Information Externalities in Corporate Governance

Konrad Raff*
SIFR & Toulouse School of Economics

November 2010

Job Market Paper

Abstract

This paper studies the interactions between monitoring choices in different firms. Shareholders gather information about a common industry shock and subsequently intervene with management. An information externality arises because intervention transmits information about industry conditions to shareholders in competing firms. Closer monitoring in one firm improves the performance of industry peers. In equilibrium, all firms free-ride and there is underprovision of monitoring. A regulatory cap on CEO compensation can increase investor welfare. Disclosure of shareholder intervention is socially valuable and inefficiently low in equilibrium. Using US data on forced CEO turnover, I provide tentative empirical evidence of information spillovers across firms. Controlling for various performance measures, I find that CEOs are significantly more likely to be dismissed after other CEOs in their industry have been fired. At the same time, CEOs' turnover risk is not related to firings outside of their industry. The predictive power of peer decisions for turnover risk is significantly stronger for firms with low institutional blockholder ownership.

JEL Classification: G30, G38, J33, J63
Keywords: Corporate Governance, Monitoring, Externality, CEO Turnover, Regulation, CEO Pay

*Contact information: konrad.raff@sifr.org. I am very thankful to Chiara Canta, Catherine Casamatta, Denis Gromb, Alexander Gümbel and Per Strömberg for many helpful discussions. I am grateful to Dirk Jenter and Martijn Cremers for generously sharing their respective data. For useful comments I would like to thank Bo Becker, Guido Friebel, Morten Sorensen, Shan Zhao and workshop participants at the ECGTN-CEMFI Summer School on Corporate Governance in Santander, University of Frankfurt, University of Oxford, Stockholm School of Economics and Toulouse School of Economics.

1 Introduction

A large literature in corporate governance studies the free-rider problem that arises when a firm's ownership is dispersed. Each individual shareholder fails to consider that gathering information and controlling management benefits all shareholders jointly. As a result, there is too little monitoring and firm performance may be inefficiently low (see Stiglitz (1985) and Admati, Pfleiderer and Zechner (1994)). While extant work emphasizes coordination problems among investors in one firm, this paper analyzes the interaction between monitoring activities in different firms. I study how a particular interaction channel, information spillovers, affects shareholders' incentives to engage in monitoring. In the model firms are exposed to a common performance shock. Shareholders in each firm gather information about the shock and subsequently intervene with management, e.g. they can replace the CEO or adopt a new strategy.[1] The main premise of the paper is that intervention in one firm transmits information about common industry conditions to shareholders in competing firms.

More specifically, I consider a simple learning framework with two firms and two periods of production. Performance is driven by two a priori unknown factors, the ability of the CEO and a common industry shock.[2] Following a first-stage performance, shareholders decide over the retention of the manager. Dismissing an unsuitable manager always increases firm value, but first-period performance is only a noisy signal of ability. In addition, shareholders can engage in costly information acquisition to filter out the industry shock. This allows them to take an efficient firing decision. While the result of monitoring cannot be directly observed by outsiders, shareholders' information is fully revealed to the peer firm through their actions. The intuition is as follows: if a manager is fired despite of good performance, monitoring must have revealed that he is incompetent and merely benefited from a favorable industry shock. The decision prompts the shareholders in the other firm to re-evaluate their own CEO. They are now more suspicious that his performance was also driven by luck rather than ability, and they discount the posterior belief about his type accordingly. Overall, they can come to a better assessment of their manager by incorporating the peer decision.

The fact that intervention transmits information about industry conditions leads to an externality. Closer monitoring in one firm raises the value of its competitor. If shareholders devote more resources to monitoring, their actions become more informative. This in turn enables more efficient intervention in the rival firm. In equilibrium, shareholders fail to exploit fully the benefits from spillovers. Free-riding behavior across firms leads to underprovision of monitoring.[3]

[1] See Klein and Zur (2009) for recent evidence of shareholder activism.

[2] Exposure to a common shock could alternatively result from operating in the same geographic area. See Section 2 for further discussion.

[3] Bris, Brisley and Cabolis (2008) provide empirical evidence in support of positive spillover effects on peer performance. They consider firms whose governance improves after being taken over by an acquirer from a country with better investor protection and higher accounting standards.

The above insight has several implications. First, learning spillovers are a potential determinant of ownership structures across firms. Monitoring externalities can be removed by putting the firms under the control of a joint monitor who maximizes combined firm values. That is, the efficiency gains from eliminating the free-rider problem establish a role for joint ownership. Examples of monitors that are engaged in multiple firms are pension funds, activist hedge funds or venture capital firms. The externality only arises if the firms face common uncertainties. Hence, the analysis predicts that a joint monitor should pursue a focused portfolio strategy and be engaged in firms that operate in the same line of business or geographic area.

A second implication of learning spillovers across firms concerns shareholders' incentives to disclose information about their actions. I introduce the notion of "governance transparency" which refers to the observability of intervention decisions. It differs from extant studies which usually focus on disclosure of financial information or proprietary technological and strategic knowledge.[4] Learning spillovers hinge on such governance transparency; it is socially desirable because it reduces wasteful duplication of monitoring efforts. However, shareholders in each firm have an incentive to hide their actions and intervene informally or privately with management.[5] The intuition is that secrecy increases the incentives to monitor at competing firms and thus facilitates free-riding. Importantly, the bias towards secrecy arises even in the absence of any direct costs of information disclosure. The result underscores the importance of regulatory reporting requirements for investors. For instance, in the US a large shareholder, who intents to influence the firm's policies, has to file a SEC Schedule 13D and declare the purpose and goal of his transactions. My model suggests that such requirements should be very broadly defined to exploit monitoring spillovers.

Third, the model establishes a role for governance regulation to foster monitoring. To study regulation I extend the baseline model by a moral hazard dimension. The CEO can raise his productivity through unobservable effort. I show that a binding regulatory cap on the level of CEO compensation can eliminate the free-rider problem and increase investor welfare. The intuition is the following: shareholders dispose of two instruments, monitoring and compensation, to incentivize the CEO to exert effort. The two instruments are substitutes, i.e., closer monitoring reduces the level of compensation that is needed to induce a given level of effort. In equilibrium, the free-rider problem discourages monitoring, forcing all firms to pay inefficiently high compensation. If the regulator limits compensation below its equilibrium level, shareholders optimally respond by monitoring their CEO more closely in order to preserve incentives. Forcing shareholders in each firm to deviate from their privately optimal mix between monitoring and compensation can make all of them jointly

They find that mergers of this kind raise the Tobin's Q of the entire industry of the target.

[4]Leuz and Wysocki (2008) survey the literature on financial disclosure. Bhattacharya and Chiesa (1995) study the incentives to reveal proprietary technological information.

[5]See Carleton, Nelson and Weisbach (1998) and Becht, Franks, Mayer and Rossi (2010) for examples of private shareholder activism.

better off because information acquisition is a public good. Intuitively, regulatory intervention in the form of a pay-cap serves as a commitment device to monitor more closely thereby eliminating the coordination failure between firms. The optimal pay-cap is lower if investor protection is strong.

I provide tentative empirical evidence for the paper's central premise of information spillovers. The model predicts a positive correlation between intervention decisions across firms. The intuition is that learning spillovers lead to imitative behavior. In particular, a CEO is ceteris paribus more likely to be dismissed after one of his peers has been fired. The correlation only arises if firms are exposed to common uncertainties. In the absence thereof, no information spillovers occur and intervention decisions are independent. Using US data on forced CEO turnover I find that CEOs are indeed significantly more likely to be dismissed after other CEOs in their industry have been fired. At the same time, CEOs' turnover risk is unrelated to past firings outside of their own industry. The results hold controlling for a variety of performance measures.

Furthermore, the model predicts that shareholders rely on peer actions for information when they failed to monitor successfully themselves. Using institutional blockownership as a proxy for shareholder informedness, I find that CEO turnover is significantly more sensitive to peer decisions in firms with low institutional blockholder ownership.

The paper is related to several strands of the literature. A few recent theoretical papers analyze the interaction between firms' choices of corporate governance. Acharya and Volpin (2010) and Dicks (2010) consider a labor market with scarce managerial talent. Their externality differs from the one in this paper: poorly governed firms overpay their managers which forces all other firms to pay their managers higher compensation in order to prevent them from leaving. In contrast to Acharya and Volpin, this paper explores the implications of spillover effects for CEO turnover and shareholders' incentives to disclose information. Dicks focuses exclusively on regulatory responses to the externality and considers as instruments the direct imposition of governance standards, subsidies and taxes. I study a different instrument and show that imposing a regulatory cap on compensation can enhance investor welfare.

The paper is also related to the literature on relative performance evaluation or benchmarking. This approach holds that peer performance provides valuable information whenever firms are exposed to common shocks (see Holmström (1979, 1982)). Shareholders can use this information to design better incentive schemes or more generally improve decision-making.[6] In so doing, their role is essentially a passive one. They simply make efficient use of freely available, informative signals. This paper goes beyond the benchmarking paradigm by allowing for a more active role of shareholders. They can engage in costly monitoring and interfere with their

[6]Many empirical studies document the use of relative-performance evaluation in CEO turnover (Warner, Watts and Wruck (1988), Morck, Shleifer and Vishny (1989), Barro and Barro (1990) and Kaplan and Minton (2008).

firm. By taking the actions at rival firms into account, shareholders can come to a more precise assessment of their firm than if they only considered the performance vector.

Finally, two recent papers provide empirical evidence of peer group effects in corporate governance. They focus on different governance mechanisms than this paper. John and Kadyrzhanova (2010) study inter-firm spillovers in the market for corporate control. They find that antitakeover provisions in one firm divert takeover pressure to substitute targets with better governance. Acharya, Gabarro and Volpin (2010) focus on interactions that arise through competition between firms in the CEO labor market. They provide evidence that executive compensation depends on other firms' governance standards.

The next section presents the setup of the baseline model. Section 3 uncovers the monitoring externality and presents tentative empirical evidence of linkages between shareholder intervention in different firms. A market-based solution to the free-rider problem in the form of common monitoring is presented in Section 4. The next section adds a moral hazard dimension to the baseline model and analyzes regulation in the form of a cap on CEO compensation. The incentives of shareholders to disclose intervention decisions to rival firms are studied in Section 6. The link between free-riding behavior across and within firms is analyzed in Section 7. The conclusion is in Section 8.

2 Model

We consider an economy with two ex-ante identical firms and two periods of production. Each firm is run by a single manager and owned by a different group of shareholders. Everyone is risk-neutral and there is no discounting. The sequence of events in each firm unfolds as follows:

At date $t = 0$, shareholders in firm $i \in \{1, 2\}$ hire a manager to run their firm. There are two types of managers, $\theta_i \in \left\{\bar{\theta}, \underline{\theta}\right\}$. Managers' types are not observed by anyone, including the managers themselves. The probability that a manager is a high ability type, $\theta_i = \bar{\theta}$, is given by $p \in (0, 1)$.[7] The manager's type refers to his ability to identify a suitable strategy or to implement a given strategy successfully, for example.

Also at $t = 0$, both firms are hit by a common performance shock s that cannot be directly observed by anyone. The shock is either positive ($s = g$) or negative ($s = b$). While s is a priori unobservable, its prior distribution is common knowledge: with probability q both firms are exposed to a positive shock, with probability $1 - q$ they suffer a negative one.

[7] The assumption that the manager does not know his own type can be justified on the grounds that no one is born knowing whether he or she will be a competent CEO. Like the board, the manager only learns from actual performance whether he or she is suitable for the tasks demanded of him.

At date $t = 1$, firm i generates a cash flow $X_{1i}(\theta_i, s)$ which depends on both the type of its manager and the common shock. The cash flow can take three possible values, $\bar{x}_1 > x_1 > \underline{x}_1 > 0$, where

$$X_{1i}(\theta_i, s) = \begin{cases} \bar{x}_1 & \text{if } \theta_i = \bar{\theta}, s = g \\ x_1 & \text{if } \theta_i = \bar{\theta}, s = b \text{ or } \theta = \underline{\theta}, s = g \\ \underline{x}_1 & \text{if } \theta_i = \underline{\theta}, s = b \end{cases}$$

Cash flow is high ($X_{1i} = \bar{x}_1$) if and only if the manager is competent and the shock is favorable. On the contrary, incompetent management and a negative shock imply a low outcome ($X_{1i} = \underline{x}_1$). Otherwise, the cash flow is intermediate.

After the first period of production, shareholders can take an action and intervene in the operation of their firm. More precisely, they decide whether to fire the manager. A new manager can be drawn from the same pool as the incumbent, i.e., he turns out to be competent with probability p. Let $D_i \in \{f, r\}$ denote the shareholders' decision in firm i, where f corresponds to intervention (firing) and r to no intervention (retention). Shareholders either take a decision immediately at date $t = 2$ or they postpone it until a later date $t = 2.5$.

At date $t = 3$, firm i produces a second-period cash flow $X_{2i}(\theta)$. To simplify the exposition, it is assumed to be fully determined by the type of the manager who is in charge at this point:

$$X_{2i}(\theta_i) = \begin{cases} 0 & \text{if } \theta_i = \underline{\theta} \\ x_2 > 0 & \text{if } \theta_i = \bar{\theta} \end{cases}$$

While a good manager succeeds in producing x_2, an incompetent one always produces a cash flow of 0.

All cash flows are publicly observable. In addition, shareholders can at a cost acquire information about the circumstances under which their manager performs. More precisely, at $t = 0$ shareholders in firm i choose a monitoring intensity $m_i \in [0, 1]$. At date $t = 1$, they can then observe directly the type θ_i of their manager with probability m_i (in addition to first-period cash flows). With probability $1 - m_i$ monitoring fails and shareholders only observe first-period cash flows. Information acquisition comes at a quadratic cost $k(m_i) = \frac{1}{2}c_i m_i^2$ that is jointly borne by all shareholders of firm i. The cost can be interpreted as the resources devoted to directly assessing managerial competence, for instance by installing an accounting system of quality m_i. Monitoring levels are chosen simultaneously and non-cooperatively in the two firms.

The setup captures in a simple manner the notion of active monitoring. After gathering information, shareholders can intervene with the operation of the firm by replacing the manager. Several features of the model deserve to be stressed: the subsequent results are insensitive to the exact timing of the monitoring activity. Firms could alternatively be assumed to choose m_i after the first period of production,

rather than before. Since the focus of the paper is on externalities and coordination failures across firms, the baseline model abstracts from any problems of this kind within firms for ease of exposition. Shareholders jointly select m_i to maximize the value of their firm and share the cost $k(m_i)$. In an extension (Section 7) I consider a more general setting with frictions at the firm level: information gathering is assumed to impose non-contractible private costs on the monitor. In this case, incentives to acquire information are provided through ownership concentration. It is shown that frictions within firms amplify the coordination problem across firms. Finally, to keep the analysis simple, I abstract from any product market interactions between the two firms.

The common performance shock admits different interpretations. First, the firms may operate in the same industry or line of business in which case s could be an industry-wide shock. Following this interpretation, the model's results should be particularly relevant for those sectors in which industry-wide shocks are important drivers of performance relative to firm-specific or macroeconomic factors.[8] Such shocks could result, for instance, from technological innovation, regulation or a change in consumers' tastes. Mitchell and Mulherin (1996) provide many examples of specific industry-wide shocks in the US, e.g. the deregulation of the airline industry in the 1970s or the adoption of an antiprotectionist foreign trade policy in the textile industry. While a regulatory or legal change itself may be observable, what matters for the analysis is that shareholders cannot a priori perfectly filter out the effect on performance. An alternative interpretation would be that the two firms operate in the same location rather than the same industry in which case s could be a shock to basic market conditions within a certain geographic area.

The above interpretation of θ_i as managerial ability and of shareholder intervention as firing the CEO is only one possibility. The model encompasses many other situations. In general, θ_i can be understood as any uncertain, idiosyncratic determinant of firm performance which is learned about over time and can be altered by shareholders' actions. For instance, the type θ_i could refer to the quality or suitability of the firm's product-market strategy or financial policy. In this case, interference by shareholders can be understood as a strategy or policy change (which may or may not be linked to the decision whether to fire the CEO).

3 Monitoring and Information Spillovers

This section shows that closer monitoring in one firm benefits its peer. The reason is that informed actions by shareholders transmit valuable information about industry conditions to outsiders. In equilibrium, firms free ride and collect too little information. Using US data on forced CEO turnover, I provide tentative empirical evidence

[8] Analysing the volatility of common stocks at the market, industry, and firm levels, Campbell et al. (2001) document a great variation across different industries. For example, they find that industry-wide uncertainty plays a much more important role in the IT sector than in the utilities sector.

of information spillovers across firms.

3.1 Learning from Intervention in other Firms

After the first period, shareholders can replace the incumbent manager. They expect a second-period cash flow of $\pi_2 \equiv px_2$ from randomly hiring a new manager in the labor market. Let $p_i(I)$ denote the posterior belief about incumbent manager i given information set I. Then the optimal firing policy corresponds to a cut-off rule of the form "retain manager i for the second period if and only if $p_i(I) \geq p$".[9] The incumbent is replaced whenever he is suspected to be less competent than a new manager.

A replacement decision can be taken either at $t = 2$ or at $t = 2.5$. The only difference between an early and a late decision concerns the available information set I. A late mover might observe the firing decision taken at $t = 2$ in the other firm. We trivially establish the following result on the timing of decisions:

Lemma 1. *There exists a symmetric timing equilibrium in which shareholders take a firing decision at $t = 2$ if and only if they are perfectly informed about their manager (i.e., $p_i(I) \in \{0; 1\}$). Otherwise, they postpone the decision until $t = 2.5$.*

Shareholders postpone the decision until $t = 2.5$ whenever they are uncertain about their manager's type (i.e., $p_i(I) \in (0, 1)$). To see that this strategy constitutes an equilibrium, note that an informed firm is always indifferent to the timing. As shown below, uncertain shareholders are strictly better off if they wait until $t = 2.5$.

In the following I suppose that the equilibrium in Lemma 1 obtains.[10] Given that the objective of the analysis is to study coordination failures and inefficiencies between firms, this assumption is conservative in the sense that it minimizes the scope for any such frictions. Moreover, a focus on the case in which an informed firm moves immediately can be motivated by the following considerations which are not modeled explicitly: if the incumbent is discovered to be incompetent, prompt dismissal may be necessary to prevent entrenchment.[11] Conversely, an instant retention decision may be advantageous to prevent a proven competent manager from leaving the firm.

It is instructive to present the optimal firing policy first for the one-firm case. After an extreme first-period outcome the optimal decision is straightforward since performance is a perfect signal of ability: in case of success there is no doubt that the manager is competent ($p_i(\bar{x}_1) = 1$). He is retained and subsequently produces x_2. Conversely, a low outcome reveals incompetence ($p_i(\underline{x}_1) = 0$). The expected future cash flow under new management is π_2. In both cases the decision is efficient in the

[9] Other theoretical papers on CEO dismissal in which the firing policy takes the form of a cut-off rule are Hermalin and Weisbach (1998, 2010).

[10] Two other pure strategy equilibria exist: one in which both an informed and an uncertain firm always postpone the decision until $t = 2.5$ and another one where informed shareholders always postpone and uncertain shareholders always decide immediately at $t = 2$.

[11] See Shleifer and Vishny (1989) for a model of entrenchment.

sense that it is based on knowledge of the manager's true type. An intermediate cash flow ($X_{1i} = x_1$) is only a noisy signal of ability.[12] Absent further information shareholders do not know whether the outcome is due to low ability or bad luck. However, with probability m_i monitoring works and shareholders take an informed decision based on θ_i. With probability $1 - m_i$ monitoring fails in which case the firing decision is based on the noisy posterior belief and inefficient in expectation: shareholders sometimes dismiss a competent manager (if $p_i(x_1) < p$) or retain an incompetent one (if $p_i(x_1) \geq p$), depending on the parameters of the model.

In the two-firm case, shareholders can use information about the rival to filter out the common underlying shock. Two distinct types of information about the other firm are valuable.

Proposition 1. *The optimal firing policy of firm i depends on both the output X_{1j} and on the firing decision D_j of firm j.*

First, the optimal firing policy entails relative-performance evaluation (or benchmarking).[13] The benefit from incorporating the rival's output is readily apparent. Let $p_1(X_{11}, X_{12})$ be the posterior belief about manager 1 given first-period cash flows in the two firms. Suppose that manager i performs mediocre and the monitoring effort of his shareholders fails. A high outcome by manager j ($X_{1j} = \bar{x}_1$) reveals that industry conditions are favorable ($s = g$). Hence, intermediate performance by manager i must be due to his incompetence and $p_i(x_1, \bar{x}_1)$ equals zero. Conversely, $p_i(x_1, \underline{x}_1)$ equals one. Benchmarking may thus substitute for a monitoring failure.

Second, the rival's hiring decision D_j can provide valuable information about the common shock beyond that contained in the performance vector. To see this, let us suppose that both managers perform mediocre ($X_{11} = X_{12} = x_1$) in which case benchmarking does not fully reveal their abilities. If monitoring fails in both firms, the firing decisions are simply based on the shared posterior belief $p_i(x_1, x_1)$:[14] if $p_i(x_1, x_1) \geq p$ both managers are retained; otherwise they are dismissed. Suppose instead that monitoring fails in, say, firm i but succeeds in firm j. Then shareholders in firm i can observe the rival's early decision $D_j \in \{f, r\}$ which prompts them to re-evaluate their own manager: if manager j was replaced, monitoring must have shown that $\theta_j = \underline{\theta}$ which implies that industry conditions are favorable ($s = g$). Manager i is thus discovered to be incompetent as well and is dismissed. Analogue reasoning applies in case of a retention decision in the informed firm ($D_j = r$). This reveals that both managers were merely unlucky ($s = b$) and prompts firm i to rehire its manager as well. Generally speaking, the superior information in firm j about the common shock can be fully extracted from its firing decision. Hence, firm i can make an efficient choice despite a monitoring failure. More formally, let $p_i(X_{11}, X_{12}, D_2)$ denote the refined posterior belief about manager i at $t = 2.5$ given observation of joint performances and of the firing decision in firm j. Then,

[12] In this case $p_i(x_1) = \frac{p(1-q)}{p(1-q)+q(1-p)}$.

[13] See the references in footnote 1 for empirical evidence in support of benchmarking.

[14] The posterior belief is $p_i(x_1, x_1) = \frac{(1-q)p^2}{(1-q)p^2+q(1-p)^2}$.

$$p_i(x_1, x_1, f) = 0 \text{ and } p_i(x_1, x_1, r) = 1.$$

These beliefs imply that firm i finds it optimal to imitate the policy of the informed rival (i.e. $D_i = D_j$) rather than base its decision on $p_i(x_1, x_1)$. Going beyond the standard benchmarking approach, the analysis shows that decisions based on costly monitoring can transmit information to peers beyond that contained in the performance vector. Firms can thereby benefit from each others monitoring efforts.

For ease of exposition I impose the following constraint on the probability of a positive industry shock q:

Assumption 1. $q \leq \min\left\{p, \frac{1}{2}\right\}$.

The constraint implies that $p_i(x_1, x_1) \geq p$. Hence, absent further information shareholders give their manager the benefit of the doubt after joint intermediate performance and retain him. Put differently, the firing policy is inefficiently lenient in expectation (compared to a decision based on the true ability) since an incompetent manager is sometimes rehired. All the following results are robust to the case where q is in $(0, 1)$ though.

3.2 Information Collection and Firm Value in Equilibrium

Shareholders in both firms simultaneously and non-cooperatively select their monitoring levels at $t = 0$. Information spillovers through firing decisions lead to strategic interactions between monitoring choices ex-ante. Given m_j, shareholders in firm i choose m_i so as to maximize firm value net of monitoring costs:

$$\max_{m_i \in [0,1]} V_i(m_i, m_j) = pq(\bar{x}_1 + x_2) + (1-p)(1-q)(\underline{x}_1 + \pi_2) + (1-q)p(x_1 + x_2)$$
$$+q(1-p)^2[x_1 + m_i\pi_2 + (1-m_i)m_j)\pi_2] + q(1-p)p[x_1 + \pi_2] - \frac{1}{2}c_i m_i^2$$

For example, with probability pq manager i is competent and industry conditions are favorable in which case total cash flows are $\bar{x}_1 + x_2$. Consider the first term in the second line which concerns the only outcome after which the assessment of ability is nontrivial: following joint intermediate performance, and absent further information, manager i is retained despite his incompetence (by Assumption 1). If successful monitoring reveals his low type, dismissal occurs and the expected second-period cash flow is π_2. If monitoring fails, shareholders look to the other firm for additional information. With probability m_j manager j is dismissed upon mediocre performance. From this shareholders in firm i can infer that $s = g$. Hence, they opt for dismissal as well at $t = 2.5$.

The FOC gives the reaction function of firm i:

$$M_i(m_j) = \frac{q(1-p)^2\pi_2(1-m_j)}{c_i} \tag{1}$$

The numerator on the RHS gives the marginal benefit of information gathering. Monitoring allows shareholders to sort out an incompetent manager and thus increases expected second-period cash flow by π_2. The firms' monitoring efforts are strategic substitutes. If firm j collects more information, firm i finds it optimal to monitor less since it can free-ride on the rival's effort. A firm will not monitor at all if it expects the rival to monitor perfectly.

Assumption 2. $\frac{2q(1-p)^2\pi_2}{c_i} < 1$ with $i \in \{1, 2\}$.

The cost of information acquisition is assumed to be sufficiently large such that monitoring levels are always lower than one. Hence, there exists a unique interior equilibrium:

$$m_i^e = \frac{q(1-p)^2\pi_2[c_j - q(1-p)^2\pi_2]}{[c_i c_j - (q(1-p)^2\pi_2)^2]}. \tag{2}$$

The equilibrium level of monitoring in firm i is decreasing in its own cost of information acquisition c_i and increasing in π_2. Higher future cash flows raise the marginal benefit from pursuing an informed firing policy. Hence, shareholders are willing to devote more resources to information acquisition ex-ante. Moreover, the following peer effects arise:

Proposition 2. *A reduction in the monitoring cost in firm i (lower c_i) is associated with an increase in the value of firm j and less monitoring in firm j (lower m_j^e).*

A change in governance in one firm influences the value of the other. Following a decrease in the cost c_i the shareholders in firm i gather more information. This adds value to firm j since it can pursue a more efficient firing policy. It can easily be checked that the positive spillover effect is greater, the weaker is the governance of the receiver (i.e., the higher is c_j). Recent research suggests that shareholder activism, for example by hedge funds, can improve the performance of target firms (Bethel, Liebeskind and Opler 1998, Clifford 2008, Klein and Zur 2009). The model predicts that the positive impact should not be confined to the target but should also arise for industry peers. Bris, Brisley and Cabolis (2008) provide empirical evidence in support of this prediction. They consider firms whose governance improves after being taken over by an acquirer from a country with better investor protection and higher accounting standards. They find that mergers of this kind raise the Tobin's Q of the entire industry of the target. In addition, the model predicts that peers with weak monitoring (e.g., with little institutional investment due to a small size) should benefit the most. Moreover, the peer group effect should be especially strong in sectors where uncertainty about common industry-wide factors is important compared to firm-specific or other factors. If the firms were exposed to purely idiosyncratic noise (i.e., two independently drawn shocks s_i and s_j), then there would be no link between the firms' monitoring choices or firing decisions.

An improvement in the monitoring technology of firm i stifles the incentives to gather information in the peer firm. Hence, monitoring efforts move in opposite

directions.[15] Proposition 2 points to a caveat when measuring the quality of corporate governance. A distinction needs to be made between the intensity of monitoring and the efficiency of decision-making: While the peer devotes fewer resources to information acquisition, it nevertheless pursues a more efficient firing policy.

3.3 Empirical Evidence

This section presents tentative empirical evidence of information spillovers across firms. To do so, I turn to data on forced CEO turnover. Spillovers translate into specific predictions about firing patterns. The main prediction is that CEO dismissal risk is positively related to firing decisions at peer firms.

According to standard economic theory, shareholders should take into account all informative signals when evaluating their CEO and deciding over his retention. In the model, each firm generates information for its peer in two ways, through its performance and through the intervention decision. While the former signal is for free, the latter is based on costly monitoring. If shareholders do not monitor ($m_i = m_j = 0$), each CEO's firing risk solely depends on the performance vector. That is, given performances the intervention decision at firm j has no predictive power for the turnover risk of manager i. More formally, let $Pr[D_i = f \mid X_{1i}; X_{1j}]$ be the probability that manager i is fired given joint first-period performance and let $Pr[D_i = f \mid X_{1i}; X_{1j}; D_j]$ be the firing probability conditional on both joint performance and on the firing decision in the other firm. Then it can easily be checked that

$$Pr[D_i = f \mid X_{1i}; X_{1j}] = Pr[D_i = f \mid X_{1i}; X_{1j}; D_j].^{16}$$

In the absence of monitoring, nothing is learned from intervention as it can be perfectly inferred from the shared information set $(X_{1i}; X_{1j})$.

Suppose now that the monitoring intensities are strictly positive. First, if shareholders in firm i monitor successfully, their decision is fully determined by the observed true type θ_i which implies:

$$Pr[D_i = f \mid \theta_i; X_{1i}; X_{1j}] = Pr[D_i = f \mid \theta_i; X_{1i}; X_{1j}; D_j].$$

Again, dismissal risk is unrelated to intervention at the peer firm. For instance, if $\theta_i = \bar{\theta}$, then manager i's turnover risk is always zero. This is no longer the case if

[15] Bris, Brisley and Cabolis (2007) attribute the aforementioned positive spillover effects on peer performance to a strategic complementarity. The improvement of governance in one firm prompts rivals to raise their standards as well. They do not, however, provide any direct evidence of their interpretation.

[16] Unless $X_{1i} = X_{1j} = x$, both managers' types can be perfectly inferred from merely observing joint performance. Consequently, any additional information does not alter the posterior belief and therefore has no effect on the firing probability. For instance, if $X_{1i} = x$ and $X_{1j} = \underline{x}$, then $Pr[D_i = f \mid x; \underline{x}] = Pr[D_i = f \mid x; \underline{x}; Dj] = 0$. If $X_{1i} = X_{1j} = x$, then each manager is retained with certainty (by Assumption 1).

shareholders in firm i fail to monitor successfully. Then they rely on peer actions for information:

$$Pr[D_i = f \mid X_{1i}; X_{1j}] \leq Pr[D_i = f \mid X_{1i}; X_{1j}; D_j = f].$$

Turnover risk depends on the decision in the other firm. More precisely, dismissal of the peer manager strictly raises manager i 's turnover risk if $X_{1i} = X_{1j} = x$:[17] if only performance is observed, manager i is rehired. Conversely, if shareholders in firm i see that manager j has been fired, they will imitate the decision and $Pr[D_i = f \mid x; x; D_j = f] = 1$.

Like benchmarking, evaluation based on peer decisions is based on the exposure to a common performance shock s. Otherwise, any information about the peer is always irrelevant. As a proxy for the exposure to a common shock I use membership within the same industry. As a proxy for the informedness of shareholders I use institutional blockholder ownership because blockholders have both the incentive and the ability to monitor. The focus is then on the following two predictions:

Prediction 1: CEO turnover is positively related to past firings in the same industry but unrelated to firings outside of the industry.

Prediction 2: The sensitivity of CEO turnover risk to peer decisions is higher in firms with low institutional blockholder ownership.

The data on CEO turnover come from Jenter and Kanaan (2010). I thank Dirk Jenter for making them available to me. CEO turnover is observed for all firms in the S&P ExecuComp database for the period from 1993 to 2001. The ExecuComp sample contains information on the top executives of all firms in the S&P 500, S&P MidCap and S&P SmallCap indexes. A CEO turnover is recognized for each year in which a change is announced of the CEO that is identified in ExecuComp. The classification of turnovers as forced follows Jenter and Kanaan (2010).[18] I use stock returns to measure performance. All stock return information comes from monthly CRSP tapes. Returns are computed over the calendar year, lagged returns are measured over the two previous calendar years. Jenter and Kanaan (2010) and Kaplan and Minton (2008) find that CEO turnover is related to three different performance components: the performance of the firm relative to the industry, the performance of the industry relative to the stock market and to the overall return in the stock market. I use the

[17] In all other cases the above condition holds with equality.

[18] They classify CEO turnover as forced if "the press reports that the CEO is fired, forced out, or retires or resigns due to policy differences or pressure. All other departures for CEOs above and including age 60 are classified as voluntary. Departures for CEOs below age 60 are reviewed further and classified as forced if either the press does not report the reason as death, poor health, or the acceptance of another position [...], or the press reports that the CEO is retiring, but does not announce the retirement at least six months before the succession. Finally, the cases classified as forced can be reclassified if the reports convincingly explain the departure as due to reasons that are unrelated to the firm's activities."

following three proxies. First, relative firm performance is equal to the firm stock return minus the return for the median firm in the same industry. Relative industry performance is measured as the difference between the stock return of the median firm and the S&P 500. Finally, I measure market performance using the annual return on the S&P 500. Industries are defined using the Fama and French (1997) classification of firms into 48 industries. Finally, as a proxy for the informedness of shareholders I use the percentage of shares held in each firm by the largest institutional blockholder (at the end of the previous year). Blockholders are defined as shareholders with greater than 5% ownership of the firm's outstanding shares. I thank Martijn Cremers for making the data available to me.[19]

The final sample has 10,392 firm-year observations from 1993 until 2001 and contains 289 forced CEO turnovers. The unconditional probability for a CEO to be forced out of office in any given year is 2.7 %. The main explanatory variable of interest is the number of CEO firings at other firms in the same industry. The average number of CEOs fired within a given industry per year equals 1.8 with a standard deviation of 2.25. The average number of CEOs fired within the market as whole in any given year is 36.21 with a standard deviation of 13.84. Mean blockownership is 7.57 %.

I estimate the sensitivity of forced CEO turnover to peer decisions using pooled annual probit regressions. The dependent variable is one if the firing of the CEO is announced and zero otherwise. In the regression (1) in Table 1, I include the number of firings in the CEO's industry in the previous calendar year (PEER FIRINGS T-1). Both firing decisions by industry peers and the CEO's industry-adjusted stock performance strongly affect turnover risk. The number of firings by peers is positively and significantly related to CEO turnover. A one standard deviation increase in the number of past firings increases the CEOs firing risk by more than 25 %. (The average probability of a forced turnover increases from 2.7 % to 3.37 %.) The second regression in Table 1 also includes the number of past firings in the market as whole (MARKET FIRINGS T-1). The coefficient is insignificant and equal to zero. Past firings outside of the CEO's industry have no predictive power for the likelihood of forced CEO turnover. The results are consistent with the hypothesis that shareholders react to firing decisions in firms that are exposed to common uncertainty, but not to decisions in unrelated firms.

The first two columns in Table 2 include blockownership as well as the interaction effects of blockownership with past firing decisions. The marginal probability associated with blockownership is insignificant. However, the interaction term with past firings is negative and statistically significant. This is consistent with the interpretation that uninformed shareholders are more sensitive to peer decisions. The result also obtains if I use a dummy variable to measure blockownership. The dummy variable in the second set of regressions in Table 2 equals one if the firm's blockownership is above the industry mean. The results in columns (3) and (4) show that the

[19] Cremers and Nair (2005) construct the measure using data from CDA Spectrum that is based on quarterly SEC 13f filings by institutional shareholders.

Table 1
Probit regressions of forced turnover on performance and past firings by other firms.

The table shows annual probit regressions of forced CEO turnover during the period 1993 to 2001 on prior firings by other firms. The dependent variable equals one if the CEO is fired and zero otherwise. Δ Prob measures the change in the probability of CEO turnover per unit change in the relevant explanatory variables. PEER FIRINGS T-1 is the number of firings in the previous year at other firms in the same industry. MARKET FIRINGS T-1 is the number of firings in the previous year in the whole market. $*$, $**$, and $***$ indicate significance at the 10 %, 5 % and 1 % levels, respectively, p-values are in parentheses.

	(1)	(2)	(3)	(4)
Variable	Δ Prob	Δ Prob	Δ Prob	Δ Prob
Peer firings t-1	0.003^{***}	0.004^{***}	0.003^{**}	0.003^{***}
	(0.008)	(0.000)	(0.010)	(0.008)
Market firings t-1	–	0.0008	–	0.0000
		(0.431)		(0.745)
Relative firm stock return t-1	-0.032^{***}	-0.032^{***}	-0.032^{***}	-0.032^{***}
	(0.000)	(0.000)	(0.000)	(0.000)
Relative firm stock return t-2	-0.008^{***}	-0.008^{***}	-0.008^{***}	-0.008^{***}
	(0.008)	(0.007)	(0.005)	(0.006)
Relative industry stock return t-1	-0.018^{**}	-0.018^{**}	-0.021^{**}	-0.021^{***}
	(0.017)	(0.016)	(0.010)	(0.009)
Relative industry stock return t-2	0.005	0.005	0.001	0.001
	(0.468)	(0.470)	(0.775)	(0.820)
Age	-0.004	-0.004	-0.004	-0.004
	(0.251)	(0.253)	(0.241)	(0.245)
Number of peer firms t-1	-0.0001^{*}	-0.0001^{*}	-0.0001^{*}	-0.0001^{*}
	(0.056)	(0.056)	(0.073)	(0.078)
Number of observations	10,392	10,392	10,392	10,392
Pseudo R^2	0.0721	0.0723	0.0676	0.0680
Year Dummies	*Yes*	*Yes*	*No*	*No*
S.E. clustered by industry	*Yes*	*Yes*	*Yes*	*Yes*

Table 2
Probit regressions of forced turnover on performance, past firings by other firms and blockholder ownership.

The table shows annual probit regressions of forced CEO turnover during the period 1993 to 2001 on prior firings by other firms and blockholder ownership. The dependent variable equals one if the CEO is fired and zero otherwise. Δ Prob measures the change in the probability of CEO turnover per unit change in the relevant explanatory variables. PEER FIRINGS T-1 is the number of firings in the previous year at other firms in the same industry. MARKET FIRINGS T-1 is the number of firings in the previous year in the whole market. BLOCK T-1 equals the ownership share of the largest institutional blockholder owning at least 5 % of outstanding shares. The dummy variable D_{Block} T-1 equals one if the share of the firm's largest institutional blockowner in the previous year is above the industry mean. The specifications in columns (1) and (2) contain but do not report Block T-1 interacted with all performance measures. The specifications in columns (3) and (4) contain but do not report D_{Block} T-1 interacted with all performance measures. $*$, $**$, and $***$ indicate significance at the 10 %, 5 % and 1 % levels, respectively, p-values are in parentheses.

	(1)	(2)	(3)	(4)
Variable	Δ Prob	Δ Prob	Δ Prob	Δ Prob
Peer firings t-1	0.005***	0.005***	0.004***	0.004***
	(0.000)	(0.000)	(0.002)	(0.002)
Block t-1	−0.0004	−0.0004	–	–
	(0.161)	(0.158)		
(Block t-1)*(Peer firings t-1)	−0.0002***	−0.0002***	–	–
	(0.003)	(0.002)		
D_{Block} t-1	–	–	0.002	0.002
			(0.655)	(0.671)
(D_{Block} t-1)*(Peer firings t-1)	–	–	−0.001*	−0.001*
			(0.086)	(0.077)
Relative firm stock return t-1	−0.029***	−0.030***	−0.035***	−0.035***
	(0.000)	(0.000)	(0.000)	(0.000)
Relative firm stock return t-2	−0.003	−0.003	−0.004	−0.005
	(0.346)	(0.287)	(0.186)	(0.146)
Relative industry stock return t-1	−0.020**	−0.022**	−0.026***	−0.029***
	(0.021)	(0.021)	(0.003)	(0.004)
Relative industry stock return t-2	0.004	0.001	0.005	0.002
	(0.593)	(0.869)	(0.496)	(0.782)
Age	−0.003	−0.003	−0.006	−0.006
	(0.504)	(0.490)	(0.162)	(0.151)
Number of peer firms t-1	−0.0001***	−0.0001***	−0.0001*	−0.0001*
	(0.001)	(0.002)	(0.055)	(0.068)
Number of observations	10,392	10,392	10,392	10,392
Pseudo R^2	0.0763	0.0718	0.0748	0.0704
Year Dummies	*Yes*	*No*	*Yes*	*No*
S.E. clustered by industry	*Yes*	*Yes*	*Yes*	*Yes*

predictive power of peer decisions for CEO dismissal risk is significantly stronger for firms with institutional blockholder ownership below the industry mean.

4 Optimal Monitoring and the Benefits of a Joint Monitor

Up to this point the analysis has assumed that the two firms have different monitors who act non-cooperatively. In the non-cooperative equilibrium, shareholders fail to exploit fully the benefits from information spillovers because they ignore that information acquisition is socially valuable. This section presents the first-best benchmark which maximizes joint investor welfare. A market-based solution to the coordination failure is discussed.

The problem of choosing the first-best monitoring levels so as to maximize combined firm values is presented in the Appendix. The first-order-condition for m_i^o, the optimal monitoring intensity in firm i, is

$$M_i^o(m_j) = \frac{q(1-p)^2\pi_2 2(1-m_j)}{c_i}. \tag{3}$$

The numerator on the RHS is the marginal social benefit from overseeing management. It is twice as large as the private benefit in (1) since it includes the marginal benefit of m_i to firm j. Any information that is acquired by one firm also becomes available to the other one. Hence, if firm j is perfectly informed ($m_j = 1$), firm i optimally collects no information. firm 2 is already perfectly informed and cannot gain from peer information. Joint investor welfare is maximized if firm i chooses a monitoring level

$$m_i^o = \frac{q(1-p)^2\pi_2 2(c_j - q(1-p)^2\pi_2 2)}{c_i c_j - (q(1-p)^2\pi_2 2)^2}. \tag{4}$$

Like m_i^e in (2), the optimal monitoring effort is increasing in future cash flows and decreasing in c_i. As the rival's monitoring cost increases, so does m_i^o. A comparison between the equilibrium and the first-best outcome yields the following result:

Proposition 3. *In the noncooperative equilibrium the risk of a total monitoring failure is inefficiently high:* $(1-m_i^o)(1-m_j^o) < (1-m_i^e)(1-m_j^e)$

Due to collective free-riding the overall monitoring activity is inefficiently low in equilibrium. Incompetent managers are replaced too seldomly compared to the first-best outcome. Hence, privately optimal choices do not maximize joint firm values. While the overall monitoring intensity is too low, the direction of the distortion for the individual firm is ambiguous and depends on the degree of cost asymmetry. If the difference between monitoring costs is sufficiently small, each firm always gathers too little information. Conversely, if the difference between their costs is large, the firm with the higher (lower) cost monitors too much (little) in equilibrium. Altogether, there is too little oversight in at least one of the two firms.

The literature on monitoring and corporate control has commonly studied free-riding at the level of a single firm: in the presence of many small shareholders an externality arises because it is not in the interest of any one of them to devote resources to controlling management. The above analysis abstracts from any coordination or incentive problems within firms. Nonetheless, underprovision arises due to inter-firm free-riding behavior.

One way to implement the first-best outcome is to put the firms under the control of a joint monitor who maximizes their combined value.[20] The efficiency gains from removing the free-rider problem provide a rationale for joint ownership. Free-riding behavior, and hence a gain from joint ownership, arises only if the firms face common uncertainties. Hence, the model suggests that the exposure to common performance shocks is a potential determinant of ownership patterns across firms.[21] For ease of exposition, the analysis has focused on the polar case where the shock is perfectly correlated across firms. More generally, monitoring externalities arise unless the firms' cash flows are exposed to purely idiosyncratic shocks.

The model predicts that an active monitor, who is engaged in multiple firms (e.g. a pension fund, venture capitalist or activist hedge fund), should follow a focused rather than a diversified strategy: portfolio firms should be exposed to common performance shocks, for instance due to operating in the same line of business or geographic area. The optimality of a focused strategy arises in the model even though the monitor is not assumed to have any industry- or sector-specific expertise. Specialization arises instead as a solution to a coordination problem. Harford, Jenter, and Li (2010) study institutional cross-holdings among firms and find that these are greater if the firms are more alike on dimensions such as size, market-to-book ratio or stock return. Another application are venture capital investors which take concentrated equity positions in young, high-risk firms and engage in monitoring. Norton and Tenenbaum (1993) study the investment strategies and portfolios of venture capital firms and provide evidence of industry specialization. Similarly, Cressy, Munari and Malipiero (2007) examine private equity (PE) backed buyouts. They provide evidence that industry specialization of private equity firms raises their positive impact on targets' operating performances. Cronqvist and Fahlenbrach (2009) study the effect of large shareholders on corporate policy and focus on blockholders that are present in multiple firms. They find that blockholders with a single decision maker have a stronger effect on firm performances. This is consistent with the above view that engagement in multiple firms is driven by learning synergies.

In a next step I generalize the analysis and consider the incentive to monitor when there are more than two firms. Let N denote the number of ex-ante identical firms in the economy. With probability q all firms experience a positive shock, with

[20] An alternative way of achieving the first-best outcome, that is not further pursued here, would be cooperation between different monitors in the two firms.

[21] The extant theoretical literature has not paid much attention to ownership patterns across firms and usually studies ownership concentration within a one-firm setting. See for example Shleifer and Vishny (1986), Huddart (1993) and Burkart, Gromb and Panunzi (1997).

probability $(1-q)$ each firm is exposed to a negative shock. It is straightforward to check that the equilibrium and the optimal level of governance are decreasing in N. The higher the number of firms, the more can be learned from merely observing the performance vector and the smaller is the need to monitor. As N becomes large, the law of large numbers applies and there is no aggregate uncertainty: a fraction p of managers turns out to be competent and the remaining fraction turns out to be incompetent. Then the observation of the performance vector always fully reveals the nature of the shock. Hence, benchmarking allows perfect inference of all managers' abilities and there is no need to monitor. We state the following result which qualifies Proposition 4:

Proposition 4. *If the number of firms becomes large ($N \to \infty$), the equilibrium outcome coincides with the first-best one:*

$$\lim_{N\to\infty} m^o = \lim_{N\to\infty} m^e = 0$$

In the limit, as N becomes large, the marginal private and social benefit from monitoring coincide and the externality disappears.[22] Note also that Proposition 4 implies that efficiency gains from common monitoring disappear in competitive industries.

5 Regulation of Corporate Governance

The previous section showed that monitoring activity is too low in equilibrium and proposed joint ownership as a solution for the free-rider problem. However, in practice many factors can impede the emergence of a joint monitor such as legal constraints or limited wealth. Also, an engagement in multiple firms might deteriorate the blockholder's monitoring technology. If market-based remedies fail to eliminate the free-rider problem, the model establishes a role for the regulation of corporate governance to increase investor welfare.[23]

If the regulator had full information about the firms, she could simply directly impose the first-best monitoring intensities. Another potential instrument in this case would be a subsidy of monitoring costs. However, this idealized case is unlikely to obtain in reality. This section focuses on an alternative form of regulation, a cap

[22] In general, an increase in N is likely to affect not only the scope for benchmarking but also the degree of product market competition. My model abstracts from product market competition and its effects on governance.

[23] According to Becht, Bolton and Roëll (2003) there are two reasons for the regulation of corporate governance. If governance arrangements impose externalities on third parties, the founders or shareholders may adopt inefficient rules. For instance, the adoption of anti-takeover devices with the goal of extracting surplus from a future raider can lead to an inefficiently low level of takeovers. A second argument in favor of regulation can be based on a commitment problem of the firm vis-à-vis its investors. Even if rules are put in place ex-ante to protect outside investors, the firm may later on (once funding has been provided) have an incentive to break those rules.

on CEO compensation, and shows that it can implement the first-best outcome. In order to study regulation of CEO pay, I extend the above pure learning-model by a moral hazard dimension. Up to this point management has played a passive role and affected cash flows only through its exogenously determined ability. Now cash flows are assumed to depend also on unobservable managerial effort. Compensation is thereby introduced into the model and can be exploited as a viable target-variable for regulation. Beyond regulation, the extension allows to study the link between monitoring externalities and agency conflicts within firms.

Suppose that each manager can invest in his human capital in order to enhance his productivity. More formally, each manager chooses an unobservable effort $e_i \in \{\bar{e}, \underline{e}\}$ which determines the distribution of his ability, $p_i(e_i)$, as follows: with high effort ($e_i = \bar{e}$), the distribution is the same as in the previous analysis, i.e., manager i turns out to be competent with probability $p_i(\bar{e}) = p$. If manager i exerts low effort ($e_i = \underline{e}$), the probability that $\theta_i = \bar{\theta}$ decreases from p to $p_i(\underline{e}) < p$. Low effort implies private benefits of $Z_1 > 0$ for the manager. Let $\delta p \equiv p - p(\underline{e})$. Furthermore, the manager enjoys private benefits $Z_2 > 0$ if he is retained after the first period.

The timing of the augmented game is as follows. As before, shareholders fix monitoring intensities at an initial date, $t = 0$. At $t = 0.5$ each firm contracts with a manager. The managers simultaneously choose their effort levels at $t = 0.75$. At $t = 1$ first-period cash flows realize and the game continues as described in Section 2. To simplify the exposition I focus on the fringe case where $\tau = 1$. We can limit attention to contracts which specify payments to a manager that are only contingent on first-period cash flows.[24] Let $(\bar{w}_i, w_i, \underline{w}_i)$ denote the rewards in case of a high, mediocre and low outcome respectively to manager i who is protected by limited liability.

The analysis proceeds on Assumptions 1 and 2. Hence, joint mediocre performance ensures retention absent further information ($p_i(x_1, x_1) \geq p$). Furthermore, we assume that shareholders find it optimal to induce high effort:

Assumption 3. $q(\bar{x}_1 + x_1) + (1 - q)(x_1 - \underline{x}_1) + x_2(1 - p) \geq \frac{Z_1}{\delta p}$.

Before tackling the full problem, it can easily be checked that the optimal compensation scheme rewards the manager only if there is no doubt about his competence, i.e., if the posterior belief about his ability equals one. This follows from the maximum likelihood ratio criterion according to which compensation should only be offered for those outcomes that are most indicative of the manager having exerted high effort. Hence, we can set $w^* = 0$ and $\underline{w}^* = 0$ and focus without loss of generality on the case where the wage is positive if and only if $X_1 = \bar{x}_1$.[25]

[24] The manager takes no further action after the first period and all information is symmetrically held. Hence, there is no need to make payments contingent on second-period outcomes. Moreover, allowing for payments that are contingent on the publicly observable firing decision does not improve upon the solution derived below.

[25] Ignoring positive payments for the case when $X_1 = x_1$ and monitoring or benchmarking per-

5.1 Moral Hazard and Monitoring: The One-Firm Case

To develop the intuition I first consider the shareholders' contracting problem in a one-firm setting. The firing policy affects firm value in two ways: as before, it determines the ability of management in the second period and thus the cash flow X_2 (ex-post effect). Moreover, in the presence of moral hazard it also influences the manager's effort choice because his payoff depends on the future private benefits Z_2 (ex-ante effect). It can easily be seen that the ex-post optimal firing policy, that was derived in Section 3.1, is also optimal from an ex-ante perspective: the incumbent is rehired (respectively fired) whenever he is known with certainty to be of type $\bar{\theta}$ (respectively $\underline{\theta}$). In case of mediocre performance he is given the benefit of the doubt and is retained unless monitoring reveals his incompetence. This policy provides maximal incentives for the manager to exert effort by punishing him with the loss of private benefits whenever he is deemed incompetent.

When determining the compensation scheme and the monitoring intensity, shareholders solve the following problem:

$$\max_{m\in[0,1],\bar{w}\geq 0} qp(\bar{x}_1 - \bar{w} + x_2) + q(1-p)(x_1 + m\pi_2)$$
$$+(1-q)p(x_1 + x_2) + (1-q)(1-p)(\underline{x}_1 + \pi_2) - \frac{1}{2}cm^2$$

subject to

$$qp\bar{w} + [p + q(1-p)(1-m)]Z_2$$
$$\geq qp(\underline{e})\bar{w} + [p(\underline{e}) + q(1-p(\underline{e}))(1-m)]Z_2 + Z_1.$$

The objective function equals expected cash-flows net of monitoring and compensation costs. The LHS of the incentive compatibility constraint gives the manager's expected payoff if he exerts high effort, the RHS if he exerts low effort. The manager is induced to work not only by the wage $\bar{w}$ but also by the private benefits Z_2 from running the firm. Future private benefits provide implicit incentives because the manager is more likely to keep his position if he exerts high effort. This can be seen from the retention probabilities given in squared brackets above. Rearranging the IC constraint yields

$$\bar{w} \geq \frac{Z_1}{q\delta p} - \frac{1}{q}[1 - q(1-m)]Z_2.$$

Monitoring relaxes the IC constraint. The term in squared brackets equals the dismissal threat faced by an incompetent manager which is increasing in m. Monitoring reduces moral hazard by making the firing decision more sensitive to the manager's true type and thus to his ex-ante effort choice.

Since the objective function is decreasing in $\bar{w}$, either the incentive compatibility constraint or the limited liability constraint determines the optimal wage. The

fectly reveals the manager's competence is without loss of generality.

following assumption ensures interior solutions for the monitoring intensity and for compensation.

Assumption 4. $\frac{2[q(1-p)\pi_2+qpZ_2]}{c} < 1$ and $\frac{Z_1}{\delta p} > Z_2$.

The first inequality states that the monitoring cost is sufficiently large to ensure an interior solution for m. The second inequality states that moral hazard, as measured by Z_1, is severe enough such that monetary incentives are needed to induce high effort. In other words, the limited liability constraint is slack.

The optimal compensation scheme which solves the above program is

$$\bar{w}_I^* = \frac{Z_1}{q\delta p} - \frac{1}{q}[1 - q(1-m)]Z_2 \tag{5}$$

and the optimal monitoring intensity is

$$m_I^* = \frac{q(1-p)\pi_2 + qpZ_2}{c}. \tag{6}$$

The optimal wage in (5) is decreasing in Z_2 which measures the strength of implicit incentives. Importantly, compensation is also decreasing in m. Monitoring reduces moral hazard by threatening an incompetent manager with dismissal.[26] Overall, shareholders therefore dispose of two instruments to induce effort - compensation and monitoring - which are substitutes.[27] (As shown below, this property generalizes to the two-firm case and plays a crucial role in the derivation of the optimal regulatory policy.)

Consider the optimal monitoring intensity in (6). The numerator on the RHS corresponds to the marginal benefit from acquiring information: the first summand reflects the ex-post benefit, i.e., the increase in second-period cash flows from replacing an incompetent incumbent. The second summand captures the ex-ante effect, the decrease in the first-period wage due to a greater dismissal threat. Thus, in the presence of moral hazard monitoring plays a dual role. It not only addresses a sorting problem at the interim date but also mitigates the agency conflict ex-ante. Importantly, in the one-firm setting there is no possibility for regulation to increase investor welfare. The firm chooses the optimal mix between compensation and monitoring.

5.2 CEO Pay and Information Spillovers

In the two-firm case the optimal firing policy plays as before a dual role, affecting both second-period profits and ex-ante incentives. The ex-post optimal firing policy, derived in Section 3.1, remains optimal in the presence of moral hazard: the

[26] It can easily be checked that firing the manager following mediocre performance and monitoring failure would harden the incentive compatibility constraint (and thus increase compensation), since, by Assumption 1, $q \leq \frac{1}{2}$.

[27] Fahlenbrach (2009) provides supportive empirical evidence that there is indeed a negative relationship between these two mechanisms.

incumbent is rehired (respectively fired) if he is known with certainty to be of type $\bar{\theta}$ (respectively $\underline{\theta}$). A manager is retained upon mediocre performance unless he is revealed to be incompetent. In addition to maximizing expected second-period cash flow, this policy also provides maximal incentives ex-ante to invest in competence.

Suppose that shareholders in firm 1 expect manager 2 to exert high effort and his shareholders to choose an intensity m_2. Then they determine the optimal monitoring intensity and compensation scheme by solving the following problem:

$$\begin{aligned} &\max_{m_1\in[0,1],\bar{w}_1\geq 0} qp(\bar{x}_1-\bar{w}_1+x_2) \\ &+q(1-p)(x_1+m_1\pi_2+(1-m_1)(p+(1-p)m_2)\pi_2) \\ &+(1-q)p(x_1+x_2)+(1-q)(1-p)(\underline{x}_1+\pi_2)-\frac{1}{2}cm_1^2 \end{aligned}$$

subject to

$$\begin{aligned} &qp\bar{w}_1+[p+q(1-p)^2(1-m_1)(1-m_2)]Z_2 \\ \geq\ &qp(\underline{e})\bar{w}_1+[p(\underline{e})+q(1-p)(1-p(\underline{e}))(1-m_1)(1-m_2)]Z_2+Z_1. \end{aligned}$$

Consider the IC constraint. As before, the manager is incentivized both by the wage $\bar{w}_1$ and by future private benefits Z_2. The crucial difference compared to the one-firm case concerns the retention probabilities in squared brackets which now depend on the monitoring intensities in both firms. Rearranging the IC constraint yields

$$\bar{w}_1\geq\frac{Z_1}{q\delta p}-\frac{1}{q}[1-q(1-p)(1-m_1^e)(1-m_2^e)]Z_2.$$

The term in squared brackets gives the dismissal threat to an incompetent manager which relaxes the IC constraint. The dismissal risk is increasing in both m_1 and m_2 due to monitoring spillovers. An incompetent manager is only rehired if there is a monitoring failure in both firms. Hence, information acquisition in firm 2 disciplines manager 1. If m_2 increases, the replacement policy in firm 1 is more sensitive to its own manager's true ability and thus provides stronger implicit incentives.

The reaction function, $M_1(m_2)$, gives m_1 as a function of the monitoring intensity in firm 2:

$$M_1(m_2)=\frac{q(1-p)^2\pi_2(1-m_2)+qp(1-p)(1-m_2)Z_2}{c}=(1-p)(1-m_2)m_I^*$$

The numerator on the RHS corresponds to firm 1's marginal benefit from monitoring: the first summand is the increase in expected second-period cash flows (ex-post benefit), the second summand is the reduction in incentive pay (ex-ante effect). Monitoring choices are strategic substitutes. If firm 2 monitors more, firm 1 finds it optimal to monitor less due to a reduction in both the ex-ante and the ex-post benefit. We obtain the following result: There exists a symmetric equilibrium in monitoring efforts such that

$$m_1^e=m_2^e=m^e=\frac{q(1-p)^2\pi_2+qp(1-p)Z_2}{c+q(1-p)^2\pi_2+qp(1-p)Z_2}. \tag{7}$$

In equilibrium, the reward for first-period success is

$$\bar{w}_1(m^e) = \bar{w}_2(m^e) = \bar{w}(m^e) = \frac{Z_1}{q\delta p} - \frac{1}{q}[1 - q(1-p)(1-m_1^e)(1-m_2^e)]Z_2. \tag{8}$$

The equilibrium monitoring intensity in (7) is increasing in Z_2 and π_2 which measure the strength of the ex-ante and the ex-post benefit from information acquisition respectively. Compensation is decreasing in the monitoring intensities chosen in both firms as they jointly determine the dismissal threat to an incompetent manager.

In a next step, I consider the first-best outcome, i.e., the monitoring intensities and compensation schemes that maximize joint firm values. The full program is presented in the Appendix. Joint investor welfare is maximized if

$$m_1^o = m_2^o = m^o = \frac{q(1-p)^2 2\pi_2 + qp(1-p)2Z_2}{c + q(1-p)^2 2\pi_2 + qp(1-p)2Z_2} \tag{9}$$

and

$$\bar{w}_1(m^o) = \bar{w}_2(m^o) = \bar{w}(m^o) = \frac{Z_1}{q\delta p} - \frac{(1-q)}{q}Z_2 - [1 - (1-p)(1-m^o)^2]Z_2 \tag{10}$$

The functional form of the first-best compensation scheme in (10) coincides with that in (8). A comparison of the first-best and the equilibrium outcome yields the following result:

Proposition 5. *In equilibrium, there is too little monitoring ($m^o > m^e$) and compensation is inefficiently high ($\bar{w}^o < \bar{w}^e$).*

Too little information is acquired in equilibrium. As in the pure learning framework, shareholders ignore that monitoring increases expected second-period cash flows in the rival firm (ex-post effect). Moreover, they do not take into account that monitoring disciplines the rival manager and thus reduces his compensation (ex-ante effect). The second distortion is new compared to the analysis in Section 3. Underprovision of monitoring translates into excessive compensation in equilibrium.

5.3 A Regulatory Cap on CEO Pay

The privately optimal governance arrangement deviates from the socially optimal one. This raises the question under which conditions, and in what form, regulatory intervention can increase investor welfare. If the regulator has full information about the firms, she can simply directly impose the socially optimal monitoring intensities. However, this idealized case is unlikely to obtain in reality. An alternative instrument is the imposition of a constraint on the compensation scheme. More specifically, let $(\bar{C}, C, \underline{C})$ denote the pay-cap set by the regulator for the wage payment in case of a high, intermediate, and low first-period outcome respectively. Hence, shareholders solve the problem presented in the previous subsection under the additional constraint that $(\bar{w}, w, \underline{w}) \leq (\bar{C}, C, \underline{C})$. We then obtain the following result:

Proposition 6. *If the regulator imposes a pay cap* $(\bar{C}^o, C^o, \underline{C}^o) = (\bar{w}(m^o), 0, 0)$*, there exists a symmetric equilibrium in which the two firms choose* m^o *and the pay-cap is binding.*

The regulator can implement the first-best outcome by constraining each firm to pay at most $\bar{w}(m^o)$, the wage given in (8), for first-period success and zero otherwise. Thus, a cap on compensation can increase investor welfare. The intuition is straightforward. Shareholders dispose of two instruments to ensure that their manager works: a monetary reward for good performance and provision of implicit incentives through costly monitoring. The two mechanisms are substitutes. If CEO pay is capped below its equilibrium level, shareholders are forced to monitor more than in equilibrium if they want to ensure high effort. Forcing each firm to deviate from its privately optimal choice can make all shareholders better off because monitoring is socially valuable.

In a recent paper Morse, Nanda and Seru (2009) study the design of incentive contracts and argue that powerful CEOs can manipulate their compensation schemes to extract rents. They advocate regulatory intervention to ensure transparency (disclosure of pay contracts) and the independence of the board. Their argument in favor of regulation is based on the power of the CEO and the impotence of internal monitoring mechanisms.[28] On the contrary, the above argument relies on the presence of a functioning monitor who becomes more vigilant when incentive pay is limited. Inderst and Mueller (2010) also address the issue of regulation and CEO compensation. The CEO in their model has private information about his suitability for running the firm. They show that a regulatory cap on "golden handshakes" can increase firm profits because it mitigates a commitment problem. In contrast, the externality in my model arises because governance choices impose externalities on third parties.

In general, the information externality - and hence the need for regulation - arises because the firms are exposed to a common shock. A regulatory cap should therefore be targeted towards firms within the same industry or geographic area. Note that Proposition 4 implies that there is no need for a pay-cap in highly competitive industries.

Corollary 1. *The pay-cap which maximizes investor welfare,* $\bar{C}^o = \bar{w}(m^o)$*, is increasing in the cost of monitoring* c*.*

As the cost c increases, the socially optimal level of monitoring decreases. Consequently, the optimal level of compensation, that is needed to ensure high effort, increases. One can interpret the cost c as the degree of shareholders' legal protection. Then the corollary suggests that a pay-cap and investor protection are complements: ceteris paribus, the introduction of a pay-cap is particularly valuable in an environment or sector with strong investor protection.

[28] Prominent proponents of the power approach to CEO compensation are Bebchuk and Fried (2004). In contrast this paper takes an "optimal contracting" or "arm's length bargaining" approach.

6 Disclosure Requirements for Active Monitors

At the core of the analysis is the idea that shareholders can learn from informed actions taken in other firms. So far, I have assumed that each firm is transparent in the sense that shareholder intervention is publicly observable. However, in practice outsiders often intervene privately or informally and may thus disseminate little information to peers (Carleton, Nelson and Weisbach 1998). I extend the baseline model and allow shareholders to choose how much information about their actions they want to make available to the rival. The focus is on the link between, on the one hand, the incentives to collect information and, on the other hand, the decision to share information.

Extant studies of disclosure commonly focus on the incentives of management to report accounting or financial information. In contrast, this section puts forward the notion of governance transparency which refers to the disclosure of intervention and actions by active monitors. More formally, suppose that shareholders in each firm can freely determine a disclosure regime before the hiring stage. Disclosure policy is modeled as the probability that third parties can observe the interference decision. More specifically, the shareholders in firm i choose $\tau_i \in [0, 1]$, which is the probability that their future actions are observable by outsiders. With probability $1 - \tau_i$ intervention cannot be observed by anyone outside firm i. Note that the previous analysis corresponds to the case of full governance transparency ($\tau_1 = \tau_2 = 1$). Low transparency is meant to capture the idea that outsiders are unsure about the underpinnings of the firing decision. For instance, dismissal may be due to reasons that are extraneous to the model rather than the assessment of ability.

We solve the game by backward induction, starting with the choice of monitoring efforts. The full problem is presented in the Appendix. Given the rival's level of governance transparency τ_j the reaction function of firm i is

$$M_i(m_j) = \frac{q(1-p)^2\pi_2(1 - m_j\tau_j)}{c_i} \tag{11}$$

which coincides with (1) if $\tau_j = 1$. The marginal benefit of monitoring on the RHS is decreasing in τ_j. If firm j shares more information for a given level of monitoring, shareholders in i have a weaker incentive to monitor on their own. Transparency facilitates free-riding and discourages monitoring. Under Assumption 2 there exists a unique interior equilibrium in monitoring efforts such that

$$m_i^e = \frac{q(1-p)^2\pi_2[c_j - \tau_j q(1-p)^2\pi_2]}{[c_i c_j - \tau_j\tau_i(q(1-p)^2\pi_2)^2]}. \tag{12}$$

As the rival's level of transparency τ_j increases, the equilibrium monitoring effort in firm i decreases. At the same time, if shareholders decide to share more information themselves, they also gather more information ($\delta m_i^e/\delta\tau_i > 0$).

It can easily be checked that greater transparency in either firm always increases joint firm values in equilibrium (since there is less wasteful duplication of monitoring

efforts). Hence, full transparency is socially optimal ($\tau_1^o = \tau_2^o = 1$). Transparency of active monitors such as pension or hedge funds is commonly advocated on the grounds that it reduces agency problems between the managers of the funds and their shareholders. The model abstracts from agency conflicts on the side of the monitor. In contrast, here opaqueness is socially costly because it does not generate information for peers monitors.

I now turn to the shareholders' equilibrium choice of transparency. It turns out that there exist multiple equilibria, (τ_1^e, τ_2^e), in disclosure policies. In equilibrium, one firm is fully opaque while the other is indifferent with regard to its disclosure policy ($\tau_i^e = 0$ and $\tau_j^e \in [0, 1]$). Hence, at least one of the firms does not share any information. Shareholders prefer to be opaque in order to foster information collection by the other firm. The model thus uncovers a new private cost of disclosure, namely the adverse effect on the monitoring incentives of peers. This idea differs from common explanations for limited transparency such as direct costs of disclosure or indirect costs from providing product-market rivals with valuable information (see Bhattacharya and Chiesa (1995)). A comparison between the equilibrium and the first-best outcome yields the following result:

Proposition 7. *In equilibrium, governance transparency is always inefficiently low.*

Even though there are no direct costs of disclosure, transparency is too low. Proposition 7 establishes a role for mandatory disclosure requirements. Suppose that the regulator can force investors to reveal their actions. Then governance regulation can increase investor welfare by requiring all investors to be fully transparent about their actions, i.e., by enforcing $\tau_1^o = \tau_2^o = 1$.

7 Dispersed Ownership as a Commitment not to Monitor

The equilibrium analysis in Section 3 abstracts from any coordination or incentive problems among shareholders *within* each firm. The cost of monitoring is assumed to be shared equally by all shareholders who jointly determine its optimal level. In a next step, I consider the case where information acquisition imposes non-contractible private costs on the monitor. The focus is on a particular institutional setting in which monitoring within each firm is undertaken by a single large blockholder whose incentives to acquire information and intervene are determined by the size of his block.[29] Starting from the assumption that the two firms have distinct blockholders, this section studies the link between, on the one hand, monitoring externalities across firms and, on the other hand, free-riding and ownership structures within firms. For

[29] La Porta, Lopez-de-Silanes and Shleifer (1999) present empirical evidence on the prevalence of concentrated ownership around the world. Black (1992) provides evidence of large shareholder activism. See also Holderness (2003) for a survey of blockholders and corporate control.

simplicity, the analysis abstracts from the moral hazard problem and reverts to the pure learning framework of Sections 2 through 4.

Suppose that each firm is associated with the following ownership structure: a fraction γ_i of shares in firm i is held by large blockholder i and the remaining shares, $1-\gamma_i$, are held by atomistic shareholders.[30] Each individual shareholder in firm i can observe manager i's ability with probability m_i at a *private* cost $\frac{1}{2}cm_i^2$.[31] Shareholders cannot coordinate or contract on m_i.

Each firm is associated with a different founder. The setup is augmented by an ex-ante stage, $t = -1$, at which each founder determines the ownership structure of his firm, i.e., he chooses the level of ownership concentration. Ownership structures are chosen simultaneously and non-cooperatively and are publicly observable.[32] We ignore the possibility that the founders sell stakes to the same blockholder.[33] Finally, to facilitate the exposition, τ is set equal to one in this section.

We solve the game by backward induction. Let us start by defining $\Pi_i(m_i, m_j)$ as the total expected cash flow in firm i given its monitoring intensity m_i and the intensity in the rival firm m_j:

$$\begin{aligned}\Pi_i(m_i, m_j) = pq(\bar{x}_1 + x_2) + (1-p)(1-q)(\underline{x}_1 + \pi_2) + (1-q)p(x_1 + x_2)\\ +q(1-p)[x_1 + m_i\pi_2 + (1-m_i)(p + (1-p)m_j)\pi_2].\end{aligned}$$

Shareholders within each firm have perfectly congruent objectives and the optimal replacement policy at the interim date is the same as in Section 3. At $t = 0$ the two blockholders simultaneously and non-cooperatively choose their monitoring efforts. Each large shareholder maximizes the value of his block. Given his block γ_i and the rival's monitoring level m_j, large shareholder i solves:

$$\max_{m_i \in [0,1]} \gamma_i \Pi_i(m_i, m_j) - \frac{1}{2}cm_i^2.$$

Under Assumption 2 the following unique, symmetric equilibrium in monitoring efforts obtains for any given pair of ownership structures:

$$m_i^e(\gamma_i, \gamma_j) = \frac{\gamma_i a(1-p)[c - \gamma_j a(1-p)]}{c^2 - \gamma_i\gamma_j a^2(1-p)^2}, \tag{13}$$

where $i, j = 1, 2$ (with $i \neq j$) and $a \equiv q(1-p)\pi_2$. It can easily be checked that the equilibrium monitoring intensity of blockholder i is increasing in γ_i, the size of his

[30] See Shleifer and Vishny (1986), Huddart (1993) and Admati, Pfleiderer, and Zechner (1994) for other models that consider a firm with one large shareholder and a set of small shareholders. Edmans and Manso (2010) provide a theory of multiple blockholders within a single firm.

[31] For instance, monitoring could refer to shareholder activism in the form of a proxy fight.

[32] Observability can be due to regulation. For example, according to the Security Exchange Act of 1934 entities that own more than 5 percent of a firm's outstanding shares have to be reported as "Principal Shareholders" in corporations' proxy statements.

[33] As before in the baseline model, this can be motivated by limited wealth or regulatory constraints.

block. A higher share of cash flows provides greater incentives to acquire information. At the same time, monitoring is decreasing in γ_j, the level of ownership concentration in the rival firm. Closer monitoring by blockholder j encourages free-riding. The model thus predicts that a blockholder's monitoring effort decreases if an exogenous shock increases the level of ownership concentration in a rival firm.

In a next step, I consider the game played by the founders of the two firms at $t = -1$. The founders anticipate the equilibrium outcome in (13) which obtains for any pair of ownership structures. When choosing γ_i, founder i maximizes the total proceeds from selling the firm which equal expected future cash-flows net of the monitoring costs. Taking γ_j as given, founder i solves the following problem:

$$\max_{\gamma_i \in [0,1]} \Pi_i(m_i^e(\gamma_i, \gamma_j), m_j^e(\gamma_i, \gamma_j)) - \frac{1}{2} c m_i^e(\gamma_i, \gamma_j)^2$$

The FOC, or reaction function, for founder i is then

$$\frac{\delta \Pi_i}{\delta m_i^e} \frac{\delta m_i^e}{\delta \gamma_i} + \frac{\delta \Pi_i}{\delta m_j^e} \frac{\delta m_j^e}{\delta \gamma_i} - c m_i^e \frac{\delta m_i^e}{\delta \gamma_i} = 0. \tag{14}$$

The first summand on the LHS is positive and equals the marginal profits in firm i from closer monitoring by blockholder i. The second summand is (weakly) negative, reflecting the negative impact of concentrated ownership in firm i on the incentives of blockholder j. This effect disappears if $\gamma_j = 0$: in this case an increase in γ_i does not deteriorate the incentives of blockholder j since m_j is zero anyhow.

Postulating a symmetric solution, the FOC in (14) yields the following result:

Lemma 2. *There exists a symmetric equilibrium in ownership structures in which each founder chooses*

$$\gamma^e = \frac{c}{c + q(1-p)^2 \pi_2} < 1.$$

In equilibrium ownership is dispersed. Dispersion serves as a commitment not to monitor vis-á-vis the blockholder in the rival firm in order to strengthen his incentive to acquire information.[34] It arises due to the information externality at the interim date and despite the absence of any direct costs of concentrated ownership (which may result from underdiversification or illiquidity).

Consider as a benchmark the socially optimal levels of ownership concentration which the two founders would choose if they maximized joint firm values. Let γ_i^o denote the socially optimal block size in firm i. Then it can easily be checked that fully concentrated ownership is optimal, i.e., $\gamma_1^o = \gamma_2^o = 1$. The analysis in Section 3 corresponds to the case of fully concentrated ownership. Even then, monitoring is inefficiently low in equilibrium (Corollary 2). The only effect of dispersion is to further weaken monitoring incentives relative to the outcome in Proposition 2. Hence, it can never be optimal from a social planner's perspective.

[34] Burkart, Gromb, and Panunzi (1997) argue that a dispersed ownership structure can serve as a commitment device vis-á-vis management: weaker monitoring incentives for the large shareholder enhance the CEO's initiative.

Proposition 8. *In equilibrium, ownership structures are inefficiently dispersed ($\gamma^e < \gamma^o = 1$).*

The monitoring externality translates into inefficiently low levels of ownership concentration which, in turn, further exacerbate the free-rider problem. In Acharya and Volpin (2009) and Bebchuk and Zingales (2000) the privately optimal ownership structure also deviates from the socially optimal one. Moreover, the direction of the deviation is the same in their models as in this paper, i.e., there is excessive dispersion. Bebchuk and Zingales argue that a founder may choose an inefficiently low level of ownership concentration in anticipation of a takeover for the purpose of extracting the surplus from a future raider: free-riding by atomistic shareholders during the tender offer process strengthens the target firm's bargaining power. In contrast, the distortion in my model does not depend on the existence of an active takeover market but obtains as long as firms are exposed to common performance shocks. In Acharya and Volpin a larger ownership stake leads to better governance which reduces CEO compensation. When firms compete for scarce managerial talent, an externality arises. Poorly governed firms overpay their managers which forces their competitors to pay their managers more. In equilibrium ownership stakes are too low because founders fail to internalize that better governance reduces compensation in rival firms. In the absence of any direct costs of concentration (e.g., costs of underdiversification) the socially optimal ownership structures would obtain in their model which is not the case in my framework.

Finally, note that dispersion gives rise to a second externality at the monitoring stage in $t = 0$. Monitoring choices are inefficiently low not only because of non-internalization of rival profits but also because a blockholder does not internalize the security benefits accruing to the small shareholders of his firm. In a nutshell, the inter-firm externality leads to an intra-firm externality.

8 Conclusion

Going beyond the standard one-firm approach, the paper studies interactions between monitoring activities in different firms. I show that shareholders free-ride on information acquisition in rival firms. An externality arises because informed intervention transmits valuable information about a common performance shock to peers.

Using US data on forced CEO turnover, I provide tentative empirical evidence of information spillovers across firms. Controlling for a variety of performance measures, I find that CEOs are significantly more likely to be dismissed after other CEOs in their industry have been fired. At the same time, CEOs' turnover risk is unrelated to (presumably uninformative) firings in firms outside of their industry. Moreover, the impact of firings at peer firms on turnover risk is significantly greater for CEOs in firms with low institutional blockholder ownership.

The paper studies two ways to mitigate the free-rider problem, a market-based

mechanism and regulation. The model predicts the emergence of investors that are engaged in multiple firms and pursue focused strategies in order to capitalize on learning synergies. I show that regulation in the form of a (binding) cap on CEO compensation can mitigate the externality by forcing all shareholders to monitor their manager more closely. Generally speaking, two basic ingredients are needed for the pay cap to be welfare enhancing. First, shareholders need to have more than one instrument at their disposal to provide incentives. If they are forced to reduce pay, they can preserve incentives by monitoring more closely. Second, the instrument towards which shareholders are forced to switch generates positive externalities.

While the analysis has focused on CEO turnover as a specific application, the model's results can also be applied to other types of shareholder intervention, for instance strategy changes. A natural extension of the model would be to let firm performance depend on different types of exogenous factors, for instance both an industry and a macroeconomic shock. In addition to determining the intensity of monitoring, shareholders would also have to choose between different information types. The question then arises how the information choices of shareholders in different firms interact. An extension along these lines might provide a fruitful avenue for future research.

Appendix

Proof of Proposition 2 (First-best outcome).

The effect of a reduction in monitoring costs in firm i on the value of firm j is:

$$\frac{\delta V_j}{\delta c_i} = q(1-p)^{\pi}[\frac{\delta m_j}{\delta c_i} + (1-m_j)\frac{\delta m_i}{\delta c_i} - \frac{\delta m_j}{\delta c_i} m_i] - c_j m_j \frac{\delta m_j}{\delta c_i}$$

Using the envelope theorem, the condition reduces to

$$\frac{\delta V_j}{\delta c_i} = q(1-p)^2 \pi (1-m_j)\frac{\delta m_i}{\delta c_i} < 0.$$

Hence, a reduction in mointoring costs raises the value of the peer firm.

Proof of Proposition 3 (First-best outcome).

The first-best monitoring intensities solve the following problem:

$$\max_{m_i, m_j \in [0,1]} qp^2 2(\bar{x}_1 + x_2) + 2q(1-p)p(\bar{x}_1 + x_1 + x_2 + \pi_2) + (1-q)(1-p)^2 2(\underline{x}_1 + \pi_2)$$
$$+q(1-p)^2[2m_i m_j \pi_2 + (1-m_i)m_j 2\pi_2 + (1-m_j)m_i 2\pi_2] + (1-q)p^2 2(x_1 + x_2)$$
$$2(1-q)p(1-p)(\underline{x}_1 + x_1 + x_2 + \pi_2) - \frac{1}{2}c_i(m_i)^2 - \frac{1}{2}c_j(m_j)^2$$

The First Order Condition with respect to m_i is

$$q(1-p)^2 \pi_2 2(1-m_j) = c_i m_i$$

which coincides with (3). Then Assumption 2 implies that there exists a unique interior solution:

$$m_i^o = \frac{q(1-p)^2 \pi_2 2(c_j - q(1-p)^2 \pi_2 2)}{c_i c_j - (q(1-p)^2 \pi_2 2)^2}$$

Comparing m_i^e and m_i^o, one finds that $m_i^o > m_i^e$ if

$$c_i(c_j - c_i) + (c_i - q(1-p)^2\pi)(c_i - 2q(1-p)^2\pi) > 0. \tag{15}$$

The second summand is always positive by Assumption 2. Hence, if firm i has lower monitoring costs than firm j, the above condition always holds. Hence, firm i's first-best monitoring intensity exceeds the equilibrium one. Conversely, if c_i is sufficiently high relative to c_j, then the LHS above is negative and $m_i^o < m_i^e$. The comparison between the risk of a total monitoring failure in the first-best case $((1-m_i^o)(1-m_j^o))$ and in equilibrium $((1-m_i^e)(1-m_j^e))$ is unambiguous:

$$(1-m_i^o)(1-m_j^o) < (1-m_i^e)(1-m_j^e). \tag{16}$$

To verify (16) two cases have to be distinguished: suppose without loss of generality that firm i has higher costs ($c_i > c_j$). If the condition in (15) holds, then

$m_i^o > m_i^e$ and $m_j^o > m_j^e$ which immediately implies (16). Suppose instead that condition (15) does not hold (i.e., $m_i^o < m_i^e$). Then it can easily be checked that $m_i^o < m_i^e < m_j^e < m_j^o$ which again implies (16).

Derivation of the first-best contract in Section 5.2. The solution to the following program maximizes joint firm values:

$$\begin{aligned}\max_{\bar{w}_i \geq 0, m_i \in [0,1], i=1,2} & qp(-\bar{w}_1) + q(1-p)(m_1\pi_2 + (1-m_1)(p+(1-p)m_2)\pi_2) \\ & +qp(-\bar{w}_2) + q(1-p)(m_2\pi_2 + (1-m_2)(p+(1-p)m_1)\pi_2) \\ & -\frac{1}{2}c(m_1)^2 - \frac{1}{2}c(m_2)^2\end{aligned}$$

subject to the two IC constraints

$$\bar{w}_i \geq \frac{Z_1}{q\delta p} - \frac{(1-q)}{q}Z_2 - [1-(1-p)(1-m_1)(1-m_2)]Z_2, \text{ with } i = 1, 2.$$

From Assumption 4 it follows that the IC constraints are binding. Hence, the FOC with respect to m_i is

$$2qp(1-p)(1-m_j)Z_2 + 2q(1-p)^2(1-m_j)\pi_2 = cm_i; \text{ with } i, j = 1, 2.$$

Assumption 4 implies that there is a unique intersection of the two reaction functions in m^o given in (9).

Proof of Proposition 6. Suppose that shareholders in, say, firm 1 expect the other firm to choose m^o (as given in (9)) and the rival manager to exert high effort. Then they solve the following problem:

$$\begin{aligned}\max_{m_1 \in [0,1], \bar{w}_1 \geq 0} & qp(\bar{x}_1 - \bar{w}_1 + x_2) \\ & +q(1-p)(x_1 + m_1\pi + (1-m_1)(p+(1-p)m^o)\pi_2) - \frac{1}{2}cm_1^2\end{aligned}$$

subject to the IC constraint

$$\begin{aligned}& qp\bar{w}_1 + [p + q(1-p)^2(1-m_1)(1-m^o)]Z_2 \\ \geq\ & qp(\underline{e})\bar{w}_1 + [p(\underline{e}) + q(1-p)(1-p(\underline{e}))(1-m_1)(1-m^o)]Z_2 + Z_1\end{aligned}$$

and subject to the pay-cap $\bar{w}_1 \leq \bar{w}(m^o)$.

Assumption 4 implies that the IC constraint is binding. To see that the pay-cap constraint is binding, suppose to the contrary that this was not the case. Then the optimal monitoring level m_1' for firm 1 is given by

$$cm_1' = qp(1-p)(1-m^o)Z_2 + q(1-p)^2(1-m^o)\pi_2$$

which is clearly lower than m^o (see the previous proof). From $m_1' < m^o$ it follows that $\bar{w}_1(m_1') > \bar{w}(m^o)$, i.e., the pay-cap constraint is violated. Hence, we have shown by contradiction that the pay-cap is binding. In a next step, the binding IC and pay-cap constraints uniquely determine the monitoring intensity of firm 1 to be m^o.

Proof of Lemma 2 ($\gamma^e < \gamma^o = 1$) Consider the reaction function in (14). Substituting the reaction function of blockholder i into (14) yields

$$(1-\gamma_i)\frac{\delta \Pi_i}{\delta m_i}\frac{\delta m_i}{\delta \gamma_i} + \frac{\delta \Pi_i}{\delta m_j}\frac{\delta m_j}{\delta \gamma_i} = 0.$$

Note that the privately optimal level of ownership concentration for firm i is strictly lower than one unless $\gamma_j = 0$ (in this case $\frac{\delta m_j}{\delta \gamma_i} = 0$ since $m_j = 0$)). This equation is equivalent to

$$(1-\gamma_i)q(1-p)^2\pi_2(1-m_j)\frac{\delta m_i}{\delta \gamma_i} + q(1-p)^2\pi_2(1-m_i)\frac{\delta m_j}{\delta \gamma_i} = 0.$$

This yields

$$(1-\gamma_i)(1-p)(1-m_j)c - q(1-p)^3(1-m_i)\gamma_j\pi_2 = 0.$$

Positing a symmetric solution for the level of ownership concentration implies that $m_i = m_j = m$. Then, using the fact that $1 - m = \frac{c}{c+\gamma q(1-p)^2\pi_2}$, the above expression reduces to

$$(1-\gamma^e)c^2(1-p) - q(1-p)^3\gamma^e\pi_2 c = 0 \Leftrightarrow \gamma^e = \frac{c}{c+q(1-p)^2\pi_2}.$$

References

[1] Acharya, Viral, Marc Gabarro and Paolo Volpin, 2010, Competition for managers, corporate governance and incentive compensation, Working paper.

[2] Acharya, Viral, and Paolo Volpin, 2010, Corporate governance externalities, *Review of Finance* 14, 1-33.

[3] Admati, Anat, Paul Pfleiderer and Josef Zechner, 1994, Large shareholder activism, risk sharing, and financial market equilibrium, *Journal of Political Economy* 102, 1097-1130.

[4] Barro, Jason, and Robert Barro, 1990, Pay, performance, and turnover of bank CEOs, *Journal of Labor Economics* 8, 448-481.

[5] Bebchuk, Lucian and Jesse Fried, 2004, Pay without Performance: The Unfulfilled Promise of Executive Compensation, *Harvard University Press*.

[6] Bebchuk, Lucian and Luigi Zingales, 2000, Ownership structures and the decision to go public: private versus social optimality, in: Randall Morck (ed.), Concentrated Corporate Ownership, University of Chicago Press.

[7] Becht, Marco, Patrick Bolton and Ailsa Röell, 2003, Corporate governance and control, in: G. Constantinides, M. Harris and R. Stulz (ed.), The Handbook of the Economics of Finance, North-Holland.

[8] Becht, Marco, Julian Franks, Colin Mayer and Stefano Rossi, 2010, Returns to shareholder activism: Evidence from a clinical study of the Hermes UK Focus Fund, *Review of Financial Studies* 23, 3093-3129.

[9] Bethel, Jennifer, Julia Porter Liebeskind and Tim Opler, 1998, Block share purchases and corporate performance, *Journal of Finance* 53, 605-634.

[10] Bhattacharya, Sudipto and Chiesa Gabriella, 1995, Proprietary information, financial intermediation, and research incentives, *Journal of Financial Intermediation* 4, 328-357.

[11] Black, Bernard, 1992, The value of institutional investor monitoring: The empirical evidence, *UCLA Law Review* 39, 895-939.

[12] Bris, Arturo, Neil Brisley and Christos Cabolis, 2008, Adopting better corporate governance, *Journal of Corporate Finance* 14, 224-240.

[13] Burkart, Mike, Denis Gromb and Fausto Panunzi, 1997, Large shareholders, monitoring, and the value of the firm, *Quarterly Journal of Economics* 112, 693-728.

[14] Campbell, John, Martin Lettau, Burton Malkiel and Yexiao Xu, 2001, Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk, *Journal of Finance* 56, 1-41.

[15] Carleton, Willard, James Nelson and Michael Weisbach, 1998, The influence of institutions on corporate governance through private negotiations: Evidence from TIAA-CREF, *Journal of Finance* 53, 1335-1362.

[16] Clifford, Christopher, 2008, Value creation or destruction? Hedge funds as shareholder activists, *Journal of Corporate Finance* 14, 323-336.

[17] Cremers, Martijn and Vinay Nair, 2005, Governance mechanisms and equity prices, *Journal of Finance* 60, 2859-2894.

[18] Cressy, Robert, Federico Munari and Alessandro Malipiero, 2007, Playing to their strengths? Evidence that specialization in the private equity industry confers competitive advantage, *Journal of Corporate Finance* 13, 647-669.

[19] Cronqvist, Henrik and Rüdiger Fahlenbrach, 2009, Large shareholders and corporate policies, *Review of Financial Studies* 22, 3941-3976.

[20] Dicks, David, 2010, Executive compensation, incentives, and the role for corporate governance regulation, University of North Carolina Working Paper.

[21] Edmans, Alex and Gustavo Manso, 2010, Governance through trading and intervention: a theory of multiple blockholders, *Review of Financial Studies* forthcoming.

[22] Ertugrul, Mine and Karthik Krishnan, 2008, Can CEO dismissal be proactive?, Working paper.

[23] Fahlenbrach, Rüdiger, 2009, Shareholder rights, boards, and CEO compensation, *Review of Finance* 13, 81-113.

[24] Harford, Jarrad, Dirk Jenter and Kai Li, 2010, Shareholder cross-holdings and their effect on acquisition decisions, *Journal of Financial Economics* forthcoming.

[25] Hermalin, Benjamin and Michael Weisbach, 1998, Endogenously chosen boards of directors and their monitoring of the CEO, *American Economic Review* 88, 96-118.

[26] Hermalin, Benjamin and Michael Weisbach, 2010, Information disclosure and corporate governance, Working paper.

[27] Holderness, Clifford, 2003, A survey of blockholders and corporate control, *FRBNY Economic Policy Review* 9, 51-64.

[28] Holmström, Bengt, 1979, Moral hazard and observability, *Bell Journal of Economics* 10, 74-91.

[29] Holmström, Bengt, 1982, Moral hazard in teams, *Bell Journal of Economics* 13, 324-340.

[30] Huddart, Steven, 1993, The effect of a large shareholder on corporate value 39, 1407-1421.

[31] Inderst, Roman and Holger Mueller, 2010, CEO replacement under private information, *Review of Financial Studies* forthcoming.

[32] Jenter, Dirk and Fadi Kanaan, 2010, CEO turnover and relative performance evaluation, *Journal of Finance* forthcoming.

[33] John, Kose and Dalida Kadyrzhanova, 2010, Relative governance, Working paper.

[34] Kaplan, Steven and Bernadette Minton, 2008, How has CEO turnover changed?, Working paper.

[35] Klein, April and Emanuel Zur, 2009, Entrepreneurial shareholder activism: Hedge funds and other private investors, *Journal of Finance* 64, 187-229.

[36] La Porta, Rafael, Florencio Lopez-de-Silanes and Andrei Shleifer, 1999, Corporate ownership around the world, *Journal of Finance* 54, 471-517.

[37] Leuz, Christian and Peter Wysocki, 2008, Economic consequences of financial reporting and disclosure regulation: a review and suggestions for future research, Working paper.

[38] Mitchell, Mark and J. Harold Mulherin, 1996, The impact of industry shocks on takeover and restructuring activity, *Journal of Financial Economics* 41, 193-229.

[39] Mookherjee, Dilip, 1984, Optimal incentive schemes with many agents, *Review of Economic Studies* 51, 433-446.

[40] Morck, Randall, Andrei Shleifer and Robert W. Vishny, 1989, Alternative mechanisms for corporate control, *American Economic Review* 79, 842-852.

[41] Morse, Adair, Vikram Nanda and Amit Seru, 2009, Are incentive contracts rigged by powerful CEOs?, *Journal of Finance* forthcoming.

[42] Norton, Edgar and Bernard Tenenbaum, 1993, Specialization versus diversification as a venture capital investment strategy, *Journal of Business Venturing* 8, 431-442.

[43] Scharfstein, David, 1988, Product market competition and managerial slack, *RAND Journal of Economics* 19, 147- 155.

[44] Schmidt, Klaus, 1997, Managerial incentives and product market competition, *Review of Economic Studies* 64, 191-213.

[45] Shleifer, Andrei and Robert Vishny, 1989, Management entrenchment: The case of manager-specific investments, *Journal of Financial Economics* 25, 123-139.

[46] Stiglitz, Joseph, 1985, Credit markets and the control of credit markets, *Journal of Money, Credit and Banking* 17, 133-152.

[47] Warner, Jerold, Ron Watts, and Karen Wruck, 1988, Stock prices and top management changes, *Journal of Financial Economics* 20, 461-492.

[48] Zwiebel, Jeffrey, 1995, Block investment and partial benefits of corporate control, *Review of Economic Studies* 62, 161-185.