields the mean revenue $\overline{V}$. We refer to this path (somewhat incorrectly) as the "trend" path, and index it with o.

It is important to bear in mind that the agents in the model are unaware that the deviations from the mean are zero on the trend path. Hence, they update their beliefs as if the reorganization revenues were genuinely random. More precisely, since $\overline{v}^t = \overline{V}$ for all t, market expectations on the trend path evolve according to

$$\overline{V}_t^o = \frac{\alpha(\gamma + M^t\overline{V})}{\tau + M^t\alpha - 1}. \tag{4.1}$$

The expectations monotonically converge to $\overline{V}$ as M^t goes to infinity. The speed of convergence decreases for large absolute values of τ and γ (keeping their ratio constant). Accordingly, one may interpret a large value of $\tau = \bar{z}\gamma$ for constant $\bar{z}$ as a low "signal-to-noise" ratio.

4.1. Waves. In $t = 0$, the economy receives news about the occurrence of the shock and forms prior expectations about the expected reorganization value. For entry to occur, these expectations must exceed $C_1/(1-\psi)$ so that at least private equity firm 1 finds it worthwhile to raise a fund (Lemma 1). Since $C_1 = 0 < \overline{V}_0$, there is initial entry and consequently some learning that can serve as impetus for future entry.

4.1.1. *Learning and attrition.* Given entry, the evolution of fund activity (on the trend path) is determined by the true $\overline{V}$. If $\overline{V}$ is small, the initial reorganizations generate modest revenues, and investment activity remains low. Indeed, for $\overline{V} < \overline{V}_0$, the revenues disappoint the market and investment activity subsides. By contrast, if $\overline{V}$ is large, the market becomes increasingly optimistic because the investments are more profitable than expected. This attracts new funds, which in turn causes the target stock to decline faster. The two effects, *learning* and *attrition*, have countervailing consequences for future fund activity. When the number of funds reaches the number of remaining targets, investment climaxes and then collapses.

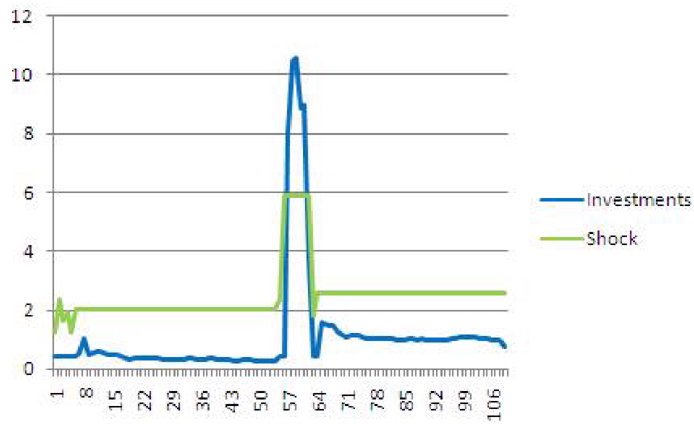

FIGURE A1. Long-run pattern.

The ultimate decline in investments is rather extreme on the trend path. Yet, it epitomizes the wave pattern inherent in any equilibrium path. Even on stochastic paths, investment booms endogenously transition to sudden busts.

PROPOSITION 1. *Expansions in fund activity follow a boom-bust pattern.*

In reality, productivity shocks occur more than once. In most cases, the shocks are probably small with little impact on overall activity. In a few cases, however, the shocks may be large, leading to a wave-like expansion in fund activity. While ex post observed, such waves are ex ante unpredictable. To illustrate such a long-run pattern, we simulate the equilibrium paths for a large number of shocks $\{\beta_k\}$ drawn from a gamma distribution with a high mean τ/γ (so that $\overline{V}_0$ is low). Figure A1 depicts a representative sequence of shocks with the fund activity that followed in their wake. As expected, lengthy periods with little fund activity are interrupted by a rare large wave. Thus, the model can plausibly produce patterns that are consistent with the documented cyclicality of private equity activity (Kaplan and Stein, 1993; Lerner, 2002; Acharya et al., 2007; Kaplan and Strömberg, 2008).[8]

4.1.2. *Inelastic supply.* The specific shape of a wave depends on the speed of entry, which in turn depends on the speed of learning and on the skill distribution $C(\cdot)$. On the one hand, when learning is slow (high $\tau = \bar{z}\gamma$), the market develops confidence

[8] For instance, venture capital activity expanded during the biotechnology boom in the early 1990s and during the information technology boom in the late 1990s. Similarly, buyout activity experienced high levels in the 1980s and in the mid-2000s.

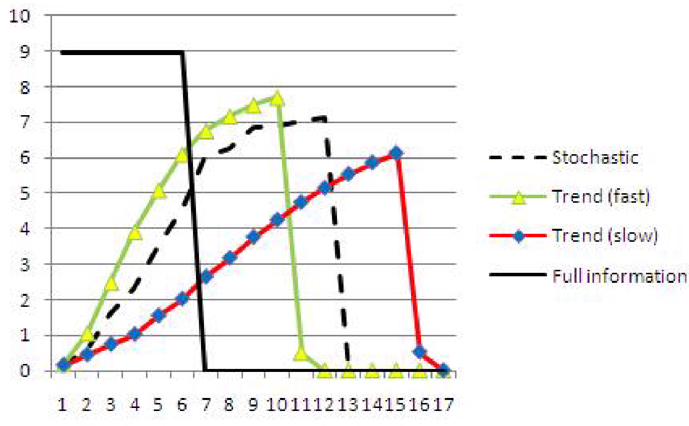

FIGURE A2. Different speeds of learning.

more slowly. On the other hand, when skill is scarce (high $C' > 0$ and low $C'' < 0$), private equity firms want to be more confident before they enter. When slow learning and skill scarcity are combined, fund activity incubates slowly, then suddenly booms, and crashes in the end. The boom occurs when market confidence reaches a level that attracts many entrants, which in turn accelerate learning and boosts confidence even further. The crash occurs because, once fund activity reaches its climax, the high rate of attrition rapidly diminishes the target stock. (The magnitude of the wave depends, of course, also on the true $\overline{V}$ and on the initial target stock $\mathcal{N}$.)

Figure A2 depicts four different equilibrium paths following a large shock ($\overline{V} \gg \overline{V}_0$). The plain solid line is the equilibrium path when $\overline{V}$ is immediately observed. The other two solid lines (marked with triangles and diamonds respectively) are trend paths that differ in the speed of learning. Finally, the dashed line depicts the stochastic path that corresponds to the trend path with faster learning. Comparing the different paths bears on the notion of inelastic supply and demand in the private equity market (Gompers and Lerner, 1999). Demand inelasticity is hard-wired into the model. Demand arises due to the exogenous productivity shock; as such, it does not respond to changes in supply. By contrast, supply inelasticity is endogenous. Supply responds slowly to changes in demand because private equity managers do not enter until they are confident enough. Accordingly, supply is less elastic when learning is slower or skill is scarcer.

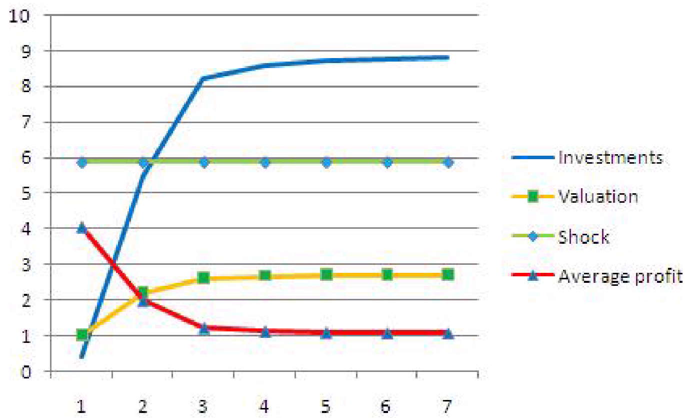

FIGURE A3. Entry, valuation and average profit.

4.2. Industry. We now describe in more detail how, at the industry level, (i) fund activity relates to valuation levels in the market, (ii) average fund performance evolves during a wave, and (iii) fund activity relates to past and future performance.

4.2.1. *Entry and valuation.* When the market grows more confident about the expected reorganization value, potential target companies increase in value. Thus, a rise in market confidence not only attracts more private equity funds to the market but also raises the price that these funds must pay to acquire companies (Lemma 1).

PROPOSITION 2. *Fund activity and valuation levels increase together.*

Proposition 2 is consistent with Kaplan and Stein (1993) who document that, during the buyout wave in the 1980s, buyout prices rose relative to fundamentals. Gompers and Lerner (2000) find similar results using a large data set comprising private equity investments in different stages and industries from 1987 to 1995. Specifically, they report that capital inflows into the private equity industry coincided with higher valuations of the funds' new investments. Both papers argue that the valuation increases were driven by fund competition rather than by improved investment prospects, suggesting that too much capital was chasing too few attractive investment opportunities.

The model can explain the observed pattern even in the absence of fund competition (which we introduce in Section 5.2). Higher entry and higher valuations are *jointly* caused by learning about the expected reorganization value. However, neither effect coincides with a concurrent or subsequent increase in the actual reorganization value.[9] The top three lines in Figure A3 illustrate these relationships for a trend path.

4.2.2. *Cross-sectional average performance.* In spite of learning, the market expectations $\overline{V}_t$ typically diverge from the true $\overline{V}$. When taking the model to empirical data, this distinction is crucial as observed fund revenues reflect the true $\overline{V}$—as opposed to the *subjective* expectations $\overline{V}_t$. Model predictions about fund performance therefore depend on $\overline{V}$. At the industry level, the *true* (data-generating) process that determines *average* per-period fund profits is $\overline{\pi}_t = \overline{v}_t - P_t - \overline{C}_t$, where $\overline{v}_t$ is the average gross return (reorganization revenue), and $\overline{C}_t = \sum_{i=1}^{M_t} (C_i/M_t)$ reflects average fund quality, in period t.

To see how average per-period fund profits evolve on the trend path, we simply need to set $\overline{v}_t = \overline{V}$, $P_t = P_t^o = \psi \overline{V}_t^o$, $M_t = M_t^o$, and $\overline{C}_t = \overline{C}_t^o = \sum_{i=1}^{M_t^o} (C_i/M_t^o)$. For $\overline{V} > \overline{V}_0$, which induces a wave, we know from (4.1) that market confidence, $\overline{V}_t^o$, monotonically increases over time. This causes prices, P_t^o, and fund activity, M_t^o, to monotonically increase (Lemma 1) but average fund quality, $1/\overline{C}_t^o$, to monotonically decrease (see Section 4.3.1). All the while, the *true* expected reorganization value, $\overline{V}$, remains *constant.* The rising prices and the declining quality thus imply that $\overline{\pi}_t^o$ decreases over time.

PROPOSITION 3. *In a wave, average per-period fund performance tends to decrease.*

The line marked with triangles in Figure A3 shows the evolution of average fund profits on a trend path. The decrease in average profits is steeper than the increase in prices because of the declining fund quality. It is noteworthy that Proposition 3 is not the result of increased fund competition. It merely requires learning and heterogeneity among private equity firms.

The decline in fund profitability across vintages appears to be at odds with the empirical finding that first-time funds underperform the industry (Kaplan and Schoar, 2005). However, this is not the case if the comparison between first-time and later-time funds is made in the cross-section, or if the comparison between first-time and later-time funds by the same private equity firm is based on the performance *relative* to the industry. Sections 4.3.2 and 4.3.3 elaborate on these points. Nevertheless, the model cannot explain *systematic* increases in the *absolute* performance of consecutive funds by the same private equity firm during a wave.

4.2.3. *Lagged entry-performance correlations.* Kaplan and Schoar (2005) find that capital flows *to* the private equity industry are positively correlated with last period's industry returns but negatively correlated with next period's fund returns. Note that

[9] If the shock to profitability is a shock to future cash flows, the increase in valuation levels corresponds to an increase in valuation multiples, such as the price-earnings ratio.

our model in general exhibits dynamics where industry growth goes together with a decline in fund profits, i.e., where high past performance precedes high(er) future entry and low(er) future performance.

To highlight such dynamics, let us consider a stochastic path for a shock that happens to coincide with the market's initial expectations, $\overline{V} = \overline{V}_0$. The dynamics on the stochastic path are driven by the exogenous random process $\{\overline{v}_t\}$, i.e., the random (average) per-period revenues. As market confidence in the next period, $\overline{V}_{t+1}$, positively depends on the average revenues in the current period, $\overline{v}_t$, the process $\{\overline{v}_t\}$ serves as a "leading" indicator. To see this, note that

$$\overline{v}^t = \frac{M^{t-1}}{M^t}\overline{v}^{t-1} + \frac{M_t}{M^t}\overline{v}_t \quad \text{and} \quad \overline{V}_{t+1} = \frac{\alpha\left(\gamma + M^t\overline{v}^t\right)}{\tau + M^t\alpha - 1}$$

and therefore

$$M_{t+1} = C^{-1}\left[\left(1-\psi\right)\overline{V}_{t+1}\right], \quad P_{t+1} = \psi\overline{V}_{t+1}, \quad \text{and} \quad \overline{C}_{t+1} = \sum_{i=1}^{M_{t+1}}\left(C_i/M_{t+1}\right)$$

are all increasing in $\overline{v}_t$. That is, via the historic average, high period-t revenues increase market confidence, fund activity, prices, and average costs in period $t+1$.

Now consider the preditive power of $\{\overline{v}_t\}$ with respect to fund i's per-period fund profits $\{\pi_t^i\}$. Given a history up to t, the mean of the true (data-generating) distribution of i's profit in $t+1$ is

$$E_{t+1}\left[\pi_{t+1}^i \middle| \overline{V}_{t+1} = \overline{V}_0\right] = \overline{V}_0 - \psi\frac{\alpha\left(\gamma + M^t\overline{v}^t\right)}{\tau + M^t\alpha - 1} - C_i,$$

which increases in the average period-t revenue $\overline{v}_t$ (via the historic average $\overline{v}^t$). Similarly, consider $\Delta\left(\overline{v}_t\right) \equiv E_{t+1}\left[\overline{\pi}_{t+1} \middle| \overline{V}_{t+1} = \overline{V}_0\right] - \overline{\pi}(t)$, which represents the expected "drop" in average fund profits from t to $t+1$ as a function of $\overline{v}_t$:

$$\Delta\left(\overline{v}_t\right) = \overline{V}_0 - \overline{v}_t + \left(P_t - P_{t+1}\right) + \left(\overline{C}_t - \overline{C}_{t+1}\right).$$

Clearly, $\Delta'\left(\overline{v}_t\right) = -1 - \left(\partial P_{t+1}/\partial\overline{v}_t\right) - \left(\partial\overline{C}_{t+1}/\partial\overline{v}_t\right) < 0$. On the one hand, a higher average revenue today both increases prices ($\partial P_{t+1}/\partial\overline{v}_t > 0$) and decreases average fund quality ($\partial\overline{C}_{t+1}/\partial\overline{v}_t > 0$) tomorrow. On the other hand, since average per-period revenues, $\{\overline{v}_t\}$, are independent draws from distributions with mean $\overline{V}_0$, any realization $\overline{v}_t > \overline{V}_0$ means that the market was "lucky" in t. In comparison, the revenues in $t+1$ are likely to be "corrected" downwards.

PROPOSITION 4. *High industry performance predicts high entry, which in turn predicts lower industry performance.*

Figure A4 illustrates the performance-entry patterns of a stochastic path. One may be tempted to view them as "bad timing" by private equity firms that choose

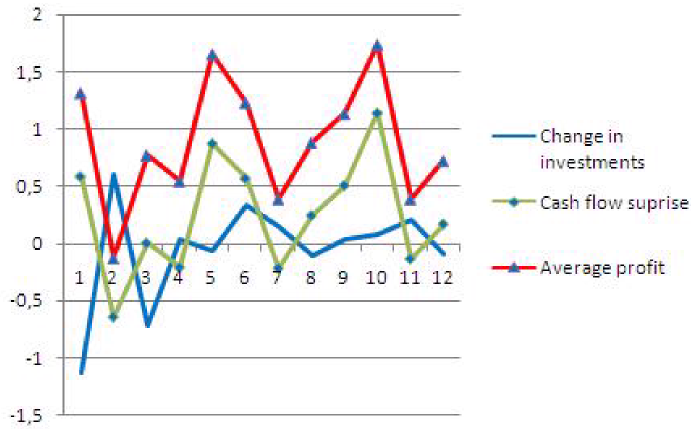

FIGURE A4. Lagged correlations.

to enter the market when profitability drops, while being absent when profitability is high. However, such patterns emerge naturally in a model with learning, where changes in perceived profitability and in actual profitability do not necessarily go in the same direction.

4.3. Funds. Given skill heterogeneity and entry timing, the model also generates both cross-sectional and time-series predictions about performance at the level of individual funds, to which we turn below.

4.3.1. *Persistent differences and last-in-first-out pattern.* While average profitability declines during a wave, performance differences among private equity firms are persistent. That is, a firm (or a particular fund) that has outperformed the industry likely continues to outperform the industry in subsequent periods. This follows directly from the assumed skill heterogeneity, and is consistent with the empirical evidence (Kaplan and Schoar, 2005).

A more interesting implication of the model is that a private equity firm's quality and its time of entry are related. By Lemma 1, all and only firms above a threshold quality level C_{i^*} enter the market, and this threshold level is increasing in the expected reorganization value $\overline{V}_t$. This implies that, if the market becomes more confident (higher $\overline{V}_t$), the funds raised by newly entering firms are of lower quality than the funds of "incumbent" firms. By the same token, if the market becomes less confident (lower $\overline{V}_t$), the firms that exit—i.e., do not raise a follow-up fund—are of lower quality than the firms that remain in the market. Thus, as $\overline{V}_t$ varies over time, entry and exit

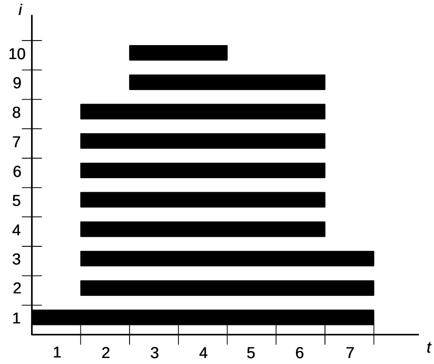

FIGURE A5. Last-in-first-out pattern.

follow a last-in-first-out pattern: the least talented are the latest to enter when market conditions improve, and the earliest to exit when the conditions deteriorate. Figure A5 illustrates this for the case of ten partnerships and a stochastic path that lasted for seven periods.

4.3.2. *First-time fund underperformance.* The last-in-first-out pattern endogenously creates a cross-sectional link between a fund's "age" and its performance relative to the industry. For example, first-time funds are run by less skilled managers than contemporaneous later-time funds.

PROPOSITION 5. *Funds with short track records tend to underperform the industry and are less likely to raise follow-on funds.*

Proposition 5 highlights that a positive relationship between the maturity of a private equity fund and its performance need not (solely) be driven by experience gains ("learning-by-doing"). Rather, it may reflect a causal relation between the fund managers' intrinsic abilities and their timing of entry and exit. Note further that many new funds are raised after highly profitable periods (Proposition 4), when valuation levels are high (Proposition 2), and during periods in which fund activity ex post turns out to have peaked (Proposition 1). Given the last-in-first-out pattern, these funds are run by the least qualified managers who are likely to exit the market soon after.

COROLLARY 1. *First-time funds that are raised in boom periods are unlikely to be suceeded by follow-on funds.*

Proposition 5 and Corollary 1 are both consistent with the evidence in Kaplan and Schoar (2005). Broadly speaking, the predicted last-in-first-out pattern says that many

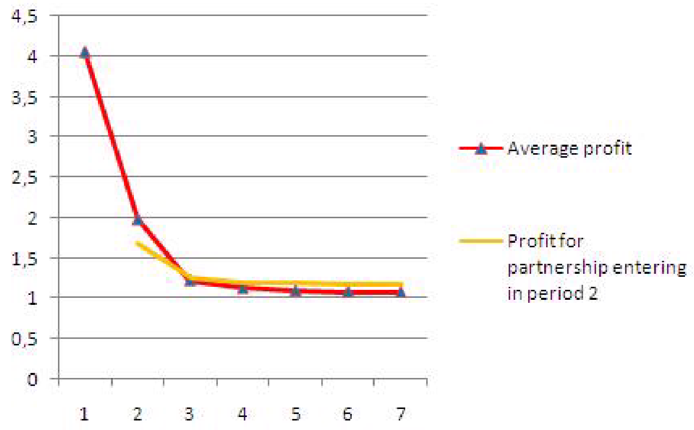

FIGURE A6. First-time fund underperformance.

"transient" private equity firms emerge during a wave, while the firms that are left at the end are those that have been around from the beginning.

4.3.3. *Relative improvement over time.* Proposition 5 is a potential explanation for why first-time funds (young private equity firms) underperform the industry in the cross-section. At the time a private equity firm i enters the market with its first fund, it belongs to the least skilled firms in the industry. However, if the boom continues, even less skilled firms enter in subsequent periods. As a result, the relative quality of firm i's follow-on funds improves over time.

PROPOSITION 6. *Consecutive funds tend to improve in relative performance.*

Figure A6 illustrates this result by comparing the average fund profit on a trend path with the profit of a private equity firm that enters the market in period 2. While its first fund performs below average, its follow-on funds outperform the average fund in the industry from period 3 onwards.

5. Fund Size and Fund Competition

5.1. Fund size. Kaplan and Schoar (2005) also study the relationship between fund size and fund profitability and report two distinct findings: the relationship is positive and concave across different funds, whereas it is negative across successive funds of the same private equity firm. Our baseline model is mute on this issue as it assumes a uniform and constant fund size. In this section, we extend the model to allow for variable fund size and show that the above relationships between size and profitability arise naturally.

For simplicity, suppose that $\mathcal{M} = 2$. Each private equity firm $i \in \mathcal{M}$ can now undertake as many investments as desired. However, we assume that a firm's per-period cost of operating a fund is increasing and convex in the number of considered investments. More specifically, let $C_{it}(M_{it}) = (M_{it} + C_i)^2$ where C_i is a constant that reflects the (inverse) quality of firm i, and M_{it} is the number of investments undertaken by firm i in period t.[10]

5.1.1. *Fund size and cross-sectional performance.* As long as $M_t \leq N_t$ is not a binding constraint, the number of investments chosen by private equity firm i in period t satisfies $C_{it}(M_{it}) = (1-\psi)\overline{V}_t$. This yields

$$M_{it} = \sqrt{(1-\psi)\overline{V}_t} - C_i.$$

Since $C_1 < C_2$, this immediately implies that the fund of firm 1 is larger than the fund of firm 2. That is, fund size increases with fund quality.

We measure a fund's profitability by its true expected profit *per investment*

$$\frac{M_{it}(\overline{V} - P_t) - C_{it}(M_{it})}{M_{it}} = \overline{V} - P_t - \frac{(1-\psi)\overline{V}_t}{\sqrt{(1-\psi)\overline{V}_t} - C_i}$$

which is decreasing in C_i. Thus, the larger fund earns a higher return per investment. The reason is that the average cost per investment is lower for the better fund, whereas the true expected revenue per investment $\overline{V} - P_t$ is the same for both funds. Rewriting the expected profit per investment as $\overline{V} - P_t - (1-\psi)\overline{V}_t/M_{it}$ and differentiating twice with respect to M_{it} furthermore shows that the relationship between fund size and fund profitability is concave.

PROPOSITION 7. *Within the cross-section of funds, performance is increasing and concave in fund size.*

This is consistent with the first of the two findings mentioned above. For given market expectations, the better private equity firm raises a larger fund. Fund size and fund profitability are jointly driven by the fund managers' quality, and hence positive correlated. This result relies on the heterogeneity among fund managers but does not exploit the dynamic properties of the model, to which we turn next.

5.1.2. *Fund size and time-series performance.* To examine how a fund size and fund profitability evolve during a wave, consider two arbitrary points in time, t'' and t', such

[10] The results also hold for $C_{it}(M_{it}) = M_{it}^2 + C_i$. In this case, a fund's marginal cost per investment is the same across all partnerships. By contrast, under the cost function in the text, a fund's marginal cost per investment decreases in the partnership's talent.

that $\overline{V}_{t''} > \overline{V}_{t'}$. From the above analysis, it follows that (as long as $M_t \leq N_t$ is not a binding constraint) a private equity firm i raises a larger fund in t'' than in t', i.e., $M_{it''} > M_{it'}$. Its true expected profit in t can be written as

$$\overline{V} - P_t - \frac{(M_{it} + C_i)^2}{M_{it}}.$$

Since $P_t = \psi \overline{V}_t$, we know that $P_{t''} > P_{t'}$. Furthermore, $(M_{it} + C_i)^2 / M_{it}$ is increasing in M_{it}. Taken together, this implies that the true expected revenue per investment $\overline{V} - P_t$ is lower in t'' (due to the higher prices), while the average cost per investment is higher in t'' (due to the larger fund size).

PROPOSITION 8. *Across consecutive funds of the same private equity firm, fund performance is decreasing in fund size.*

During a wave, market expectations tend to rise over time. Proposition 8 says that, as a result, private equity firms will raise larger but less profitable funds in the course of a wave. In fact, the decrease in profitability across consecutive funds will be proportional to the increase in size, consistent with the second finding by Kaplan and Schoar (2005).

5.2. Fund competition. One approach to modeling fund competition is to incorporate search frictions into the model. With search frictions, the more parties enter one side of the market, the more difficulty they have in finding alternative trading partners. As a result, bargaining power shifts to the other side of the market.[11]

Such "congestion" effects arising from fund competition tend to reinforce many of the conclusions of our model. To illustrate this, we split the bargaining game in stage 3 into three substages. In substage 3-1, each fund is paired with a company. As before, they bargain over the price at which the fund can acquire the company. Each pair that successfully negotiates the price moves immediately to stage 4. If a negotiation fails, the pair moves to substage 3-2, in which the fund tries to find another target company. The probability of finding a new target is given by the matching function $\phi(m, n)$, where m is the number of funds contemporaneously searching for a new target, and n is the number of available target companies. In substage 3-3, the fund bargains with a newfound target or, when the search fails, resumes negotiations with the previous one. In either case, successful negotiations lead to stage 4. A failure to agree moves the pair to the next period.

We make standard assumptions about the matching function: $\partial\phi/\partial m < 0$ and $\partial\phi/\partial n > 0$. For a fund, the probability of being matched with a (new) company is

[11] Several papers have used this approach to model venture capital markets and merger markets (Inderst and Mueller, 2001; Michelacci and Suarez, 2004; Rhodes-Kropf and Robinson, 2008).

lower when there are many other funds on the search, and higher when there are many available target companies. To simplify matters, we further assume that the target companies have all the bargaining power in substage 3-3. Let $\rho \in (0, \delta)$ denote the *intra*-period discount factor between substages 3-1 and 3-3.

We solve the bargaining game for an arbitrary fund i in period t by backward induction. In substage 3-3, any company negotiating with the fund offers the price $\overline{V}_t$, and the fund accepts the offer. In substage 3-1, the initial fund-company pair bargains under the conjecture that all contemporaneous negotiations are successful (which is true in equilibrium). Thus, the fund's and the company's outside options are given by $O_t^i = 0$ and $O_t^c = \rho\left[1 - \phi\left(1, N_t - M_t\right)\right]\overline{V}_t$. (The pool of alternative target companies excludes the $M_t - 1$ companies that are conjectured to successfully negotiate with the other funds and the current negotiation partner.) The Nash bargaining solution (3.1) is then given by $P_t = \psi\overline{V}_t$ where

$$\psi = (1 - \omega) + \omega\rho\left[1 - \phi\left(1, N_t - M_t\right)\right].$$

Given the properties of the matching function, the price is increasing in the number of funds and decreasing in the number of potential targets. In reduced form, we can therefore define the sharing rule as a function $\psi(N_t, M_t)$ where $\partial\psi/\partial N_t > 0$ and $\partial\psi/\partial M_t < 0$. Note that $\psi(N_t, M_t)$ measures the degree of fund competition. It is worth emphasizing that, along with N_t and M_t, the degree of competition endogenously varies over time. For example, by attracting more entry, an increase in market confidence, $\overline{V}_t$, will increase fund competition.

Thus, the key difference to the basic model is that the sharing rule ψ is not time-invariant but increases with entry and attrition. In a model with fund competition, prices therefore increase—and fund profitability deteriorates—faster as more funds enter the market and the target stock is depleted, capturing the idea that profits drop when "more money chases fewer deals." This slows down entry and precipitates exit so that the fund activity both builds up and declines more gradually than in the basic model. In other words, fund competition neither undermines the boom-bust pattern nor the last-in-first-out pattern of fund activity; it merely "smoothes" the wave.

6. Concluding Remarks

The paper presents a model of the private equity market in which heterogenous private equity firms learn about investment profitability from past outcomes and the stock of potential target companies is depletable. We derive the optimal entry and exit strategies of private equity firms as a function of their ability and market expectations. A characteristic feature of the model is that large expansions in private equity activity occur in waves with endogenous transitions from booms to busts. In addition, the model matches a wide range of stylized facts regarding the dynamics of investment, prices and performance *during* a wave.

Appendix

Derivation of $\overline{V}_t$. Let X be a gamma-distributed random variable with shape parameter α and scale parameter θ. It is convenient to define $\beta = \theta^{-1}$ as the inverse scale parameter. The expected value of X is equal to $\alpha\theta$ or equivalently $\alpha\beta^{-1}$.

In Bayesian probability theory, a class of prior probability distributions $p(\zeta)$ is said to be conjugate to a class of likelihood functions $p(x|\zeta)$ if the posterior distributions $p(\zeta|x)$ belong to the same family as the prior probability distributions.

The gamma distribution is a conjugate prior to itself whenever the likelihood function is a gamma distribution with known shape parameter α and unknown inverse scale parameter β. Suppose that we have a random sample $\{x_i\}_{i=1}^n$ of a gamma-distributed random variable X with known shape parameter α and unknown inverse scale parameter β. The likelihood function is then a gamma distribution with known shape parameter α and unknown inverse scale parameter β. If the prior probability distribution for β is a gamma distribution with known shape and inverse scale parameters τ and γ respectively, then the posterior distribution is a gamma distribution and has a shape parameter equal to $\tau + n\alpha$ and an inverse scale parameter equal to $\gamma + \sum_{i=1}^{n} X_i$.

In addition, if a random variable X is gamma distributed with shape parameter α and scale parameter θ then the random variable X^{-1} is inverse gamma distributed with shape parameter α and scale parameter $\theta^{-1} = \beta$. The expected value of the random variable X^{-1} is then $\beta/(\alpha - 1)$. Finally if the random variable X is inverse gamma distributed with shape parameter α and scale parameter $\theta^{-1} = \beta$ then the random variable cX, where $c \in R^+$, is inverse gamma distributed with shape parameter α and scale parameter $c\theta^{-1} = c\beta$.

In our particular case, this is all we need to derive the conditional expectation of the magnitude of the shock. Simply let $\alpha \to \alpha$, $\beta \to \beta$, $\tau \to \tau$, $\gamma \to \gamma$, $n \to M^t$, $X_i \to v_j$ and it immediately follows that

$$\overline{V}_t = E(V\,|\mathcal{H}_t) = E(\alpha\beta^{-1}\,|\mathcal{H}_t) = \frac{\alpha\gamma + \alpha\sum_{i=1}^{M^t} v_j}{\tau + M^t\alpha - 1},$$

where $\mathcal{H}_t = \{v_j\}_{j=1}^{M^t}$.

References

Acharya, Viral V., Julian Franks and Henri Servaes, 2007, "Private equity: Boom and bust?," *Journal of Applied Corporate Finance*, 19(4), 1-10.

Acharya, V. and C. Kehoe, 2008, "Corporate governance and value creation: Evidence from private equity," London Business School.

Berk, Jonathan and Richard C. Green, 2004, "Mutual fund flows and performance in rational markets," *Journal of Political Economy*, 112, 1269-1295.

Busse, Jeffrey A., Amit Goyal, Sunil Wahal, 2008, "Performance and persistence in institutional investment management," Emory University and Arizona State University.

Cornelli, Francesca and Oğhuzan Karakas, 2008, "Private equity and corporate governance: Do LBOs have more effective boards?" London Business School.

Gertner, Robert and Steven N. Kaplan, 1996, "The value maximizing board", University of Chicago.

Gompers, Paul and Josh Lerner, 1999, *The Venture Capital Cycle*, The MIT Press.

Gompers, Paul and Josh Lerner, 2000, "Money chasing deals? The impact of fund inflows on private equity valuations," *Journal of Financial Economics*, 55, 281-325.

Glode, Vincent, and Richard C. Green, 2008, "Information Spill-Overs and Performance Persistence in Private Equity Partnerships," Carnegie-Mellon University.

He, Ping, 2007, "A theory of IPO waves," *Review of Financial Studies*, 20(4), 983-1020.

Hochberg, Yael, Alexander Ljunqvist and Annette Vissing-Jørgensen, 2008, "Informational hold-up and performance persistence in venture capital," Northwestern University and New York University.

Inderst, Roman and Holger M. Müller, 2004, "The effect of capital market characteristics on the value of start-up firms," *Journal of Financial Economics*, 72, 319-356.

Jovanovic, Boyan and Peter L. Rousseau, 2002, "The Q-theory of mergers," *American Economic Review*, 92(2), 198-204.

Kaplan, Steven N. and Antoinette Schoar, 2005, "Private equity returns: Persistence and capital flows," *Journal of Finance*, 60, 1791-1823.

Kaplan, Steven N. and Jeremy C. Stein, 1993, "The evolution of buyout pricing and financial structure in the 1980s," *Quarterly Journal of Economics*, 108, 313-357.

Kaplan, Steven N. and Per Strömberg, 2009, "Leveraged buyouts and private equity," Journal of Economic Perspectives, 23(1), 121-146..

Lerner, Josh, 2002, "Boom and bust in the venture capital industry and the impact on innovation," Harvard University.

Malkiel, Burton G., 1995, "Returns from Investing in Equity Mutual Funds 1971-1991," *Journal of Finance*, 50(2), 549-572.

Michelacci, Claudio and Javier Suarez, 2004, "Business creation and the stock market," *Review of Economic Studies*, 71, 459-481.

Pastor, Lubos, Lucian Taylor and Pietro Veronesi, 2008, "Entrepreneurial learning, the IPO decision, and the post-IPO drop in firm profitability," *Review of Financial Studies*, forthcoming.

Pastor, Lubos and Pietro Veronesi, 2008, "Technological revolutions and stock prices," *American Economic Review*, forthcoming.

Persons, John and Vincent Warther, 1997, "Boom and bust patterns in the adoption of financial innovations," *Review of Financial Studies*, 10(4), 939-967.

Rhodes-Kropf, Matthew and David T. Robinson, 2008, "The market for mergers and the boundaries of the firm," *Journal of Finance*, 63(3), 1169-1211.

Shleifer, Andrei and Robert W. Vishny, 2002, "Stock market driven acquisitions," *Journal of Financial Economics*, 70, 295-311.

Timmermann, Allan, 1993, "How learning in financial markets generates excess volatility and predictability in stock prices," *Quarterly Journal of Economics*, 108, 1135-1145.

Timmermann, Allan, 1996, "Excess volatility and predictability of stock prices in autoregressive dividend models with learning," *Review of Economic Studies*, 63, 523-557.

Van Nieuwerburgh, Stijn and Laura Veldkamp, 2006, "Learning asymmetries in real business cycles," *Journal of Monetary Economics*, May 2006, 53(4), 753-772.

Veldkamp, Laura, 2005, "Slow boom, sudden crash," *Journal of Economic Theory*, 124(2), 230-257.

Veronesi, Pietro, 1999, "Stock market overreaction to bad news in good times: A rational expectations equilibrium model," *Review of Financial Studies*, 12, 975-1007.

PAPER 5

Reputable Friends as Watchdogs: Social Ties and Governance

with Petra Persson

Abstract. What happens when a monitor befriends the person to be monitored? We embed a delegated monitoring problem in a social structure: the supervisor and the agent are mutually altruistic, and the supervisor desires to be socially recognized as a person of integrity. Strengthening their friendship undermines the supervisor's monitoring incentives, but it also makes the agent more reluctant to take an action which harms the supervisor. The principal's preferences over friendship are therefore non-monotonic. When friendship is weak, the supervisor exerts a high monitoring effort to earn recognition. When friendship is strong, the supervisor need not monitor because the agent refrains from actions that put the supervisor in a bad light. The strength of friendship required for the latter outcome decreases in the supervisor's desire for recognition. This suggests that overlapping personal and professional ties can enhance delegated governance in cultures or contexts where social recognition is important. We discuss the implications of our analysis for crony capitalism, corporate governance, and organizational culture.

1. Introduction

Economic action is embedded into structures of social relations (Granovetter, 1985). Examples are ubiquitous and reach far beyond the realms of alumni networks, family firms, and close-knit corporate partnerships. Indeed, most of us mix business with pleasure as our professional and personal networks overlap. For the design of efficient governance mechanisms, it is therefore crucial to understand when social ties may be beneficial and when they may be harmful.

However, the desirability of mingling professional and personal relationships is not well understood; social ties between economic agents are often believed to undermine governance, yet outcomes vary considerably depending on the context.[1] For example,

Thanks to Patrick Bolton, Navin Kartik, and Bernard Salanie for helpful discussions.

[1] Social ties have been shown to promote cooperation, coordination and trade in various settings, including regional governments (Putnam, 1993), bank lending (Petersen and Rajan, 1994; Uzzi, 1999), and job search (Granovetter, 1974; Bian, 1997). Skeptics counter that, on the contrary, social ties may constrain a person's cooperation to her immediate network, thereby impairing adaptation and growth (Olson, 1982; Portes and Landolt, 1996). This view is supported by studies documenting that social ties can invoke favoritism in bank lending (La Porta et al., 2003; Charumilind et al., 2006),

many are critical towards the mélange of social ties and corporate control evident in family firms. Yet, family ownership is the dominant form of corporate ownership around the world, and the empirical evidence on the performance consequences of family ownership is inconclusive.[2] In particular, there are substantial differences across countries; in Thailand and Korea, for example, economic growth has been driven by the success of large family business groups (Claessens et al., 2000).

At a broader level, the ambiguous implications of social ties for governance can be illustrated by the apparent tension between—on the one hand—the recurrent critique of *crony capitalism*, self-serving friendships and family ties among elites pervading the economy, and—on the other hand—the persistently strong economic performance of several emerging economies where such structures are highly persistent. Critics refer to countries in which entrenched networks of friends in business and government have lead to expropriation, corruption, and consequently economic stagnation (Morck et al., 2005). Others argue that such social ties, or *family values*, can serve as informal institutions that facilitate economic transactions.[3]

The idea that social ties between agents can be both harmful and beneficial is not novel. Granovetter (1973) distinguishes between *strong ties* within a close-knit group and *weak ties* across different groups: "[W]eak ties [...] are [...] indispensable to individuals' opportunities and to their integration into communities; strong ties, breeding local cohesion, lead to overall fragmentation" (p.1378). That is, as paraphrased by many others, increasing the social connections inside a group can come at the expense of relations to outsiders (Portes and Sensenbrenner, 1993; Portes, 1998). Similarly, Putnam (1993) speaks of social capital as *bonding* (group members) or *bridging* (different groups).

However, from a governance perspective, the crucial question is what determines whether *a given social connection between economic agents is harmful or beneficial.* In essence, if family firms are more prevalent and perform better in Asia than in the Western world, if crony capitalism impedes growth in some countries but *guanxi*[4] is

---

discrimination and nepotism (Becker, 1971; Fershtman et al., 2005), and corruption (Callahan, 2005; Harris, 2007).

[2] See, for example, La Porta et al., 1999; Faccio and Lang, 2002; Villalonga and Amit, 2008; Franks et al., 2008; Cronqvist et al., 2003; Anderson and Reeb, 2003; Khanna and Palepu, 2000.

[3] Sebastian Mallaby in his book, *The World's Banker*, narrates an anecdote about a discussion between Wolfensohn, the then president of the World Bank, Suharto, Indonesia's former dictator and Zhu Rongji, Chinese vice Premier, in which Suharto asked Zhu: "Don't you think we should tell the president of the World Bank about corruption in this part of the world?" Then Suharto looked at Wolfensohn, and said: "You know, what you regard as corruption in your part of the world, we regard as family values."

[4] The Wikipedia entry retrieved on February 28, 2009, at http://en.wikipedia.org/wiki/Guanxi, begins with "Guanxi describes the basic dynamic in the complex nature of personalized networks of

reconcilable with China's exceptional economic performance, and if what is viewed as corruption in one context may be perceived as family values in another, it is imperative to understand *what makes the difference*.

Therefore, in this paper, we revisit a classic governance problem: delegated monitoring (Diamond, 1984; Holmstrom and Tirole, 1997). A principal (P, she) delegates a decision to an agent (A, he), and hires a supervisor (S, she) to monitor the agent and, if necessary, to intervene. To be able to assess whether friendship undermines governance, we model a social connection between the agent and the supervisor; they are mutual friends, who partly internalize each other's well-being, and the strength of their friendship can vary (ϕ). This setting is analyzed in different contexts. Specifically, we vary how important it is for the supervisor to be perceived by the principal as a person of high integrity, i.e., as having been faithful to her monitoring duty. That is, we vary the supervisor's care for social esteem (ρ). Formally, caring for social esteem means that the supervisor values others' assessment of her (interim) type, as in Benabou and Tirole (2006) and Ellingsen and Johannesson (2008).

As discussed in detail in Section 2.4, considerable empirical evidence suggests that desire for social esteem affects economic behavior. Intuitively, "[approval] makes us proud and happy while disapproval causes embarrassment and shame and makes us unhappy. These social rewards and punishments are a basic 'currency' that induces [us] to perform certain activities and avoid others" (Fehr and Falk, 2002, p.705). Further, the section also discusses experimental evidence suggesting that people's sensitivity to such approval incentives is heterogenous. For this reason, distinguishing between contexts in which pride and shame constitute strong incentives and contexts in which image concerns are not very important may be essential in analyzing governance mechanisms.

Our analysis generates three main results. First, (stronger) friendship between the supervisor and the agent has an ambiguous effect on the quality of the (ultimate) decision. Consistent with common sense, friendship erodes the supervisor's incentives to monitor and to intervene against her friend's, i.e. the agent's, interest. As a result, she becomes increasingly reluctant to perform her duty as monitor, so that fewer bad decisions are detected and corrected. We refer to this effect as *capture*. Capture is a bonding effect that forges an alliance between the two subordinates against the principal. However, because friendship is mutual, it also affects the agent's behavior. Crucially, a bad decision lowers the principal's (ex post) esteem for the supervisor. With friendship, the agent is averse to such a loss in esteem by the supervisor; making

influence and social relationships, and is a central concept in Chinese society [...]." Guanxi is discussed in detail in section 5.1.1.

him reluctant to misbehave. Intuitively, friendship creates *loyalty* on the part of the agent. Unlike capture, loyalty is a bridging effect that alleviates the conflict between all parties, including the agent and the principal.

Second, the principal always prefers—if feasible—a *full loyalty equilibrium* in which the supervisor never monitors and the agent, on his own accord, consistently acts in the principal's interest. The feasibility constraint stems from the fact that, for a given level of desire for esteem, this equilibrium requires a sufficiently high level of friendship, which may be difficult to obtain. In addition, we find that if social esteem is sufficiently important, the second most preferred equilibrium is a traditional *full monitoring equilibrium* in which the supervisor always monitors the agent. If so, the principal has a preference for extremes: she wants friendship either to be very high or to be very low. With low friendship, decision quality is ensured by the supervisor's own efforts to keep face. By contrast, high friendship ensures that the agent behaves well to prevent the supervisor from losing face. In the latter case, discipline obtains because the supervisor's desire for esteem, via friendship, acts as *social collateral* for the agent.

In both of these equilibria, it is the supervisor's reputability, i.e. her susceptibility to praise or shame, that makes her valuable as a bearer of responsibility.[5] In the full monitoring equilibrium, she perfectly oversees the agent's actions for fear of being held morally responsible for a bad outcome.[6] A motivation for instating delegated monitors may hence be to relay responsibility to persons that are responsive to image concerns, i.e., to combine moral responsibility and reputability. The significance of assigning moral responsibility is even more apparent in the full loyalty equilibrium. There, the supervisor's role does not involve any monitoring. It is rather the role of a designated scapegoat, standing in to bear the blame in the event of a bad outcome. Due to social ties, the supervisor's sense of responsibility proves contagious and is inherited by the agent. Thus, to be effective, delegated monitoring need not be active; it may suffice to appoint the right person to assume moral responsibility.

Third, the full loyalty equilibrium becomes more feasible when the supervisor's desire for esteem is high. Formally, for high(er) levels of desire for esteem, the equilibrium can be supported for low(er) levels of friendship. The intuition behind this result is

[5] Bartling and Fischbacher (2008, p.1), studying the incentives to delegate to shift blame, point to the following definition of moral responsibility (from the Stanford Encyclopedia of Philosphy): "To be morally responsible for something, say an action, is to be worthy of a particular kind of reaction—praise, blame, or something akin to these—for having performed it." They further write that "the question of moral responsibility is of economic relevance because praise and blame can constitute effective incentives."

[6] Whenever necessary for her to break even from an *ex ante* point of view, she is also compensated for her expected monitoring costs with a fixed wage.

that a stronger desire for esteem leverages the bridging effect of friendship. That is, friendship produces loyalty more readily in circumstances in which the fear of losing face is salient. Hence, *a given* social connection between an agent and a supervisor can have very different implications for the welfare of the principal, depending on the supervisor's concern for esteem. This yields an important empirical prediction: in order to understand whether a given social connection is likely to be desirable, it is crucial to consider the effects of friendship and desire for esteem *in interaction*.

For example, our model suggests that the inconclusive empirical evidence on the performance consequences of family ownership may reflect that whether family managers (A) supervised by family directors (S) run the firm in the best interest of its non-family investors (P) depends on whether the governance structure is placed in a cultural context where integrity and social esteem play an important role. If so, delegating management to family members may be a way of ensuring not only that the managers further the family's material interests, but also that they eschew actions that jeopardize the family's (brand) name. Thus, the loyalty created by family bonds may also benefit non-family shareholders. In such contexts, the intermingling of business and personal relationships can hence be a solution to, rather than a manifestation of, agency problems. Further, our results suggest that family firms should be more prevalent, and perform better, precisely in countries where such cultural traits are more salient (inter alia, in a number of Asian countries). On the contrary, our model suggests that family ownership may exacerbate agency problems in contexts where social esteem is not so important. These results confirm the view that a given social connection is a double-edged sword, generating loyalty as well as capture, while *also making precise what should determine which effect dominates.*

Another empirical implication is that when assessing the effect of social ties between a company's non-executive directors and its executive management, it may be crucial to interact the social connection variable with the reputational concerns of non-executive directors. In light of our model, appointments of directors who are both close friends of the CEO and highly reputable can be understood as attempts at implementing the full loyalty equilibrium: If possible, the shareholders (P) prefer to appoint a board (S) which has close social ties to the management (A), and which cares considerably about its public esteem. This would implement a full loyalty equilibrium such that the firm is well-managed although the board is passive.[7] Otherwise, the shareholders might prefer a board which has no ties to, and hence actively monitors, the management. The model thus endorses the view that independent directors are more active monitors,

[7] One possible example is the appointment of Bill Gates as a director of Berkshire Hathaway; Bill Gates is both highly reputable and a close friend of the CEO, Warren Buffett.

but challenges the view that they are necessarily the best advocates of shareholder interests.[8]

Furthermore, our model suggests that in organizational design, it may be important to consider how social ties and desire for esteem can be promoted, and how such social capital may affect incentives and decisions. For example, public performance rewards and transparency may stimulate concern for esteem, and vertical social ties (promoted through, e.g., joint get-togethers and team-work activities) may cause this concern to trickle down within an organization. Hence, organizational culture can be thought of as a way to endogenously influence the extent of social connections as well as the strength of approval incentives in order to achieve a desired governance structure.

This paper is closely related to the literature on delegated monitoring (Diamond, 1984; Holmstrom and Tirole, 1997; Tirole, 1986). We extend the standard analysis of this two-layered moral hazard problem by embedding the formal relationships in a social structure that comprises two informal relationships: the supervisor and the agent are mutually altruistic, and the supervisor cares about the principal's social esteem for her. The effects of altruism and social esteem in interaction have been studied in different settings (Benabou and Tirole, 2006; Ellingsen and Johannesson, 2008; Fehr et al., 2008). Our contribution is to show how the interaction can bear on the efficacy, and even the nature, of delegated monitoring. More broadly, since we model economic behavior embedded in social structure, this paper relates to literature in sociology and economics (e.g., Coleman, 1990; Granovetter, 1973, 1985; Putnam, 1993). Moreover, in Section 5, we discuss various applications of our model to, and relate to the literature on, family firms, crony capitalism, corporate governance, and organizational culture.

The remainder of the paper is organized as follows. Section 2 presents the basic model and discusses the central assumptions; social esteem and directed altruism. Section 3 analyzes the equilibrium behavior of the supervisor and the agent. Section 4 explores the principal's preferences over friendship and derives the paper's main results. Section 5 discusses various applications. Section 6 concludes the paper. All mathematical proofs are relegated to the Appendix.

2. Governance and Social Structure

2.1. Agency and delegated monitoring. The principal (P) can hire an agent to implement an action on her behalf. There are two possible actions, g and b. The

[8] The theoretical analyses of Adams and Ferreira (2007) and Harris and Raviv (2008) reach similar conclusions, albeit for different reasons. In their models, management-friendly directors can elicit more information from corporate insiders. This induces a trade-off between monitoring and information aggregation. In our model, close ties may render monitoring obsolete, rather than infeasible.

principal cannot distinguish g from b, but the agent has this ability. Each action generates benefits for the principal, π_P, and for the agent, π_A, given by

$$\begin{aligned}\pi_P(g) &= X_g > 0, \quad \pi_P(b) = X_b = 0 \\ \pi_A(g) &= B_g > 0, \quad \pi_A(b) = \tilde{B}_b > B_g,\end{aligned}$$

where $\tilde{B}_b$ is a random variable with p.d.f.

$$f(\tilde{B}_b) = \begin{cases} \alpha & \text{for } \tilde{B}_b = B_b^L \\ 1-\alpha & \text{for } \tilde{B}_b = B_b^H > B_b^L \end{cases}.$$

The agent privately observes the realization of $\tilde{B}_b$ prior to making his action choice, which the principal cannot observe. The (commonly known) payoff structure induces a conflict of interest over the choice between g and b: the agent strictly prefers b, which provides her with additional perks at the expense of the principal's benefits, over g.

To mitigate this agency problem, the principal can hire a supervisor (S). The supervisor cannot freely observe whether the agent's proposed action is of type b or g. However, she can engage in monitoring, which reveals the true action type with probability 1, at a random cost $\tilde{c}$, with p.d.f.

$$g(\tilde{c}) = \begin{cases} \beta & \text{for } \tilde{c} = \underline{c} \\ 1-\beta & \text{for } \tilde{c} = \overline{c} > \underline{c} \end{cases}.$$

If the supervisor monitors ($e = 1$) and discovers that the agent proposed b, the supervisor can convert b to an action of type g at no cost. If she does not monitor ($e = 0$), the agent's action choice is implemented.

The supervisor privately observes the realization of $\tilde{c}$ prior to making her monitoring choice, which the principal cannot observe. For simplicity, we assume that the random variables $\tilde{c}$ and $\tilde{B}_b$ are independent. The principal may have to pay the supervisor a fixed wage $w \geq 0$.

To isolate the problem of interest, we further assume that the agent and the supervisor have no wealth (limited liability), and that the actual outcome X is observable to the three players but not verifiable by a third party (incomplete contracts). Thus, without further constraints, the principal can neither induce the agent to choose g, nor the supervisor to monitor. For now, we also abstract from the possibility of communication, including side transfers, between the agent and the supervisor. All players are risk-neutral.

Finally, we assume that $X_g > B_b^H + \overline{c}$. This implies that, *in terms of material payoffs*, the efficient choice for the agent is to implement g (and for the supervisor not to monitor); and that the efficient choice for the supervisor is to monitor if (and only if) the agent chooses b. This assumption is standard. Due to the players' social preferences, introduced below, utilitarian statements are more ambiguous. Though, for reasonable parameters, it remains true that choosing g and monitoring, if needed, maximize aggregate welfare. For our results, specifying the exact parameter region where this holds is not crucial.[9]

2.2. Social esteem and directed altruism. In addition to the *formal* relationships outlined above, there are two *informal bilateral* relationships.

First, the supervisor cares about what the principal thinks about her. More precisely, she derives (more) utility from being perceived by the principal as a person of high(er) integrity.[10] We suppose that the supervisor *ex ante* promises to monitor the agent's project proposal. *Ex post*, the principal evaluates the integrity of the supervisor by forming a (Bayesian) belief about the probability that the supervisor kept her word and *de facto* monitored the project, which is given by

$$\Pr(e = 1 \,|X\,) = \begin{cases} \theta_g & \text{if } X = X_g \\ \theta_b & \text{if } X = X_b \end{cases} .$$

Let this integrity assessment denote the principal's (ultimate) social esteem for the supervisor. In the spirit of Ellingsen and Johannesson (2008), to capture the notion that the supervisor values social esteem, let the supervisor's payoff function π_S be given by

$$\pi_S = \begin{cases} \rho\theta_g + w - ec & \text{if } X = X_g \\ -\rho\,(1-\theta_b) + w - ec & \text{if } X = X_b \end{cases} .$$

Here, ρ reflects the strength of the supervisor's desire for esteem. Intuitively, if the good outcome, X_g, is realized, the principal may give some credit for this to the supervisor, which is captured by θ_g. We refer to $\rho\theta_g$ as the *pride* that supervisor experiences from getting this credit. Conversely, if the bad outcome, X_b, is realized, the principal may place some blame for this on the supervisor, which is captured by $(1-\theta_b)$. We refer to $-\rho\,(1-\theta_b)$ as the supervisor's *shame* from bearing this blame.[11]

Second, we assume that the agent and the supervisor can exhibit directed altruism towards each other (but not vis-à-vis the principal). Formally, recalling that π_A and π_S denote the agent's and the supervisor's (non-altruistic) payoff functions, their utility functions are given by

$$u_A = \pi_A + \phi_{AS}\pi_S \quad \text{and } u_S = \pi_S + \phi_{SA}\pi_A$$

---

[9] In general, it suffices to assume that X_g is sufficiently large relative to $B_b^H + \overline{c}$.

[10] For our results, it does not matter whether the value of social esteem is purely affective (being a purely hedonic good) or instrumental (conferring deferred material consequences).

[11] The agent does not care about the principal's esteem for him. This is merely a simplification. If the agent cares enough about what the principal thinks about him, this will mitigate the primary agency problem. However, to the extent that it does not fully resolve this problem, the need for a supervisor remains and the current analysis characterizes the residual interaction. There is also no role for social esteem between the supervisor and the agent.

For simplicity, let $\phi_{SA} = \phi_{AS} = \phi \in [0, 1]$, where ϕ reflects the intensity of the personal relationship, or *friendship*, between the agent and the supervisor.

The parameters c, ρ and ϕ are commonly known and, for now, exogenously given.

2.3. Sequence of events. The basic analysis takes the presence of the supervisor as given. In stage 0, the supervisor promises to monitor and to convert a detected type b action. In stage 1, the agent privately observes the realization of $\tilde{B}_b$, and the supervisor privately observes the realization of $\tilde{c}$. In stage 2, the agent makes an action proposal to the supervisor, and the supervisor decides whether or not to monitor. If she monitors and detects a type b action, the supervisor chooses whether or not to convert it. In stage 3, the action is implemented, the principal's benefits and the agent's perks are publicly observed, and the principal updates her beliefs about the supervisor's integrity. In stage 4, all utilities are realized, and the game ends.

2.4. Motivating the social structure. Adam Smith's motivation in writing *The Theory of Moral Sentiments* (1790) was to describe the moral bonds that maintain social harmony. Smith believed that an agreeable social order emerges primarily because every individual has a natural inclination to strive for the approval of his or her peers. For example, he writes[12]

> "Nature, when she formed man for society, endowed him with an original desire to please, and an original aversion to offend his brethren. She taught him to feel pleasure in their favourable, and pain in their unfavourable regard. She rendered their approbation most flattering and most agreeable to him for its own sake; and their disapprobation most mortifying and most offensive."

Social (dis)approval can occur when a person's observed behavior conforms to, or violates, a social norm. In the above model, the relevant social norm is integrity. Showing integrity invokes admiration, whereas lack of integrity invokes criticism. These social responses in turn invoke emotions that affect the person's felicity, and thereby her behavior. That is, "[approval] makes us proud and happy while disapproval causes embarrassment and shame and makes us unhappy. These social rewards and punishments are a basic currency that induces [us] to perform certain activities and avoid others" (Fehr and Falk, 2002, p.705).[13]

[12] Ellingsen and Johannesson (2008) use this as their opening quote. Fehr and Falk (2002) use a similar quote by Smith: "We are pleased to think that we have rendered ourselves the natural objects of approbation,...mortified to relect that we have justly merited the blame of those we live with."

[13] They cite additional references that endorse the idea that social esteem matters for behavior, including the following quote by Harsanyi (1969): "People's behavior can largely be explained in terms of two dominant interests: economic gain and social acceptance."

Gächter and Fehr (1999), Masclet et al. (2001), and Rege and Telle (2004) present experimental evidence that desire for social esteem affects economic behavior. Studying public goods experiments that are plagued by free-riding problems, their basic finding is that the revelation of players' identities and their individual contributions, or the possibility of expressing disapproval, increases average contributions. Their results further suggest that people's sensitivity to approval incentives is heterogeneous, increases with familiarity, and increases with (the strength of) the association to social norms.[14] Studying a natural experiment, Funk (2005) finds that the introduction of mail voting failed to increase voter turnout in Switzerland. Her explanation that voting by mail dilutes the social incentives of "showing up" to cast a vote is supported by her finding that turnout *decreased* in small communes. In a labor context, Mas and Moretti (2008) find that a worker's effort is positively related to the productivity of workers who face her (but not of those whom she faces but who do not face her), suggesting the presence of social pressure.

There is also a plethora of casual evidence. Benabou and Tirole (2006) name examples of "public displays and private mementos conveying honor and shame" intended to influence behavior: "Nations award medals and honorific titles, charitable organizations send donors pictures of 'their' sponsored child, nonprofit organizations give bumper stickers and T-shirts with logos, and universities award honorary 'degrees' to scholars. Conversely, the ancient practice of the pillory has been updated in the form of televised arrests, posting on the Internet the names of parents who are delinquent on child support and those of sexual offenders, and publishing in local newspapers the license plate numbers of cars photographed in areas known for drug trafficking or prostitution" (p.1663).[15] That the media can serve to pillory wrongdoers is also documented by Dyck and Zingales (2002) and Dyck et al. (2008) in the context of corporate governance. Specifically, they study the activities of a hedge fund that successfully used shaming in the press to force Russian companies to reverse decisions that harmed the (minority) shareholders.

A recent example of social disapproval is the public outrage over bonuses paid to bank managers during the ongoing banking crisis. Paying those bonuses struck a moral nerve. Bank of America chairman and CEO Ken Lewis urged those in the banking industry to find some *humility* (*Telegraph*, 9 February 2009). Invoking a sense of civic

[14] More experimental and field evidence, in particular from the "social facilitation" literature in social psychology, is cited in Ellingsen and Johannesson (2008, fn.3).

[15] Gächter and Fehr (1999) and Rege and Telle (2004) also discuss circumstancial evidence, including (i) anti-littering norms in cities, (ii) norms against absenteeism or tardiness in factories, (iii) norms among union members during the British miners' strike in 1984, and (iv) British government ads using social pressure to get young men to subscribe to the army during World War II.

duty, Goldman Sachs CEO Lloyd Blankfein said, "Many people believe—and, in many cases, justifiably so—that Wall Street lost sight of its larger *public obligations*." In a similar spirit, Morgan Stanley CEO John Mack sympathized: "I know the American people are *outraged* about some compensation practices on Wall Street. I can understand why." Partly in response, top executives in several banks have voluntarily forgone (millions worth of) bonuses.[16]

Our modeling of social esteem closely follows Benabou and Tirole (2006) and Ellingsen and Johannesson (2008), where esteem for a player is a posterior *belief* about the unobserved type of that player, and where the player's utility from esteem depends linearly on these beliefs. However, we depart from their framework in that the *shape* (or more precisely, the intercept) of the supervisor's utility function is *conditional on the material outcome*. In particular, esteem takes the form of either praise or blame, and the player hence experiences either pride or shame, depending on whether the outcome for the principal is good or bad (Figure 1). That is, success breeds glory and heroes, whereas failure breeds accusations and scapegoats.[17] While intuitive, the assumption is not crucial.

An important implication of our specification is that the supervisor's utility from social esteem can be negative; that is, shame does not merely confer a lower utility but a *dis*utility. In hindsight, a person may therefore wish that she had refused the task that caused her to suffer shame; like Freddie Mac's CEO Richard F. Syron who, after leading the mortgage company to the brink of collapse, said: "If I had perfect foresight, I would never have taken this job in the first place" (*New York Times*, 5 August 2008). Following the Enron scandal, Patrick McGurn, vice president at Institutional Shareholder Services, commented in a similar vein, "The directors of Enron are going to carry this stigma with them" (*New York Times*, 16 December 2008).[18]

Altruism is a form of *un*conditional kindness. It motivates a person to sacrifice own resources to improve the well-being of others because of empathy; i.e., neither as a response to, nor in anticipation of, a favor. The experimental evidence (primarily from dictatorship games) on altruism is well-established (Fehr and Schmidt, 2006,

[16] The quotes from John Mack and Lloyd Blankfein are taken from *Reuters Blogs*, 11 February 2009. The emphases are added.

[17] A BBC article with the title "The blame game starts at Davos" from January 29, 2009, reports widespread criticism that blames the US economic model for the latest financial crisis. Not so long ago, the model was credited for the spectacular growth of the US economy, and served as the blueprint for many policy initiatives by the World Bank and the IMF. Similarly, the Asian model was long admired as an engine of growth before it was widely accused of having led to the Asian crisis in the 1990s.

[18] Relatedly, Neilson (2008) and Dillenberger and Sadowski (2008) present models in which "shame" may not only prevent selfish behavior but may also induce reluctance to enter a game in which shame might be inflicted. A central aim of both papers is to explain why some people prefer not having to make a choice in a dictator game (Dana et al., 2006).

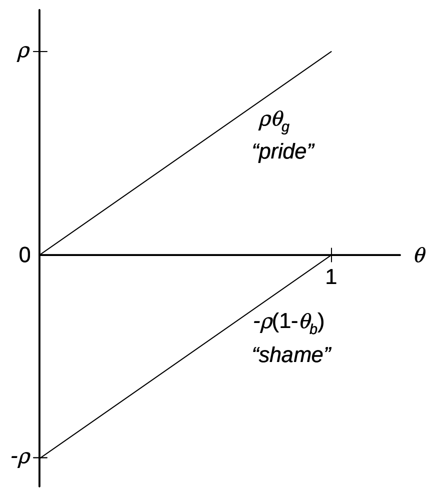

FIGURE A1. Praise/pride and blame/shame.

and the references therein). Recent field experiments in real-world social networks, conducted by Leider et al. (2008a,b), document that altruism can be decomposed into *baseline* altruism (vis-à-vis strangers) and *directed* altruism (that favors friends over strangers).[19] This confirms Adam Smith's (1790, p.23) view: "We expect less sympathy from a common acquaintance than from a friend. We expect still less sympathy from an assembly of strangers." Similar conclusions are drawn by neuroeconomists who use brain imaging techniques in behavioral experiments to trace empathic reactions. Studying couples in love, Singer et al. (2004a,b) find that painful stimulation applied to one subject induces pain-related brain activity in his or her partner. In another experiment, they also document that the neural response is stronger, the more the subject "likes" the object of empathy (Singer et al., 2006). It is therefore natural to assume some form of directed altruism when modeling interactions among friends (Karlan et al., 2008) or family members (Becker, 1974; Alger and Weibull, 2009).

Since both desire for social esteem and directed altruism cause agents to take into account the consequences of their behavior on others, we should point out some principal differences. If a person is altruistic towards another person, she partly, though genuinely, internalizes the material well-being of the latter. If she cares about social

[19] They find that subjects increase giving to friends by 52 percent relative to random strangers, suggesting that directed altruism is a significant component of prosocial behavior. Apart from altruism, they also find that the prospect of future interaction (or reciprocity concerns) further increases prosocial behavior. Similarly, Freeman (1997) finds that people volunteer more help when the request comes from a friend, a colleague, or family.

esteem, she is merely interested in what the other person *thinks* about her. Unlike in the case of altruism, it therefore matters crucially whether her behavior is observable or not. Also, social esteem—especially, public esteem—need not always be mutual. For instance, Barack Obama may be very concerned about how his decisions are received by voters, but few of the voters worry about the approval of Barack Obama when making their own decisions. Finally, an agent's social esteem for another fluctuates with her beliefs about the latter's type or behavior. In comparison, friendship is a more stable property of relationships. The effects of altruism and social esteem in interaction have been studied in different settings (Benabou and Tirole, 2006; Ellingsen and Johannesson, 2008; Fehr et al., 2008). We show how the interaction can bear on the efficacy, and even the *nature*, of delegated monitoring.

3. Does Friendship Undermine Governance?

3.1. Strategies and equilibrium concept. A pure strategy of the agent is a function specifying an action choice for each realization of $\tilde{B}_b$. Three pure strategies are relevant: never choose g, choose g when the benefits from b are low, and always choose g.[20] Each strategy is characterized by the probability that the good action is chosen, $p \in \{0, \alpha, 1\}$.

A pure strategy of the supervisor is a function specifying a monitoring decision e for each realization of $\widetilde{c}$. Three pure strategies are relevant: always monitor, monitor when the cost is low, and never monitor.[21] Each strategy is characterized by the probability of monitoring, $m \in \{1, \beta, 0\}$. The supervisor's decision whether or not to correct the agent's choice is subsumed in her monitoring decision. It is simple to show that the supervisor does not monitor unless she prefers g over b.

The principal's beliefs about the probability that the supervisor monitored the agent's action choice are a function of her observed outcome X, and given by

$$\theta(X) = \frac{\hat{m}(X)}{\hat{m}(X) + \hat{p}(X)(1 - \hat{m}(X))} \equiv \begin{cases} \theta_g & \text{if } X = X_g \\ \theta_b & \text{if } X = X_b \end{cases},$$

where $\widehat{\sigma}$ denotes a conjecture about a variable σ.

The game is solved for pure strategy Perfect Bayesian Equilibria (PBE). A PBE (henceforth equilibrium) consists of strategies p and m, and Bayesian beliefs θ_g and θ_b, wherever defined, such that the strategies are optimal given the beliefs, and the beliefs

[20] Clearly, if the agent chooses b when $\widetilde{B}_b = B_b^L$, he will never choose g when the incentives for choosing b are stronger, i.e., when $\widetilde{B}_b = B_b^H$.

[21] Clearly, if the supervisor chooses to monitor when $\widetilde{c} = \bar{c}$, he will never choose not to monitor when the cost of monitoring is lower, i.e., when $\tilde{c} = \underline{c}$.

are consistent given the strategies. Wherever applicable, we use the intuitive criterion to refine off-equilibrium beliefs.

3.2. Existence and multiplicity. As the agent and the supervisor have three pure strategies each, nine strategy profiles (p, m) can potentially be supported in equilibrium. However, if the agent always proposes the good action $(p = 1)$, the supervisor will never incur any monitoring cost in equilibrium. Thus, the profiles $(p, m) = (1, 1)$ and $(p, m) = (1, \beta)$ can be ruled out, and seven candidates remain.

LEMMA 1. *For each strategy profile* (p, m) *such that* $(p, m) \notin \{(1, 1), (1, \beta)\}$*, there exist pairs* (ρ, ϕ) *for which the profile can be supported in equilibrium.*

Each of the seven equilibria inhabits a certain region in the ρ-ϕ-space, i.e., it can be supported for certain combinations of desire for esteem and friendship. Some equilibrium regions overlap, so that their intersections admit multiple strategy profiles as equilibria. Nevertheless, their locations in the ρ-ϕ-space broadly determine how ρ and ϕ affect the agent's and the supervisor's behavior.

3.3. Capture vs. loyalty. We focus on the comparative statics of ϕ for given levels of ρ, reflecting our primary interest in the impact of friendship between the supervisor and the agent on equilibrium behavior (for given levels of the supervisor's desire for esteem).

Recall that the supervisor's strategies are $m \in \{1, \beta, 0\}$. For any (ϕ, ρ), let $M^* (\phi, \rho)$ be the set of S' strategies that can be supported in equilibrium.

LEMMA 2 (Capture). *Both* $\min M^* (\phi, \rho)$ *and* $\max M^* (\phi, \rho)$ *are decreasing in* ϕ.

Lemma 2 describes how friendship affects the supervisor: equilibria at higher levels of friendship tend to involve lower levels of monitoring. The reason is that, as friendship increases, the supervisor internalizes more of the agent's well-being, which effectively increases her cost of intervening against the agent's interest. As a result, she becomes increasingly reluctant to perform her duty as monitor. We refer to this effect of friendship as *capture*. By means of capture, friendship creates a bond between the agent and the supervisor, forging an alliance against the principal; it affiliates the supervisor with the interests of the agent, which are in direct conflict with those of the principal.

Recall that the agent's strategies are $p \in \{0, \alpha, 1\}$. For any (ϕ, ρ), let $P^* (\phi, \rho)$ be the set of A's strategies that can be supported in equilibrium.

LEMMA 3 (Loyalty). *Both* $\min P^* (\phi, \rho)$ *and* $\max P^* (\phi, \rho)$ *are increasing in* ϕ.

Lemma 3 describes how friendship affects the agent: in equilibria at higher levels of friendship, he tends to voluntarily choose g more often. As friendship increases, the agent internalizes more of the supervisor's well-being and thus experiences more disutility when the supervisor feels shame in case of a bad outcome. Therefore, he becomes increasingly reluctant to put the supervisor's reputation at risk by choosing b. We refer to this effect of friendship as *loyalty*. By creating loyalty, friendship bridges the conflict between the principal and the agent; it affiliates the agent with the interests of the supervisor and hence, in turn, with those of the principal.

4. Loyalty, Social Esteem, and Social Collateral

4.1. Preferences over friendship. The key questions in this paper concern the principal's preferences over friendship between the agent and the supervisor, when the supervisor cares for social esteem. Since different equilibria can be supported for different combinations of friendship and desire for esteem, the principal's preferences over the equilibria translate into preferences over friendship, for given levels of desire for esteem. The principal's *ex ante* expected utility in equilibrium (p, m)[22] is given by

$$E(u_P) = X_g(m + (1 - m)p) + X_b(1 - m)(1 - p) - w_{(p,m)},$$

where $w_{(p,m)} \geq 0$, and $w_{(p,m)} > 0$ whenever the supervisor's *ex ante* expected utility would be negative if $w_{(p,m)} = 0$. Hence, the principal must take into account that, in some equilibria, she may have to pay the supervisor a wage.[23]

The following auxiliary result is important for the subsequent analysis.

LEMMA 4. *For a given ρ, there exist levels of friendship $\underline{\phi}(\rho)$ and $\overline{\phi}(\rho)$ such that $(p, m) = (0, 1)$ is an equilibrium outcome for all $\phi < \underline{\phi}(\rho)$, and $(p, m) = (1, 0)$ is an equilibrium outcome for all $\phi > \overline{\phi}(\rho)$.*

We are now ready to state our first main result.

PROPOSITION 1. *Equilibria with full loyalty, $(p, m) = (1, 0)$, are the principal's preferred equilibria, for all ρ. Furthermore, there exist $\hat{\rho}$ and $\bar{\rho} > \hat{\rho}$ such that*

- *for all $\rho \in [\hat{\rho}, \bar{\rho})$, equilibria with full monitoring, $(p, m) = (0, 1)$, are the principal's second most preferred equilibria, and*

[22] We will henceforth denote an equilibrium by its strategy profile (p, m) alone, hence suppressing the equilibrium beliefs derived in the proof of Lemma 1.

[23] The two equilibria with full monitoring, $(p, m) = (0, 1)$ and $(p, m) = (\alpha, 1)$, yield the same payoffs for all players. Moreover, whenever $(p, m) = (\alpha, 1)$ is an equilibrium, $(p, m) = (0, 1)$ can also be supported as an equilibrium. Therefore, without loss of generality, we henceforth neglect $(p, m) = (\alpha, 1)$.

– *for all $\rho \geq \bar{\rho}$, the principal is indifferent between equilibria with full loyalty and equilibria with full monitoring.*

Proposition 1 describes the principal's preferences over the equilibria, and hence over friendship between the supervisor and the agent, for given levels of desire for esteem. In general, her preferences over friendship are non-monotonic; an increase in friendship can sometimes benefit the principal, and sometimes harm her. Intuitively, there are two effects at play, and the net effect on the principal's expected utility depends on which effect is more important: she may suffer from the incremental capture, but simultaneously benefit from the strengthening of loyalty. To be more specific, a higher ϕ may (i) harm the principal by reducing monitoring, (ii) benefit her because the agent chooses g more often, or (iii) allow her to save on the supervisor's wage because working with "friends" has intrinsic value.[24]

Despite these countervailing effects, three results are unambiguous. First, regardless of the level of desire for social esteem, the equilibrium with full loyalty is always the principal's most preferred equilibrium. Intuitively, the principal obtains the good outcome with probability one, but since the supervisor never monitors, there is no need to compensate her for any monitoring costs (i.e., $w_{(1,0)} = 0$).

Second, for intermediate and high levels of desire for social esteem ($\rho \geq \hat{\rho}$), the non-monotonicity takes the form of a preference for extremes: the principal's second most preferred equilibrium entails full monitoring, in contrast to the most preferred equilibrium described above, in which the supervisor never monitors. In the full monitoring equilibrium, the principal also obtains the good outcome with probability one; however she must in general pay the supervisor a wage to compensate for the monitoring costs.[25] Notice that, when $\rho > \bar{\rho}$, the supervisor cares so much about social esteem that she is willing to monitor at no wage. In this case, the principal is clearly indifferent between the equilibrium in which the agent always proposes the good project and the equilibrium in which the supervisor always monitors.

Third, by Lemma 4, the principal's equilibrium preferences translate into preferences over friendship: In order for the full monitoring equilibrium to exist, friendship must be low or moderate ($\phi < \underline{\phi}(\rho)$). Intuitively, in this equilibrium, the supervisor's care for esteem induces her to engage in full monitoring, so as to save face before the principal. If her friendship with the agent is too high, however, the effect of capture will erode her incentives to monitor. Even within the range of levels of friendship for which the full monitoring equilibrium exists, any marginal increase in friendship harms

[24] When ϕ increases, the supervisor *internalizes* more of the agent's private benefits. In this sense, their friendship creates "intrinsic" value. This benefits the principal because the wage that is required for the supervisor to break even becomes smaller.

[25] This wage is decreasing in ρ, since the supervisor's (utility from) pride is increasing in ρ.

the principal: because of the stronger bond between the supervisor and the agent, the supervisor requires a larger compensation for disciplining the agent.

In contrast, in order for the full loyalty equilibrium to exist, the principal must find an agent who is close enough a friend of the supervisor ($\phi > \overline{\phi}(\rho)$).[26] Intuitively, in the principal's most preferred equilibrium, the agent cares so much for the supervisor that he always proposes the good project, so as to save the face *of the supervisor*. Put differently, if friendship is strong enough, the agent inherits the supervisor's desire for esteem, which disciplines him to behave in accordance with the principal's preferences. In this equilibrium, a marginal decrease in friendship can harm the principal, since it weakens the loyalty that bridges the conflict between the agent and the principal.

The main message of the above proposition is that regardless of the level of care for esteem, the equilibrium with full loyalty, $(p, m) = (1, 0)$, is the principal's preferred equilibrium. Importantly, the feasibility of this equilibrium relies upon the ability of the principal to find an agent who is close enough a friend of the supervisor ($\phi > \bar{\phi}(\rho)$); if no such friend can be found, it is unattainable.

PROPOSITION 2. *The smallest level of friendship required for the principal's preferred equilibrium* $(p, m) = (1, 0)$ *to exist is decreasing in* ρ.

Proposition 2 states the main result of this paper: the more the supervisor values esteem, the more feasible is the principal's preferred equilibrium $(p, m) = (1, 0)$, because the range of levels of friendship for which it is attainable is increasing in ρ. Intuitively, holding the level of friendship constant, if the supervisor cares more about esteem, the agent suffers more when the supervisor is blamed for a bad outcome. As a consequence, the agent is more inclined to propose the good project. Thus, desire for esteem leverages the bridging effect of friendship.

4.2. Reinterpreting delegated monitoring. The full loyalty equilibrium drastically alters the nature of delegated monitoring. We discuss this change from three different angles, focusing in turn on the relationship between the agent and the supervisor (4.2.1), on the cohesion of the entire social network (4.2.2), and on the supervisor's *raison d'être* (4.2.3).

4.2.1. *Trust and trustworthiness.* The standard view is that a successful supervisor watches over the agent because she distrusts him, i.e., suspects that he misbehaves. Contrary to this view, the supervisor in the full loyalty equilibrium does not keep an eye on the agent, even though she cannot be sure about the agent's behavior. Rather,

[26] This equilibrium is not feasible if $\bar{\phi}(\rho) > 1$. However, there exists a wide range of parameter values such that $\bar{\phi}(\rho) \leq 1$, and we restrict attention to these.

she *trusts* him in that she makes herself "vulnerable by taking a course of action that creates incentives for the other party [...] to exploit [her]" (Ashraf et al., 2008, p.1).[27]

Her trust is based on the *confidence* that the agent will voluntarily choose the good project, a confidence that imposes an informal *obligation* on the agent.[28] According to Coleman (1988, p.S102), such conditionality defines trust: "If A does something for B and trusts B to reciprocate in the future, this establishes an expectation in A and an obligation on the part of B. This obligation can be conceived as a credit slip held by A for performance by B. [...] This form of social capital depends on two elements: trustworthiness of the social environment, which means that obligations will be repaid, and the actual extent of obligations held."

In the full loyalty equilibrium, the agent fulfills this informal obligation. Importantly, his *trustworthiness* originates from his altruism towards the supervisor. That is, he does not literally reciprocate the favor. Rather, he acts in accordance with his intrinsic desire to behave in the interest of his friend. Thus, here, trustworthiness arises from the friendship between the agent and the supervisor.[29]

A core proposition of our model is therefore that a monitoring relationship need not be characterized by distrust to be functional. On the contrary, such a relationship may successfully operate on the basis of trust and trustworthiness that arises from the social structure into which the professional duties are embedded.

4.2.2. *Indirect closure of social networks.* The root problem in our model is the presence of moral hazard between the agent and the principal. Furthermore, we assume that their conflict cannot be resolved through immediate social ties. According to Coleman (1988, Fig. 1 and adjacent text), the social network lacks *closure*, because the agent and the principal have no direct connection. For Coleman, the role of closure is to overcome conflicts of interest among network members: "Closure of the social structure is important not only for the existence of effective norms but also for another form of social capital: the trustworthiness of social structures that allows the proliferation of obligations and expectations. Defection from an obligation is a form of imposing a negative externality on another. Yet, in a structure without closure, it can be effectively sanctioned, if at all, only by the person to whom the obligation is owed."

[27] The full loyalty equilibrium also exists if the supervisor moves first (making herself openly vulnerable), so that the agent knows whether he will be monitored when he chooses his action.

[28] The absence of monitoring in the full loyalty equilibrium is not a result of capture (and, hence, not a result of her altruism towards the agent). First, it may well be that the supervisor would monitor if she knew for sure that the agent chooses b. Second, if the agent indeed proposes g, the supervisor's decision to remain passive is orthogonal to capture; in this case, she would not monitor even if (her) ϕ were zero.

[29] This is consistent with the findings of Ashraf et al. (2008) who conduct experiments to study whether trustworthiness in trust games arises from unconditional kindness or from reciprocity. They find strong support for the former but only weak support for the latter.

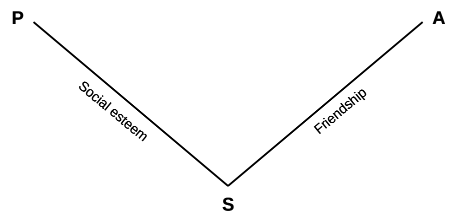

FIGURE A2. Network without direct closure.

Figure 2 depicts the lack of closure in our model: the principal lacks the means to directly sanction the agent for "defections" which harm her or the supervisor. Coleman's argument is that, in a network with closure, the supervisor and the principal could *each* pressure the agent to honor his obligation to choose g; and that the threat of double sanctions might deter defections by the agent, thus maintaining harmony in the network.

Our model suggests a different way to achieve the same outcome. Due to the nature of the connection between the supervisor and the agent (friendship), the principal can affect the agent's utility via her connection to the supervisor (social esteem). In particular, while the agent is not directly susceptible to shame, he is susceptible to shame inflicted on the supervisor. The threat of shaming the supervisor thus affects the agent's behavior. It is, in effect, *as if* the principal is able to punish the agent directly. In this sense, a full loyalty outcome can be seen as a network with *in*direct closure.

The "degree" of indirect closure in the network depends on the strength of the dyadic ties, P-S and S-A. This is because the bridging effect increases in ϕ and in ρ (Lemma 3 and Proposition 2). In fact, ϕ and ρ are complementary in this respect: a given level of friendship creates more loyalty when the supervisor cares more about esteem, and a given level of desire for esteem creates more loyalty when the friendship is stronger. Thus, the stronger the existing connections, the stronger the harmony created between unconnected members.

4.2.3. *Assigning moral responsibility.* Bartling and Fischbacher (2008, p.1), studying the incentives to delegate to shift blame, point to the following definition of moral responsibility (from the Stanford Encyclopedia of Philosophy): "To be morally responsible for something, say an action, is to be worthy of a particular kind of reaction—praise, blame, or something akin to these—for having performed it." They further write that "the question of moral responsibility is of economic relevance because praise and blame can constitute effective incentives."

In our model, we assume that the agent does not care about social esteem, so that such incentives are mute in his case. By contrast, if properly selected, the supervisor may respond strongly to these incentives. It is therefore the supervisor's reputability, i.e. her susceptibility to praise or to shame, that makes her valuable as a bearer of responsibility. In the full monitoring equilibrium, she perfectly oversees the agent's actions for fear of being held morally responsible for a bad outcome. A motivation for instating delegated monitors may hence be to relay responsibility to persons that are responsive to image concerns, i.e., to combine moral responsibility and reputability.

The significance of assigning moral responsibility is even more apparent in the full loyalty equilibrium. There, the supervisor's role does not involve any monitoring. It is rather the role of a designated scapegoat, standing in to bear the blame in the event of a bad outcome. Due to social ties, however, the supervisor's sense of responsibility proves contagious and is inherited by the agent. Thus, to be effective, delegated monitoring need not be active; it may suffice to appoint the right person to assume moral responsibility.

5. Cronyism, Boards, and Organizational Culture

The key result of this paper is that a given social connection between an agent and a supervisor can have very different implications for the welfare of the principal, depending on the supervisor's concern for esteem. Hence, in order to understand whether a given social connection is likely to be desirable, it is crucial to consider the effects of friendship and desire for esteem *in interaction*. This section highlights some of the contexts in which this insight is valuable.

5.1. Familism, nepotism and cronyism. Family ownership is the dominant form of corporate ownership around the world, although there are substantial differences across countries (La Porta et al., 1999; Faccio and Lang, 2002; Claessens et al., 2000; Villalonga and Amit, 2008; Franks et al., 2008). In certain countries, such as Thailand and Korea, economic growth was—and perhaps still is—driven by the success of large family business groups.[30]

Many are critical towards the mélange of social ties and corporate control evident in family firms. Their main concern is that entrenched networks of families and friends in business and government can lead to expropriation, corruption, and consequently economic stagnation (Morck et al., 2005). That is, they are concerned about the

[30] In their sample of 2,980 publicly traded companies in nine East Asian countries (Hong Kong, Indonesia, Japan, South Korea, Malaysia, the Philippines, Singapore, Taiwan, and Thailand), Claessens et al. (2000) find that in Indonesia, the Philippines and Thailand, about half of the corporate assets are controlled by ten families, while in Hong Kong and Korea about a third of the corporate sector is controlled by ten families.

dark side of social capital: members of closely-knit groups, who favor their social connections, may harm the broader interests of people outside of these groups. Or to be specific, family managers (A) supervised by family directors (S) may not run the firm in the best interest of its non-family investors (P).

The empirical evidence on the performance consequences of family ownership is inconclusive (Claessens et al., 2002; Cronqvist et al., 2003; Anderson and Reeb, 2003; Khanna and Palepu, 2000). The extant literature points to some channels through which families may impact firm performance: the incompetence of family heirs (Peréz-González, 2006; Villalonga and Amit, 2006), feuds among multiple heirs (Bertrand et al., 2008), and the value of political connections (Morck et al., 2000; Fisman, 2001).[31] Nonetheless, the key questions remain: Is family ownership a cause of, or a second-best solution for, agency problems? Does the dominance of family ownership persist because it is an efficient form of organization or because the inefficiencies that it creates are persistent?

5.1.1. *Importance of culture.* Most studies on family firms view family ties within the firm as a *solution* to the conflict between controlling shareholders and management, but as a *source* of the conflict between the controlling family shareholders and non-controlling shareholders. We argue that this trade-off need not be there.

We enrich the view on family ties by analyzing them embedded in a social context, rather than in isolation. If placed in a cultural context where integrity and social esteem play an important role, we argue that family ties can have bridging, rather than mere bonding, effects through which their benefits spill over to outsiders. This is because in such settings, social ties may induce agents to behave well for fear of causing family members or friends to lose face. The intermingling of business and personal relationships can thus be a solution to, rather than a manifestation of, agency problems.

Such cultural traits are salient inter alia in a number of Asian countries. Consider, for example, the Chinese institution of *guanxi*. At a basic level, *guanxi* refers to a network of personal relationships which can be used to facilitate transactions and cooperation. However, as the following description reveals, it works on the basis of complex interactions between trust, goodwill and social esteem:

> "*Guanxi* describes the basic dynamic in the complex nature of personalized networks of influence and social relationships, and is a central concept in Chinese society [...], becoming more widely used instead of

[31] Bertrand and Schoar (2007) also mention a family's long-term horizon, its concerns for legacy, and the nurture of firm-specific human capital within the family.

> the two common translations – 'connections' and 'relationships' – as neither of those terms sufficiently reflect the wide cultural implications that *guanxi* describes.
>
> Closely related concepts include that of *ganqing*, a measure which reflects the depth of feeling within an interpersonal relationship, *renqing*, the moral obligation to maintain the relationship, and "face", divided into the concepts of *mian* (or *mianzi*) meaning social status and prestige, and *lian*, the idea of being perceived as a morally correct actor within society."[32]

To illustrate *guanxi*, Standifird and Marshall (2000) give a trade example, cited also by Karlan et al. (2008). We can interpret their example within our framework: A seller (P) and a buyer (A) want to engage in a trade with deferred payment. A third person (S), who has ties to both parties, acts as a social intermediary (*zhongjian ren*). Her role is to vouch for the buyer, informally assuring the supplier that payment will be made. Thus, if the buyer fails to meet his obligation, the intermediary loses her face before the seller (and, by word of mouth, potentially before others). Given the tie between the buyer and the intermediary, the buyer will take into account that non-payment would damage the intermediary's reputation. Conversely, the intermediary presumes this implicitly: "[He] is aware of my wants/needs and will take them into account when deciding [his] course of future actions which concern or could concern me without any specific discussion or request."[33]

5.1.2. *Beyond the 'controlling shareholder trade-off'.* Consider again the family firm managed by members of the controlling family (A), who are accountable to the head of the family (S) and the firm's non-family investors (P). It is typically assumed that the presence of the controlling family leads to what corporate governance scholars refer to as the controlling shareholder trade-off (Gilson, 2006): Family ownership, on the one hand, may police the management better because of proximity and lower information costs. On the other hand, it creates a conflict between the family shareholders and non-family shareholders over the extraction of private benefits of control—benefits to the family not provided to the non-family shareholders.

We claim that, in a culture in which *guanxi* or similar notions are important, family ownership may overthrow this logic and resolve the trade-off. Consider a culture in which the head of the family bears the duty to safeguard the family's "name", i.e. its status in society; at the same time, she is entrusted with the fate of the family's firm. Delegating management to family members may then be a way of ensuring not only that

[32] Wikipedia entry retrieved on February 28, 2009, at http://en.wikipedia.org/wiki/Guanxi.
[33] Quote is taken (and adapted) from Wikipedia (see previous footnote).

the managers further the family's material interests but also that they eschew actions that jeopardize the family's name. Thus, the loyalty created by family bonds may also benefit non-family shareholders. Note that this turns the controlling shareholder trade-off on its head: a full loyalty equilibrium involves no monitoring, and it resolves *all* conflicts of interest through indirect closure (see 4.2.2).[34]

5.1.3. *Substitution or entrenchment?* It is sometimes suggested that informal institutions, such as *guanxi*, serve as a substitute for weak, or inexistent, formal institutions (Fock and Woo, 1998). In our model, we can view formal mechanisms as (non-modeled) legal instruments, or rights, that make monitoring more effective. Suppose that the expected cost of monitoring, $\hat{c} \equiv E(\tilde{c})$, reflects how weak these formal institutions are.

PROPOSITION 3. *For any ρ, the range of ϕ for which full monitoring equilibria exist decreases in $\hat{c}$. For a given $\phi < \underline{\phi}(\rho)$, the principal's utility in a full monitoring equilibrium also decreases in $\hat{c}$. By contrast, neither the range for which full loyalty equilibria exist nor the principal's utility in a full loyalty equilibrium depend on $\hat{c}$.*

In short, Proposition 3 states that weak(er) formal institutions reduce both the feasibility and the profitability of the full monitoring equilibrium, whereas they affect neither the feasibility nor the profitability of the full loyalty equilibrium. Thus, our model suggests that the use of social ties to create loyalty is a more prevalent, and also a more attractive, governance mechanism than active monitoring when formal institutions are ineffective. While offering a different explanation, this consents with the view that family businesses or conglomerates may be a constrained-optimal form of organization in markets suffering from weak legal protection (Leff, 1978; Burkart et al., 2003; Almeida and Wolfenzon, 2006; Khanna and Yafeh, 2007).

Having said this, our model also supports the opposite claim that family ownership may exacerbate agency problems. It can create capture as well as loyalty, and which effect dominates depends on the circumstances. Apart from the efficacy of active monitoring mechanisms ($\tilde{c}$), part of such circumstances is the importance of private benefits of control (B_b).

---

[34] Although referring to a slightly different mechanism, Gilson (2007) in a recent paper also stresses the potential importance of *family reputation* in markets with weak legal protection: "When commerce must take place in a reputation market [...], the distribution of shareholdings, and particularly the presence of family ownership, facilitate the development and maintenance of the reputation necessary for a corporation's commercial success. More speculatively, the role of reputation [...] may help explain why we observe publicly held minority shares in the capital market even though poor shareholder protection does not impose a formal limit on the amount of private benefits that a controlling shareholder can extract" (pp.4-5). Later he adds, "The critical point is that family ownership substitutes for internal incentive and transfer mechanisms as an assurance of the corporation's commitment to long-term reputation" (pp.15-16).

PROPOSITION 4. *For any* (ϕ, ρ), $\min P^* (\phi, \rho)$, $\max P^* (\phi, \rho)$, *and* $\max M^* (\phi, \rho)$ *are weakly decreasing in* B_b^H *or* B_b^L*; the same is true for* $\min M^* (\phi, \rho)$ *if* $\frac{\rho - c/(1-\alpha)}{B_b^H - B_g} \leq \frac{B_b^H - B_g}{\rho}$.

Proposition 4 says that larger private benefits weaken loyalty and, in general, strengthen capture. One can interpret high B_b's as a situation in which it is less costly for the controlling shareholders to divert resources from the firm. For example, this can be the case when control-enhancing mechanisms (pyramid structures, dual-class shares) allow them to maintain control while owning a very small equity stake. Interestingly, in their study of Chilean family business groups, Silva et al. (2006) find a positive impact of family ties on firm performance as long as the family's interests are sufficiently aligned with the interests of other shareholders. By contrast, they find a negative impact for the subset of firms in which the separation of ownership and control is very large.

To conclude, social ties among controlling shareholders and firm managers can benefit or harm minority shareholders. We believe that this ambiguity reflects the reality that not all family firms are alike. Moreover, we view the full loyalty equilibrium as an example of Gilson's (2006, p.1) claim that, for a deeper understanding of the diversity of ownership structures, one needs to "complicate the prevailing analysis of controlling shareholders and corporate governance by developing a more nuanced taxonomy."

5.2. Boards, networks and reputation. Much of the corporate governance research on boards of directors focuses on the distinction between *outside* and *inside* directors.[35] In addition, there is an emerging literature on the role of *gray*, or affiliated, directors who are independent according to the regulatory definition but have *social* ties to the management (Kramarz and Thesmar, 2006; Cohen et al., 2008; Horton et al., 2008; Hwang and Kim, 2008; Schmidt, 2009; Fracassi and Tate, 2009). The social connections between the CEO and other board members, or among all board members, are construed from background information, such as Alma mater, military service, regional origin, and academic discipline. On average, the findings suggest that strong social ties between the CEO and the rest of the board are associated with less

[35] The Investor Responsibility Research Center (IRRC) defines an outside (or independent) director as an individual who is not (i) a current or former employee, (ii) a relative of an executive officer, (iii) a customer of or a supplier to the firm, (iv) a provider of professional services, (v) a recipient of charitable funds, (vi) interlocked with an executive of the firm, and (vii) affiliated in any other way that poses a potential conflict of interest. In 2003, the SEC adopted a new, stricter definition of independence, which is similar and details are given at http://www.sec.gov/rules/sro/34-48745.htm for more details.

active boards, higher management compensation, or in some cases lower market values.[36] The natural conclusion seems to be that social ties lead to less effective boards, which resonates with the popular view that *independent* non-executive directors are the ideal advocates of the shareholders' interests.

As in the case of family firms, we contend that more attention needs to be given to the precise mechanisms through which social ties may operate. For instance, our model predicts that the desirability of appointing a (non-executive) director who has social ties to the management depends on the image concerns of the director as well as the strength of the social connection between the director and the management: The shareholders (P) prefer to appoint a board (S) which has close social ties to the management (A), and which cares a lot about public esteem. This would implement a full loyalty equilibrium such that the firm is well-managed although the board is passive. If this is not possible, the shareholders prefer a board which has few ties to, and hence actively monitors, the management. The model thus concurs with the view that independent directors are more active monitors, but challenges the view that they are necessarily the best advocates of shareholder interests.[37]

This insight has two implications for empirical studies. First, less active monitoring by affiliated directors need not imply that the board is ineffective; indeed, the absence of monitoring reflects the virtue of a full loyalty equilibrium. Second, *interacting* the level of social ties with the directors' desire for social esteem may help to identify *positive* effects of social ties, as appointments of directors who are both close to the CEO and reputable are conducive to a full loyalty equilibrium. One possible example is the appointment of Bill Gates as a director of Berkshire Hathaway; he is not only highly reputable, but also a close friend of the CEO, Warren Buffett. To strengthen our claim, we discuss below the importance of reputational incentives for boards of directors and the prevalence of social networks among corporate directors. The section concludes with a note on modeling social ties in boards of directors.

5.2.1. *Reputation and shaming of corporate boards.* Can "shaming" sanctions constitute effective penalties for corporate directors? To be an effective disciplining device, shaming requires that directors' responsibilities are governed by a social norm and that directors care about social esteem. Both requirements seem to be satisfied in reality. On the one hand, investors, regulators and public media increasingly demand that

[36] Two exceptions are Klein (1998) who finds that affiliated directors seem to fill strategic needs of the firms, and Kim (2005) who documents a non-monotonic relationship between the density of connections among board members and a firm's market value in a sample of Korean firms.

[37] Adams and Ferreira (2007) and Harris and Raviv (2008) reach similar conclusions in their models, albeit for a different reason: management-friendly directors can elicit more information from corporate insiders, which induces a trade-off between monitoring and information aggregation. In our model, close ties may render monitoring obsolete, rather than infeasible.

boards exercise meaningful oversight and accordingly condemn directorial sloth, negligence or misconduct. On the other hand, corporate directors are often under public scrutiny and enmeshed in communities in which reputation matters. The well-known shareholder activist and corporate governance expert Nell Minow even believes that "[t]he directors of large U.S. corporations are the most reputationally sensitive people in the world."[38]

The following examples illustrate the use of shaming as a disciplining device. In April 1992, shareholder activist and founder of Institutional Shareholder Services (ISS) Robert Monks, in order to promote shareholder proposals against the management of Sears, paid for a full-page advertisement in the *Wall Street Journal* with the title: "The non-performing assets of Sears." The page showed a silhouette outline of the nine Sears directors, listing each by name and by position as responsible for the company's poor performance. Although none of the shareholder proposals were approved by vote, the embarrassed directors—in response to the ad—voluntarily adopted several proposals endorsed by Robert Monks and Nell Minow. Another example is CalPERS' *focus list*, a list of underperforming corporations in which the Californian public pension fund owns a non-negligible stake. While CalPERS also takes other measures to effect change in those corporations, it believes that the shame associated with being published on the list is sufficiently embarrassing to make directors more responsive to shareholder demands.[39] As a former CalPERS' general counsel argues: "The shaming aspect is very important. Firms hate to be on the list." Finally, business media like the *Wall Street Journal*, *Business Week*, or *Fortune* also draw attention to firms and the quality of their boards through occasional investigative reports or periodical special reports on the "best and worst boards" of the year.

5.2.2. *Market for directors and relevant audience.* The idea that directors care about their reputation dates back at least to Fama (1980) and Fama and Jensen (1983), who hypothesized that a director's reputation has (instrumental) value in a market for directors. A number of empirical studies on directors' reputational incentives has found support for this hypothesis (Gilson, 1990; Kaplan and Reishus, 1990; Brickley et al., 1999; Coles and Hoi, 2003; Harford, 2003; Fich, 2005). These studies typically proxy a director's reputation through a measure of his or her "popularity" on the market for directors. Yermack (2004), for example, identifies reputational incentives by looking at the number of additional board seats that are awarded to directors of firms with

[38] The quotes and examples in this subsection are taken from Skeel (2001, p.1812 and Section III).

[39] It should be pointed out that the past evidence on the success of activist pension or mutual funds in general, and on the "CalPERS' effect" in particular, is weak (Black, 1998). More recent studies, however, report that hedge funds successfully implement activist strategies (Becht et al., 2008; Brav et al., 2008), including shaming strategies (Dyck et al., 2008).

good performance. These appointments are interpreted as rewards to the directors' reputations. Similarly, Fich and Shivdasani (2007) find that outside directors of firms facing a financial fraud lawsuit are likely to lose their board seats in *other* firms; moreover, this effect is stronger when the fraud allegation is more severe and when the outside director bears greater responsibility for monitoring fraud. Finally, Masulis and Mobbs (2008) find evidence suggesting that inside directors have a positive effect on firm performance *if* they simultaneously hold board seats in other firms.

Clearly, reputational incentives are ineffective in anonymous markets. Part of their effectiveness in markets for directors stems from the fact that directors of large corporations typically belong to a small world of corporate elites. However, this social structure enhances image concerns not only in the market for directors but also in the directors' personal environment. As Robert Monks explains, his ad in the Wall Street Journal aimed at a wider audience than just the board members and Sears' shareholders: "We were speaking to their friends, their families, their professional associates. Anyone seeing the ad would read it. Anyone reading it would understand it. Anyone understanding it would feel free to ask questions of any board member they encountered." As Nell Minow recounts the episode, Sears' directors were indeed ridiculed at their country clubs as a result of the advertisement (Dyck and Zingales, 2002, p.4).

5.2.3. *Modeling affiliated directors.* Most corporate governance theory analyzes self-interested individuals, engaged in exclusively professional relationships, focusing on monetary rewards to induce the right incentives. However, in practice, there exist extensive social networks among capital providers, management, directors, and employees. That is, the formal relationships, which allocate legal control rights, are embedded in a social structure.

In the theoretical literature on boards of directors, several papers have modeled differences in board members' alignment with the management or with the shareholders (e.g., Hermalin and Weisbach, 1998; Adams and Ferreira, 2007; Harris and Raviv, 2008). In most cases, a board's sympathy towards the CEO has been incorporated by assuming that its preferred action is more similar to the CEO's (and thereby more dissimilar to the shareholders'). This can be modeled as an additional cost that the supervisor incurs when reversing the CEO's decision, or alternatively as an additional benefit that the supervisor receives when the CEO's preferred action is implemented.[40] Compared to our approach (of modeling social preferences directly), these assumptions only permit the capture effect. Put differently, because the sympathy is assumed to be

[40] In their recent survey of board models, Adams et al. (2008) write: "Independence is a complex concept. With respect to monitoring the CEO, one imagines that directors who have close ties to the CEO (e.g., professionally, socially, or because the ceo has power over them) would *find monitoring him more costly* than directors with fewer ties" (fn.13, emphasis added).

one-directional, those models preclude the possibility of the loyalty effect. We believe that this is an important omission, because sympathy arising from social ties is often *mutual*.

As Adams et al. (2008) point out: "Much of the research on boards ultimately touches on the question 'what is the role of the board?' Possible answers range from boards' being simply legal necessities, something akin to the wearing of wigs in English courts, to their playing an active part in the overall management and control of the corporation. No doubt the truth lies somewhere between these extremes; indeed, there are probably multiple truths when this question is asked of different firms, in different countries, or in different periods." In accordance with this view, our theory suggests that explicitly taking into account the role of social preferences can shed light on previously overlooked roles that boards may play in practice; and that the categorization into outsiders, insiders and affiliated directors may have to be further refined with respect to reputational concerns and the strength of social ties.

5.3. Organizational design and corporate culture. As illustrated above, the theory can inform a principal about which exogenous characteristics (desire for esteem and strength of friendship with the agent) to look for when appointing a supervisor.[41] The model can also be applied to any hierarchical organization in which a high-level manager (P) can put mechanisms in place to *endogenously* shape the degree of social ties between, and the care for esteem of, lower-level units (A) and their mid-level managers (S). In light of the current theory, organizational culture can be viewed as a mechanism to promote the desired social connections and mold the desired values within an organization. Milgrom and Roberts (1992) acknowledge this point in their widely known textbook *Economics, Organizations and Management* where they write that "important features of many organizations can best be understood in terms of deliberate attempts to change preferences of individual participants."

For example, team-building activities for a lower-level unit and its mid-level manager can be viewed as an active measure to increase the strength of social ties (ϕ) between each employee and his direct superior. Designing infrastructure and work procedures in a way that creates opportunities for social interaction between the unit's employees and their mid-managers may serve the same purpose. In the words of Homans

[41] Clearly, the theory also informs a principal about which exogenous characteristics (strength of friendship with the supervisor) to look for when hiring an agent. For example, Fehr and Falk (2002, p.693) write: "If it is true that some people are more self-interested than others, then choosing the 'right' people is one way of affecting the preferences of a firm's workforce. For this reason employers have a strong interest in recruiting employees who have favourable preferences and whose preferences can be affected in favourable ways. There is circumstantial evidence for this because the testing and screening of employees is often as much about the employee's willingness to become a loyal firm member as it is about the employee's technical abilities."

(1950, p.133), "the more frequently persons interact with one another, the stronger their sentiments of friendship for one another" (1950, p.133).

In this context, ρ can be reinterpreted as a measure of transparency within the organization. To see this, consider an organizational policy such as naming the 'mid-manager of the month' and praising her within the entire organization. Intuitively, the more recognition given to this mid-manager of the month, i.e., the more transparent, or visible, the praise for good performance, the stronger are the incentives for the mid-level managers to earn this distinction.

The theory predicts that if measures to boost the social ties between a unit's employees and their direct superior are combined with initiatives to increase transparency and to publicly praise the direct superior in case of a good outcome, an organization may be able promote a corporate culture in which the unit's employees work hard even though the monitoring effort of their superior is low, so as not to put their superior's reputation with the senior management at risk, and, potentially, to even earn her public praise.

6. Concluding Remarks

This paper studies a governance problem which is embedded in a social structure: a principal hires a supervisor to monitor an agent, yet the supervisor and the agent may be connected through social ties, i.e., be "friends". We analyze how this friendship affects the supervisor's and the agent's behavior, as well as the principal's utility, across different settings where we vary the supervisor's desire for social esteem. On the one hand, friendship can bond the supervisor and the agent together against the principal, so that the supervisor monitors less (capture). On the other hand, it can induce the agent to voluntarily behave well, thereby bridging the conflict between the agent and the principal (loyalty). The effect on the principal is therefore ambiguous. We show that the principal's most preferred equilibrium is one in which the agent is fully loyal and the supervisor need not actively monitor the agent. Furthermore, we show that the full loyalty equilibrium is easier to achieve when the supervisor's desire for esteem is high(er). Thus, we conclude that the effect of social ties crucially depends on the salience of reputational concerns. As we discuss, this has important implications for family ownership and crony capitalism, for corporate governance and boards of directors, and for the impact of organizational design on corporate culture.

Appendix

Proof of Lemma 1. We can abstract from the participation constraints and only take the incentive compatibility constraints into account. This is because by the proof of Proposition 1 below, S will be given a fixed wage w in any equilibrium where this is necessary to satisfy her participation constraint, which in turn will guarantee that A's participation constraint is satisfied. Since the fixed wage does not affect the parameter values for which the respective equilibria can be sustained, the participation constraints can be abstracted from here.

After having observed the realization of $\tilde{B}_b$, B_b, the agent chooses g iff

$$\begin{aligned} B_g + \phi\left(\rho\theta_g - mE\left(\tilde{c}\,|\text{S monitors}\right)\right) \quad &\geq \quad m\left[B_g + \phi\left(\rho\theta_g - E\left(\tilde{c}\,|\text{S monitors}\right)\right)\right] \\ &\quad + (1-m)\left[B_b + \phi\left[-\rho\left(1-\theta_b\right)\right]\right] \end{aligned} \qquad (\text{IC}_A)$$

After having observed the realization of $\tilde{c}$, c, the supervisor monitors iff

$$\begin{aligned} \rho\theta_g - c + \phi B_g \quad &\geq \quad p\left(\rho\theta_g + \phi B_g\right) + (1-p)\left(-\rho\left(1-\theta_b\right) + \phi E\left(\tilde{B}_b\,|\text{A proposed } b\right)\right) \qquad (\text{IC}_S) \\ &\Leftrightarrow \quad c \leq (1-p)\,\rho\left(1-\theta_b+\theta_g\right) - \phi\left(1-p\right)\left(E\left(\tilde{B}_b\,|\text{A proposed } b\right) - B_g\right). \end{aligned}$$

Equilibria in which $m = 1$*.* Conjecture that $m = 1$. Then, the agent is indifferent between g and b, since IC_A reduces to

$$B_g + \phi\left(\rho\theta_g - E\left(\tilde{c}\,|\text{S monitors}\right)\right) = B_g + \phi\left(\rho\theta_g - E\left(\tilde{c}\,|\text{S monitors}\right)\right).$$

IC_S is given by

$$(\text{IC}_S\text{'}) \qquad \bar{c} \leq (1-p)\,\rho\left(1-\theta_b+\theta_g\right) - \phi\left(1-p\right)\left(E\left(\tilde{B}_b\,|\text{A proposed } b\right) - B_g\right)$$

Consistent beliefs must satisfy $\theta_g = 1$, $\theta_b \in [0,1]$. We know that $m \neq 1$ if $p = 1$, so only two cases are distinguishable:

Case 1: Conjecture that $p = 0$. Substituting the beliefs and conjectures into (IC_S') yields that $m = 1$ is supported in equilibrium iff

$$(6.1) \qquad \bar{c} \leq \rho\left(2-\theta_b\right) - \phi\left[E\left(\tilde{B}_b\right) - B_g\right] \Leftrightarrow \rho \geq \frac{\bar{c} + \phi\left[E\left(\tilde{B}_b\right) - B_g\right]}{\left(2-\theta_b\right)}$$

Applying the intuitive criterion yields $\theta_b = 0$, since the supervisor would never choose to incur the monitoring cost unless she intended to convert a detected type b action. Formally, given that she has monitored and detected a type b action, she obtains a higher benefit from converting it to a type g action (LHS), than from not doing so (RHS):

$$\rho\theta_g + \phi B_g \geq -\rho\left(1-\theta_b\right) + \phi E\left(\tilde{B}_b\right) \Leftrightarrow 0 \leq \rho\left(2-\theta_b\right) - \phi\left[E\left(\tilde{B}_b\right) - B_g\right],$$

which is implied by (6.1).

Substituting $\theta_b = 0$ into (6.1) yields that $m = 1$ is supported in equilibrium iff

$$(6.2) \qquad \rho \geq \frac{\bar{c} + \phi\left[E\left(\tilde{B}_b\right) - B_g\right]}{2}$$

Hence, for (ρ, ϕ) satisfying (6.2), $\{(p, m) = (0, 1), \theta_g = 1, \theta_b = 0\}$ is a PBE.

Case 2: Conjecture that $p=\alpha$. Substituting the beliefs and conjectures into (IC$_S$') yields that $m=1$ is supported in equilibrium iff

$$\bar{c}\leq\rho\left(1-\alpha\right)\left(2-\theta_b\right)-\phi\left(1-\alpha\right)\left(B_b^H-B_g\right)\Leftrightarrow\rho\geq\frac{\bar{c}+\phi\left(1-\alpha\right)\left(B_b^H-B_g\right)}{\left(1-\alpha\right)\left(2-\theta_b\right)}\tag{6.3}$$

Applying the intuitive criterion yields $\theta_b=0$, since the supervisor would never choose to incur the monitoring cost unless she intended to convert a detected type b action. Formally, given that she has monitored and detected a type b action, she obtains a higher benefit from converting it to a type g action (LHS), than from not doing so (RHS):

$$\rho\theta_g+\phi B_g\geq-\rho\left(1-\theta_b\right)+\phi B_b^H\Leftrightarrow 0\leq\rho\left(2-\theta_b\right)-\phi\left[B_b^H-B_g\right],$$

which is implied by (6.3).

Substituting $\theta_b=0$ into (6.3) yields that $m=1$ is supported in equilibrium iff

$$\rho\geq\frac{\bar{c}/\left(1-\alpha\right)+\phi\left(B_b^H-B_g\right)}{2}\tag{6.4}$$

Hence, for (ρ,ϕ) satisfying (6.4), $\{(p,m)=(\alpha,1),\theta_g=1,\theta_b=0\}$ is a PBE.

As will be shown in the proof of Proposition 1 below, the two equilibria with full monitoring, $\{(p,m)=(0,1),\theta_g=1,\theta_b=0\}$ and $\{(p,m)=(\alpha,1),\theta_g=1,\theta_b=0\}$, yield the same payoffs for all players. Moreover, the above conditions yield that whenever $\{(p,m)=(\alpha,1),\theta_g=1,\theta_b=0\}$ is an equilibrium, $\{(p,m)=(0,1),\theta_g=1,\theta_b=0\}$ can also be supported as an equilibrium. Therefore, without loss of generality, we henceforth neglect $\{(p,m)=(0,1),\theta_g=1,\theta_b=0\}$.

Equilibria in which $m=\beta$. Conjecture that $m=\beta$. Then, IC$_S$ is given by

$$\underline{c}\leq\left(1-p\right)\rho\left(1-\theta_b+\theta_g\right)-\phi\left(1-p\right)\left(E\left(\tilde{B}_b\,|\text{A proposed } b\right)-B_g\right)\leq\bar{c}\tag{IC$_S$”}$$

Whenever $m\neq1$, (IC$_A$) reduces to

$$B_g+\phi\rho\left(\theta_g+1-\theta_b\right)\geq B_b,\tag{IC$_A$”}$$

where B_b is a realization of $\tilde{B}_b$.

We know that $m\neq\beta$ if $p=1$, so only two cases are distinguishable:

Case 1: Conjecture that $p=0$. Then consistent beliefs must satisfy $\theta_g=1$, $\theta_b=0$. Substituting the beliefs and conjectures into (IC$_S$") yields that $m=\beta$ is supported in equilibrium iff

$$\frac{\underline{c}+\phi\left(E\left(\tilde{B}_b\right)-B_g\right)}{2}\leq\rho\leq\frac{\bar{c}+\phi\left(E\left(\tilde{B}_b\right)-B_g\right)}{2}.\tag{6.5}$$

Substituting the beliefs and conjectures into (IC$_A$") yields that $p=0$ is supported in equilibrium iff

$$\phi\rho\leq\frac{B_b^L-B_g}{2}\tag{6.6}$$

Hence, for (ρ,ϕ) satisfying (6.5) and (6.6), $\{(p,m)=(0,\beta),\theta_g=1,\theta_b=0\}$ is a PBE.

Case 2: Conjecture that $p=\alpha$. Then consistent beliefs must satisfy $\theta_g=\frac{\beta}{\beta+\alpha(1-\beta)},\theta_b=0$. Substituting the beliefs and conjectures into (IC$_S$") yields that $m=\beta$ is supported in equilibrium iff

$$\frac{\underline{c}/\left(1-\alpha\right)+\phi\left(B_b^H-B_g\right)}{\left(2\beta+\alpha\left(1-\beta\right)\right)/\left(\beta+\alpha\left(1-\beta\right)\right)}\leq\rho\leq\frac{\bar{c}/\left(1-\alpha\right)+\phi\left(B_b^H-B_g\right)}{\left(2\beta+\alpha\left(1-\beta\right)\right)/\left(\beta+\alpha\left(1-\beta\right)\right)}\tag{6.7}$$

Substituting the beliefs and conjectures into (IC_A") yields that $p=\alpha$ is supported in equilibrium iff

$$\frac{B_b^L - B_g}{\left(2\beta+\alpha\left(1-\beta\right)\right)/\left(\beta+\alpha\left(1-\beta\right)\right)} \leq \phi\rho \leq \frac{B_b^H - B_g}{\left(2\beta+\alpha\left(1-\beta\right)\right)/\left(\beta+\alpha\left(1-\beta\right)\right)}. \tag{6.8}$$

Note that in both 6.7 and 6.8, the denominator is in $(1,2)$ since it is equal to $(1+\theta_g)$.

Hence, for (ρ,ϕ) satisfying (6.7) and (6.8), $\left\{(p,m)=(\alpha,\beta),\theta_g=\frac{\beta}{\beta+\alpha(1-\beta)},\theta_b=0\right\}$ is a PBE.

Equilibria in which $m=0$*.* Then, (IC_S) is given by

$$\underline{c} \geq (1-p)\,\rho\left(1-\theta_b+\theta_g\right) - \phi\left(1-p\right)\left(E\left(\tilde{B}_b\,|\text{A proposed } b\right)-B_g\right), \tag{IC_S" '}$$

and as $m\neq 1$, (IC_A) reduces to (IC_A"). Three cases are distinguishable.

Case 1: Conjecture that $p=0$. Then consistent beliefs must satisfy $\theta_g\in[0,1]$, $\theta_b=0$. Substituting the beliefs and conjectures into (IC_S"') yields that $m=0$ is supported in equilibrium iff

$$\rho \leq \frac{\underline{c}+\phi\left(E\left(\tilde{B}_b\right)-B_g\right)}{\left(1+\theta_g\right)} \tag{6.9}$$

Substituting the beliefs and conjectures into (IC_A") yields that $p=0$ is supported in equilibrium iff

$$\phi\rho \leq \frac{B_b^L-B_g}{\left(1+\theta_g\right)} \tag{6.10}$$

Here, we cannot apply the intuitive criterion to select off-equilibrium beliefs, since the supervisor, if she had monitored and detected a type b action, would not necessarily have obtained a higher benefit from converting it to a type g action. Formally, given that she has monitored and detected a type b action (i.e., when the monitoring cost is sunk), she obtains a benefit from converting it to a type g action (LHS), and a benefit from not doing so (RHS):

$$\rho\theta_g+\phi B_g \geq -\rho+\phi E\left(\tilde{B}_b\right) \Leftrightarrow \rho\left(1+\theta_g\right) \geq \phi\left(E\left(\tilde{B}_b\right)-B_g\right),$$

which does not necessarily hold.

Hence, for (ρ,ϕ) satisfying (6.9) and (6.10), $\{(p,m)=(0,0),\theta_g:\theta_g\in[0,1],\theta_b=0\}$ constitute PBE.

Case 2: Conjecture that $p=\alpha$. Then consistent beliefs must satisfy $\theta_g=0$, $\theta_b=0$. Substituting the beliefs and conjectures into (IC_S"') yields that $m=0$ is supported in equilibrium iff

$$\rho \leq \underline{c}/\left(1-\alpha\right)+\phi\left(B_b^H-B_g\right) \tag{6.11}$$

Substituting the beliefs and conjectures into (IC_A") yields that $p=\alpha$ is supported in equilibrium iff

$$B_b^L-B_g \leq \phi\rho \leq B_b^H-B_g \tag{6.12}$$

Hence, for (ρ,ϕ) satisfying (6.11) and (6.12), $\{(p,m)=(\alpha,0),\theta_g=0,\theta_b=0\}$ is a PBE.

Case 3: Conjecture that $p=1$. Then consistent beliefs must satisfy $\theta_g=0$, $\theta_b\in[0,1]$. Substituting the beliefs and conjectures into (IC_S"') yields that $m=0$ is supported in equilibrium iff

$$\underline{c}\geq 0,$$

which hold by assumption. Substituting the beliefs and conjectures into (IC_A") yields that $p=1$ is supported in equilibrium iff

$$\phi\rho \geq \frac{B_b^H-B_g}{\left(1-\theta_b\right)} \tag{6.13}$$

Applying the intuitive criterion yields $\theta_b = 0$, since the supervisor, if she had monitored and detected a type b action, would have obtained a higher benefit from converting it to a type g action (LHS), than from not doing so (RHS):

$$\rho\theta_g + \phi B_g \geq -\rho\left(1-\theta_b\right) + \phi E\left(\tilde{B}_b\right) \Leftrightarrow \rho \geq \frac{\phi\left(E\left(\tilde{B}_b\right) - B_g\right)}{\left(1-\theta_b\right)},$$

which is implied by (6.13) for $\phi \in (0,1]$:

$$\rho \geq \frac{B_b^H - B_g}{\phi\left(1-\theta_b\right)} > \frac{E\left(\tilde{B}_b\right) - B_g}{\phi\left(1-\theta_b\right)} \geq \frac{E\left(\tilde{B}_b\right) - B_g}{\left(1-\theta_b\right)} \geq \frac{\phi\left(E\left(\tilde{B}_b\right) - B_g\right)}{\left(1-\theta_b\right)}.$$

Substituting $\theta_b = 0$ into (6.13) yields that $p = 1$ is supported in equilibrium iff

$$\phi\rho \geq B_b^H - B_g \tag{6.14}$$

Hence, for (ρ, ϕ) satisfying (6.14), $\{(p,m) = (1,0), \theta_g = 0, \theta_b = 0\}$ is a PBE. ■

Proof of Lemma 2. Rearranging the conditions in the proof of Lemma 1 yields that the supervisor's equilibrium strategies, for a given ρ, can be supported for the following ϕ:

$$m^* = \begin{cases} 1 & if \quad \phi \leq \phi_1 \\ \beta & if \quad \phi_2 \leq \phi \leq \phi_3 \\ 0 & if \quad \phi \geq \phi_4 \end{cases}, \tag{6.15}$$

where $\phi_1 \equiv \frac{2\rho - \bar{c}}{\left[E\left(\tilde{B}_b\right) - B_g\right]}$, $\phi_2 \equiv \frac{\left(1 + \frac{\beta}{\beta + \alpha(1-\beta)}\right)\rho - \bar{c}/(1-\alpha)}{\left(B_b^H - B_g\right)}$, $\phi_3 \equiv \frac{2\rho - \underline{c}}{E\left(\tilde{B}_b\right) - B_g}$, and $\phi_4 \equiv \min\left\{\frac{\rho - \underline{c}/(1-\alpha)}{B_b^H - B_g}, \frac{B_b^H - B_g}{\rho}\right\}$, and hence where $\phi_2 < \phi_1 < \phi_3$ and $\phi_4 < \phi_3$. This implies that there are three cases: (i) $\phi_2 < \phi_1 < \phi_4 < \phi_3$, (ii) $\phi_2 < \phi_4 < \phi_1 < \phi_3$, and (iii) $\phi_4 < \phi_2 < \phi_1 < \phi_3$. In either case, (6.15) yields that

$$\max M^*(\phi, \rho) = \begin{cases} 1 & if \quad \phi \leq \phi_1, \\ \beta & if \quad \phi_1 < \phi \leq \phi_3 \\ 0 & if \quad \phi > \phi_3 \end{cases}.$$

As $\phi_1 < \phi_3$, $\max M^*(\phi, \rho)$ is a decreasing function of ϕ, for a given ρ.

For $\min M^*(\phi, \rho)$, it is necessary to distinguish between the cases when $\phi_2 < \phi_4$ and when $\phi_2 \geq \phi_4$. From (6.15), we obtain that

$$\begin{array}{ll} \min M^*(\phi, \rho) = \begin{cases} 1 & if \quad \phi < \phi_2, \\ \beta & if \quad \phi_2 \leq \phi < \phi_4 \\ 0 & if \quad \phi \geq \phi_4 \end{cases} & if \quad \phi_2 < \phi_4 \\ \min M^*(\phi, \rho) = \begin{cases} 1 & \phi < \phi_4, \\ 0 & \phi \geq \phi_4 \end{cases} & if \quad \phi_2 \geq \phi_4 \end{array}.$$

In either case ($\phi_2 < \phi_4$ or $\phi_2 \geq \phi_4$), $\min M^*(\phi, \rho)$ is a decreasing function of ϕ, for a given ρ. ■

Proof of Lemma 3. Rearranging the conditions in the proof of Lemma 1 yields that the agent's equilibrium strategies, for a given ρ, can be supported for the following ϕ:

$$p^* = \begin{cases} 0 & if \quad \phi \leq \min\{\phi_5, \phi_6\} \\ \alpha & if \quad \phi_7 \leq \phi \leq \phi_8 \\ 1 & if \quad \phi \geq \phi_8 \end{cases}, \tag{6.16}$$

where $\phi_5 \equiv \frac{B_b^L - B_g}{\rho}$, $\phi_6 \equiv \frac{2\rho - \bar{c}}{E\left(\tilde{B}_b\right) - B_g}$, $\phi_7 \equiv \frac{B_b^L - B_g}{\rho(2\beta + \alpha(1-\beta))/(\beta + \alpha(1-\beta))}$, $\phi_8 \equiv \frac{B_b^H - B_g}{\rho}$, and hence where $\phi_7 < \phi_5 < \phi_8$, and where ϕ_6 cannot be directly compared to ϕ_5, ϕ_7, or ϕ_8. Regardless of the size of

ϕ_6 relative to ϕ_5, ϕ_7, or ϕ_8, (6.16) yields that

$$\max P^*(\phi,\rho) = \begin{cases} 0 & if \quad \phi < \phi_7 \\ \alpha & if \quad \phi_7 \leq \phi < \phi_8 \\ 1 & if \quad \phi \geq \phi_8 \end{cases}.$$

As $\phi_7 < \phi_8$, $\max P^*(\phi,\rho)$ is an increasing function of ϕ, for a given ρ.

For $\min P^*(\phi,\rho)$, it is useful to distinguish between the cases when $\phi > \phi_6$ and $\phi \leq \phi_6$. From (6.16) we obtain that

$$\begin{array}{ll} \min P^*(\phi,\rho) = \begin{cases} 0 & if \quad \phi \leq \phi_5 \\ \alpha & if \quad \phi_5 < \phi \leq \phi_8 \\ 1 & if \quad \phi > \phi_8 \end{cases} & for \quad \phi > \phi_6 \\ \min P^*(\phi,\rho) = 0 & for \quad \phi \leq \phi_6 \end{array}.$$

As $\phi_5 < \phi_8$, $\min P^*(\phi,\rho)$ is an increasing function of ϕ in both cases. Furthermore, for $\phi \leq \phi_6$, $\min\{P^*(\phi,\rho)\} = 0$ and as ϕ increases beyond ϕ_6, $\min\{P^*(\phi,\rho)\}$ will take any value greater than or equal to 0, and hence $\min\{P^*(\phi,\rho)\}$ is an increasing function of ϕ, for a given ρ. ■

Proof of Lemma 4. Comparing the conditions in the proof of Lemma 1 yields $\underline{\phi}(\rho) \equiv \frac{2\rho-\bar{c}}{\left[E\left(\tilde{B}_b\right)-B_g\right]}$ and $\overline{\phi}(\rho) = \frac{B_b^H - B_g}{\rho}$. ■

Proof of Proposition 1. First, the principal's and the supervisor's *ex ante* expected utilities in each equilibrium, before any wages are paid to the supervisor, are calculated. In an equilibrium where strategies (p,m) are played, denote these expected utilities $E(u_P)_{(p,m)}^{-w}$ and $E(u_S)_{(p,m)}^{-w}$, respectively.

In the equilibria $\{(p,m) = (0,1), \theta_g = 1, \theta_b = 0\}$ and $(p,m) = \{(\alpha,1), \theta_g = 1, \theta_b = 0\}$,

$$E(u_P)_{(p,m)}^{-w} = X_g, \quad E(u_S)_{(p,m)}^{-w} = -\left[\beta\underline{c} + (1-\beta)\bar{c}\right] + \rho + \phi B_g.$$

In the equilibrium $\{(p,m) = (0,\beta), \theta_g = 1, \theta_b = 0\}$,

$$E(u_P)_{(0,\beta)}^{-w} = \beta X_g + (1-\beta) X_b, \quad E(U_S)_{(0,\beta)}^{-w} = \beta(\rho - \underline{c} + \phi B_g) + (1-\beta)(\phi E(B_b) - \rho).$$

In the equilibrium $\left\{(p,m) = (\alpha,\beta), \theta_g = \frac{\beta}{\beta+\alpha(1-\beta)}, \theta_b = 0\right\}$,

$$E(u_P)_{(\alpha,\beta)}^{-w} = \beta X_g + (1-\beta)\left[\alpha X_g + (1-\alpha) X_b\right]$$

$$E(u_S)_{(\alpha,\beta)}^{-w} = (\beta\rho - \underline{c}\alpha\beta - \rho(\alpha-1)(\beta-1) + \underline{c}\beta(\alpha-1)) + \phi\left(\alpha B_g + B_b^H(\alpha-1)(\beta-1) - \beta B_g(\alpha-1)\right).$$

In the equilibria $\{(p,m) = (0,0), \theta_g : \theta_g \in [0,1], \theta_b = 0\}$,

$$E(u_P)_{(0,0)}^{-w} = X_b, \quad E(u_S)_{(0,0)}^{-w} = -\rho + \phi\left(\alpha B_b^L + (1-\alpha) B_b^H\right).$$

In the equilibrium $\{(p,m) = (\alpha,0), \theta_g = 0, \theta_b = 0\}$,

$$E(u_P)_{(\alpha,0)}^{-w} = \alpha X_g + (1-\alpha) X_b, \quad E(u_S)_{(\alpha,0)}^{-w} = \rho(\alpha-1) + \phi\left(\alpha B_g - B_b^H(\alpha-1)\right).$$

In the equilibrium $\{(p,m) = (1,0), \theta_g = 0, \theta_b = 0\}$,

$$E(u_P)_{(1,0)}^{-w} = X_g, \quad E(u_S)_{(1,0)}^{-w} = \phi B_g.$$

Second, the principal's *ex ante* utility, after having paid the supervisor a wage when needed, is calculated. Henceforth, we refer to the equilibrium $\{(p,m) = (0,1), \theta_g = 1, \theta_b = 0\}$ as $(p,m) = (0,1)$, and so on. Let $w_{(p,m)}$ denote the wage that the principal pays to the supervisor in equilibrium (p,m).

The principal's maximization problem yields that

$$w_{(p,m)} = \begin{cases} -E(\pi_S)_{(p,m)} & if \quad E(\pi_S)_{(p,m)} < 0 \\ 0 & \text{otherwise} \end{cases}.$$

Hence, the principal's ex ante utility is given by $E(u_P)_{(p,m)} = E(u_P)^{-w}_{(p,m)} - w_{(p,m)}$.

We note that here when $w_{(p,m)} > 0$, the supervisor's participation constraint is binding, and when $w_{(p,m)} = 0$, we have that $E(u_S) \geq 0$. This implies that the agent's participation constraint is always satisfied for $\phi \leq 1$, as $E(u_A) = \pi_A + \phi\pi_S = E(B) + \phi(E(u_S) - \phi E(B)) = \left(1 - \phi^2\right) E(B) + \phi E(u_S) \geq 0$, where $E(B)$ denotes the agent's expected benefits.

Claim 1. The claim that the equilibrium with full loyalty, $(p,m) = (1,0)$, is the principal's preferred equilibrium for all ρ is proven.

The above calculations yield that in all equilibria, $E(U_P)^{-w}_{(p,m)}$ is a convex combination of $X_b = 0$ and $X_g > 0$. Hence, the equilibria $(p,m) = (0,1)$, and $(p,m) = (1,0)$ yield the principal's highest *ex ante* pre-wage expected utility, X_g. Since $E(U_S)^{-w}_{(1,0)} = \phi B_g > 0$, $w_{(1,0)} = 0$ for all ρ. Hence, $E(U_P)_{(1,0)} = X_g$. This is strictly preferred to $(p,m) = (0,1)$ whenever $w_{(0,1)} > 0$, and weakly preferred to $(p,m) = (0,1)$ whenever $w_{(0,1)} = 0$. Since $E(U_P)^{-w}_{(p,m)}$ are non-degenerate convex combinations of X_b and X_g for all equilibria except $(p,m) = (1,0)$ and $(p,m) = (0,1)$, the equilibrium $(p,m) = (1,0)$ is always strictly preferred to all equilibria (p,m) such that $(p,m) \neq (0,1)$ when $w_{(p,m)} = 0$.

If the equilibrium $(p,m) = (1,0)$ is strictly preferred to another equilibrium (p,m) when $w_{(p,m)} = 0$, it is clearly strictly preferred to this equilibrium when $w_{(p,m)} > 0$.

Claim 2. The claim that there exists a $\hat{\rho} < \bar{\rho}$ such that for all $\rho \in [\hat{\rho}, \bar{\rho})$, the principal's second most preferred equilibrium is $(p,m) = (0,1)$ is proven.

When $w_{(p,m)} = 0$ in all equilibria, the above calculations (under Claim 1) show that $E(U_P)^{-w}_{(p,m)} = E(U_P)_{(p,m)}$ are convex combinations of $X_b = 0$ and $X_g > 0$ in all equilibria. Now let $\rho < \bar{\rho}$, i.e., $\rho < E(\tilde{c}) - \phi B_g$. Then, $w_{(0,1)} = E(\tilde{c}) - \phi B_g - \rho > 0$, and

$$E(U_P)_{(0,1)} = E(U_P)^{-w}_{(0,1)} - w_{(0,1)} = X_g - (E(\tilde{c}) - \phi B_g - \rho) < X_g.$$

As $w_{(0,1)}$ is decreasing in ρ,

$$\rho \to (E(\tilde{c}) - \phi B_g)^{-} \Rightarrow E(U_P)_{(0,1)} \to X_g. \tag{6.17}$$

Since $E(U_P)^{-w}_{(p,m)}$ are non-degenerate convex combinations of X_b and X_g for all equilibria except $(p,m) = (1,0)$ and $(p,m) = (0,1)$, by (6.17) there exists some $\hat{\rho} < \bar{\rho}$ such that for all $\rho \in [\hat{\rho}, \bar{\rho})$, the equilibrium $(p,m) = (0,1)$ is strictly preferred to all other equilibria (p,m) when $w_{(p,m)} = 0$ (except $(p,m) = (1,0)$, which follows directly from Claim 1, since we are concerned only with ρ such that $\rho < \bar{\rho}$).

If the equilibrium $(p,m) = (0,1)$ is strictly preferred to another equilibrium (p,m) when $w_{(p,m)} = 0$, it is clearly strictly preferred to this equilibrium when $w_{(p,m)} > 0$. This concludes the proof.

Claim 3. The claim that for $\rho \geq \bar{\rho}$, the principal is indifferent between $(p,m) = (1,0)$ and $(p,m) = (0,1)$ is proven.

If $\rho \geq \bar{\rho} \Leftrightarrow \rho \geq E(\tilde{c}) - \phi B_g$, then $w_{(0,1)} = 0$, and hence

$$E(U_P)^{-w}_{(0,1)} = E(U_P)^{-w}_{(1,0)} = E(U_P)_{(0,1)} = E(U_P)_{(1,0)} = X_g. \blacksquare$$

Proof of Proposition 2. $\frac{\partial\bar{\phi}(\rho)}{\partial\rho} = \frac{\partial\left(\frac{B_b^H - B_g}{\rho}\right)}{\partial\rho} = -\left(\frac{B_b^H - B_g}{\rho^2}\right) < 0.$ ■

Proof of Proposition 3. First, we need to show that, for any ρ, the range of ϕ for which the full monitoring equilibrium exists decreases in $\hat{c}$, i.e., that $\partial\underline{\phi}(\rho)/\partial\hat{c} < 0$. Since $\hat{c} \equiv E(\tilde{c}) = \beta\underline{c} + (1-\beta)\bar{c}$, we have that $\underline{\phi}(\rho) \equiv \frac{2\rho - \bar{c}}{\left[E(\tilde{B}_b) - B_g\right]} = \frac{2\rho - \left(\frac{\hat{c} - \beta c}{(1-\beta)}\right)}{\left[E(\tilde{B}_b) - B_g\right]} = \frac{2\rho + \frac{\beta c}{(1-\beta)}}{\left[E(\tilde{B}_b) - B_g\right]} - \hat{c}\frac{1}{\left[E(\tilde{B}_b) - B_g\right](1-\beta)}$, which yields that

$$\frac{\partial\underline{\phi}(\rho)}{\partial\hat{c}} = \frac{\partial\left(\frac{2\rho + \frac{\beta c}{(1-\beta)}}{\left[E(\tilde{B}_b) - B_g\right]} - \hat{c}\frac{1}{\left[E(\tilde{B}_b) - B_g\right](1-\beta)}\right)}{\partial\hat{c}} = -\frac{1}{\left[E\left(\tilde{B}_b\right) - B_g\right](1-\beta)} < 0.$$

Second, we need to show that, for a given $\phi < \underline{\phi}(\rho)$, the principal's utility in a full monitoring equilibrium decreases in $\hat{c}$, i.e., that $\partial E(u_P)/\partial\hat{c} \leq 0$ for $(p,m) = (0,1)$. For $(p,m) = (0,1)$, by the proof of Proposition 1, we have that

$$E(u_P)_{(0,1)} = E(u_P)_{(0,1)}^{-w} - w_{(0,1)} = \begin{cases} X_g - \left[-\left(-\hat{c} + \rho + \phi B_g\right)\right] & if \quad \left(-\hat{c} + \rho + \phi B_g\right) < 0 \\ X_g & \text{otherwise} \end{cases}.$$

If $E(u_P)_{(0,1)} = X_g - \left[-\left(-\hat{c} + \rho + \phi B_g\right)\right] = X_g - \hat{c} + \rho + \phi B_g$, then $\frac{\partial E(u_P)}{\partial\hat{c}} = -1 < 0$. If $E(u_P)_{(0,1)} = X_g$, then $\frac{\partial E(u_P)}{\partial\hat{c}} = 0$. Hence, $\frac{\partial E(u_P)}{\partial\hat{c}} \leq 0$ for $(p,m) = (0,1)$.

Third, we need to show that neither the range for which the full loyalty equilibrium exists nor the principal's utility in a full loyalty equilibrium depend on $\hat{c}$, i.e., that $\partial\bar{\phi}(\rho)/\partial\hat{c} = 0$ and that $\partial E(u_P)/\partial\hat{c} = 0$ for $(p,m) = (1,0)$. Since $\bar{\phi}(\rho) = \left(\frac{B_b^H - B_g}{\rho}\right)$, we have that $\frac{\partial\bar{\phi}(\rho)}{\partial\hat{c}} = \frac{\partial\left(\frac{B_b^H - B_g}{\rho}\right)}{\partial\hat{c}} = 0$. Furthermore, for $(p,m) = (1,0)$, by the proof of Proposition 1, we have that $E(u_P)_{(0,1)} = X_g$, so clearly $\frac{\partial E(u_P)}{\partial\hat{c}} = 0$ for $(p,m) = (1,0)$. ■

Proof of Proposition 4. For any given pair (ϕ,ρ), from the proof of Lemma 3, we recall that

$$\begin{array}{ll} \min P^*(\phi,\rho) = \begin{cases} 0 & if \quad \phi \leq \phi_5 \\ \alpha & if \quad \phi_5 < \phi \leq \phi_8 \\ 1 & if \quad \phi > \phi_8 \end{cases} & for \quad \phi > \phi_6 \\ \min P^*(\phi,\rho) = 0 & for \quad \phi \leq \phi_6 \end{array},$$

and that

$$\max P^*(\phi,\rho) = \begin{cases} 0 & if \quad \phi < \phi_7 \\ \alpha & if \quad \phi_7 \leq \phi < \phi_8 \\ 1 & if \quad \phi \geq \phi_8 \end{cases},$$

where $\phi_5 \equiv \frac{B_b^L - B_g}{\rho}$, $\phi_6 \equiv \frac{2\rho - \bar{c}}{E(\tilde{B}_b) - B_g}$, $\phi_7 \equiv \frac{B_b^L - B_g}{\rho(2\beta + \alpha(1-\beta))/(\beta + \alpha(1-\beta))}$, $\phi_8 \equiv \frac{B_b^H - B_g}{\rho}$, and $E\left(\tilde{B}_b\right) \equiv \alpha B_b^L + (1-\alpha)B_b^H$. We restrict attention to the (relevant) cases when $\phi_i > 0 \; \forall i \in \{5,8\}$.

We want to show that $\min P^*(\phi,\rho)$ is weakly decreasing in B_b^H or B_b^L. Thus we want to show that for $\phi > \phi_6$, $\frac{\partial\phi_5}{\partial B_b^H} \geq 0$, $\frac{\partial\phi_5}{\partial B_b^L} \geq 0$, $\frac{\partial\phi_8}{\partial B_b^H} \geq 0$, and $\frac{\partial\phi_8}{\partial B_b^L} \geq 0$. We obtain

$$\begin{aligned} \frac{\partial\phi_5}{\partial B_b^H} &= \frac{\partial\left(\frac{B_b^L - B_g}{\rho}\right)}{\partial B_b^H} = 0, \; \frac{\partial\phi_5}{\partial B_b^L} = \frac{\partial\left(\frac{B_b^L - B_g}{\rho}\right)}{\partial B_b^L} = \frac{1}{\rho} > 0, \\ \frac{\partial\phi_8}{\partial B_b^H} &= \frac{\partial\left(\frac{B_b^H - B_g}{\rho}\right)}{\partial B_b^H} = \frac{1}{\rho} > 0, \; \frac{\partial\phi_8}{\partial B_b^L} = \frac{\partial\left(\frac{B_b^H - B_g}{\rho}\right)}{\partial B_b^L} = 0. \end{aligned}$$

Hence, $\min P^*(\phi,\rho)$ is weakly decreasing in B_b^H or B_b^L.

We also want to show that $\max P^*(\phi,\rho)$ is weakly decreasing in B_b^H or B_b^L. Thus, in addition to $\frac{\partial\phi_8}{\partial B_b^H}\geq 0$ and $\frac{\partial\phi_8}{\partial B_b^L}\geq 0$, we want to show that $\frac{\partial\phi_7}{\partial B_b^H}\geq 0$ and $\frac{\partial\phi_7}{\partial B_b^L}\geq 0$. We obtain

$$\begin{aligned}\frac{\partial\phi_7}{\partial B_b^H} &= \frac{\partial\left(\frac{B_b^L-B_g}{\rho(2\beta+\alpha(1-\beta))/(\beta+\alpha(1-\beta))}\right)}{\partial B_b^H}=0\\ \frac{\partial\phi_7}{\partial B_b^L} &= \frac{\partial\left(\frac{B_b^L-B_g}{\rho(2\beta+\alpha(1-\beta))/(\beta+\alpha(1-\beta))}\right)}{\partial B_b^L}=\frac{1}{\rho\left(2\beta+\alpha\left(1-\beta\right)\right)/\left(\beta+\alpha\left(1-\beta\right)\right)}>0.\end{aligned}$$

Hence, $\max P^*(\phi,\rho)$ is weakly decreasing in B_b^H or B_b^L.

For any given pair (ϕ,ρ), from the proof of Lemma 2, we recall that

$$\begin{array}{ll}\min M^*(\phi,\rho)=\left\{\begin{array}{lll}1 & if & \phi<\phi_2,\\ \beta & if & \phi_2\leq\phi<\phi_4\\ 0 & if & \phi\geq\phi_4\end{array}\right. & if\quad \phi_2<\phi_4\\ \min M^*(\phi,\rho)=\left\{\begin{array}{ll}1 & \phi<\phi_4,\\ 0 & \phi\geq\phi_4\end{array}\right. & if\quad \phi_2\geq\phi_4\end{array},$$

and that

$$\max M^*(\phi,\rho)=\left\{\begin{array}{lll}1 & if & \phi\leq\phi_1,\\ \beta & if & \phi_1<\phi\leq\phi_3\\ 0 & if & \phi>\phi_3\end{array}\right.,$$

where $\phi_1\equiv\frac{2\rho-\overline{c}}{E(\tilde{B}_b)-B_g}$, $\phi_2\equiv\frac{\left(1+\frac{\beta}{\beta+\alpha(1-\beta)}\right)\rho-\overline{c}/(1-\alpha)}{\left(B_b^H-B_g\right)}$, $\phi_3\equiv\frac{2\rho-\underline{c}}{E(\tilde{B}_b)-B_g}$, $\phi_4\equiv\min\left\{\frac{\rho-\underline{c}/(1-\alpha)}{B_b^H-B_g},\frac{B_b^H-B_g}{\rho}\right\}$, and $E\left(\tilde{B}_b\right)\equiv\alpha B_b^L+\left(1-\alpha\right)B_b^H$. We restrict attention to the (relevant) cases when $\phi_i>0\ \forall i\in\{1,4\}$.

We want to show that $\max M^*\left(\phi,\rho\right)$ is weakly decreasing in B_b^H or B_b^L. Thus, we want to show that $\frac{\partial\phi_1}{\partial B_b^H}\leq 0$, $\frac{\partial\phi_1}{\partial B_b^L}\leq 0$, $\frac{\partial\phi_3}{\partial B_b^H}\leq 0$, and that $\frac{\partial\phi_3}{\partial B_b^L}\leq 0$. We obtain

$$\begin{aligned}\frac{\partial\phi_1}{\partial B_b^H} &= \frac{\partial\left(\frac{2\rho-\overline{c}}{\alpha B_b^L+(1-\alpha)B_b^H-B_g}\right)}{\partial B_b^H}=\frac{-\left(1-\alpha\right)\left(2\rho-\overline{c}\right)}{\left(\alpha B_b^L+\left(1-\alpha\right)B_b^H-B_g\right)^2}<0\\ \frac{\partial\phi_1}{\partial B_b^L} &= \frac{-\alpha\left(2\rho-\overline{c}\right)}{\left(\alpha B_b^L+\left(1-\alpha\right)B_b^H-B_g\right)^2}<0\\ \frac{\partial\phi_3}{\partial B_b^H} &= \frac{\partial\left(\frac{2\rho-\underline{c}}{\alpha B_b^L+(1-\alpha)B_b^H-B_g}\right)}{\partial B_b^H}=\frac{-\left(1-\alpha\right)\left(2\rho-\underline{c}\right)}{\left(\alpha B_b^L+\left(1-\alpha\right)B_b^H-B_g\right)^2}<0\\ \frac{\partial\phi_3}{\partial B_b^L} &= \frac{-\alpha\left(2\rho-\underline{c}\right)}{\left(\alpha B_b^L+\left(1-\alpha\right)B_b^H-B_g\right)^2}<0.\end{aligned}$$

where the inequalities hold whenever $\phi_1>0$, since the fact that $\left(E\left(\tilde{B}_b\right)-B_g\right)>0$ yields that $(2\rho-\overline{c})>0$ whenever $\phi_1>0$. Hence, $\max M^*\left(\phi,\rho\right)$ is weakly decreasing in B_b^H or B_b^L.

We also want to show that if $\min\left\{\frac{\rho-\underline{c}/(1-\alpha)}{B_b^H-B_g},\frac{B_b^H-B_g}{\rho}\right\}=\frac{\rho-\underline{c}/(1-\alpha)}{B_b^H-B_g}$, then $\min M^*(\phi,\rho)$ is weakly decreasing in B_b^H or B_b^L. Thus, under the condition $\phi_4\equiv\frac{\rho-\underline{c}/(1-\alpha)}{B_b^H-B_g}$, we want to show that $\frac{\partial\phi_4}{\partial B_b^H}\leq 0$, $\frac{\partial\phi_4}{\partial B_b^L}\leq 0$, and (if $\phi_2<\phi_4$), that $\frac{\partial\phi_2}{\partial B_b^H}\leq 0$ and $\frac{\partial\phi_2}{\partial B_b^L}\leq 0$. We obtain

$$\frac{\partial\phi_4}{\partial B_b^H}=\frac{\partial\left(\frac{\rho-\underline{c}/(1-\alpha)}{B_b^H-B_g}\right)}{\partial B_b^H}=\frac{-\left(\rho-\underline{c}/\left(1-\alpha\right)\right)}{\left(B_b^H-B_g\right)^2}<0,\ \frac{\partial\phi_4}{\partial B_b^L}=0,$$

where the inequality holds whenever $\phi_4 > 0$, since the fact that $B_b^H - B_g > 0$ yields that $\rho - \underline{c}/\left(1-\alpha\right) > 0$ whenever $\phi_4 > 0$. Furthermore, we obtain

$$\frac{\partial \phi_2}{\partial B_b^H} = \frac{\partial \left(\frac{\left(1+\frac{\beta}{\beta+\alpha(1-\beta)}\right)\rho - \bar{c}/(1-\alpha)}{\left(B_b^H - B_g\right)} \right)}{\partial B_b^H} = \frac{-\left(\rho \frac{2\beta+\alpha(1-\beta)}{\beta+\alpha(1-\beta)} - \bar{c}/\left(1-\alpha\right)\right)}{\left(B_b^H - B_g\right)^2} < 0, \frac{\partial \phi_2}{\partial B_b^L} = 0,$$

where the inequality holds whenever $\phi_2 > 0$, since the fact that $B_b^H - B_g > 0$ yields that $\rho \frac{2\beta+\alpha(1-\beta)}{\beta+\alpha(1-\beta)} - \bar{c}/\left(1-\alpha\right) > 0$ whenever $\phi_2 > 0$. Hence, if $\min\left\{\frac{\rho - \underline{c}/(1-\alpha)}{B_b^H - B_g}, \frac{B_b^H - B_g}{\rho}\right\} = \frac{\rho - \underline{c}/(1-\alpha)}{B_b^H - B_g}$, then $\min M^*(\phi, \rho)$ is weakly decreasing in B_b^H or B_b^L. ■

References

Adams, Renée and Daniel Ferreira, 2007, "A Theory of Friendly Boards," *Journal of Finance*, 62(1), 217-250.

Alger, Ingela and Jörgen W. Weibull, 2009, "Kinship, Incentives and Evolution," working paper.

Almeida, Heitor V. and Daniel Wolfenzon, 2006, "A Theory of Pyramidal Ownership and Family Business Groups," *Journal of Finance*, 61(6), 2637-2679.

Ashraf, Nava, Iris Bohnet and Nikita Piankov, 2008, "Decomposing Trust and Trustworthiness," working paper.

Becht, Marco, Julian Franks, Colin Mayer, and Stefano Rossi, 2008, "Returns to Shareholder Activism. Evidence from a Clinical Study of the Hermes U.K. Focus Fund," *Review of Financial Studies*, forthcoming.

Becker, Gary, 1974, "A Theory of Social Interactions," *Journal of Political Economy*, 82(6), 1063-1093.

Bertrand, Marianne and Antoinette Schoar, 2006, "The Role of Family in Family Firms," *Journal of Economic Perspectives*, 20(2), 73-96.

Bertrand, Marianne, Simon Johnson, Krislert Samphantharak, and Antoinette Schoar, 2008, "Mixing family with business: A study of Thai business groups and the families behind them," *Journal of Financial Economics*, 88, 466-498.

Bertrand, Marianne, Francis Kramarz, Antoinette Schoar, and David Thesmar, 2007, "Politicians, Firms and the Political Business Cycle: Evidence from France," working paper.

Bian, Yanjie, 1997, "Bringing Strong Ties Back In: Indirect Ties, Network Bridges, and Job Searches in China," *American Sociological Review*, 62, 366-85.

Becker, Gary, 1971, *The Economics of Discrimination*, 2nd ed., Chicago: University of Chicago Press.

Black, Bernard S., "Shareholder Activism and Corporate Governance in the United States, in: Peter Newman (ed.), 1998, *The New Palgrave Dictionary of Economics and the Law*, Vol. 3, 459-463.

Brav, Alon, Wei Jiang, Frank Partnoy, and Thomas Randall, 2008, "Hedge Fund Activism, Corporate Governance, and Firm Performance," *Journal of Finance*, 63(4), 1729-1775.

Brickley, J.A., Coles, J.L., Linck, J., 1999, "What happens to CEOs after they retire? New evidence on career concerns, horizon problems, and CEO incentives," *Journal of Financial Economics*, 52, 341-377.

British Broadcasting Corporation (BBC), 2009, "The blame game starts at Davos," January 29, accessed online on February 28, 2009.

Burkart, Michael, Fausto Panunzi and Andrei Shleifer, 2003, "Family Firms," *Journal of Finance*, 58(5), 2167–2202.

Callahan, William A., 2005, "Social Capital and Corruption: Vote Buying and the Politics of Reform in Thailand," *Perspectives on Politics*, 3(3), 495-508.

Charumilind, Chutatong, Raja Kali, and Yupana Wiwattanakantang, 2006, "Connected Lending: Thailand before the Financial Crisis," *Journal of Business*, 79(1), 181-218.

Claessens, Stijn, Simeon Djankov, Larry H.P. Lang, 2000, "The separation of ownership and control in East Asian Corporations," *Journal of Financial Economics*, 58, 81-112.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2008a, "The Small World of Investing: Board Connections and Mutual Fund Returns" *Journal of Political Economy*, 116, 951-979.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2008b, "Sell Side School Ties," working paper.

Coleman, James, 1990, *Foundations of Social Theory*, Cambridge: Harvard University Press.

Coles, J.L., Hoi, C-K., 2003, "New evidence on the market for directors: Board membership and Pennsylvania senate bill 1310," *Journal of Finance*, 58, 197-230.

Dana, Jason D., Dalian M. Cain and Robin M. Dawes, 2006, "What you don't Know Won't Hurt me: Costly (but quiet) Exit in a Dictator Game," *Organizational Behavior and Human Decision Processes*, 100(2), 193-201.

Diamond, Douglas W., 1984, "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393-414.

Dillenberger, David and Philipp Sadowski, 2008, "Ashamed to be Selfish," PIER Working Paper 08-037.

Dyck, Alexander, and Luigi Zingales, 2002, "The corporate governance role of the media," in Roumeen Islam, ed., *The Right to Tell: The Role of the Media in Development*, Washington D.C.: The World Bank

Dyck, Alexander, Natalya Volchkova, and Luigi Zingales, 2008, "The Corporate Governance Role of the Media: Evidence from Russia," *Journal of Finance*, 63(3), 1093-1135.

Fama, Eugene, 1980, "Agency problems and the theory of the firm," *Journal of Political Economy*, 88, 288-303.

Fama, Eugene, and Michael Jensen, 1983, "The separation of ownership and control," *Journal of Law and Economics*, 26, 301-325.

Fan, Joseph P.H., Yupana Wiwattanakantang, and Pramuan Bunkanwanicha, 2008, "Why Do Shareholders Value Marriage?" ECGI-Finance Working Paper No. 227.

Fehr, Ernst and Armin Falk, 2002, "Psychological Foundations of Incentives," *European Economic Review*, 46, 687-724.

Fehr, Ernst and Klaus M. Schmidt, 2006, "The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories," in: Serge-Christophe Kolm and Jean Mercier Ythier (eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity*, Vol. 1, Elsevier B.V.

Fehr, Ernst, Martin Brown, and Christian Zehnder, 2008, "On Reputation - A Microfoundation of Contract Enforcement and Price Rigidity," *Economic Journal*, forthcoming.

Fershtman, Chaim, Uri Gneezy, and Frank Verboven, 2005, "Discrimination and Nepotism: The Efficiency of the Anonymity Rule," Journal of Legal Studies, 34, 371-394.

Fich, Eliezer M., 2005, "Are some outside directors better than others? Evidence from director appointments by Fortune 1000 firms," *Journal of Business*, 78, 1943-1972.

Fich, Eliezer M. and Anil Shivdasani, 2007, "Financial fraud, director reputation, and shareholder wealth," *Journal of Financial Economics*, 86, 306-336.

Fock, Henry K.Y. and Kashing Woo, 1998, "The China Market: Strategic Implications of Guanxi," *Business Strategic Review*, 9(3), 33–43.

Freeman, Richard B., 1997 "Working for Nothing: The Supply of Volunteer Labor," *Journal of Labor Economics*, 15(1), S140-S166.

Gächter, Simon and Armin Falk, 1999, "Collective actions as a social exchange," *Journal of Economic Behavior and Organization*, 39, 341-369.

Gilson, Ronald J., 2006, "Controlling Shareholders and Corporate Governance: Complication the Comparative Taxonomy," *Harvard Law Review*, 119(6), 1641-1679.

Gilson, Ronald J., 2007, "Controlling Family Shareholders in Developing Countries: Anchoring Relational Exchange," Stanford Law and Economics Olin Working Paper, No. 333 / Columbia Law and Economics Working Paper, No. 311.

Gilson, S., 1990, "Bankruptcy, boards, banks and blockholders: Evidence on changes in corporate ownership and control when firms default," *Journal of Financial Economics*, 27, 355-387.

Granovetter, Mark, 1974, *Getting a Job: A Study of Contacts and Careers*, Cambridge, MA: Harvard University Press.

Granovetter, Mark, 1973, "The Strength of Weak Ties," American Journal of Sociology, 78(6), 1360-1380.

Granovetter, Mark, 1985, "Economic Action and Social Structure: The Problem of Embeddedness," *American Journal of Sociology*, 91(3), 481-510.

Harford, J., 2003, "Takeover bids and target directors' incentives: The impact of a bid on directors' wealth and board seats," *Journal of Financial Economics*, 69, 51–83.

Harris, Donna, 2007, "Bonding Social Capital and Corruption: A Cross-National Empirical Analysis," working paper.

Harris, Milton and Artur Raviv, 2008, "A Theory of Board Control and Size," *Review of Financial Studies*, 21(4), 1797-1832.

Harsanyi, John C., 1969, "Rational choice models of political behavior vs. functionalist and conformist theories," *World Politics*, 21, 513-538.

Holmstrom, Bengt and Jean Tirole, 1997, "Financial Intermediation, Loanable Funds, and the Real Sector," *Quarterly Journal of Economics*, 112, 663-691.

Homans, George, 1950, *The Human Group*, New York: Harcourt, Brace & World.

Hwang, Byoung-Hyoun and Seoyoung Kim, 2008, "It pays to have friends," *Journal of Financial Economics*, forthcoming.

Olson, Mancur, 1982, *The Rise and Decline of Nations*, New Haven: Yale University Press.

Kaplan, S. and D. Reishus, 1990, "Outside directorships and corporate performance," *Journal of Financial Economics*, 27, 389-410.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl, 2008, "Trust and Social Collateral," *Quarterly Journal of Economics*, forthcoming.

Khanna, Tarun and Yishay Yafeh, 2007, "Business Groups in Emerging Markets: Paragons or Parasites?," Journal of Economic Literature, 45, 331–372.

Kramarz, Francis and David Thesmar, 2007, "Social Networks in The Boardroom," working paper.

La Porta, Rafael, Florencio Lopez-De-Silanes, and Andrei Shleifer, 1999, "Corporate Ownership Around the World," *Journal of Finance*, 54(2), 471–518.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Guillermo Zamarripa, 2003, "Related Lending", *Quarterly Journal of Economics*, 118(1), 231-268.

Leff, Nathaniel, 1978, "Industrial organization and entrepreneurship in the developing countries: The economic groups," *Economic Development and Cultural Change*, 26, 661–675.

Mallaby, S., 2004, *The World's Banker: A Story of Failed States, Financial Crises, and the Wealth and Poverty of Nations,* Penguin Press, New York, NY.

Masclet, D., Noussair, C., Tucker, S., Villeval, M.-C., 2001, "Monetary and Non-Monetary Punishment in the Voluntary Contributions Mechanism," Purdue University, Krannert Graduate School of Management, USA, Working paper.

Milgrom, Paul R., Roberts, John, 1992, *Economics, Organization and Management*, Prentice-Hall, Englewood Cliffs, NJ.

Morck, Randall, Daniel Wolfenzon, and Bernard Yeung, 2005, "Corporate Governance, Economic Entrenchment, and Growth," *Journal of Economic Literature*, XLIII, 655-720.

Neilson, William S., 2008, "A theory of kindness, reluctance, and shame for social preferences," *Games and Economic Behavior*, forthcoming.

New York Times, 2001, "Enron Board Comes Under a Storm of Criticism," December 16, accessed online on February 28, 2009.

New York Times, 2008, "At Freddie Mac, Chief Discarded Warning Signs," August 5, accessed online on February 28, 2009.

Petersen, Mitchell A. and Raghuram G. Rajan, 1994, "The Benefits of Lending Relationships: Evidence from Small Business Data." *Journal of Finance*, 49(1), 3-37.

Portes, Alejandro, 1998, "Social Capital: its origins and applications in modern sociology," *Annual Review of Sociology*, 24, 1-24.

Portes, Alejandro and Patricia Landolt, 1996, "The downside of social capital," *American Prospect*, 26, 18-21, 94.

Portes, Alejandro and Julia Sensenbrenner, 1993, "Embeddedness and immigration: notes on the social determinants of economic action," *American Journal of Sociology*, 98(6), 1320-1350.

Putnam, Robert, 1993, *Making Democracy Work*, Princeton: Princeton University Press.

Rege, Mari and Kjetil Telle, 2004, "The impact of social approval and framing on cooperation in public good situations", *Journal of Public Economics*, 88, 1625-1644.

Reuters Blogs, 2009, "Wall Street bankers — so humble, so frugal!!!," February 11, accessed online on February 28, 2009.

Skeel, Jr., David A., 2001, "Shaming in Corporate Law," *University of Pennsylvania Law Review*, 149, 1811-1868.

Silva, Francisca Nicolás Majluf, and Ricardo D. Paredes, 2006, "Family ties, interlocking directors and performance of business groups in emerging countries: The case of Chile," Journal of Business Research, 59, 315-321.

Standifird, Steve S., and R. Scott Marshall, 2000, "The Transaction Cost Advantage of Guanxi-Based Business Practices," *Journal of World Business*, 35(1), 21–42.

Telegraph, 2009, "Ken Lewis calls for Wall Street bankers to find some humility before Congress," February 9, accessed online on February 28, 2009.

Uzzi, Brian, 1999, "Embeddedness in the Making of Financial Capital: How Social Relations and Networks Benefit Firms Seeking Finance," *American Sociological Review*, 64, 481–505.

Yermack, David, 2004, "Remuneration, retention, and reputation incentives for outside directors," *Journal of Finance*, 59, 2281-2308.

EFI THE ECONOMIC RESEARCH INSTITUTE

EFI Mission

EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution that works independently of economic, political and sectional interests. It conducts theoretical and empirical research in the management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics. At EFI, the researchers select their projects based on the need for theoretical or practical development of a research domain, on their methodological interests, and on the generality of a problem.

Research Organization

The research activities at the Institute are organized into 20 Research Centres.
Centre Directors are professors at the Stockholm School of Economics.

EFI Research Centre:	*Centre Director:*
Management and Organization (A)	Sven-Erik Sjöstrand
Centre for Entrepreneurship and Business Creation (E)	Carin Holmquist
Public Management (F)	Nils Brunsson
Information Management (I)	Mats Lundeberg
Centre for People and Organization (PMO)	Andreas Werr
Centre for Innovation and Operations Management (T)	Pär Åhlström
Centre for Media and Economic Psychology (P)	Richard Wahlund
Centre for Consumer Marketing (CCM)	Magnus Söderlund
Centre for Information and Communication Research (CIC)	Per Andersson
Marketing, Distribution and Industry Dynamics (D)	Björn Axelsson
Centre for Strategy and Competitiveness (CSC)	Örjan Sölvell
Accounting and Managerial Finance (B)	Johnny Lind
Centre for Financial Analysis and Managerial Economics in Accounting (BFAC)	Kenth Skogsvik
Finance (FI)	Clas Bergström
Centre for Health Economics (CHE)	Magnus Johannesson
International Economics and Geography (IEG)	Mats Lundahl
Economics (S)	Paul Segerstrom
Economic Statistics (ES)	Anders Westlund
Centre for Business Law (RV)	Johnny Herre
Centre for Tax Law (SR)	Bertil Wiman

Chair of the Board: Professor Carin Holmquist
Director: Associate Professor Filip Wijkström

Address

EFI, Box 6501, SE-113 83 Stockholm, Sweden • Website: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

EFI, The Economic Research Institute

Published in the language indicated by the title.
A complete publication list can be found at *www.hhs.se/efi*
Books and dissertations can be ordered from EFI via e-mail: EFI.Publications@hhs.se

Reports since 2006

2009

Dissertations

Hellström, Katerina. *Financial accounting quality in a European transition economy: the case of the Czech republic.*
Hernant, Mikael. *Profitability performance of supermarkets: the effects of scale of operation, local market conditions, and conduct on the economic performance of supermarkets.*
Lakomaa, Erik. *The economic psychology of the welfare state.*
Melén, Sara. *New insights on the internationalisation process of SMEs: a study of foreign market knowledge development.*
Törn, Fredrik. *Challenging consistency: effects of brand-incongruent communications.*
Wennberg, Karl. *Entrepreneurial Exit.*

2008

Books

Breman, Anna. *Forskning om filantropi. Varför skänker vi bort pengar?* Forskning i Fickformat.
Einarsson, Torbjörn. *Medlemskapet i den svenska idrottsrörelsen: En studie av medlemmar i fyra idrottsföreningar.* EFI Civil Society Reports.
Helgesson, Claes-Fredrik and Hans Winberg (eds). *Detta borde vårddebatten handla om.*
Jennergren, Peter, Johnny Lind, Walter Schuster and Kenth Skogsvik (eds). *Redovisning i fokus.* EFI:s Årsbok 2008. EFI/Studentlitteratur.
Kraus, Kalle. *Sven eller pengarna? Styrningsdilemman i äldrevården.* Forskning i Fickformat.
Petrelius Karlberg, Pernilla. *Vd under press: om medialiseringen av näringslivets ledare.* Forskning i Fickformat.
Portnoff, Linda. *Musikbranschens styrningsproblematik.* Forskning i Fickformat.
Sjöstrand, Sven-Erik. *Management: från kontorsteknik till lednings- och organisationsteori: utvecklingen på Handelshögskolan under 100 år: 1909–2009.*
Östman, Lars. *Den finansiella styrningens realiteter och fiktioner: de finansiella styrformernas svenska historia, berättelser om Petersson och "Ericsson", finansiell styrning – en ansats till generell teori.*
Östman, Lars. *Mycket hände på vägen från Buchhaltung till Accounting: delar av Handelshögskolan under 100 år.*

Dissertations

Axelson, Mattias. *Enabling knowledge communication between companies: the role of integration mechanisms in product development collaborations.*
Benson, Ilinca. *Organisering av övergångar på arbetsmarknaden: en studie av omställningsprogram.*
Elhouar, Mikael. *Essays on interest rate theory.*
Farooqi Lind, Raana. *On capital structure and debt placement in Swedish companies.*
Granström, Ola. *Aid, drugs, and informality: essays in empirical economics.*
Hvenmark, Johan. *Reconsidering membership: a study of individual members' formal affiliation with democratically governed federations.*
Höglin, Erik. *Inequality in the labor market: insurance, unions, and discrimination.*

Johansson, Marjana. *Engaging resources for cultural events: a performative view.*
Kallenberg, Kristian. *Business at risk. Four studies on operational risk anagement.*
Kviselius, Niklas Z. *Trust-building and communication in SME internationalization: a study of Swedish-Japanese business relations.*
Landberg, Anders. *New venture creation: resistance, coping and energy.*
Pemer, Frida. *Framgång eller fiasko? En studie av hur konsultprojekt värderas i klientorganisationer.*
Rosengren, Sara. *Facing clutter: on message competition in marketing communication.*
Schilling, Annika. *Kan konsulter fusionera?: en studie av betydelsen av identitet vid en fusion mellan konsultföretag.*
Schriber, Svante. *Ledning av synergirealisering i fusioner och förvärv.*
Sjödin, Henrik. *Tensions of extensions: adverse effects of brand extension within consumer relationship.*
Strandqvist, Kristoffer. *Kritiska år: formativa moment för den svenska flygplansindustrin 1944–1951.*
Strömqvist, Maria. *Hedge funds and international capital flow.*
Söderström, Johan. *Empirical studies in market efficiency.*
Sölvell, Ingela. *Formalization in high-technology ventures.*
Thorsell, Håkan. *The pricing of corporate bonds and determinants of financial structure.*
Ulbrich, Frank. *The adoption of IT-enabled management ideas: insights from shared services in government agencies.*
Östling, Robert. *Bounded rationality and endogenous preferences.*

2007

Books

Andersson, Per, Ulf Essler and Bertil Thorngren (eds). *Beyond mobility.* EFI Yearbook 2007. EFI/Studentlitteratur.
Einarsson, Torbjörn and Filip Wijkström. *Analysmodell för sektorsöverskridande statistik: fallet vård och omsorg.* EFI Civil Society Reports.
Ericsson, Daniel. *Musikmysteriet: organiserade stämningar och motstämningar.*
Samuelson, Lennart (ed). *Bönder och bolsjeviker: den ryska landsbygdens historia 1902–1939.*

Dissertations

Ahlersten, Krister. *Empirical asset pricing and investment strategies.*
Alexius, Susanna. *Regelmotståndarna: om konsten att undkomma regler.*
Andersson, Magnus. *Essays in empirical finance.*
Berg, Bengt Åke. *Volatility, integration and grain bank: studies in harvests, rye prices and institutional development of the parish magasins in Sweden in the 18th and 19th centuries.*
Bianchi, Milo. *Of speculators, migrants and entrepreneurs: essays on the economics of trying your fortune.*
Brodin, Karolina. *Consuming the commercial break: an ethnographic study of the potential audiences for television advertising.*
Elger, Max. *Three essays on investment-specific technical change.*
Hagberg, Axel. *Bankkrishantering: aktörer, marknad och stat.*
Hinnerich, Mia. *Derivatives pricing and term structure modeling.*
Hjalmarson, Hanna. *En växande marknad: studie av nöjdheten med konsumtionsrelaterade livsområden bland unga konsumenter.*
Hjelström, Tomas. *The closed-end investment company premium puzzle: model development and empirical tests on Swedish and British data.*
Kraus, Kalle. *Sven, inter-organisational relationships and control: a case study of domestic care of the elderly.*
Lindqvist, Erik. *Essays on privatization, identity, and political polarization.*
Macquet, Monica. *Partnerskap för hållbar utveckling: systrar av Oikos och guvernanten som blev diplomat.*
Melian, Catharina. *Progressive open source.*

Nilsson, Daniel. *Transactions in cyberspace: the continued use of Internet banking.*
Petrelius Karlberg, Pernilla. *Den medialiserade direktören.*
Portnoff, Linda. *Control, cultural production and consumption: theoretical perspectives, empirical dilemmas, and Swedish music industry practices.*
Sköld, Martin. *Synergirealisering: realisering av produktsynergier efter företagssammanslagningar.*
Sonnerby, Per. *Contract-theoretic analyses of consultants and trade unions.*
Tyrefors, Björn. *Institutions, policy and quasi-experimental evidence.*
Valiente, Pablo. *Re-innovating the existing: a study of wireless IS capabilities to support mobile workforces.*

2006

Books

Lundeberg, Mats, Pär Mårtensson and Magnus Mähring (eds) *IT & business performance: a dynamic relationship.* EFI Yearbook 2006. EFI / Studentlitteratur.
Thodenius, Björn. *Organisering av kunskap: en studie av Wallenberg Consortium North.* EFI Civil Society Reports.
Wijkström, Filip and Torbjörn Einarsson. *Från nationalstat till näringsliv?: det civila samhällets organisationsliv i förändring.*
Wijkström, Filip, Stefan Einarsson and Ola Larsson. *Staten och det civila samhället: idétraditioner och tankemodeller i den statliga bidragsgivningen till ideella organisationer.*
Östman, Lars. *Lysande ögonblick och finansiella kriser: Dramaten under ett sekel.*

Dissertations

Argenton, Cedric. *Quality provision in duopoly.*
Beckerman, Carina. *The clinical eye: constructiong and computerizing an anesthesia patient record.*
Borglund, Tommy. *Aktievärden i fokus: internationell påverkan på intressentrelationer vid förvärv och fusion.*
Breman, Anna. *The Economics of altruism, paternalism and self-control.*
Edquist, Harald. *Technological breakthroughs and productivity growth.*
Eklund, Jana. *Essays on forecasting and bayesian model averaging.*
Frostenson, Magnus. *Legitimitetskontrollen: en studie av etiska värderingars roll i gränsöverskridande förvärv och fusioner.*
Gaspar, Raquel M. *Credit risk and forward price models.*
Gustafsson, Peter. *Essays on trade and technological change.*
Hopkins, Elisabeth. *Is a higher degree of local currency pricing associated with lower exchange rate pass-through?: a study of import pricing in 51 Swedish industries.*
Kling, Ragnar. *Developing product development in times of brutal change.*
Langenskiöld, Sophie. *Peer influence on smoking: causation or correlation?*
Lychnell, Lars-Olof. *"Och fungerar det inte, gör vi på något annat sätt": en klinisk fallstudie av IT-relaterat förändringsarbete i småföretag*
Meitz, Mika. *Five contributions to econometric theory and the econometrics of ultra-high-frequency data.*
Mendicino, Caterina. *Financial market imperfections, business cycle fluctuations and economic growth.*
Ovanfors, Anna. *Essays on nonlinear time series analysis and health economics.*
Paltseva, Elena. *Essays on commitment and inefficiency in political economy.*
Rogberg, Martin. *Den modeföljande organisationen: om acceptansen av TQM och andra populära managementmodeller.*
Silvennoinen, Annastiina. *Essays on autoregressive conditional heteroskedasticity.*
Sjögren, Ebba. *Reasonable drugs: making decisions with ambiguous knowledge.*
Slinko, Irina. *Essays in option pricing and interest rate models.*
Wilander, Fredrik. *Essays on exchange rates and prices.*