Advising and Monitoring CEOs: The Dual Role of Boards*

Tim Baldenius, Nahum Melumad and Xiaojing Meng

Columbia Business School

May 13, 2010

*We would like to thank seminar participants at Carnegie Mellon University, Columbia University, and the University of Houston, as well as Wouter Dessein, Jon Glover, Navin Kartik, Ken Koppelmann, and Shiva Sivaramakrishnan for comments and suggestions.

Abstract

Advising and Monitoring CEOs: The Dual Role of Boards

The board of directors performs the dual role of monitoring and advising the firm's management, and at times makes certain key decisions itself. We study the optimal board composition (of monitoring and advisory "types") within a cheap-talk framework where the CEO and the board each may have private information about an impending investment decision, and their incentives are imperfectly aligned. When shareholders choose both the board composition and the allocation of decision rights between CEO and board, a non-monotonic relationship between CEO bias and board composition emerges. Counter to conventional wisdom, we show that powerful CEOs who nominate board members themselves may in fact prefer a greater degree of monitoring intensity on the board than the shareholders do. As a result, regulatory interventions (such as the Sarbanes-Oxley Act) that attempt to strengthen the compliance role of boards, will in fact be harmful in precisely those cases where agency problems are the most severe. Lastly, CEOs may be able to entrench themselves by choosing "complex" projects involving greater information advantage. In response to the threat of CEO entrenchment, shareholders may commit to a board composed of mostly advisory types.

1 Introduction

Boards of directors perform the dual function of monitoring the firm's management and advising it on key decisions. In light of recent corporate governance failures, commentators have questioned the effectiveness of boards in performing these tasks—in particular, the monitoring task. Hermalin and Weisbach (1998), among others, have highlighted the potential problem of CEOs having significant influence over their board's composition. The Sarbanes-Oxley Act (SOX) of 2002 was designed in part to strengthen the monitoring function of the board possibly at the expense of the advisory role. In this paper we develop a model to study the link among the optimal board composition (holding the board size constant), the associated information environment, the assignment of decision-making authority, and the information transmission process inside a firm.

We consider a firm that faces an investment decision. The CEO of the firm has private information about the optimal investment choice but is an empire builder. Our model permits two types of board members: *monitoring* and *advisory* types. The primary task of the former is to monitor the CEO, whereas that of the advisory types is to discover decision-relevant information incremental to that privately held by the CEO. The information collectively available to the players, thus, is endogenously determined by the board composition: the greater is the fraction of advisory types on the board, the more likely it is that "new" information about the optimal decision will be discovered, but the less likely it is that the board learns the private information of the CEO (as there will be fewer monitors on the board). If the board learns the CEO's information, then the CEO cannot bias the investment decision, even if the decision is *formally* delegated to him—i.e., the board will then assume *real* authority (Aghion and Tirole 1997).

Consistent with earlier models that have treated the information environment as exogenous (e.g., Dessein 2002, Harris and Raviv 2005, Adams and Ferreira 2007), the board in our setting is more willing to delegate decision-making power, the less biased is the CEO.[1] This holds regardless of the board composition. Exploiting the endogenous nature of information in our model, we show that centralized firms tends to emphasize monitoring less than firms that delegate decisions to the CEO; conversely, boards become more willing to delegate, the more monitor-heavy they are. Under centralization the board can perfectly adjust the investment decision to its own information, whereas under delegation there is an information loss in the process of noisy cheap-talk communication—hence the board's information is more valuable under centralization. A more monitor-heavy board is less likely to discover such information, which lowers the opportunity cost associated with delegation.

In reality, of course, firms will choose board structure and the allocation of decision rights jointly. In that case, we establish a non-monotonic relationship between CEO bias and board composition. For small biases, as those increase, firms will stick to delegation, but will add monitors to the board. As the CEO's bias increases further, the firm will eventually centralize investment decisions. This in turn increases the value of board-generated information, which results in a drop in the fraction of monitoring types on the board. As CEO bias increases even further, the firm will again add more monitoring types to the board, because a highly biased CEO will submit very noisy signals about his information to the board. Thus, the board has greater incentive to uncover the CEO's information directly via monitoring. Hence, the relationship between agency problems and monitoring intensity of boards may be *locally decreasing* in the data, if researchers fail to control for the allocation of decision rights within firms.

[1] Harris and Raviv (2008) also endogenize the board's information, but only in regard to discovering new information. The board in their model never makes an attempt at uncovering the CEO's information.

An often voiced concern is that CEOs, who can influence the board composition, nominate few monitoring types to the board in order to have an agreeable board. One key aspect of the SOX act of 2002 was an attempt to enhance the monitoring role of boards. Ironically, our results suggest that for SOX to increase shareholder value, CEO bias has to be small. Only then does the conventional wisdom prevail that the CEO prefers fewer monitors on the board than the shareholders. In contrast, more biased CEOs may have incentives to nominate *more* monitors to the board than the shareholders would, because a monitor-heavy board is more likely to delegate. In other words, the CEO prefers to be in control of the decision (even if under greater scrutiny) over having the decision centralized by an "activist" board that may acquire valuable information of its own. SOX therefore may have exacerbated the distortions relative to the shareholder value-maximizing board composition.[2]

Short of choosing the board, CEOs have more subtle means of entrenching themselves and securing real authority within the firm, e.g., by choosing investment projects that confer them an information advantage.[3] We refer to such projects as "complex" ones. To prevent such entrenchment, shareholders will be better off—under certain conditions, strictly so—if they can commit to a board composed of mostly advisory types before the CEO can determine the project complexity. Given the stickiness of boards, this constitutes a credible commitment to centralize the investment decision at a later stage. Still, even if the CEO's "noising up" of investment projects can be avoided in equilibrium,

[2]Numerous studies have considered the welfare effects of SOX, focusing for the cost side mostly on the direct compliance costs associated with certain SOX provisions, such as Section 404 on internal controls, e.g., Zhang (2007). Our result demonstrate an additional opportunity cost of strengthening the monitoring role of boards, namely that the board becomes less productive as a source of strategic advice.

[3]This argument has been made by Shleifer and Vishny (1989) and Edlin and Stiglitz (1995) in connection with CEO replacements, and mentioned by Harris and Raviv (2005) in a similar setting to ours.

the specter of such off-equilibrium behavior may well create additional costs to shareholders.

Related literature. The view of the dual role of boards as advisors and monitors dates back to Mace (1971). The main dichotomy regarding board member types studied in the recent literature, however, has been that of insiders versus outsiders (e.g., Raheja 2005, Adams and Ferreira 2007, Harris and Raviv 2008, Linck et al. 2008). Yet, Linck et al. (2009) have brought the dual role (a la Mace) back to the fore by documenting a significant shift between pre- and post-SOX boards towards more monitoring types (see also Becht et al. 2008). Jensen (1993) has observed that insiders make for bad monitors, which suggests equating insiders with monitors, and outsiders with advisors.[4]

However, this mapping is far from robust. For instance, outsiders may be former entrepreneurs (more likely to be advisors), or bankers, former regulators/public sector officials or CPAs (more likely to be monitors). Hence, caution is in order when linking our results to empirical findings along the insiders/outsiders dichotomy. A paper that directly speaks to our framework is Baker and Gompers (2003) who have shown that venture capitalist representation results in a more powerful board and thereby in more centralization. Since VCs typically tend to perform advisory roles more than other board members, this result is in line with our model predictions.

In a dynamic model, Hermalin and Weisbach (1998) have shown that CEO

[4]This approach is taken in numerous empirical studies. For instance, Coles et al. (2008) hypothesize, and ultimately find, that firms with greater reliance on R&D (and hence more advisory needs) will nominate more insiders on the board. Bathala and Rao (1995) and Lehn et al. (2004) show that growth options are positively associated with the percentage of insiders on the board; growth options are viewed as a proxy for a firm's demand for advice. More generally, Lehn et al. (2004) show that over the life cycle of firms boards become more outsider-heavy, which is consistent with the view that mature firms are plagued by more severe agency problems. On the other hand, Bhagat and Black (1999) have argued that outside directors may be particularly valuable as a source of advice, rather than monitoring, for high-growth firms.

power may arise endogenously as a result of delivered performance. Shleifer and Vishny (1989) and Edlin and Stiglitz (1995) have demonstrated how mangers can entrench themselves by choosing complex strategies or products. We will take the ability of the CEO to nominate the board (or entrench himself) as given in some parts of our analysis.

Technically, our paper is most closely related to Dessein (2002), Harris and Raviv (2005, 2008), and Adams and Ferreira (2007). All of these studies employ cheap-talk models to analyze the strategic information transmission between the board (supervisor) and the CEO (agent). Our paper adds to this stream of literature by investigating the monitoring role of the board in terms of attempting to uncover the CEO's information—the information the parties are endowed with is therefore endogenous. In Adams and Ferreira (2007) board monitoring is phrased in terms of wrestling control away from the CEO, and an independent board has lower disutility from doing so than a dependent board. Harris and Raviv (2008) model the board's acquisition of "new" information (incremental to the CEO's information), but they do not consider board monitoring.

The remainder of the paper is organized as follows: Section 2 lays out the model setup. Section 3 looks at special cases of the model where either the allocation of decision rights or the board composition is exogenously specified. Section 4 solves for the equilibrium outcome if shareholders choose both these design variables endogenously. In Section 5 we consider the case where a powerful CEO can determine the board composition himself. In Section 6 we allow for the CEO to choose the complexity of the project as a way to entrench himself. Section 7 concludes. All proofs are contained in the appendix.

2 Model

We model the strategic interaction between a privately informed CEO of a firm and its board of directors. We take the board size as given and focus on the interaction between board composition and the allocation of decision-making authority.[5] To capture the dual role of boards, we let $x \in [0,1]$ denote the fraction of board members that are engaged in monitoring (the "monitoring" types) and $1-x$ that of board members engaged in generating ideas (the "advisory" types). The firm faces an investment decision, denoted by $y \in \mathbb{R}$. In line with earlier studies by Harris and Raviv (2005, 2008) and Adams and Ferreira (2007), both the CEO and the board aim to adjust the investment decision to some state of the world, θ. The board aims to maximize shareholder value (i.e., the NPV), which equals (ignoring lump sum terms):

$$-(y-\theta)^2. \tag{1}$$

The CEO on the other hand is biased towards larger investments, possibly due to empire benefits,[6] i.e., the CEO aims to maximize

$$-(y-\theta-b)^2, \quad b>0. \tag{2}$$

We assume that the state variable is the sum of two signals, each independently uniformly distributed: $\theta = a+p$, where $a \sim U[0,A]$ and $p \sim U[0,P]$. The CEO observes a for sure, whereas the board *may* observe a and p. We assume that it is common knowledge whether the board is informed about p. The probability of the board learning these information parameters depends on

[5] Several empirical studies have shown negative correlation between board size and firm performance beyond a certain board size, among which see Yermack (1996) and Eisenberg et al. (1998).

[6] We do not explicitly model compensation schemes as means to align the CEO's incentives with that of the shareholders. In practice, the higher the pay-performance sensitivity in the compensation plan, the lower one would expect the bias b to be.

the board composition. In particular, for given fraction of monitoring types, x, the board observes a with probability $q_a(x)$ and p with probability $q_p(x)$. The higher is x, the higher is $q_a(x)$ (better monitoring) but the lower is $q_p(x)$ (less incremental information gathering). That is, engaging in more monitoring comes at an opportunity cost of hampered board creativity, and vice versa. Specifically:

Assumption 1 $q_a(\cdot) \in [0,1),\ q_p(\cdot) \in [0,1),\ q_a'(\cdot) > 0,\ q_a''(\cdot) \leq 0,\ q_p'(\cdot) < 0,\ q_p''(\cdot) \leq 0.$

Throughout the paper we assume that $A \geq P$, i.e., the CEO's private information is weakly "more important" to the investment decision than is the information that may or may not be discovered by the board. Moreover, we assume that the decision-making authority is allocated before the parties observe their respective information.[7] This seems descriptive given the fact that the choice between delegation and centralization often entails certain communication (and other) infrastructure and that firms tend to be reluctant to redesign this infrastructure on an ongoing basis.[8] However, we do allow below for the board to "veto" the CEO's investment choice under delegation if the board has learned a and the CEO chooses an investment level that is inconsistent with this information. Clearly, in this case of common information, the CEO would not be able to indulge in his bias.

2.1 Payoffs for Given Allocation of Decision Rights

If decision-making is centralized, it is the board that chooses y so as to maximize its objective specified in (1), given its available information. Denote the board's

[7]Thus, our model is in broadly line with the "ex-ante environment" in Harris and Raviv (2005), with the modification that the board will veto an investment choice made by the CEO that is inconsistent with any commonly known information.

[8]For example, centralized decision making is likely to require more frequent and longer board meetings (e.g., meetings with investment bankers, lawyers, outside consultants) compared with decentralized decision making where much of this effort is taken by the CEO and his direct reports. Also, allowing for an ex-post "re"-assignment of authority will delay decision making.

information set by Ω_B and the resulting investment by $y_B(\Omega_B)$. Recall that the board learns a with probability $q_a(x)$ and p with probability $q_p(x)$. Therefore,

$$
\begin{aligned}
y_B(\Omega_B) &\in \arg\max_y \; E\{-(y-\theta)^2 \mid \Omega_B\} \\
&= \begin{cases} a+p & \text{if board learns } a \text{ and } p \\ a+E[p] & \text{if board only learns } a \\ E[a \mid r_a]+p & \text{if board only learns } p \\ E[a \mid r_a]+E[p] & \text{if board learns neither } a \text{ nor } p \end{cases}
\end{aligned} \tag{3}
$$

We denote by $\overline{\sigma}_p^2$ the prior variance of p, and by σ_a^2 the posterior variance of a conditional on receiving a cheap-talk message r_a sent by the CEO, in case the board has not learned a. The board's (and hence the shareholders') expected utility under centralization then equals

$$
\begin{aligned}
\pi^C &= -\left\{q_a(x)(1-q_p(x))\overline{\sigma}_p^2 + (1-q_a(x))\left[q_p(x)\sigma_a^2 + (1-q_p(x))(\sigma_a^2+\overline{\sigma}_p^2)\right]\right\} \\
&= -\left[(1-q_p(x))\overline{\sigma}_p^2 + (1-q_a(x))\sigma_a^2\right] \\
&\equiv -Var^C,
\end{aligned} \tag{4}
$$

where Var^C denotes the *residual variance* (post-cheap talk) under centralization.

Consider now the case of delegation. As we have argued above, if the board learns a, the CEO cannot indulge in his bias but instead has to choose the investment level which is optimal for the shareholders.[9] That amounts to effectively replicating the first two cases in (3). If the board does not learn a, on the other hand, the CEO has *real* decision-making authority and chooses y so as to

[9] That is, in case a becomes common knowledge, the board can veto the CEO's decision, e.g., by way of an implicit (off-equilibrium) threat of dismissal.

maximize his objective in (2) given his available information set, Ω_{CEO}:

$$
\begin{aligned}
y_{CEO}(\Omega_{CEO}) &\in \arg\max_{y} \; E\{-(y-\theta-b)^2 \mid \Omega_{CEO}\} \\
&= \begin{cases} a+p & \text{if board learns } a \text{ and } p \\ a+E[p] & \text{if board only learns } a \\ a+E[p \mid r_p]+b & \text{if board learns only } p \\ a+E[p]+b & \text{if board learns neither } a \text{ nor } p \end{cases}
\end{aligned} \tag{5}
$$

That is, with probability q_p the CEO can incorporate at least a noisy cheap-talk report r_p, sent by the informed board, into the investment decision. Denoting by σ_p^2 the posterior variance of p conditional on this report, the board's expected utility under delegation reads:

$$
\begin{aligned}
\pi^D &= -\left\{q_a(x)(1-q_p(x))\overline{\sigma}_p^2 + (1-q_a(x))\left[q_p(x)(\sigma_p^2+b^2) + (1-q_p(x))(\overline{\sigma}_p^2+b^2)\right]\right\} \\
&= -\left[(1-q_p(x))\overline{\sigma}_p^2 + (1-q_a(x))q_p(x)\sigma_p^2 + (1-q_a(x))b^2\right] \\
&\equiv -Var^D - (1-q_a(x))b^2,
\end{aligned} \tag{6}
$$

where Var^D gives the residual variance under delegation. Delegation imposes an additional cost on the shareholders: with probability $(1-q_a)$ the board's monitoring is unsuccessful; the CEO then retains effective control over the investment decision and biases it in line with his preferences.

Now consider the CEO's payoff under the respective regimes. Notice that, all else equal, the CEO, too, aims to minimize the residual variance. At the same time, he has a preference for being in control:[10]

$$
\pi_{CEO}^C = \pi^C - b^2 \quad \text{and} \quad \pi_{CEO}^D = \pi^D + [1-2q_a(x)]b^2. \tag{7}
$$

As one would expect, if the board prefers to delegate, then *a fortiori* so does the CEO, as $\pi^D \geq \pi^C$ implies $\pi_{CEO}^D > \pi_{CEO}^C$.

[10]It is easy to show that π^C, π_{CEO}^C and π^D, respectively, each are concave in x. The same does not hold for π_{CEO}^D.

2.2 Communication Equilibrium

In this subsection, we investigate the communication equilibria and characterize the posterior variance terms σ_a^2 and σ_p^2. Since there is no explicit cost of misreporting, the communication between CEO and board is of the form of "cheap talk." The sender remits a (possibly noisy) signal to a receiver, who then updates his belief according to Bayes' rule and makes the investment decision. As shown by Crawford and Sobel (1982), all equilibria of this communication game are characterized by a partitioning of the signal space, where the sender only specifies to which partition the true realized state of nature belongs. In line with the standard cheap talk literature, we select the most informative equilibrium, corresponding to the finest feasible partitioning, as our focal equilibrium because it Pareto dominates all other equilibria.[11]

We compute the equilibrium using the Crawford and Sobel (1982) approach. Under centralization, if the board does not observe a, it will form its beliefs about a based on the report r_a sent by CEO. The posterior variance is then given by

$$\sigma_a^2 = \frac{A^2}{12(N(A,b))^2} + \frac{b^2((N(A,b))^2 - 1)}{3},$$

where $N(A,b)$ is the maximum number of the equilibrium partitions under centralization (see "Preliminaries" in the Appendix). The posterior variance σ_p^2 under delegation is computed analogously.[12]

[11]The most informative equilibrium is also supported by the NITS ("No Incentive to Separate") refinement proposed by Chen, Kartik, and Sobel (2008). Note however, as pointed out in Melumad and Shibano (1991), the most informative equilibrium may not be the Pareto-dominant one when the sender's bias depends on his private information.

[12]It is readily verified that both σ_a^2 and σ_p^2 are increasing in b everywhere. Moreover, σ_a^2 and σ_p^2 are continuous in b, even at those b points where $N(A,b)$ and $N(P,b)$ change in a discrete fashion (of course, σ_a^2 and σ_p^2 are not differentiable at those thresholds).

3 Board Composition and Allocation of Decision Rights

In this section we investigate the optimal allocation of decision rights for given board composition, and the optimal board composition for given allocation of decision rights. We later integrate the insights from this preliminary analysis and solve for the optimal endogenous choice for both design variables.

3.1 Exogenous Board Composition

We begin by studying the effect of the given board composition, x, on the allocation of decision rights. Define $\Delta(x,b)$ as the profit differential between delegation and centralization holding x fixed:[13]

$$\Delta(x,b) \equiv \pi^D(x,b) - \pi^C(x,b) = [1 - q_a(x)]Z(x,b) \tag{8}$$

Here, $Z(x,b) \equiv \sigma_a^2 - [b^2 + q_p(x)\sigma_p^2]$ captures the *differential* costs between the two structures: under centralization the shareholders incur the information loss due to the noisy reporting of a by the CEO; under delegation they incur the bias cost and the information loss due to the noisy communication of p (with probability q_p).[14] Then:

Lemma 1 *When the CEO bias, b, is sufficiently small, delegation dominates centralization for any $x \in [0,1]$. When b is sufficiently large, centralization dominates delegation for any $x \in [0,1]$. For intermediate values of b, there exists a cutoff $\hat{x}(b) \in [0,1]$ such that the shareholders prefer centralization for $x \leq \hat{x}(b)$, and delegation for $x > \hat{x}(b)$.*

---

[13]To avoid clutter we suppress functional arguments that are not essential at this point and introduce them only as needed.

[14]Note that these costs are incurred only with probability $(1 - q_a)$, so this common terms cancels out.

Proof. All proofs are found in the Appendix.

Extending the results in Dessein (2002) in a straightforward fashion, if the CEO's bias b is sufficiently small, the cost differential between centralization and delegation, $Z(x, b)$, is always positive for all x; the reverse holds for b sufficiently large. The more novel insight of Lemma 1 relates to intermediate values of b. For those we can show that the board is indifferent between centralization and delegation for some cutoff level of board composition, $\hat{x} \in [0, 1]$. For more advisor-heavy boards (i.e., $x < \hat{x}(b)$), centralization is preferred, because such a board is more likely to discover p, and this information is more valuable under centralization. Under centralization, p is perfectly impounded in the board's choice of y_B, whereas under delegation, y_{CEO} only reflects the noisy report about p submitted by the board. This result is consistent with Baker and Gompers (2003) who have shown that adding venture capitalists (whose primary role is to advise the CEO) to the board results in more centralization.[15]

3.2 Exogenous Allocation of Decision Rights

We now analyze the optimal board composition for given allocation of decision rights. Specifically, we aim to compare $x^C(b)$ and $x^D(b)$, where $x^S(b) \in \arg\max_x \pi^S(x, b)$, $S = C, D$. The optimal interior solutions (if they exist) are

[15] Note that Lemma 1 also applies to the special case where $A = P$. This contrasts with the results in Harris and Raviv (2005); in their model centralization is always preferred for $A = P$ because the board *always* observes p. Then for $A = P$, the information loss under centralization due to a noisy report of a is the same as that under delegation given the noisy report of p; yet the board also incurs the bias cost under delegation. In our model, the board observes p with probability $q_p(x) < 1$, so the information loss due to noisy communication is lower under delegation than under centralization. Thus, in our setting, delegation is preferred for b sufficiently small even when $A = P$. If we were to assume $A > P$, we could relax Assumption 1 such that $q_i(x) \in [0, 1]$, $i = a, p$ and still get a meaningful tradeoff between centralization and delegation for small biases.

determined, respectively, by the following first-order conditions:

$$\frac{\partial \pi^C(x^C(b), b)}{\partial x} = q_p'(x^C(b))\overline{\sigma}_p^2 + q_a'(x^C(b))\sigma_a^2 = 0 \tag{9}$$

and

$$\frac{\partial \pi^D(x^D(b), b)}{\partial x} = q_p'(x^D(b))\overline{\sigma}_p^2 + q_a'(x^D(b))b^2 + B(x^D(b))\sigma_p^2 = 0, \tag{10}$$

where $B(x) \equiv q_a'(x)q_p(x) - (1 - q_a(x))q_p'(x) > 0$. For some of the following results it will be useful to invoke the following specific exponential form for the probability functions of learning a or p, respectively:

Assumption 2 $q_a(x) = \eta x^z$, $q_p(x) = \eta(1-x)^z$, *where* $\eta \in (0,1)$ *and* $z \in (0,1)$.

Lemma 2 *Suppose Assumption 2 holds with η sufficiently large. Then, if $A = P$, the optimal board composition weighs monitoring more heavily under delegation than under centralization: $x^D > x^C$.*

Starting out with the optimal board composition under centralization, the shareholders would always prefer more weight on monitoring when switching to delegation. The reason for this is that the opportunity cost of raising x is lower under delegation. As x goes up, the board is less likely to discover p, and as argued above, the value for the shareholders of knowing p is lower under delegation. Therefore, the monitoring role of boards is deemphasized in firms characterized by centralized boards, and vice versa.

The primary benefit of board monitoring is to mitigate the inefficiencies arising from biased decision-making (under delegation) or noisy reporting (under centralization) by the CEO. As a result, one would expect the optimal monitoring intensity to go up as b increases. The next result confirms this intuition irrespective of the allocation of decision rights.

Lemma 3 *Holding constant the allocation of decision rights $S = C, D$, $x^S(b)$ is non-decreasing in b.*

This result follows directly from the fact that $\pi^S,\ S = C, D$, has increasing differences in x and b, because both posterior variances σ_a^2 and σ_p^2 are non-decreasing in b. The more biased is the CEO, the more important it is for the board to learn a, regardless of the allocation of decision rights. Conversely, if b is small, there is little value to monitoring; instead the board should devote its resources to generating more information about p (i.e., x is small).

4 Equilibrium Choice of Board Composition and Decision Rights

We now consider the board's simultaneous choice of x and S as a function of the economic environment of the firm. Specifically, we allow for the bias parameter, b, to vary, while keeping the uncertainty parameters (A, P) fixed. Earlier literature has linked the optimal allocation of decision rights to the magnitude of the CEO's bias (Dessein 2002, Harris and Raviv 2005). We aim to generalize these findings to our setting with endogenously chosen boards, as characterized in Lemma 3.

The sequence of events is as follows:

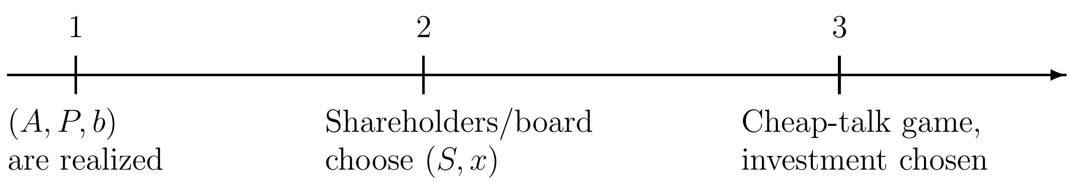

Figure 1: Timeline I—Shareholders/board choose x and S

In this base scenario, the shareholders choose both S and x. Since S is a discrete choice variable, the shareholders' problem can be decomposed as follows. First, for any allocation of decision rights, S, find the optimal board composition,

$x^S(b) \in \arg\max_x \pi^S(x,b)$.[16] Then, pick $S = C$, if $\pi^C(x^C(b),b) > \pi^D(x^D(b),b)$, and $S = D$ otherwise. For notational convenience we introduce the indicator variable $\mathbb{1}_D \in \{0,1\}$, where $\mathbb{1}_D = 1$ is equivalent to $S = D$ (delegation), while $\mathbb{1}_D = 0$ indicates $S = C$ (centralization). The shareholders' optimization problem at Date 2 can be stated as follows:

$$\mathcal{P}_0: \max_{\mathbb{1}_D \in \{0,1\}} (1-\mathbb{1}_D)\pi^C(x^C(b),b) + \mathbb{1}_D\pi^D(x^D(b),b), \quad \text{for any } b,$$

where $x^S(b) \in \arg\max_x \pi^S(x,b),\ S = C, D$ is the conditionally optimal board composition characterized in Section 3.2. Denote the solution to this program by $(S^*(b), x^*(b))$ where $x^*(b) = x^{S^*(b)}(b)$.

Our first main result characterizes the solution to the shareholders' optimization problem $\mathcal{P}_0$.

Proposition 1 (Timeline I)

(i) For sufficiently small levels of CEO bias, b, the shareholders choose delegation, i.e., $S^*(b) = D$, and board composition $x^*(b) = x^D(b)$.

(ii) For b sufficiently large, the shareholders choose centralization, i.e., $S^*(b) = C$, and $x^*(b) = x^C(b)$.

(iii) There exists at least one cutoff $\hat{b}$ at which the optimal allocation of decision rights switches from delegation to centralization.

The fraction of monitoring types on the board, $x^*(b)$, is increasing in b almost everywhere, except at such cutoffs, $\hat{b}$, where $\lim_{\varepsilon\to 0} x^*(\hat{b}-\varepsilon) = \lim_{\varepsilon\to 0} x^D(\hat{b}-\varepsilon) > \hat{x}(\hat{b}) > \lim_{\varepsilon\to 0} x^*(\hat{b}+\varepsilon) = \lim_{\varepsilon\to 0} x^C(\hat{b}+\varepsilon)$.

Consistent with earlier studies, the more biased is the CEO, the more decision-making authority will be retained by the board. The last part of the proposition

[16] Since A and P are fixed for now, we suppress them as functional arguments here.

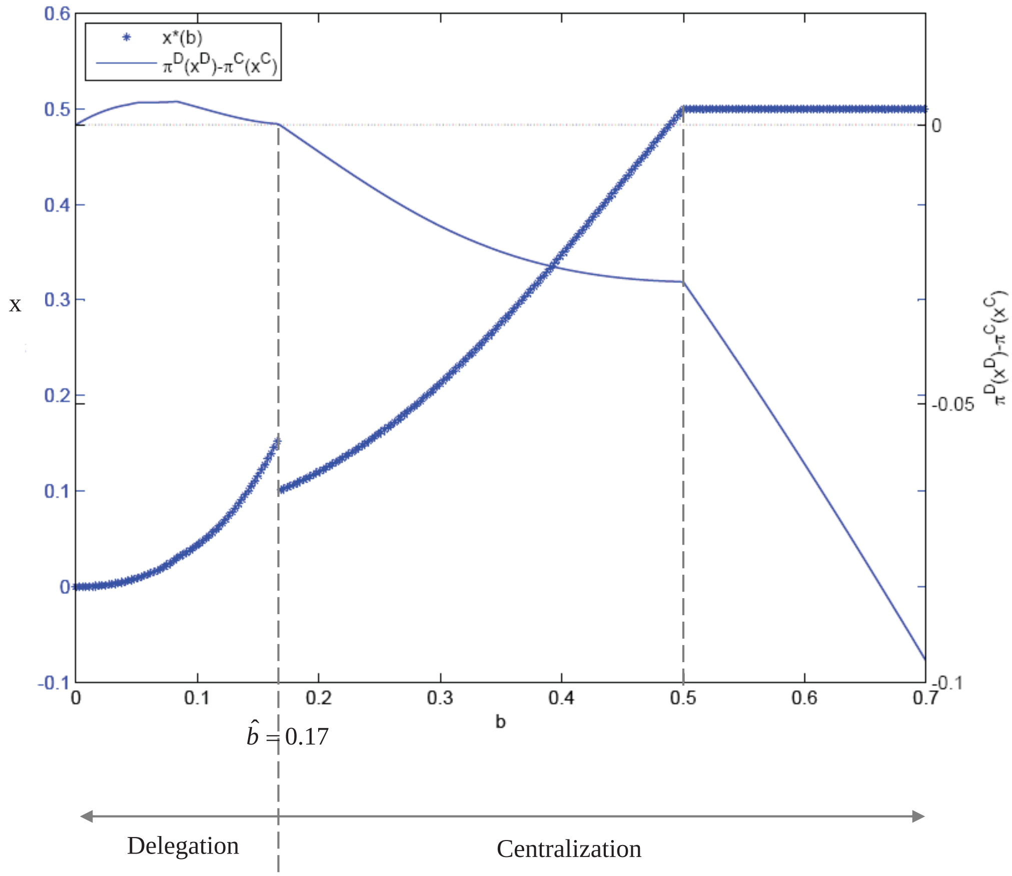

Figure 2: Board composition as a function of CEO bias under Timeline I

For this figure, $q_a(x)$ and $q_p(x)$ take the exponential function form as in Assumption 2. The parameter values are $A = P = 2$, $\eta = 0.8$, $z = 0.5$.

is more surprising as it shows that x^* is non-monotonic in b. Starting from small levels of b, the firm optimally delegates decision-making authority to the CEO and it weighs heavily idea generation on the part of the board (low x). As b increases, by Lemma 3, x^D will increase also. At some point, the CEO bias is sufficiently severe such that the board will switch from delegation to centralization. Under centralization, however, the value of learning p is higher than under delegation, as argued above. Thus, the discrete downward jump in $x^*(b)$ at the cutoff $\hat{b}$ where the board switches to centralization (see Figure 2).[17]

Our result has empirical implications regarding the life cycle of firms. As firms mature, agency problems (as measured by b) tend to increase, hence one would expect to see more centralization at the board level and more monitor-heavy boards (e.g., Lehn et al. 2004). However, if tests do not control for the allocation of decision rights, this general trend may be obscured in the data due to a missing correlated variable problem. Specifically, between two firms operating in similar informational environments, but differing in their severity of agency problems, the one with the less severe CEO bias (say, $b < \hat{b}$) may have more monitors on the board than the other firm that has CEO bias of $b > \hat{b}$. The reason is that the former firm will delegate decisions to the CEO, whereas the latter will choose centralization.

[17] As suggested by Proposition 1, there may be more than one such threshold $\hat{b}$ at which the optimal allocation of decision rights changes. This is due to the fact that $\pi^C(x^C(b), b)$ and $\pi^D(x^D(b), b)$ may intersect more than once, see also Harris and Raviv (2005). To abstract from such technicalities and focus on the main economic tensions, we will henceforth only consider scenarios where such a threshold is unique. The threshold can be shown to be unique for plausible restrictions on the primitives of the model.

5 CEO Power: Influence Over Board Composition

Critics of corporate governance practice have argued that one reason why boards exert insufficient control over management is that CEOs have significant say over the board selection process. This issue was analyzed by Hermalin and Weisbach (1998). In their model the CEO's bargaining power over board composition evolves dynamically as a result of delivered performance. To address the issue of CEO power—and how this affects the benchmark solution of the preceding section—we now consider an alternative timeline where the CEO chooses the board composition, x, followed by the board allocating the decision rights, S. We will maintain the earlier assumption that the board, even if selected by the CEO, will act in the shareholders interest when choosing S. In the Concluding Remarks section, below, we speculate on the possible effects of biased boards.

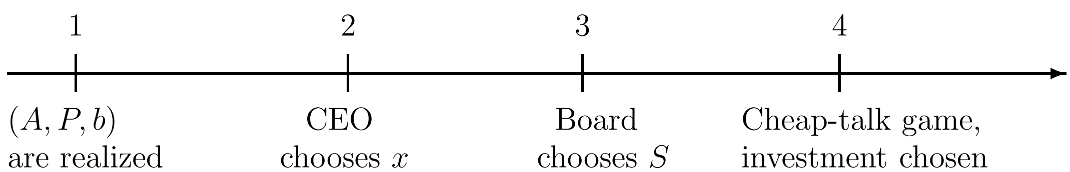

Figure 3: Timeline II—CEO influences board composition

The main question here is whether the CEO will choose a lower or higher percentage of monitoring types on the board compared with the level of x preferred by the shareholders. A key concern in the practitioners' literature is that a powerful CEO might prefer a less vigilant board because, holding constant the allocation of decision rights, that would make it more likely he can indulge in his bias. Specifically, under delegation the CEO can act according to his bias with probability $1 - q_a(x)$, hence he prefers to have more advisors on the board (i.e., x to be small). The main result of this section shows that this intuition may not always hold; in fact, the CEO may sometimes prefer more monitors on the board than the shareholders.

The modified optimization program with the CEO choosing x at Date 2 reads:

$$\begin{aligned}\mathcal{P}_1 \quad : \quad & \max_{x\in[0,1]} [1 - \mathbb{1}_D(x,b)]\pi^C_{CEO}(x,b) + \mathbb{1}_D(x,b)\pi^D_{CEO}(x,b), \quad \text{for any } b,\\ & \text{subject to:}\\ & \mathbb{1}_D(x,b) \in \arg\max_{\mathbb{1}_D\in\{0,1\}} (1 - \mathbb{1}_D)\pi^C(x,b) + \mathbb{1}_D\pi^D(x,b).\end{aligned}$$

Denote the solution to this program by $x_{CEO}(b)$. Our next result compares the CEO's choice of board composition with that selected by the shareholders in the benchmark model of Section 4. (Recall that $\hat{b}$ denotes the threshold bias value at which a shareholder-nominated board under Timeline I is indifferent between centralizing and delegating, i.e., $\pi^C(x^C(\hat{b}),\hat{b}) \equiv \pi^D(x^D(\hat{b}),\hat{b})$.)

Proposition 2 (Timeline II) *Suppose the CEO chooses the board composition, x, before the board allocates the decision rights, S. Then:*

(i) For $b < \hat{b}$, $x_{CEO}(b) < x^*(b)$ and $S(x_{CEO}(b),b) = S^*(b) = D$.

(ii) For b sufficiently large, $x_{CEO}(b) = x^*(b)$ and $S(x_{CEO}(b),b) = S^*(b) = C$.

(iii) For $b = \hat{b}+\varepsilon$ with $\varepsilon \to 0$, $x_{CEO}(b) > x^*(b)$ and $S(x_{CEO}(b),b) = D$ whereas $S^*(b) = C$.

The key to understanding the first part of the result is to notice that by (7), whenever the board prefers delegation in the benchmark model of Section 4, then *a fortiori* so does the CEO, as the CEO can simply set $x_{CEO}(b) = x^D(b)$ to induce delegation and act according to his bias. In addition, under delegation the CEO prefers a *lower* level of monitoring than the shareholders do, as argued above, as that would reduce $q_a(x)$, the probability that the board assumes real authority upon learning a. Therefore, for $b < \hat{b}$, delegation is preferred by the board in the benchmark model; thus the CEO has no incentive to influence the board's decision and simply sets $x_{CEO}(b) < x^*(b) = x^D(b)$. On the other hand, when the bias b is sufficiently large (part (ii)), the board will choose centralization irrespective of the CEO's choice of board composition.[18] Therefore, the CEO will set $x_{CEO}(b) = x^*(b) = x^C(b)$, since, conditional on the decision being centralized, the CEO's preference over board composition coincides with that of the shareholders.

The main insight of Proposition 2 is that, in certain cases, the CEO prefers *more* board monitoring than would be chosen by the shareholders (part (iii) of this result). When the optimal solution in the benchmark model of Program $\mathcal{P}_0$ is such that the shareholders marginally prefer centralization over delegation (i.e., $b = \hat{b}+\varepsilon$, for small ε), then the CEO has incentives to influence the shareholders' choice of S towards delegation by nominating more monitors to the board. Recall from Lemma 1, that shareholders prefer to delegate for high values of x, as the probability of the board learning p will then be smaller, and the opportunity cost associated with delegation will be reduced (see Figure 4). Put differently, the CEO prefers delegation—even if monitored by a vigilant board—rather than having the decision rights taken away from him by an advisor-heavy board.

The Sarbanes-Oxley Act of 2002 (SOX) has sought to strengthen the monitoring role of boards, with particular emphasis on the audit committee (see SOX, Section 301). In the context of our model, we interpret SOX as imposing a lower bound on x, i.e., $x \geq x_{SOX}$. Clearly, as long as shareholders determine the board composition, any regulatory intervention regarding the latter would weakly decrease the shareholders' payoff. However, if the CEO nominates the board, a

[18] Technically, the critical value $\hat{x}(b)$ at which the board is indifferent to delegate the investment decision would exceed one.

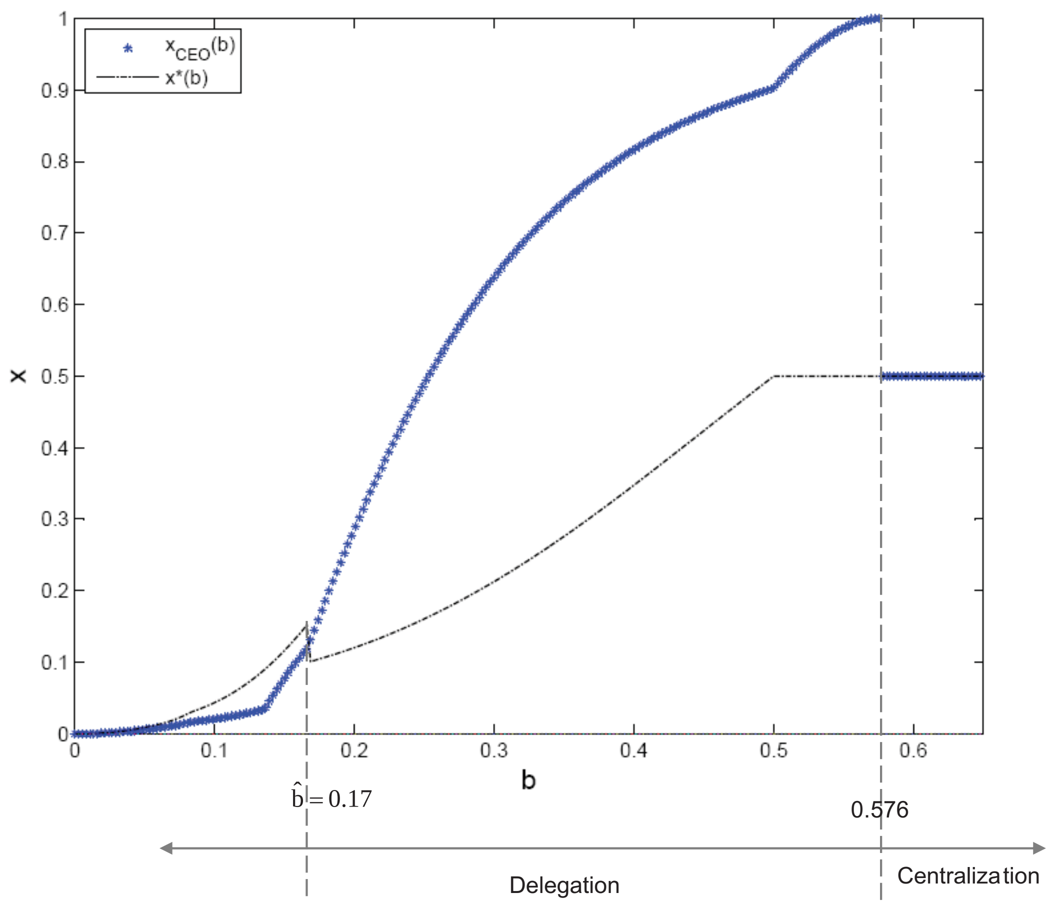

Figure 4: Board composition as a function of CEO bias under Timeline II

The dashed line is the optimal board composition under the benchmark model of Timeline I. For this figure, $q_a(x)$ and $q_p(x)$ take the exponential function form as in Assumption 2. The parameter values are the same as in Figure 2, i.e., $A = P = 2$, $\eta = 0.8$, $z = 0.5$.

potentially welfare-increasing role for regulation arises because, as Proposition 2 has shown, the CEO's choice of x at times diverges from that preferred by the shareholders.

Corollary 1 *Suppose the CEO chooses the board composition as in Timeline II. Then:*

(i) A necessary condition for SOX to improve shareholder value is that the CEO's bias is sufficiently small, i.e., $b < \hat{b}$.

(ii) SOX (weakly) lowers shareholder value for $b > \hat{b}$*, and strictly so if SOX is effective in that* $x_{CEO}(b) < x_{SOX}(b)$.

For small levels of CEO bias, the conventional wisdom applies in that the CEO prefers a "monitor-light" board. In that case, SOX may indeed improve shareholders value by constraining the CEO's choice. However, for greater biases, SOX, if binding, will result in a board composition that departs even further from the benchmark solution. The irony therefore is that SOX, which was intended to mitigate incentive costs arising from CEO bias and power, will increase shareholder value only if the bias is small.

There is a growing literature on the questionable welfare effects of SOX (e.g., Zhang 2007). This literature has focused, for the most part, on the direct costs associated with the implementation of internal controls required by Section 404 of SOX. Our analysis complements this literature by pointing out additional opportunity costs associated with strengthening the monitoring function of boards, namely that for given board size, boards that become more focused on monitoring will become less effective advisors to the firm's management. Stepping outside the scope of our model, Corollary 1 also speaks (albeit indirectly) to the recent trend of an increase in board size post-SOX (Linck et al 2008, 2009). Assuming firms had chosen their board composition optimally pre-SOX, then an exogenous shock that forces firms to add more monitoring types to the board, should be followed by an enlarged board as firms try to re-balance the dual functions to the new optimal monitoring/advising weights.

6 CEO Power: Entrenchment

A somewhat more subtle way in which CEOs can exert power and obtain *de facto* authority is by choosing investment projects that endow them with greater information advantage (Shleifer and Vishny 1989, Edlin and Stiglitz 1995). To address this issue in the context of our model, suppose the CEO can choose the level of "complexity" of the project endogenously. Specifically, a remains uniformly distributed over $[0, A]$, but we now assume the CEO can choose the upper bound $A \in [\underline{A}, \bar{A}]$.[19] This choice is assumed to be observable but not contractible. In the following, for given $\underline{A}$, we interpret $\bar{A}$ as a measure of the CEO's discretion.

We first address the case where the CEO chooses A before the shareholders assemble the board and allocate the decision rights (Section 6.1), and then the case where the shareholders can preempt the CEO's choice of project complexity by committing to a certain board composition at the outset (Section 6.2).

6.1 Endogenous Project Complexity

Consider the revised sequence of events as depicted in Timeline III.

[19]The possibility of CEO entrenchment via project choice is alluded to in Harris and Raviv (2005).

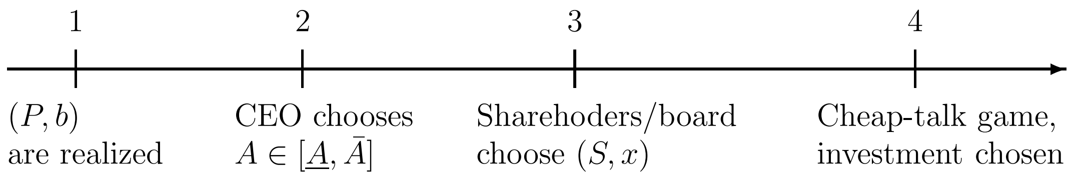

Figure 5: Timeline III—CEO chooses A endogenously

When allocating the decision rights at Date 3, the board trades off the information loss regarding a under centralization against the information loss regarding p *plus* the bias cost under delegation. As the information advantage of the CEO becomes larger (i.e., A increases), the board is more willing to delegate authority to the CEO. Clearly, the CEO will only want to "noise up" the project, if absent such action the board would choose to centralize, i.e., if $b \geq \hat{b}(\underline{A})$. In that case, there exists a threshold $\hat{A}(b)$ given by

$$\pi^C(x^C(\hat{A}(b), b), \hat{A}(b), b) \equiv \pi^D(x^D(b), b), \tag{11}$$

such that the board prefers to delegate the investment decision for any $A \geq \hat{A}(b)$, and to centralize it otherwise.

Note that the shareholders' payoff under delegation does not depend on the project complexity $\hat{A}$ because the realization of $a \in [0, A]$ for $A \in [\underline{A}, \bar{A}]$ will always be impounded without noise in the investment decision by the CEO. In other words, a more complex project (a higher A) adversely affects the payoff of the shareholders (and of the CEO) only under centralization. If by choosing $A \geq \hat{A}(b)$ the CEO succeeds in influencing the board's decision towards delegating, this will come without any additional informational loss. On the other hand, if the equilibrium path is such that the investment will be made by the board at Date 3, then the CEO's and the board's objectives regarding project complexity are fully aligned in that both prefer the least complex project, $\underline{A}$, at Date 2. We assume without loss of generality that, whenever the CEO is indifferent among various levels of A, he will chose the lowest one.[20]

For the CEO to be able to entrench himself, $\hat{A}(b) \leq \bar{A}$ has to hold.[21] In that case, the CEO trades the control benefits that come with delegation against the expected information loss regarding p. That is, entrenchment may be feasible for the CEO, but not necessarily optimal. To evaluate this tradeoff, we define $g(b) \equiv \pi_{CEO}^{D}(x^{D}(b), b) - \pi_{CEO}^{C}(x^{C}(\underline{A}, b), \underline{A}, b)$ as the payoff differential between the two regimes for the CEO, and b^* such that $g(b^*) \equiv 0$.[22] It is easy to show that $b^* > \hat{b}(\underline{A})$ since $g(b) > 0$ for all $b \leq \hat{b}(\underline{A})$. The reason is straightforward: for any $b \leq \hat{b}(\underline{A})$, the board prefers delegation, which by (7) is *a fortiori* also the preferred choice for the CEO. Therefore:

Observation 1 (Timeline III) *Suppose the CEO chooses $A \in [\underline{A}, \bar{A}]$ before the shareholders/board choose S and x. Then:*

(i) For any $b < \hat{b}(\underline{A})$, the CEO chooses $A = \underline{A}$, followed by the board delegating the investment decision and setting $x = x^{D}(b)$.

(ii) For any $b \in [\hat{b}(\underline{A}), \min\{\hat{b}(\bar{A}), b^*\}]$, the CEO chooses $A = \hat{A}(b)$, followed by the board delegating the investment decision to the CEO and setting $x = x^{D}(b)$.

(iii) For any $b > \hat{b}(\bar{A})$, the CEO chooses $A = \underline{A}$, followed by the board centralizing the investment decision and setting $x = x^{C}(\underline{A}, b)$.

---

[20] This would be the unique equilibrium choice in a more general model where the CEO observes A with probability less than one.

[21] Note that $\hat{A}(b) \leq \bar{A}$ is equivalent to $b \leq \hat{b}(\bar{A})$.

[22] Note that b^* may not exist. Whether b^* exists depends on the specific functional form of $q_a(\cdot)$ and $q_p(\cdot)$. Given the exponential specification of Assumption 2 with $z = 0.5$ and $A = P = 2$, we can show that b^* exists for $\eta = 0.999$, but not for $\eta = 0.8$. In the following observation, we assume b^* exists to simplify the exposition. If there does not exist any b^* as defined here, the qualifier in Observation 1 part (ii) would instead read: "For any $b \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$,...".

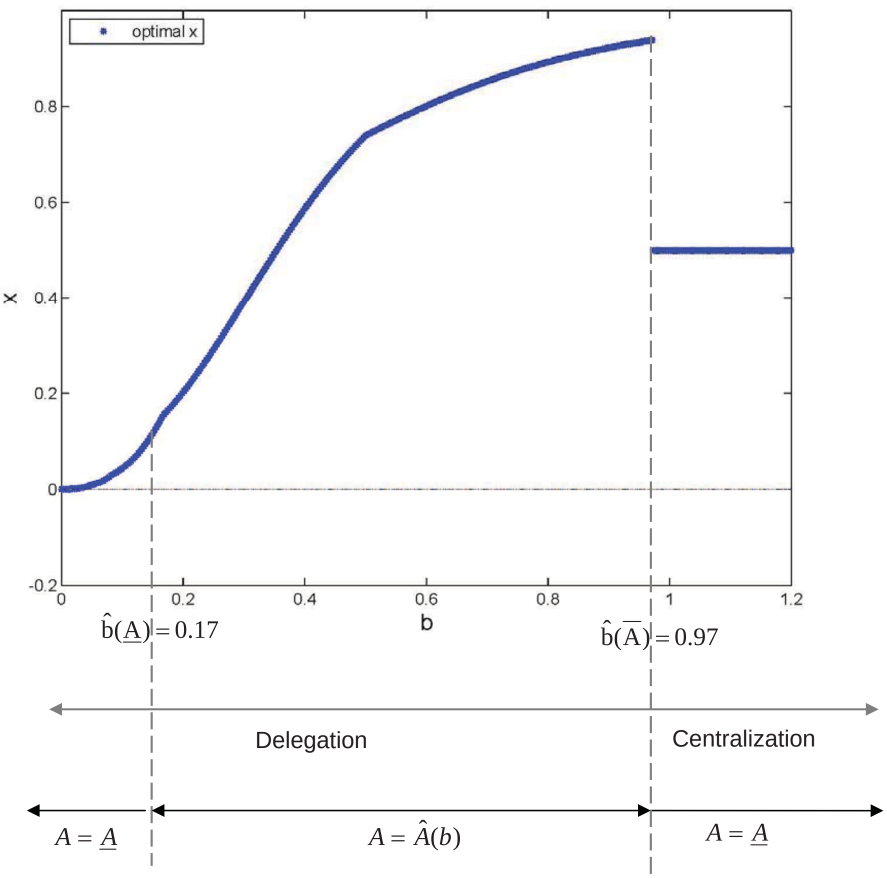

Figure 6: Board composition as a function of CEO bias under Timeline III

For this figure, $q_a(x)$ and $q_p(x)$ take the exponential function form as in Assumption 2. The parameter values are $\underline{A} = P = 2$, $\bar{A} = 3.5$, $\eta = 0.8$, $z = 0.5$.

For small bias levels there is no need for the CEO to entrench himself, as the board will delegate the decision anyway. For intermediate bias levels (part (ii) of the observation), it is both feasible and profitable for the CEO to entrench himself. However, for sufficiently high bias parameters (part (iii)), the board is so strongly inclined to centralize that the level of complexity required to trigger delegation exceeds the maximum complexity—CEO entrenchment then is not feasible. Instead, the CEO will choose the least complex project, $\underline{A}$, as that will maximize his own payoff conditional on the investment decision being centralized (see Figure 6 for illustration).

We now ask if the shareholders can preempt CEO entrenchment by nominating the board strategically at the outset.

6.2 Preempting CEO Entrenchment via Board Composition

Board composition in practice is fairly sticky, in the sense that board members usually serve for long periods of time. Hence, a CEO, when considering to entrench himself by adding to project complexity, often has to take the existing board and its structure as given. Shareholders, in turn, when nominating the board at the outset, should optimally do so with an eye to the possibility of future CEO entrenchment. Therefore, we now consider the sequence of events depicted in Timeline IV where the shareholders first choose the board composition, followed by the CEO choosing the project complexity and, finally, the board deciding whether to delegate the investment decision.

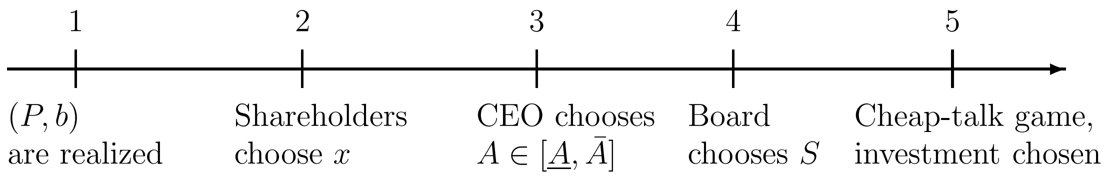

Figure 7: Timeline IV—Shareholders may preempt CEO entrenchment

To ensure CEO entrenchment is a meaningful threat, we shall maintain the following assumption throughout this section:

Assumption 3 $b \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$.

This assumption postulates that absent CEO entrenchment the board would centralize the decision (as $b \geq \hat{b}(\underline{A})$). At the same time, it is feasible for the CEO to entrench himself by exacerbating the project complexity (as $b \leq \hat{b}(\bar{A})$ is equivalent to $\hat{A}(b) \leq \bar{A}$).

The shareholders may now be able to preempt the CEO's entrenchment action by nominating an advisor-heavy board (i.e., a low level of x) as way to commit to centralization at Date 4 in Timeline IV. Recall from Lemma 1 that the lower is x, the more likely the board will discover p, and the value of this information is higher under centralization. Denote by $\tilde{A}(b, x)$ the level of project complexity that makes the board indifferent between centralizing and delegating for given bias and board composition, i.e.,

$$\pi^D(x, b) \equiv \pi^C(x, b, \tilde{A}(x, b)). \tag{12}$$

By the above logic, $\tilde{A}(x, b)$ is decreasing in x—the more monitors are on the board, the more the board is willing to delegate the decision, hence a lower level of project complexity would be required for the CEO to gain control. As a consequence, by setting x below a threshold, denoted $\hat{x}(\bar{A}, b)$, the shareholders would credibly commit the board to centralize *for any* level of feasible project complexity, given b.[23]

The shareholders' decision problem at Date 2 then boils down to a discrete profit comparison between two alternative courses of action:

(i) to *preempt* the CEO's entrenchment behavior by setting $x < \hat{x}(\bar{A}, b)$ (or, equivalently, $\tilde{A}(x, b) > \bar{A}$), versus

(ii) to *acquiesce* to the CEO's entrenchment behavior by accepting the fact that, along the equilibrium path, the investment decision will ultimately be delegated, and to adapt x optimally.

More formally, this choice can be expressed as comparing the values of the following two optimization programs:

$\mathcal{P}_{\mathcal{C}}$: *(Preempt CEO entrenchment)*

$$\max_{x \in [0,1]} \pi^C(x, \underline{A}, b)$$
$$\text{subject to: } \tilde{A}(x, b) > \bar{A} \tag{C}$$

$\mathcal{P}_{\mathcal{D}}$: *(Acquiesce to CEO entrenchment)*

$$\max_{x \in [0,1]} \pi^D(x, b)$$
$$\text{subject to: } \tilde{A}(x, b) \leq \bar{A} \tag{D}$$

Holding $A = \underline{A}$ fixed for now, by definition of $\hat{b}(\underline{A})$, the board is indifferent between centralization and delegation for $b = \hat{b}(\underline{A})$, and strictly prefers centralization for any $b > \hat{b}(\underline{A})$. Thus, if both constraints (C) and (D) in the respective optimization programs were slack, the analysis would revert to the benchmark

[23] In Section 3.1 we suppressed the functional dependance of $\hat{x}(\cdot)$ on A and P. Since we now allow for A to vary, we explicitly write $\hat{x}(\cdot)$ as a function of A and b.

setting of Timeline I, above, so that for any $b \geq \hat{b}(\underline{A})$, as postulated by Assumption 3, the value of $\mathcal{P}_{\mathcal{C}}$ would exceed that of $\mathcal{P}_{\mathcal{D}}$; hence the board would prefer centralization. As a result, a necessary condition for delegation to obtain in equilibrium is that constraint (C) has to be binding. More generally, the shadow prices of the constraints (C) and (D) indicate whether there is a cost to the shareholders associated with the threat of CEO entrenchment under the respective organizational modes.

Lemma 4 *Suppose Assumption 3 holds.*

(i) Constraint (D) in program $\mathcal{P}_{\mathcal{D}}$ is slack for any $b \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$.

(ii) For $\bar{A}$ sufficiently small, constraint (C) in program $\mathcal{P}_{\mathcal{C}}$ is slack for any $b \geq \hat{b}(\underline{A})$. On the other hand, for $\bar{A}$ sufficiently large, constraint (C) is binding at $b = \hat{b}(\underline{A})$, and there exists at least one $b^C \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$ such that (C) holds with equality at b^C, i.e., $x^C(\underline{A}, b^C) = \hat{x}(\bar{A}, b^C)$.

Recall that $\tilde{A}(x, b) > \bar{A}$ is equivalent to $x < \hat{x}(\bar{A}, b)$. Consider first the "Preempt" program $\mathcal{P}_{\mathcal{C}}$. Suppose the CEO has limited discretion in that $\bar{A}$ is close to $\underline{A}$. From Proposition 1 we know that at the indifference bias level $\hat{b}(A)$, $\hat{x}(A, \hat{b}(A)) > x^C(A, \hat{b}(A)$ holds for any A; thus constraint (C) in program $\mathcal{P}_{\mathcal{C}}$ will be satisfied. This implies the following:

Observation 2 (Timeline IV) ***[Preempt at no cost]*** *Suppose Assumption 3 holds and the shareholders choose x before the CEO chooses $A \in [\underline{A}, \bar{A}]$. For $\bar{A}$ sufficiently small, the shareholders choose $x = x^C(\underline{A}, b)$ at Date 2, the CEO chooses $A = \underline{A}$ at Date 3, and the board centralizes the decision at Date 4.*

On the other hand, if the CEO's discretion is sufficiently large ($\bar{A}$ is large), then, eventually, $\hat{x}(\bar{A}, \hat{b}(\underline{A}))$ will drop below $x^C(\underline{A}, \hat{b}(\underline{A})$, because the board is more willing to delegate if the CEO has significant private information. In that case, (C) is a binding constraint at $b = \hat{b}(\underline{A})$. The threat of entrenchment then is costly for the shareholders. As the CEO bias increases beyond $\hat{b}(\underline{A})$, the shareholders become less willing to delegate. More precisely, in the limit as b approaches $\hat{b}(\bar{A})$, we find that

$$x^C(\underline{A}, \hat{b}(\bar{A}) < x^C(\bar{A}, \hat{b}(\bar{A}) < \hat{x}(\bar{A}, \hat{b}(\bar{A}),$$

where the first inequality holds because $x^C(\cdot)$ is strictly increasing in A, and the second inequality holds by Proposition 1. That is, in case of high CEO discretion, the threat of entrenchment constitutes a binding constraint for moderate levels of CEO bias (b close to $\hat{b}(\underline{A})$). For severe conflicts of interest (b close to $\hat{b}(\bar{A})$), however, constraint (C) will be slack, i.e., CEO entrenchment does not impose additional costs on the shareholders.

Now turn to the "Acquiesce" program $\mathcal{P}_{\mathcal{D}}$. Given $x^D(b)$, constraint (D) may or may not be satisfied without CEO entrenchment, i.e., at $A = \underline{A}$. If (D) is satisfied without entrenchment, the CEO will have the decision delegated even at the minimum level of project complexity. If (D) is violated at $A = \underline{A}$, i.e., $x^D(b) < \hat{x}(b, \underline{A})$, then we show in the Appendix that for any b satisfying Assumption 3, there exists a feasible level of project complexity, $A = \tilde{A}(x^D(b), b) \in (\underline{A}, \bar{A}]$, that induces the board to delegate, given $x^D(b)$. To understand the intuition behind this finding, note that $x^D(b)$ is independent of A, whereas $\hat{x}(b, A)$ is decreasing in A—the more complex the project, the more willing is the board to delegate.

In summary, CEO entrenchment is a credible threat only if the CEO has sufficient discretion, i.e., $\bar{A}$ large enough. Whether it is optimal (or even feasible) for the shareholders to preempt CEO entrenchment in that case, depends on the CEO's discretion and his bias. More specifically, we now build on Lemma 4 to evaluate the shareholders' optimal choice of board composition at Date 2. To that end, if there exist multiple thresholds $b^C \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$, as defined in Lemma 4, we will focus on the smallest one.

Proposition 3 (Timeline IV) *Suppose Assumption 3 holds and the shareholders choose x before the CEO chooses $A \in [\underline{A}, \bar{A}]$. Then, for $\bar{A}$ sufficiently large:*

(i) ***[Acquiesce]*** *If $b = \hat{b}(\underline{A}) + \varepsilon$, with ε "small", then the shareholders choose $x = x^D(b)$ at Date 2, the CEO chooses $A = \tilde{A}(x^D(b), b) > \underline{A}$ at Date 3, and the board delegates the decision to the CEO at Date 4.*

(ii) ***[Preempt at a cost]*** *If $b = b^C - \delta$, with δ "small", then the shareholders choose $x = \hat{x}(\bar{A}, b)$ at Date 2, the CEO chooses $A = \underline{A}$ at Date 3, and the board centralizes the decision at Date 4.*

(iii) ***[Preempt at no cost]*** *If $b = b^C + \delta$, with δ "small", then the shareholders choose $x = x^C(\underline{A}, b)$ at Date 2, the CEO chooses $A = \underline{A}$ at Date 3, and the board centralizes the decision at Date 4.*

Suppose the CEO has significant discretion in adding to project complexity. Then, to commit to centralization (if feasible), the board would have to deviate from $x^C(\cdot)$. When the board only has a marginal preference for centralization over delegation absent CEO entrenchment—i.e., for CEO bias slightly above $\hat{b}(\underline{A})$—it is not worthwhile for the board to deviate from $x^C(\cdot)$ and bear the shadow price of constraint (C). Instead, the board is better off acquiescing to the CEO's entrenchment threat, and therefore will set $x = x^D(b)$.

However, as the CEO's bias increases beyond $\hat{b}(\underline{A})$, then in the benchmark setting of Timeline I (without any threat of CEO entrenchment, i.e., $A = \underline{A}$), the shareholders would strictly prefer centralization. As the bias approaches b^C, the shadow price of constraint (C) under centralization becomes small. Hence, the shareholders prefer to preempt CEO entrenchment by deviating from $x^C(\cdot)$. By distorting the board composition in favor of adding more advisory types, the shareholders effectively commit to centralization. While this causes an additional

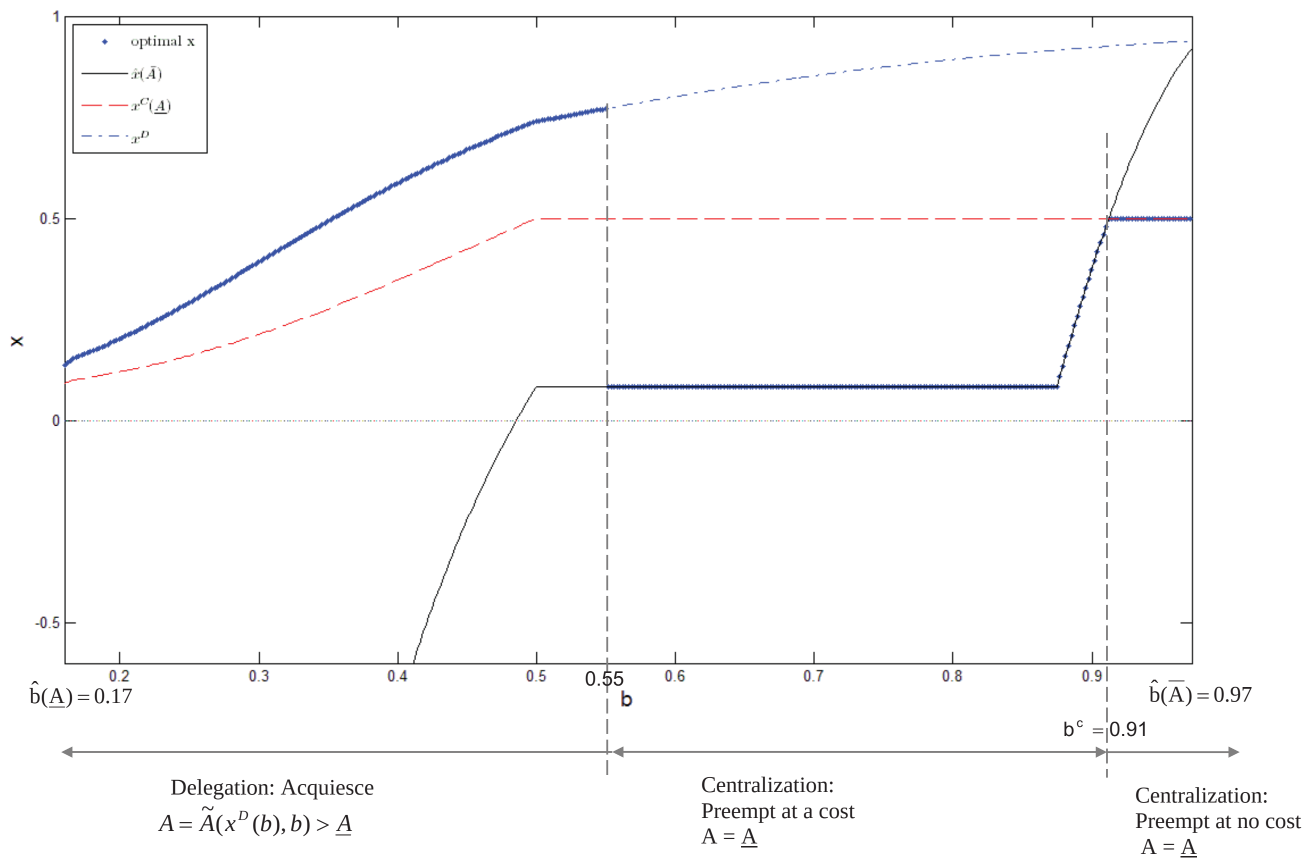

Figure 8: Board composition as a function of CEO bias under Timeline IV

For this figure, $q_a(x)$ and $q_p(x)$ take the exponential function form as in Assumption 2. The parameter values are the same as in Figure 6, i.e., $\underline{A} = P = 2$, $\bar{A} = 3.5$, $\eta = 0.8$, $z = 0.5$.

cost compared with the benchmark case of Timeline I, this cost is lower than that of acquiescing to the CEO's entrenchment (as under Timeline III).

On the other hand, when the CEO's bias exceeds b^C, constraint (C) becomes slack. Then, the unconstrained optimal board composition under centralization $x^C(\cdot)$ under Timeline I constitutes a credible commitment to centralization. Therefore, the shareholders can preempt CEO entrenchment at no cost (see Figure 8 for illustration).

As shown in Corollary 1, a SOX-like requirement of a lower bound for monitoring types on the board can make shareholders worse off by exacerbating distortions in the board composition if the latter is effectively determined by the CEO. Proposition 3 reveals another potential downside of such regulation: it might jeopardize the shareholders' ability to commit to centralized decision-making by nominating an advisor-heavy board.

7 Concluding Remarks

Shareholder activists, regulators, and the professional and academic literatures have been increasingly interested in, and at times concerned about, CEO power. Our paper examines two aspects of CEO power. The first is the power to appoint members to the board of directors. In that case, we show, somewhat surprisingly, that the CEO may find it desirable to appoint a more monitor-heavy board as compared with that the shareholders would have chosen. The second aspect is CEO entrenchment by way of strategically exacerbating project complexity. In this case we show that when shareholders appoint the board, they sometimes choose a more advisor-heavy board as compared with the board they would have chosen in the absence of entrenchment.

Throughout the paper we assumed that board members internalize shareholders' preferences. In reality, board members' preferences might diverge from those of shareholders, or even differ amongst themselves. An interesting extension would be to incorporate divergent preferences into our model. Similar to an argument in Dessein (2002), in certain situations, shareholders might be able to choose among board member candidates with different (known) biases. The research question then is whether shareholders might prefer to choose a biased rather than unbiased board. The reason why choosing a biased board might be appealing to shareholders is that it could serve as a commitment device for the shareholders in their (cheap-talk) relationship with the CEO. This could prove especially useful if the CEO bias is relatively large. In that case, communication between the CEO and shareholders is likely to be of little value. Electing a board with an intermediate bias level could result in more informative communication between CEO and board. This gain would have to be traded off against additional bias cost.

Our model assumes that a board member can be either an advisor or a monitor. Often, however, board members could perform—with different levels of expertise—either role, so the type choice boils down to prioritizing time spent between these activities. Even then, some members are likely to be better at monitoring (e.g., compliance and accounting experts) while others are better at advising (e.g., industry experts). We speculate that our qualitative findings will continue to hold in those cases.

Other possible model extensions include allowing for the probability of successful monitoring by the board to depend on the complexity of the project. Another possible extension would be to consider CEO compensation. That might mitigate the associated bias cost and align better the objectives of the CEO and the board, but in general no compensation arrangement (short of effectively "selling" the firm to the CEO) could fully avoid the agency cost of CEO bias. The extent to which these extensions affect our results is yet to be explored.

Appendix

Preliminaries. The prior variance of p given that p is uniformly distributed over $[0, P]$ is $\overline{\sigma}_p^2 = P^2/12$. As mentioned in the text, we focus on the most informative equilibrium, i.e., the one with the finest partitioning. The posterior variance of the random variable $a \sim U[0, A]$, conditional on a cheap talk report made by a privately informed party equals

$$\sigma_a^2 = \frac{A^2}{12(N(A,b))^2} + \frac{b^2((N(A,b))^2 - 1)}{3}. \tag{13}$$

Here, $N(A, b)$ is the maximum size of the equilibrium partition under centralization, as given by

$$N(A,b) = \left\langle -\frac{1}{2} + \frac{1}{2}\left(1 + \frac{2A}{b}\right)^{\frac{1}{2}} \right\rangle,$$

where $\langle t \rangle$ denotes the smallest integer greater or equal to t. The number of partitions, $N(A, b)$, jumps from $n+1$ to n at $b = b_n(A) \equiv \frac{2A}{(2n+1)^2-1}$. In particular, there will be no information conveyed if and only if $b > b_1(A) = A/4$. Following the arguments in Dessein (2002), in the limit, $N(A,b) \to \sqrt{\frac{A}{2b}}$ as $b \to 0$. Therefore,

$$\sigma_a^2 \to \frac{Ab}{3} - \frac{b^2}{3} \text{ as } b \to 0, \quad \text{and} \quad \left.\frac{\partial \sigma_a^2}{\partial b}\right|_{b=0} = \frac{A}{3}.$$

The posterior variance σ_p^2 for p is computed analogously and a corresponding comparative statics result at $b = 0$ obtains for σ_p^2.

Proof of Lemma 1.

Simple algebra establishes the profit differential between delegation and centralization,

$$\Delta(x,b) \equiv \pi^D(x,b) - \pi^C(x,b) = (1 - q_a(x))Z(x,b),$$

as stated in (8), where[24]

$$Z(x,b) \equiv \sigma_a^2 - [b^2 + q_p(x)\sigma_p^2].$$

Consider first the case of small CEO biases. If $b \to 0$, then $\frac{\partial \sigma_a^2}{\partial b} \to \frac{A}{3}$ and $\frac{\partial \sigma_p^2}{\partial b} \to \frac{P}{3}$. Hence, $\frac{\partial Z}{\partial b} \to (\frac{A}{3} - q_p(x)\frac{P}{3}) > 0$ since $A \geq P$ and $q_p(x) < 1$. Applying similar arguments as in Dessein (2002), it follows that when b tends to 0, delegation converges to first-best faster than centralization. So when b is sufficiently small, delegation dominates centralization for all $x \in [0,1]$.

Now consider the case of high CEO biases. When $b \geq \frac{A}{2\sqrt{3}}$, $b^2 \geq \frac{A^2}{12} \geq \sigma_a^2$. Then $Z(\cdot) = \sigma_a^2 - b^2 - q_p(x)\sigma_p^2 < 0$, i.e., centralization dominates delegation for all $x \in [0,1]$.

Lastly, if b is in the intermediate region, define $\hat{x}(b)$ by $Z(\hat{x}(b), b) = 0$. Because $q_p(\cdot)$ is decreasing in x, $Z(\cdot)$ is increasing in x. As a result, when $x > \hat{x}(b)$, $Z(x,b) > Z(\hat{x}(b), b) = 0$, with the consequence that shareholders prefer delegation, and vice versa.

Now we show that $\hat{x}(b)$ is increasing in b for $A = P$. At $\hat{x}(b)$,

$$Z(\hat{x}(b), b) \equiv 0$$

Applying the implicit function theorem:

$$\frac{d\hat{x}}{d(b^2)} = \frac{\frac{\partial \sigma_a^2}{\partial b^2} - \left[1 + q_p(\hat{x}(b))\frac{\partial \sigma_p^2}{\partial b^2}\right]}{q_p'(\hat{x})\sigma_p^2}.$$

The denominator is negative. For the numerator, when $A = P$, $\sigma_a^2(b) = \sigma_p^2(b) \equiv \sigma^2(b)$, then by $Z(\hat{x}(b), b) \equiv 0$, we have

$$1 - q_p(\hat{x}(b)) = \frac{b^2}{\sigma^2(b)}.$$

[24] As in the main text we shall suppress the functional dependence of key functions on A and P, if there is no risk of confusion.

Then, the numerator of $d\hat{x}/d(b^2)$ almost everywhere (ignoring values of b at which $N(\cdot)$ changes value) equals

$$\frac{N^2-1}{3} - \left[1 + q_p(\hat{x}(b))\frac{N^2-1}{3}\right] = \frac{N^2-1}{3}[1-q_p(\hat{x}(b))] - 1 = \frac{N^2-1}{3}\frac{b^2}{\sigma^2(b)} - 1.$$

Using (13), we have

$$\frac{N^2-1}{3}\frac{b^2}{\sigma^2(b)} = \frac{\frac{N^2-1}{3}b^2}{\left(\frac{\overline{\sigma}}{N(b)}\right)^2 + \frac{(N(b))^2-1}{3}b^2} < 1.$$

Hence, $\hat{x}(b)$ is increasing almost everywhere. Finally, note that $\hat{x}(b)$ is continuous at the values of b at which $N(\cdot)$ changes value. Hence, we can conclude that $\hat{x}$ is globally increasing in b for $A = P$. ∎

Proof of Lemma 2. The optimal $x^C(b)$ and $x^D(b)$ are obtained, respectively, by inspecting the following first-order derivatives:

$$\frac{\partial \pi^C(x,b)}{\partial x} = q_p'(x)\overline{\sigma}_p^2 + q_a'(x)\sigma_a^2, \quad (14)$$

$$\frac{\partial \pi^D(x,b)}{\partial x} = q_p'(x)\overline{\sigma}_p^2 + q_a'(x)b^2 + B(x)\sigma_p^2, \quad (15)$$

where

$$B(x) \equiv q_a'(x)q_p(x) - (1-q_a(x))q_p'(x) > 0.$$

Given Assumption 2, when $b > 0$, we get interior solutions for both regimes. The optimal interior solutions $x^C(b)$ and $x^D(b)$, respectively, are obtained by setting (14) and (15) each equal to zero. Therefore,

$$\begin{aligned}\left.\frac{\partial \pi^D(x,b)}{\partial x}\right|_{x=x^C} &= q_p'(x^C)\overline{\sigma}_p^2 + q_a'(x^C)b^2 + B(x^C)\sigma_p^2 \\ &= -q_a'(x^C)\sigma_a^2 + q_a'(x^C)b^2 + B(x^C)\sigma_p^2\end{aligned}$$

If $A = P$, then $\sigma_a^2 = \sigma_p^2$, and therefore

$$\left.\frac{\partial \pi^D(x,b)}{\partial x}\right|_{x=x^C} = -\sigma_p^2\{q_a'(x^C)(1-q_p(x^C)) + (1-q_a(x^C))q_p'(x^C)\} + q_a'(x^C)b^2.$$

Let

$$h(x) \equiv q_a'(x)(1-q_p(x)) + (1-q_a(x))q_p'(x).$$

Given Assumption 2,

$$h(x) = \eta z(1-x)^{z-1}x^{z-1}[(1-x)^{1-z} - x^{1-z} - \eta(1-2x)].$$

Let $w(x) \equiv (1-x)^{1-z} - x^{1-z} - \eta(1-2x)$. Then, $\lim_{\eta\to 1} w(0) = \lim_{\eta\to 1}(1-\eta) = 0$ and $w(\frac{1}{2}) = 0$. Moreover, $w''(x) = (1-z)z(x^{-z-1} - (1-x)^{-z-1}) \geq 0$ for $x \in [0, 1/2]$, i.e., $w(x)$ is a convex function for $x \in [0, 1/2]$. Therefore, by Jensen's inequality, $w(x) \leq 0$ for $x \in [0, 1/2]$. Hence $h(x) \leq 0$ for $x \in [0, 1/2]$.

Next, we show that $x^C(b) \in (0, 1/2]$, where (the assumed interior) $x^C(b)$ is again obtained by setting (14) equal to zero. When $b > A/4$ and $A = P$, then $\sigma_a^2 = \overline{\sigma}_p^2$. In that case, x^C is given by $q_p'(x^C) + q_a'(x^C) = 0$. Under Assumption 2, $q_p'(x^C) = -q_a'(1-x^C)$, which directly leads to $x^C(b) = 1/2$ for any $b > A/4$. Since $x^C(b)$ is non-decreasing in b (Lemma 3), $x^C(b) \in (0, 1/2]$ for all b. Therefore,

$$\left.\frac{\partial \pi^D}{\partial x}\right|_{x=x^C(b)} = -\sigma_p^2 h(x^C) + q_a'(x^C)b^2 > 0 = \left.\frac{\partial \pi^D}{\partial x}\right|_{x=x^D(b)}.$$

Then, by the global concavity of $\pi^D(x, b)$, we have $x^D(b) > x^C(b)$. ■

Proof of Proposition 1. This proof provides the solution to the shareholders's optimization program $\mathcal{P}_0$.

Part (i). For b sufficiently small, Lemma 1 shows that $\pi^D(x, b) \geq \pi^C(x, b)$ for any x. Then, $\pi^D(x^D) \geq \pi^D(x^C) \geq \pi^C(x^C)$. The first inequality is by revealed preference and the second inequality follows from $\pi^D(x, b) \geq \pi^C(x, b)$ for any x. Hence, for b sufficiently small, the shareholders chooses delegation.

Part (ii). The proof just reverses the chain of arguments in part (i).

Part (iii). Define $\hat{b}$ by $\pi^D(x^D(\hat{b}), b) = \pi^C(x^C(\hat{b}), b)$. By continuity of $\pi^D(x^D(b), b) - \pi^C(x^C(b), b)$, and the fact that $\pi^D(x^D(b), b) - \pi^C(x^C(b), b) > 0$

for b small and $\pi^D(x^D(b),b) - \pi^C(x^C(b),b) < 0$ for b large, there exists at least one $\hat{b}$ such that $\pi^D(x^D(\hat{b}),\hat{b}) - \pi^C(x^C(\hat{b}),\hat{b}) = 0$.

We now show that $x^C(\hat{b}) < \hat{x}(\hat{b}) < x^D(\hat{b})$. We proceed in two steps: we first show that this chain of inequalities holds in weak form, and then rule out equality. First, suppose that $x^C(\hat{b}) > \hat{x}(\hat{b})$. By Lemma 1, then

$$\pi^D(x^C(\hat{b}),b) > \pi^C(x^C(\hat{b}),b), \quad \text{for any } b.$$

Yet, by definition of $\hat{b}$,

$$\pi^D(x^D(\hat{b}),\hat{b}) \equiv \pi^C(x^C(\hat{b}),\hat{b}).$$

Thus, $\pi^D(x^C(\hat{b}),\hat{b}) > \pi^D(x^D(\hat{b}),\hat{b})$, which contradicts the optimality of $x^D(\cdot)$. As a result, $x^C(\hat{b}) \leq \hat{x}(\hat{b})$. A similar argument shows that $\hat{x}(\hat{b}) \leq x^D(\hat{b})$. Therefore, $x^C(\hat{b}) \leq \hat{x}(\hat{b}) \leq x^D(\hat{b})$.

Now, we show that the preceding chain of inequalities holds in a strict sense. Suppose that $x^D(\hat{b}) = x^C(\hat{b}) \equiv \bar{x}(\hat{b})$. Then it must also be true that $\hat{x}(\hat{b}) = \bar{x}(\hat{b})$. Then, by virtue of $x^S(b) \in \arg\max_x \pi^S(x,b)$, the following must hold:

$$\frac{\partial}{\partial x}\pi^C(\bar{x}(\hat{b}),\hat{b}) = \frac{\partial}{\partial x}\pi^D(\bar{x}(\hat{b}),\hat{b}) \quad = \quad 0$$

Using the proof of Lemma 2, the preceding condition implies that

$$\begin{aligned}
q_a'(\bar{x}(\hat{b}))\left[\sigma_a^2 - \left(\hat{b}^2 + \frac{B(\bar{x}(\hat{b}))}{q_a'(\bar{x}(\hat{b}))}\sigma_p^2\right)\right] &= 0 \\
\iff \quad \underbrace{Z(\bar{x}(\hat{b}))}_{=0} + \left[q_p(\bar{x}(\hat{b})) - \frac{q_a'(\bar{x}(\hat{b}))q_p(\bar{x}(\hat{b})) - (1-q_a(\bar{x}(\hat{b})))q_p'(\bar{x}(\hat{b}))}{q_a'(\bar{x}(\hat{b}))}\right]\sigma_p^2 &= 0 \\
\iff \quad (1-q_a(\bar{x}(\hat{b})))q_p'(\bar{x}(\hat{b})) &= 0,
\end{aligned}$$

a contradiction. Thus, $x^D(\hat{b}) > x^C(\hat{b})$. Using a simple geometrical argument shows that $\hat{x}(\hat{b}) \in (x^C(\hat{b}), x^D(\hat{b}))$ must then hold.

Hence, at the cutoff $\hat{b}$ where the optimal organization structure switches from D to C, $x^*(b)$ jumps down from $x^D(b)$ to $x^C(b)$. Other than at such discontinuities, $x^*(b)$ is increasing in b, by Lemma 3. ■

Proof of Proposition 2. We now solve the CEO's program $\mathcal{P}_1$ and compare its solution, denoted by $x_{CEO}(b)$, with that of the benchmark program $\mathcal{P}_0$, denoted $x^*(b)$.

The CEO's program $\mathcal{P}_1$ can be decomposed as follows. First, find the CEO's optimal payoff $\pi^{S^*}_{CEO}(b)$ when a certain S is induced:

$$\begin{aligned} \pi^{D^*}_{CEO}(b) &= \max_x \ \pi^D_{CEO}(x,b), \quad \text{s.t.} \ \ x \geq \hat{x}(b), \\ \pi^{C^*}_{CEO}(b) &= \max_x \ \pi^C_{CEO}(x,b), \quad \text{s.t.} \ \ x < \hat{x}(b). \end{aligned}$$

In a second step, compare $\pi^{D^*}_{CEO}(b)$ and $\pi^{C^*}_{CEO}(b)$ and pick the larger one.

Denote by $x^{D^*}_{CEO}(b)$ and $x^{C^*}_{CEO}(b)$, respectively, the solutions to the above constrained maximization problems. Analogously, denote by $x^D_{CEO}(b)$ and $x^C_{CEO}(b)$, respectively, the solutions to the corresponding unconstrained maximization problems. It is immediate that $x^C_{CEO}(b) = x^C(b)$, because under centralization the CEO's preference over the board composition coincides with that of the shareholders (except for the lump-sum bias cost b^2).

Part (i): $b < \hat{b}$. First, we show that $\hat{x}(b) < x^D(b)$ for $b < \hat{b}$. Suppose on the contrary that $\hat{x}(b) \geq x^D(b)$. Then,

$$\pi^C(x^C(b), b) \geq \pi^C(x^D(b), b) \geq \pi^D(x^D(b), b).$$

The first inequality follows by revealed preference and the second inequality follows from $\hat{x}(b) \geq x^D(b)$ together with Lemma 1. However, by the definition of $\hat{b}$ and the fact that we only allow a unique $\hat{b}$ (footnote 17), $\pi^D(x^D(b), b) > \pi^C(x^C(b), b)$ for $b < \hat{b}$. Contradiction. Thus, $\hat{x}(b) < x^D(b)$ for $b < \hat{b}$.

Secondly, since $x^D(b) > \hat{x}(b)$, a revealed preference argument shows that $\pi^{D*}_{CEO}(b) \geq \pi^D_{CEO}(x^D(b), b)$. At the same time $\pi^{C*}_{CEO}(b) \leq \pi^C_{CEO}(x^C_{CEO}(b), b)$ can also be established by revealed preference, because the first term is the optimum of a constrained optimization problem, while the latter is the optimum of the corresponding unconstrained problem. Therefore

$$
\begin{aligned}
\pi^{D*}_{CEO}(b) - \pi^{C*}_{CEO}(b) &\geq \pi^D_{CEO}(x^D(b), b) - \pi^C_{CEO}(x^C_{CEO}(b), b) \\
&= \pi^D_{CEO}(x^D(b), b) - \pi^C_{CEO}(x^C(b), b) \qquad \text{(by } x^C_{CEO}(b) = x^C(b)\text{)} \\
&= \pi^D(x^D(b), b) + [1 - 2q_a(x^D(b))]b^2 - \pi^C(x^C(b), b) + b^2 \qquad \text{(by (7))} \\
&= \pi^D(x^D(b), b) - \pi^C(x^C(b), b) + 2[1 - q_a(x^D(b))]b^2 \\
&> 0 \qquad \text{(since } \pi^D(x^D(b), b) > \pi^C(x^C(b), b) \text{ when } b < \hat{b}\text{)}
\end{aligned}
$$

Hence, when $b < \hat{b}$, the CEO will choose $x_{CEO}(b) = x^{D*}_{CEO}(b)$, followed by the board choosing $S^*(x_{CEO}(b)) = D$.

Next, we show that $x^{D*}_{CEO}(b) < x^D(b)$. To see this, notice that

$$\frac{\partial \pi^D_{CEO}(x, b)}{\partial x} < \frac{\partial \pi^D(x, b)}{\partial x},$$

for any x. Then

$$\left.\frac{\partial \pi^D_{CEO}(x, b)}{\partial x}\right|_{x \geq x^D(b)} < \left.\frac{\partial \pi^D(x, b)}{\partial x}\right|_{x \geq x^D(b)} \leq 0$$

The last inequality comes from the global concavity of $\pi^D(x, b)$ and the optimality of $x^D(b)$. Therefore, $\hat{x}(b) \leq x^{D*}_{CEO}(b) < x^D(b)$. (Recall that $\hat{x}(b) < x^D(b)$ when $b < \hat{b}$.)

Part (ii): b is sufficiently large. In this case, centralization dominates delegation for all x (by Lemma 1). Then manipulating x to get decentralization is not feasible for CEO. Then CEO will choose $x_{CEO}(b) = x^C(b)$ and $S^*(x_{CEO}(b)) = C$.

Part (iii): $\hat{b} \leq b \leq \hat{b} + \varepsilon$, where $\varepsilon \to 0$. Proposition 1 shows that $x^C(\hat{b}) < \hat{x}(\hat{b}) < x^D(\hat{b})$. By continuity, $x^C(b) < \hat{x}(b) < x^D(b)$ holds also for $\hat{b} \leq b \leq \hat{b} + \varepsilon$.

Then the same revealed preference arguments as in part (i) show that $\pi_{CEO}^{D^*}(b) \geq \pi_{CEO}^{D}(x^D(b), b)$ and $\pi_{CEO}^{C^*}(b) \leq \pi_{CEO}^{C}(x_{CEO}^{C}(b), b)$. Therefore

$$\begin{aligned}
\pi_{CEO}^{D^*}(b) - \pi_{CEO}^{C^*}(b) &\geq \pi_{CEO}^{D}(x^D(b), b) - \pi_{CEO}^{C}(x_{CEO}^{C}(b), b) \\
&= \pi_{CEO}^{D}(x^D(b), b) - \pi_{CEO}^{C}(x^C(b), b) \qquad \text{(by } x_{CEO}^{C}(b) = x^C(b)\text{)} \\
&= \pi^D(x^D(b), b) - \pi^C(x^C(b), b) + 2[1 - q_a(x^D(b))]b^2
\end{aligned}$$

Clearly, when $\varepsilon \to 0$, then $\pi^D(x^D(b), b) - \pi^C(x^C(b), b) \to 0$, and therefore $\pi_{CEO}^{D^*}(b) - \pi_{CEO}^{C^*}(b) > 0$ because $q_a(x^D(b)) < 1$.

So when b is "slightly above" $\hat{b}$, CEO will choose $x_{CEO}(b) = x_{CEO}^{D^*}(b) \geq \hat{x}(b) > x^C(b)$ followed by the board choosing $S^*(x_{CEO}(b)) = D$. ■

Proofs of Lemma 4

Part (i). Using the definitions of $\hat{A}(\cdot)$ and $\tilde{A}(\cdot)$ in the main text, we have:

$$\begin{aligned}
\pi^C(x^D(b), \hat{A}(b), b) &\leq \pi^C(x^C(\hat{A}(b), b), \hat{A}(b), b), \qquad \text{by revealed preference} \\
&= \pi^D(x^D(b), b), \qquad \text{by (11)} \\
&= \pi^C(x^D(b), \tilde{A}(x^D(b), b), b), \qquad \text{by (12)}
\end{aligned}$$

(Note that π^D does not directly depend on A, as a will always be perfectly known to the decision maker.) Therefore, $\tilde{A}(x^D(b), b) \leq \hat{A}(b)$, because $\partial\pi^C/\partial A < 0$. Furthermore, $b \leq \hat{b}(\bar{A})$ implies $\hat{A}(b) \leq \bar{A}$, where $\hat{A}(b)$ is defined by (11). As a result, $\tilde{A}(x^D(b), b) \leq \hat{A}(b) \leq \bar{A}$ for any $b \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$.

Part (ii). Recall from the main text that $\tilde{A}(x, b) > \bar{A}$ is equivalent to $x < \hat{x}(\bar{A}, b)$. To prove that for $\bar{A}$ sufficiently small, constraint (C) in program $\mathcal{P}_{\mathcal{C}}$ will be slack for any $b \geq \hat{b}(\underline{A})$, we first show that $\hat{x}(\underline{A}, b) \geq x^C(\underline{A}, b)$ for any $b \geq \hat{b}(\underline{A})$. Suppose not. Then, by Lemma 1, $\pi^D(x^C(\underline{A}, b), b) > \pi^C(x^C(\underline{A}, b), \underline{A}, b)$ since $\hat{x}(\underline{A}, b) < x^C(\underline{A}, b)$. By Proposition 1, $\pi^C(x^C(\underline{A}, b), \underline{A}, b) \geq \pi^D(x^D(b), b)$ for any $b \geq \hat{b}(\underline{A})$. Therefore, $\pi^D(x^C(\underline{A}, b), b) > \pi^D(x^D(b), b)$, which contradicts the optimality of $x^D(b)$. As a result, $\hat{x}(\underline{A}, b) \geq x^C(\underline{A}, b)$ for any $b \geq \hat{b}(\underline{A})$. Therefore $\lim_{\bar{A} \to \underline{A}} \hat{x}(\bar{A}, b) = \hat{x}(\underline{A}, b) \geq x^C(\underline{A}, b)$ for any $b \geq \hat{b}(\underline{A})$, i.e., for $\bar{A}$ sufficiently small, constraint (C) in program $\mathcal{P}_{\mathcal{C}}$ is slack for any $b \geq \hat{b}(\underline{A})$.

On the other hand, when $\bar{A} > \tilde{A}(x^C(\underline{A}, \hat{b}(\underline{A})), \hat{b}(\underline{A}))$, which is equivalent to $x^C(\underline{A}, \hat{b}(\underline{A})) > \hat{x}(\bar{A}, \hat{b}(\underline{A}))$, constraint (C) will be binding at $\hat{b}(\underline{A})$. At the same time,

$$\begin{aligned} x^C(\underline{A}, \hat{b}(\bar{A})) &< x^C(\bar{A}, \hat{b}(\bar{A})), && \text{since } \partial x^C/\partial A > 0 \\ &\leq \hat{x}(\bar{A}, \hat{b}(\bar{A}))), && \text{by Proposition 1} \end{aligned}$$

Then by continuity of $x^C(\cdot)$ and $\hat{x}(\cdot)$, there must exist at least one $b^C \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$ such that $x^C(\underline{A}, b^C) = \hat{x}(\bar{A}, b^C)$. ■

Proof of Proposition 3

As shown in Lemma 4, part (ii), when $\bar{A} > \tilde{A}(x^C(\underline{A}, \hat{b}(\underline{A})), \hat{b}(\underline{A}))$, constraint (C) will be binding at $\hat{b}(\underline{A})$ and there must exist at least one $b^C \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$ such that $x^C(\underline{A}, b^C) = \hat{x}(\bar{A}, b^C)$. If there exist multiple thresholds b^C, we only focus on the smallest one. Denote by $V_C(\cdot)$ as the value of program $\mathcal{P}_{\mathcal{C}}$ and by $V_D(\cdot)$ the value of program $\mathcal{P}_{\mathcal{D}}$.

*Part (i). **[Acquiesce]*** Since constraint (C) is binding at $\hat{b}(\underline{A})$, then $V_C(\cdot) < \pi^C(x^C(\underline{A}, b), \underline{A}, b)$ due to the shadow cost of (C). On the other hand, $V_D(\cdot) = \pi^D(x^D(b), b)$ since constraint (D) is always slack. Therefore

$$\begin{aligned} \lim_{\varepsilon \to 0} V_C(\hat{b}(\underline{A}) + \varepsilon) &< \lim_{\varepsilon \to 0} \pi^C(x^C(\underline{A}, \hat{b}(\underline{A}) + \varepsilon), \underline{A}, \hat{b}(\underline{A}) + \varepsilon) \\ &= \lim_{\varepsilon \to 0} \pi^D(x^D(\hat{b}(\underline{A}) + \varepsilon), \hat{b}(\underline{A}) + \varepsilon) \\ &= \lim_{\varepsilon \to 0} V_D(\hat{b}(\underline{A}) + \varepsilon). \end{aligned}$$

Hence for b "slightly above" $\hat{b}(\underline{A})$, the shareholders will prefer to acquiesce to CEO entrenchment and choose $x = x^D(b)$. In response, the CEO will choose

$A = \tilde{A}(x^D(b), b)$ to noise up the project complexity by just enough so as to ensure delegation.

Part (ii). ***[Preempt at a cost]*** At the threshold b^C, constraint (C) is just binding, therefore $V_C(b^C) = \pi^C(x^C(\underline{A}, b^C), \underline{A}, b^C)$. For $b = b^C - \delta$,

$$\begin{aligned}\lim_{\delta \to 0} V_C(b^C - \delta) = V_C(b^C) &= \pi^C(x^C(\underline{A}, b^C), \underline{A}, b^C) \\ &> \pi^D(x^D(b^C), b^C), \qquad \text{since } b^C > \hat{b}(\underline{A}) \\ &= \lim_{\delta \to 0} \pi^D(x^D(b^C - \delta), b^C - \delta) \\ &= \lim_{\delta \to 0} V_D(b^C - \delta)\end{aligned}$$

Hence for $b = b^C - \varepsilon$, the constraint (C) is still binding, yet the shareholders are better off choosing $x = \hat{x}(\bar{A}, b)$ to preempt CEO entrenchment. In response, the CEO will choose $A = \underline{A}$.

Part (iii). ***[Preempt at no cost]*** For $b = b^C + \delta$, similarly as Case (ii), we can show that $\lim_{\delta \to 0} V_C(b^C + \delta) > \lim_{\delta \to 0} V_D(b^C + \delta)$. Hence for b slightly above b^C, constraint (C) becomes slack, the shareholders will choose $x = x^C(\underline{A}, b)$ to preempt CEO entrenchment. In response, the CEO will choose $A = \underline{A}$. ■

References

Adams, R. (2009), Asking directors about their dual roles, working paper, University of Queensland.

Adams, R., and D. Ferreira (2007), A theory of friendly boards, *Journal of Finance*, 62 (1), 217-250.

Adams, R., B. E. Hermalin, and M. S. Weisbach, (2009), The role of boards of directors in corporate governance: a conceptual framework & survey, *Journal of Economic Literature*, forthcoming.

Aghion, P., and J. Tirole (1997), Formal and real authority in organizations, *Journal of Political Economy* 105(1), 1-29.

Baker, M., and P. Gompers (2003), The determinants of board structure at the initial public offering, Journal of Law and Economics, October 46(2), 569-597.

Bathala, C. and R. Rao (1995), The determinants of board composition: an agency theory perspective, *Managerial and Decision Economics* 16, 59-69.

Becht, M., J. Franks, C. Mayer, and S. Rossi (2008), Returns to shareholder activism: evidence from a clinical study of the Hermes U.K. Focus Fund, *Review of Financial Studies*, forthcoming.

Bhagat, S. and B. Black (1999), The uncertain relationship between board composition and firm performance, *Business Lawyer* 54, 921-963.

Chen, Y., N. Kartik, and J. Sobel (2008), Selecting Cheap-Talk Equilibria, *Econometrica* 76(1), 117-136.

Coles, J., N. Daniel, N., L. Naveen (2008), Boards: does one size fit all? *Journal of Financial Economics* 87, 329-356.

Crawford, V. P., and J. Sobel (1982), Strategic information transmission, *Econometrica* 50(6), 1431-1451.

Dessein, W. (2002), Authority and communication in organizations, *Review of Economic Studies* 69, 811-838.

Edlin, A., and J. Stiglitz (1995), Discouraging rivals: managerial rent-seeking and economic inefficiencies, *American Economic Review*, 85(5), 1301-1312.

Eisenberg, T., S. Sundgren, and M. Wells (1998), Larger Board Size and Decreasing Firm Value in Small Firms, *Journal of Financial Economics* 48, 35-54.

Harris, M., and A. Raviv (2005), Allocation of decision-making authority, *Review of Finance* 9, 353-383.

Harris, M., and A. Raviv (2008), A theory of board control and size, *Review of Financial Studies* 21(4), 1797-1832.

Jensen, M., (1993), The modern industrial revolution, exit, and the failure of internal control systems, *Journal of Finance*, 48(3), 831-880.

Lehn, K., S. Patro, and M. Zhao (2004), Determinants of the size and structure of corporate boards, working paper, University of Pittsburgh.

Linck, J., J. Netter, and T. Yang (2008), The determinants of board structure, Journal of Financial Economics 87, 308-387.

Linck, J., J. Netter, and T. Yang (2009), The effects and unintended consequences of the Sarbanes-Oxley act on the supply and demand for directors, Review of Financial Studies 22(8), 3287-3328.

Melumad, N., and Shibano, T. (1991), Communication in Settings with No Transfers, *The RAND Journal of Economics* 22(2), 173-198.

Raheja, C. (2005), Determinants of board size and composition: a theory of corporate boards, Journal of Financial and Quantitative Analysis, 40(2), 283-306.

Shleifer, A., and R. Vishny (1989), Management entrenchment: the case of manager-specific investments, *Journal of Financial Economics* 25, 123-139.

Yermack D. (1996) , Higher market valuation of companies with a small board of directors, *Journal of Financial Economics* 40, 185-211.

Zhang, I. (2007), Economic consequences of the SarbanesOxley Act of 2002, Journal of Accounting and Economics 44 (1-2), 74-115.