Looking the Other Way: The Screening Role of (Weak) Internal Monitoring*

Felix Zhiyu Feng† Wenyu Wang‡ Yufeng Wu§

February 2021

Abstract

Internal monitoring is a cardinal responsibility of the board of directors, many of which have been subject to relentless criticism of conceding too much power to the managers and allowing them too much latitude. We analyze the role of internal monitoring in a dynamic adverse selection model, in which managers have unobservable ability and must be given rents through a compensation contract in exchange for revealing their private information. Monitoring has two effects: an ex-post disciplining effect, whereby more intense monitoring limits the information advantage and the rents for managers with higher ability; and an ex-ante screening effect, whereby weaker monitoring allows firms to attract better managers. We show an optimal level of monitoring intensity that balances the disciplining effect and the screening effect exists, even when monitoring is intrinsically costless and can be made arbitrarily strong. We empirically test these predictions and find that less monitoring by the board is indeed associated with on average more capable managers, and the relationship between monitoring intensity and firm value is hump-shaped in the data. Although the lack of monitoring is often viewed as a failure of internal governance, our analysis suggests it is a plausible tactic for attracting and retaining talented managers.

JEL Classification: G32, G34, D82, D86, M12
Key Words: internal monitoring, board of directors, dynamic adverse selection, persistent private information, screening

*We are indebted to Curtis Taylor, Alexei Tchistyi and Mark Westerfield for helpful discussions at the early stage of this paper.

†Michael G. Foster School of Business, University of Washington. Email: ffeng@uw.edu
‡Kelley School of Business, Indiana University. Email: wenywang@indiana.edu
§Gies College of Business, University of Illinois Urbana-Champaign. Email: yufengwu@illinois.edu

1 Introduction

Internal monitoring by the board of directors is widely regarded as indispensable for protecting investor interests and enhancing firm value. When hefty and embarrassing company losses make the headlines, the finger is often pointed at the board for its lack of internal oversight. For example, the board of Groupon came under fire when the company had to revise its past quarter's earnings downward (Boivie, Bednar, and Andrus, 2016); the board of JP Morgan was stigmatized as "blatantly irresponsible" and "asleep at the wheel" when the bank accumulated over \$6 billion in trading losses (Feeley and Bathon, 2015). Indeed, a vast literature in academic research documents that close monitoring by the board is effective in disciplining managers and mitigating the *moral hazard* problem arising from their unobservable actions. When monitoring is lax, managers tend to exert less effort, enjoy more perks, and engage in self-aggrandizing but value-destroying activities.

In this paper, we propose a complementary view that highlights a new, positive role of weak monitoring: it can facilitate the ex-ante screening of better managers, thus mitigating the *adverse selection* problem arising from unobservable managerial ability. This ex-ante screening effect, combined with the well-studied ex-post disciplining effect, leads to a hump-shaped relationship between monitoring intensity and firm value, implying an optimum level of monitoring even when monitoring is intrinsically costless and can be made arbitrarily strong.

We formalize this view in a dynamic agency model with persistent private information. In the model, a firm needs to hire a manager with unobservable innate ability. The output of the firm is determined by the manager's (costly) effort and the quality of the firm's project, the latter of which depends on three factors: the manager's ability, the firm's monitoring intensity, and a series of random shocks. Higher ability managers leads to higher quality projects on average, while more intense monitoring reduces (but does not eliminate) the time-series variation of project quality. Output is publicly observable while effort, managerial ability, and the project quality are all privately known by the manager. In other words, the firm can see the output but does not see how the output is produced. To hire a manager and start production, the firm offers a contract specifying the series of performance targets and the corresponding compensation, based on the manager's reported ability and the history of reported project quality. Given the private knowledge of their abilities, the managers willing to accept the contract form a managerial pool, from which a randomly selected manager will be matched with the firm and begin the production process.[1]

Without any information asymmetry, the firm's contract only needs to compensate the manager's cost of effort. When managerial ability and project quality is private information of the manager, an adverse selection problem arises, because the manager can conceal his true ability by misreporting the realization of project quality and meet any resulting performance target through adjusting his effort. Consequently, the firm must use a menu of output targets and compensation to solicit truthful reporting from the manager. In particular, managers with higher ability is given higher output targets and more compensation in addition to their cost of effort. This additional compensation is commonly known as the manager's (information) *rent* and is driven by both managerial ability and the firm's monitoring intensity.

The optimal contract under asymmetric information maximizes the firm's expected output minus the information rent to the manager. It is embedded with two sets of mechanisms, both related to monitoring. First, monitoring reduces the *marginal* rent the high-ability managers can extract from the contract: the stronger the monitoring, the more persistent the firm's project quality, and thus the less information rent needed to solicit the manager's private information. This is consistent with the conventional, disciplining effect of monitoring, which occurs after the manager is hired.

However, our model features a second mechanism, which takes effect before the manager is hired. Because strong monitoring results in a less sensitive relationship between managerial rent and managerial ability, low-ability managers are more willing to accept contracts with strong monitoring, while high-ability managers are more willing to work for firms with weak monitoring. Consequently, by varying the monitoring intensity, the firm can alter the distribution of managers it can be potentially matched with. In particular, while weaker monitoring renders more rents to the manager once he is hired, it also helps the firm attract

[1]We assume random matching to eliminate confounding mechanisms such as dynamic signaling or bargaining. The implications of this assumption are discussed in Section 2

managers with on average higher ability. This is the ex-ante screening effect of monitoring we highlight in this study.

The overall impact of monitoring on firm value is determined by the tradeoff between the ex-ante screening effect and the ex-post disciplining effect. We illustrate that even in the absence of any explicit monitoring cost, it is not optimal to maximize monitoring intensity. When monitoring is too weak, the firm concedes too much rent to the manager once he is hired. If monitoring is too strong, the managerial pool consists of too many low-ability managers, leading to low output and low firm value in expectation. Therefore, the relationship between firm value and monitoring intensity is hump-shaped. An optimal level of monitoring exists as a balance of the disciplining versus screening effects.

The predictions of our model are testable, which we proceed to examine in the data. Since the disciplining effect of monitoring has been well studied (Yermack, 1996; Hallock, 1997; Core, Holthausen, and Larcker, 1999), we focus on testing the screening effect and the humped-shaped relationship between firm value and monitoring intensity, which are the novel predictions of our model.

We first test the screening effect by uncovering the relationship between monitoring intensity and managerial ability. We measure shareholder monitoring using three proxies: director *co-option*, *non-coopted-independence*, and *busyness*. Prior literature shows that board monitoring decreases as a larger fraction of the directors become co-opted or busy, and these two measures have more explanatory power for monitoring effectiveness than the conventional measure of board independence (Coles, Daniel, and Naveen, 2014; Fich and Shivdasani, 2012). Compared to measuring monitoring intensity, measuring managerial ability is empirically more challenging. While CEO compensation should reflect managerial ability in equilibrium, it is also confounded with the endogenous information rent we model in the paper. To overcome this difficulty, we infer managerial ability using a specific type of corporate investment, that is, innovation activities. Arguably, innovation has profound effects on firms' long-term growth, and it relies heavily on managers' insight, judgment, and commitment (Chen, Podolski, and Veeraraghavan, 2015; Custódio, Ferreira, and Matos, 2019). We thus use a firm's R&D expenditure (innovation input), patent citations, and the market value of patents (innovation outputs) as our proxies for the manager's ability.

We find that both firm innovation inputs and outputs are negatively correlated with the measures of monitoring intensity. Specifically, firms with a smaller fraction of busy board, co-opted board members, or more independent board members have significant lower R&D expenditure, and they generate patents that are less influential and valuable in the long-run (i.e., fewer patent citations and lower market value of the patents). To the extent that managerial ability is a crucial determinant of firm innovation success, our findings lend strong support to the model's prediction that weak monitoring can increase a firm's chance of attracting more capable managers.

Next, we examine how monitoring intensity affects the overall firm value, proxied by the market-to-book ratio. We find that the effect of monitoring intensity on firm value is non-monotonic and hump-shaped in the data. When monitoring intensity increases from a relatively low level, it generates a positive effect on firm value by improving ex-post discipline and reducing managerial rents. When it further increases to above a given threshold, it begins to hurt firm value by diminishing its ex-ante screening power and generates a negative overall impact among firms who are already closely monitored. These findings are consistent with our model predictions regarding the tradeoff of tightening monitoring.

Literature review: The paper is most closely related to the dynamic contracting literature on monitoring, such as Piskorski and Westerfield (2016), Chen, Sun, and Xiao (2020), Orlov (2020), Zhu (2020), etc. These studies typically feature dynamic moral hazard models in which a contract is used to ensure that the agent takes the desired private actions. In contrast, our model features a dynamic adverse selection problem in which a contract is used to solicit private information. Moreover, in the existing models, monitoring is usually ex post inefficient, and the principal has to commit to a monitoring technology in order to provide the agent sufficient ex ante incentives not to deviate from the desired actions.[2] In our model, monitoring is ex post efficient, as it reduces the time-series variation of the agent's private information and thus the rent the agent can extract. The cost of monitoring, however, is the ex ante average quality of the manager that the firm can be matched with.

Dynamic adverse selection problems with persistent private information are known to be

[2]For example, in Piskorski and Westerfield (2016), after the contract is signed, the principal would like to make the agent believe that she has been activating the monitoring technology without actually doing so.

difficult to analyze. Our solution methodology follows the *revenue-maximizing, direct mechanism* developed in Bergemann and Strack (2015), which builds on the general Myersonian mechanism of Eső and Szentes (2007) and Pavan, Segal, and Toikka (2014) but extended to continuous time. A critical advantage of this approach is that it converts the dynamic adverse selection problem into a static mechanism design problem. Consequently, the model can be solved without the usual dynamic programming techniques involving differential equations or keeping track of the continuation utility. Despite its success in microeconomics, the direct mechanism technique has not seen much utilization in finance studies.[3] The only exception is Gao and Wong (2017), who adopts this method on a capital budgeting problem. Although we intentionally make certain assumptions to ensure that the main technical results of Bergemann and Strack (2015) apply, our model is different in that the principal in Bergemann and Strack (2015) has only one control: the allocation of resources/goods to the agent. The principal sets a price (e.g. a two-part tariff) for different levels of allocation in order to achieve the screening purpose. In our model, the principal (i.e., the firm) has two controls: the output target (which is equivalent to the "goods" in Bergemann and Strack (2015)), and the monitoring intensity. The introduction of the second control implies an endogenous distribution of the agent's type. As a result, the firm faces a tradeoff between the ex ante and ex post effects of monitoring when designing the contract as a screening device.

More broadly speaking, this paper adds to the literature of (discrete-time) mechanism design problems with persistent information, such as Fernandes and Phelan (2000), Battaglini (2005), Zhang (2009), Kapička (2013), Tchistyi (2016), etc. See Bergemann and Välimäki (2019) for a survey of this topic. In most of these studies, the persistence of the agency friction mainly arises from the serial correlation among the shocks to a noisy signal. Our model allows a fairly general form of persistence and can be solved without using the usual techniques involving ODE, PDE, or dynamic programming.

Finally, our paper is related to the literature that explores the impact of board monitoring on firms' real and financial decisions, and their implications on shareholder value, e.g., Mehran (1992); Hermalin (2005); Burns, Kedia, and Lipson (2010); Baldenius, Melumad, and

[3] Examples of the application in microeconomics are Garrett (2017), Gershkov, Moldovanu, and Strack (2018), Krasikov and Lamba (2019), etc.

Meng (2014), etc. These studies in general focus on the positive value of monitoring, such as how it contributes to curbing managers' empire-building incentives, facilitating external financing, lowering excessive compensation, and improving the overall information quality. Our paper illustrates that more intensive monitoring is not always beneficial to the firm, a view also shared with Raheja (2005), Adams and Ferreira (2007), and Harris and Raviv (2008). However, those studies assume the board has multiple roles (e.g. as monitors and advisors of the managers), and over-committing to one role (e.g. intensity of monitoring) can impair the effectiveness of the other. In contrast, our paper shows that even if the board has a single role as the monitor, monitoring itself has multiple effects. In particular, more lax monitoring can facilitate the ex-ante screening of better managers when managers have persistent private information.

2 Model

In this section we develop a dynamic screening model with persistent private information and demonstrate the basic mechanisms through which monitoring can affect the value of screening. All proofs are in the Appendix.

2.1 The Basic Environment

Time is continuous. A firm (the principal) needs to hire a manager (the agent) to manage a profitable project. Both the firm and the manager are risk-neutral, with reservation utility normalized to 0. The output of the project, denote by π_t, is given by:

$$\pi_t = e_t + q_t \tag{1}$$

Here, e_t is the effort of the manager. The cost of effort is $h(e_t)$, where

$$h(e_t) = \frac{1}{2}e_t^2 \tag{2}$$

Meanwhile, q_t is the *quality* of the project, which evolves stochastically over time based on three factors: the manager's *ability* α, the firm's *monitoring intensity* m, and an exogenous Brownian motion Z_t. That is,

$$q_t = \phi(t, Z_t, m, \alpha) \tag{3}$$

where ϕ is an aggregator that summarizes q_t as a function of α, m and the paths of the Brownian shocks Z_t. This aggregator captures the influence of both the manager and the firm on the evolution of project quality in a general form. The manager's influence stems from his ability α, which can be more specifically modeled as the initial quality of the project q_0, the drift (growth rate) of dq_t, or other characteristics of the q_t process. We assume $\alpha > 0$ and $\phi_\alpha > 0$. That is, given any realized path of the Brownian shocks, higher managerial ability leads to higher quality of the project. Meanwhile, although the firm does not observe q_t directly, the former can nevertheless affect the evolution of the latter by monitoring the manager's project selection or experimental explorations. For example, a board that keeps the manager on a "short-leash" through frequent interventions can limit the manager's ability to make risky investments or undertake intensive R&D activities, thus limiting the time-series variations in the firm's project quality (i.e. reducing the variance of dq_t). In Section 3, we solve one example of the model given an explicit functional form of ϕ.

The agency friction arises because both e_t and q_t are private information of the manager. The firm can observe the output π_t but not how it is produced, and a manager overseeing a low-quality project can always mimic the output of another manager overseeing a high-quality project by exerting higher effort. In particular, the firm does not observe the evolution path of q_t, which is equivalent to say that both manager's ability α and the realization of the Brownian shocks Z_t are private information of the manager. We thus refer to α as the manager's *type*. The distribution of α is given by a CDF $F(\alpha)$ and the associated PDF $f(\alpha)$. Section 3 solves an example in which the aggregator ϕ and the distribution $F(\alpha)$ are all explicitly defined.

The timing of the contracting relationship between the firm and the manager is as follows: at $t = 0$, the firm offers a contract to the entire population of managers. Each manager, knowing their own ability α, sees the contract and decides if he is willing to accept the job or not. Then, the firm is matched with a manager randomly selected from the set of managers that are willing to accept the contract. The manager then exerts effort, produces the output, and receives compensation according to the terms of the contract. The contract is terminated at $t = T > 0$. For simplicity, we assume that T is fixed and finite, and that both the firm and the manager are perfectly patient (i.e. zero discounting).

Our model setup warrants some discussions:

i) Unlike the typical dynamic moral hazard models (e.g. Sannikov (2008), DeMarzo and Sannikov (2006), Biais, Mariotti, Plantin, and Rochet (2007) and their various extensions), we assume no noise in the output but introduce noise (Z_t) in the evolution of project quality instead. This is mainly for technical convenience, as will be explained later. It does not affect the nature of the agency problem because output is still an imperfect signal for managerial effort and project quality.

ii) We assume monitoring intensity m is a one-time decision of the firm before the firm meet a specific manager. In that regard, monitoring in our setting can be best understood as the structure of the board that cannot be altered on a daily basis. For example, whether the majority of the board members were appointed before the CEO assumed office, whether board members are also members of the boards at other firms, etc. In contrast to the role of monitoring in dynamic moral hazard models (e.g., Piskorski and Westerfield (2016)), the role of monitoring we highlight in this paper does not require it to be dynamically adjusted. In Section 4 we explore several empirical proxies for monitoring that are consistent with the assumption that they are highly persistent decisions made before the manager is hired.

iii) We assume a one-time random matching between the firm and the pool of managers willing to accept the firm's contract. This is mainly for simplifying the structure of the managerial labor market. It rules out the competition among managers and any potential repeated signaling or bargaining games. However, the fact that each manager can decide whether to join the potential pool of managers that the firm can be matched with after seeing the terms of the contract is crucial. As we explain later, this allows the firm to use the contract terms to shape the distribution of managers that it can potentially be matched with. i.e. the contract serves as a screening device for the firm.

iv) We assume T, the contracting horizon, or the tenure of the manager, is fixed. This rules out endogenous managerial turnover, which is an important subject in dynamic moral hazard models often because managers are assumed to have limited liability. In our model, limited liability is absent except for the manager's participation constraint at $t = 0$. Therefore, ex post turnover after the production process begins is not necessary. One can also interpret this assumption of a fixed tenure as imposing a sufficiently high managerial replacement cost, so that the firm retains the manager until he naturally retires (due to foreseeable reasons, such as age).

v) For the purpose of tractability, we assume that managerial effort and project quality are perfect substitutes in the firm's production technology. Our main mechanism is qualitatively unchanged if effort and project quality also exhibit some degree of complementarity (e.g., if $\pi_t = e_t q_t$). As long as a manager with a low-quality project can always mimic the output of another manager with a high-quality project by exerting higher effort, the nature of the agency friction in our model remains intact.

2.2 The Contract

The firm possesses three instruments in the contract offered to the manager: the level of monitoring intensity, a sequence of target level of output, and the compensation associated with producing that output. Formally, we define a contract as the following:

Definition 1 (Contract) *A contract $\mathcal{C}$ consists of the monitoring intensity m, a sequence of wage $\{w_t\}_{t\geq 0}$ and the output target $\{\pi_t\}_{t\geq 0}$. m is set at the beginning of the contract while both w_t and π_t are contingent on the reported private information ($\hat{\alpha}$ and $\{\hat{q}_t\}_{t\in[0,T]}$).*

Under this definition, after the contract is initiated, the sequence of events during any $[t, t + dt)$ interval is:

1. The Brownian shock Z_t is realized. The manager privately observes Z_t and q_t, the current quality of the project under his management.

2. The manager reports $\hat{q}_t$ to the firm.

3. Given the report (and the history of past reports), the firm imposes an output target π_t.

4. The managers chooses effort e_t to produce the output.

5. Given the output, the firm makes the promised wage payment w_t.

At $t = 0$, the manager's objective is to maximizes his expected life-time utility – his *managerial rent* – under the contract, which is given by

$$R(\hat{\alpha}) = \max_{\hat{\alpha}} \mathrm{E}^{\mathcal{C}(\hat{\alpha})} \left[\int_0^T (w_t - h(e_t)) dt \right] \tag{4}$$

subject to the constraint (1). That is, at any time t, if his reported ability is $\hat{q}_t$, he must exert effort $e_t = \pi_t - \hat{q}_t$ to produce the required level of output π_t. In other words, the manager faces the tradeoff between private information (q_t) and private action (e_t).[4] The expectation is taken under $\hat{\alpha}$, the manager's initial announcement of his type, and the subsequent contract he accordingly receives. We show later that the initial announcement $\hat{\alpha}$ is a sufficient statistic for the expected payoff from the contract.

Following the literature, we restrict our attention to incentive-compatible contracts:

Definition 2 (Incentive compatible contract) *A contract is incentive-compatible if the manager finds it optimal to announce his true initial information and the subsequent realizations of his private information: i.e., $\hat{\alpha} = \alpha$ and $\hat{q}_t = q_t$ for all $t \in [0, T]$.*

Using (4), the incentive-compatibility (IC) condition can be written as

$$R(\alpha) \geq R(\hat{\alpha}), \text{ for all } \hat{\alpha} \tag{IC}$$

The manager must also be willing to accept an incentive compatible contract, thus we require the following *participation constraint* (PC)

$$R(\alpha) \geq 0 \tag{PC}$$

[4] By imposing (1) as a constraint, we effectively assume that the firm can punish the manager *as hard as necessary* if the manager does not produce the required level of output, which is a verifiable breach of the contract.

The firm's objective is to maximize its expected payoff from the contract – the *firm value* – which, given the assumption of random matching, is

$$V(m) = \int_{\mathcal{A}(m)} \mathrm{E}^{\mathcal{C}(\hat{\alpha})} \left[\int_0^T (\pi_t - w_t) dt \right] \tag{5}$$

subject to (IC) and (PC). $\mathcal{A}$ denotes the *managerial pool*, i.e., the set (distribution) of managers that can be matched with the firm.[5] Importantly, $V(m)$ contains two sources of uncertainties: first, the uncertainty from the stochastic evolution of q_t (inside the expectation sign in 5); second, the uncertainty from the quality of the manager hired (the integral over $\mathcal{A}$). The latter is also a function of the monitoring intensity m, because the managers whose (PC) constraint are violated will be excluded from the managerial pool $\mathcal{A}$. The is a crucial feature of the *optimal contract*, which we can now define as the following:

Definition 3 (Optimal contract) *A contract is optimal if it is incentive compatible, maximizes (5), and satisfies (PC) for some* α*.*

Before we proceed, it is useful to first present the *first-best* contract when the project quality q_t is public information. Under this assumption of full information, the first-best level of output target solves

$$\pi_t^{FB} = \arg\max_{\pi_t} \ \pi_t - h(e_t) = \arg\max_{\pi} \ \pi_t - \frac{(\pi_t - q_t)^2}{2} \tag{6}$$

The solution is

$$\pi_t^{FB} = 1 + q_t \tag{7}$$

That is, the first-best level of output is increasing in project quality q_t; the first-best level of effort $e_t^{FB} = \pi_t^{FB} - q_t = 1$ is a constant, and the first-best wage equals the manager's cost of effort (i.e. $w_t^{FB} = h(e_t^{FB})$), implying that the manager earns zero rent.

[5] The integral should be understood as the integration of α over the support $\mathcal{A}$ weighted by the distribution density $F(\alpha)$. E.g., if $\mathcal{A} = [0, +\infty)$, then $\int_{\mathcal{A}}(\cdot) = \int_0^{+\infty}(\cdot)dF(\alpha)$.

2.3 The Solution

We now proceed to characterize the optimal contract. Comparing to the existing literature, a theoretical innovation of this our model is that the manager's private information α is *persistent*. For example, if $\alpha = q_o$, then the project quality of a manager with higher ability will always be higher than that of another manager with low ability, following the same realization of Z_t. While this is a natural assumption based on the firm-manager relationship in practice, it imposes a non-trivial analytical hurdle for solving the optimal contract, because persistent private information generates an *dynamic adverse selection* problem that is known to be much more difficult to deal with compared to dynamic moral hazard problems.[6] Moreover, the solutions, even if they could be found, are often intractable and generate few testable predictions.

To resolve this challenge, we follow Bergemann and Strack (2015) and consider a *revenue-maximizing direct mechanism*. In general, when a_t is not observable, the adverse selection with persistent private information requires complicated IC conditions to prevent all kinds of deviations from the manager. The novelty of the direct mechanism is that under mild technical conditions (which are satisfied in this model and discussed later), it is without the loss of generality to establish the IC condition for a particular type of deviation: if a type-$\tilde{\alpha}$ manager misreports his type to be $\hat{\alpha} \neq \tilde{\alpha}$, his follow-up strategy is to continue misreporting the project quality and exerting effort *as if his type was* $\hat{\alpha}$ *and he had reported that truthfully*. That is, at any time, his reported project quality and effort satisfies

$$\tilde{q}_t + \tilde{e}_t = \pi_t^{\mathcal{C}(\hat{\alpha})} \tag{8}$$

where the right-hand-side indicates the output that the contract requires for a manager whose type is $\hat{\alpha}$ and who always reports the project quality truthfully.

This result has two critical implications: first, although the manager's private information is persistent, it is without the loss of generality to label each manager by his *type* α. The only

[6] As written in Sannikov (2007): "the incentive constraints under adverse selection are nonstandard and significantly more complicated than under pure moral hazard." Recent studies of dynamic moral hazard problems with persistent private information include Williams (2011, 2015), Marinovic and Varas (2019), Feng (2020), etc.

deviation that we need to rule out is the manager misreporting his type at the beginning of the contract. After that initial report, he will behave consistently as if that misreported type was his true type. This effectively resolves the adverse selection problem at time-0, which greatly simplifies the dynamic contracting problem into a static mechanism design problem.

Second, using (8), we can calculate the *information rent* that a manager with type $\tilde{\alpha}$ earns under the optimal contract. Recall that under the first-best, the manager's wage equals exactly his cost of effort. When the manager's type and effort are unobservable and the firm faces an adverse selection problem, the manager receives utility from the contract in addition to his cost of effort. This extra utility is a rent he can extract because of his private information and its value is summarized in the following lemma:

Lemma 1 *Under an incentive-compatible contract, the information rent of a manager with ability $\tilde{\alpha} \in \mathcal{A}$ is given by*

$$R(\tilde{\alpha}) = \int_{\underline{\alpha}}^{\tilde{\alpha}} \mathrm{E}\left[\int_0^T (\pi_t - q_t)\phi_\alpha dt\right] dF(\alpha) + R(\underline{\alpha}) \tag{9}$$

where $\underline{\alpha} \equiv \inf \mathcal{A}$ represents the lowest type in the managerial pool $\mathcal{A}$, and ϕ_α represents the marginal effect of managerial ability α on the firm's project quality q_t.

The information rent derived in Lemma 1 allows us to impose the following assumption, which ensures to the validity of the Bergemann and Strack (2015) direct mechanism in our model:

Assumption 1 *The aggregator ϕ and the distribution of the initial information $F(\alpha)$ has the following property: for any given level of monitoring intensity m, $R'(\alpha) > 0$ for all α and $R(\alpha) > 0$ for some α.*

This assumption ensures that the information rent, calculated based (9), is strictly increasing in the manager's type α, and the participation constraint is slack for some managers with sufficiently high types.[7] This allows us to extend the technique in Bergemann and Strack

[7] Assumption 1 is not defined with model primitives mainly for the ease of exposition. It can be easily verified ex-post and is not restrictive. In Section 3, we present an example in which we assign ϕ and $F(\alpha)$ fairly general, exogenous functional forms, and verify that Assumption 1 is indeed satisfied. More technical assumptions for the sufficiency of the direct mechanism based on model primitives only can be found in Bergemann and Strack (2015).

(2015) and solve the firm's optimal contract, which is summarized as follows:

Proposition 1 *Under an incentive-compatible direct mechanism, $V(m)$, the firm's expected payoff for a given level of monitoring intensity, solves the following problem:*

$$V(m) = \max_{\pi_t} \int_{\mathcal{A}(m)} \mathrm{E}\left[\int_0^T \left(\pi_t - \frac{(\pi_t - q_t)^2}{2} - (\pi_t - q_t)\phi_\alpha(m)g(\alpha)\right) dt\right] - R(\underline{\alpha}) \tag{10}$$

The firm's optimal output target π_t^* is given by

$$\pi_t^* = 1 + [q_t - \phi_\alpha(m)g(\alpha)] \tag{11}$$

and

$$g(\alpha) \equiv \frac{1 - F(\alpha)}{f(\alpha)} \tag{12}$$

is the inverse hazard rate for the distribution of α.

Equation (10) is also known as the firm's *dynamic virtual surplus* in the literature. It equals the output minus the cost of effort and the manager's information rent. The optimal output target π_t^*, given by (11), is still increasing with project quality q_t. However, comparing (11) with the first-best level of output π^{FB} in (7), the adverse selection results in a distortion in the form of the information rent conceded to the manager (the $-\phi_\alpha g(\alpha)$ term). A larger ϕ_α implies a stronger impact the manager's ability has on the firm's project quality, and thus more information rent that the manager can extract. Meanwhile (11) implies that equilibrium effort under the optimal contract

$$e_t^* = \pi_t^* - q_t = 1 - \phi_\alpha(m)g(\alpha) < e^{FB}$$

is a constant lower than the first-best and depending on the manager's type.[8]

[8]Whether the distortion of effort is higher or lower for the high-ability managers depends on specific assumptions on ϕ and $F(\alpha)$ (which determines $g(\alpha)$). Also, although e_t^* is a constant, equilibrium wage under the optimal contract is time-varying. However, unlike the output target and the managerial rent – both of which can be exactly pinned down for a given set of parameters – the wage scheme that implements the optimal contract is not necessarily unique, a property resembling that of the security implementations of the optimal contracts for dynamic moral hazard problems (e.g. DeMarzo and Sannikov (2006), Biais,

Proposition 1 highlights the two forces through which monitoring intensity m affects the firm's value. First, ϕ (and thus ϕ_α) is a function of m. Therefore, monitoring intensity affects how project quality evolves over time, which in turn has an impact on the output target specified in the optimal contract (through Eq. 11). Secondly, as previously mentioned, monitoring intensity affects $\mathcal{A}$, the pool of managers that the firm can potentially be matched with due to the manager's participation constraint. These forces can be countervailing, implying an optimal level of monitoring intensity.

In sum, using the direct mechanism technique in Bergemann and Strack (2015), we convert a dynamic adverse selection problem due to unobservable managerial ability and project quality into a static mechanism problem. The solution thus does not require the usual dynamic programming techniques involving differential equations or keeping track of the manager's continuation utility. Like Bergemann and Strack (2015), the firm uses a contract that specifies different combinations of wage and output targets as a screening device to elicit the private information from the managers. Unlike Bergemann and Strack (2015), the firm also has the additional instrument of monitoring at its disposal. The intensity of monitoring affects the rent that a manager can expect to extract. Therefore, it offers another layer of screening for the firm, who can use different combinations of monitoring intensity and the associated managerial rents to attract managers with specific types. We illustrate this second layer of screening via an explicitly solved example in the next section.

3 Optimal Monitoring Intensity: An Example

To highlight the firm's optimal choice of monitoring intensity in a more transparent manner, in this section, we solve Proposition 1 explicitly by replacing the general but abstract Assumptions 1 with two specific assumptions using the model primitives.

3.1 The Assumptions

The first specific assumption pertains to the aggregator ϕ and the evolution of q_t:

Mariotti, Plantin, and Rochet (2007)). A standard pay-performance scheme involving a fixed salary plus a variable bonus depending on the output (thus, the reported quality of the project) is one of the common and intuitive implementations. See Gao and Wong (2017) for more detailed discussions in a comparable setting.

Assumption 2 *Project quality q_t follows a geometric Brownian motion process given by*

$$dq_t = q_t \left(\mu dt + \sigma_m dZ_t \right) \tag{13}$$

where

$$\sigma_m \equiv \sigma_1 + \frac{\sigma_2}{m} \tag{14}$$

where $\mu, \sigma_1, \sigma_2 \geq 0$ are public information. The manager's ability (type) α determines the initial value of the project quality, i.e., $\alpha = q_0$, and the monitoring intensity m takes a positive finite value (i.e. $0 < m < +\infty$).

Assumption 2 can be understood as an example of the general discussions made in Section 2.1. In particular, monitoring intensity can be modeled by the diffusion of project quality because, as argued in Section 2.1, a board that keeps the manager on a "short-leash" through frequent interventions can limit the manager's ability to make risky investments or undertake intensive R&D activities, thus limiting the time-series variations in the firm's project quality. Here, σ_1 represents the variations associated with firm operation, and σ_2 represents the variations associated with firm innovation. More frequent monitoring of the manager's innovative activities can reduce the latter but not the former.

Critically, we assume μ, the drift of project quality, is independent of the firm's monitoring intensity. This is to tease out the potential confounding effect of allowing the firm to move along a risk-return frontier through varying its monitoring intensity, and isolate the purpose of monitoring to be screening only.[9]

Assumption 2 implies a simple form of ϕ_α, which plays a critical role in the manager's information rent (9), the firm's dynamic virtual surplus (10), and the optimal output target

[9] As we demonstrate later, in this model the volatility of the project quality is a pure noise that benefits the manager only in the absence of the screening purpose of the contract. Holding everything else constant, managers receive higher information rent if the volatility of dq_t is higher. However, despite that monitoring can reduce such volatility, a firm can opt for a low level of monitoring due to the screening benefits it provides.

(11). Under this assumption, the aggregator ϕ is given by

$$\phi(t, Z_t, m, \alpha) = q_t = q_0 \exp\left[\left(\mu - \frac{1}{2}\sigma_m^2\right) t + \sigma_m Z_t\right] \tag{15}$$

and $\alpha = q_0$ implies

$$\phi_\alpha = \frac{q_t}{q_0} = \exp\left[\left(\mu - \frac{1}{2}\sigma_m^2\right) t + \sigma_m Z_t\right] \tag{16}$$

The next assumption pertains to $F(a_0)$, the distribution of the manager's ability:

Assumption 3 *The manager's ability α follows a generalized exponential distribution with a scale parameter $\lambda > 0$ and a shape parameter $k > 0$.*[10] *That is*

$$F(\alpha) = 1 - e^{-(\alpha/\lambda)^k} \tag{17}$$

This simplifies the inverse hazard rate $g(\alpha)$, which enters both the dynamic virtual surplus (10) and the optimal output target (11). Under this assumption,

$$g(\alpha) \equiv \frac{1 - F(\alpha)}{f(\alpha)} = \frac{\lambda^k}{k\alpha^{k-1}} \tag{18}$$

3.2 The Analysis

Combining these two assumptions, we can solve $R(a_0)$, the manager's information rent, and $\mathcal{A}(m)$, the potential managerial pool based on the manager's PC constraint. The results are summarized as follows:

Proposition 2 *Under Assumptions 2 and 3, manager's information rent (9) is given by:*

$$R(\alpha) = \left(\frac{e^{\mu T} - 1}{\mu}\right)(\alpha - \underline{\alpha}(m)) + \frac{\lambda^k}{2-k}\gamma(m)\left(\alpha^{2-k} - \underline{\alpha}^{2-k}(m)\right) \tag{19}$$

[10] A generalized exponential distribution is also known as a Weibull distribution, with the standard exponential distribution corresponding to a special case: $k = 1$. This choice is made purely due to its analytical tractability, as evident in equation (18). Other distributions (esp., the log-normal distribution) can generate qualitatively similar results as long as they have at least two degrees of freedom and an unbounded support.

for all $\alpha > \underline{\alpha}(m)$*, where*

$$\underline{a}(m) \equiv \inf \mathcal{A}(m) = \left[\frac{\mu\lambda^k}{e^{\mu T} - 1}\gamma(m)\right]^{\frac{1}{k-1}} \tag{20}$$

is the lowest type of manager in the set $\mathcal{A}(m)$ *(i.e.,* $R(\underline{\alpha}) = 0$*), and*

$$\gamma(m) \equiv \frac{e^{(2\mu+\sigma_m^2)T} - 1}{2\mu + \sigma_m^2} \tag{21}$$

is a decreasing function of the monitoring intensity m*.*

Proposition 2 explicitly illustrates the two forces through which the monitoring intensity m affects on the distribution of managerial rents. From equation (19), m affects the level and curvature $R(\alpha)$. in particular, how fast the managerial rent increases as the manager's type α increases. From equation (20), m affects $\underline{\alpha}(m)$, the lower bound of the pool of managers that the firm can match with. These two forces can be countervailing, as demonstrated below:

Corollary 2 *If* $k > 2$*, then for any given* $m_2 > m_1 > 0$*,* R *and* $\mathcal{A}$ *have the followings properties:*

i). $\underline{\alpha}(m_2) < \underline{\alpha}(m_1)$

ii). For any given $\alpha > \underline{\alpha}(m_1)$*,* $dR(\alpha; m_2)/d\alpha < dR(\alpha; m_1)/d\alpha$*.*

iii). There exists $\tilde{\alpha} > \underline{\alpha}(m_1)$ *such that* $R(\alpha; m_2) < R(\alpha; m_1)$ *for all* $\alpha > \tilde{\alpha}$*.*

Figure 1 illustrates the above properties graphically. These properties imply that the firm faces a tradeoff when choosing between two levels of monitoring intensity. The advantage of a higher level of monitoring intensity (m_2) is the less information rent that must be given to a high-type manager, both in the marginal term (Property ii above) and in the level term (Property iii above). The disadvantage is the expansion of the left tail of the distribution of the managers that the firm can potentially be matched with (Property i above). Put differently, firms with a higher degree of monitoring intensity can *ex-post* concede less information rent if they are matched with managers with high initial ability. However, they *ex-ante* reduce the firm's chance of matching with managers with high ability.[11]

[11]Proposition 2 and Figure 1 also help demonstrate how the observablity of α affects the optimal contract

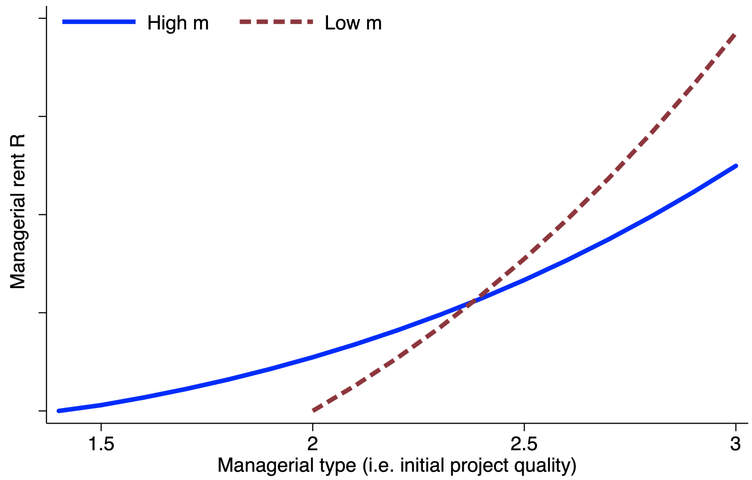

Figure 1: This figure plots the distribution of managerial rent R over managerial type α (i.e., q_0) for different levels of monitoring intensity. Parameter values: $T = \sigma_1 = \sigma_2 = \lambda = 1$, $k = 3$. Blue solid line: $m = 0.6$ (strong monitoring); red dashed line $m = 0.2$ (weak monitoring).

We can now combine Propositions 1 and 2 to derive the firm's optimal monitoring intensity: $m^* \equiv \arg\max_m V(m)$ where $V(m)$ is given by (10). Under Assumptions 2 and 3, $V(m)$ can be calculated in a closed albeit cumbersome form, which we leave to the Appendix. However, numerical solutions to $V(m)$ (thus m^*) are easy to obtain, and are plotted in Figure 2 for several sets of parameter values.

The left panel of Figure 2 shows that $V(m)$ is hump-shaped. This is a result of the tradeoff discussed following Proposition 2 and Corollary 2 above: on the one hand, stronger monitoring reduces the marginal rent a high-type manager can extract. One the other hand, weaker monitoring allows the firm to attract on average higher quality managers.

To visualize such tradeoff, the right panel of Figure 2 plots the expected net profit, defined as the expected output minus the effort cost (i.e., $\mathrm{E}[\pi_t^* - h(e_t)]$ averaged over $\mathcal{A}(m)$), and the expected managerial rent ($R(\alpha)$ averaged over $\mathcal{A}(m)$) for different levels of monitoring intensity. Both are increasing functions of m. However, the expected net profit is concave while the rent is convex, which explains the humped shape of S. Intuitively, if a firm sets the

this model. Suppose α is observable to the board, but q_t (or Z_t) is not. The resulting contract must still concede some rents to the managers because they still possess private information. However, exactly how the compensation and rents differ from when α is not observable to the board depends on how α affects the evolution of q_t. In the example used in this section, q_t evolves according to a geometric Brownian motion, and α is the initial project quality q_0. Consequently, q_t is always proportional to q_0, as in (15). If α (i.e. q_0) is observable, then the optimal contract can be scaled proportionally for different levels of α, and the managerial rent will be a linear function in α. In contrast, when α is not observable, managerial rent is convex in α, as shown in Figure 1, and is particularly large for managers with high ability.

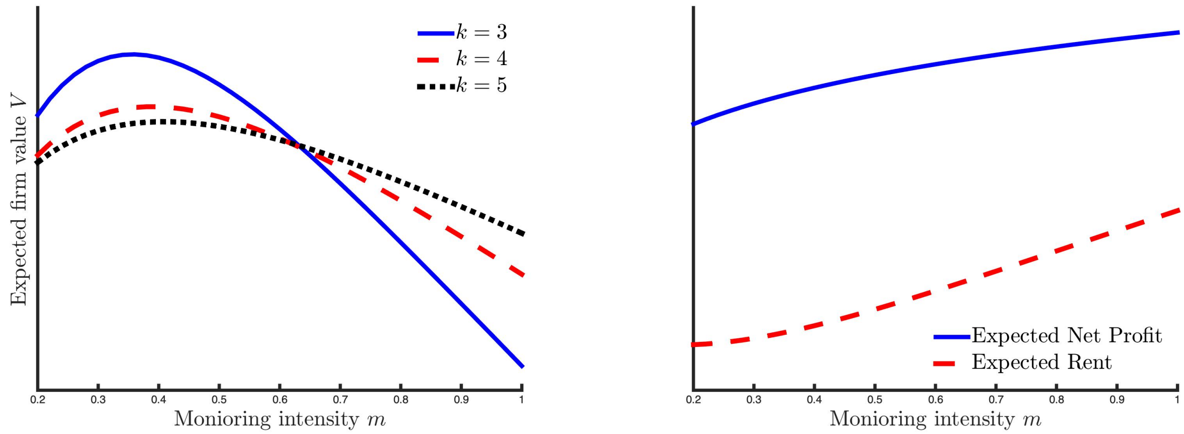

Figure 2: **Expected Firm Value $V(m)$ and the Cost and Benefit of Monitoring**
The left panel plots the expected firm value V as a function of m (monitoring intensity) and k (the shape parameter of the distribution of initial managerial type). The right panel plots the firm's expected net profit (left y-axis) and the expected managerial rent (right y-axis). Parameter values are $T = \lambda = \sigma_1 = \sigma_2 = 1$ for both panels and $k = 4$ for the right panel.

monitoring intensity too low, it will attract too many low-ability managers who produce on average low outputs. If a firm sets the monitoring intensity too high, the cost of compensating the high-ability managers outweighs the net profit of hiring them. Consequently, firm value is hump-shaped and a unique, interior optimal level of intensity m^* exists.

4 Empirical Evidence

The theoretical analysis in the previous sections can be summarized into testable hypotheses. First, strong monitoring reduces managerial rent after the manager has been hired, which implies the following hypothesis:

H1: *Monitoring intensity is negatively correlated with managerial compensation.*

Secondly, weaker monitoring facilitates the screening of managerial ability before the manager is hired, which implies the following hypothesis:

H2: *Monitoring intensity is negatively correlated with managerial ability.*

Finally, the tradeoff described in H1 and H2 implies the following:

H3: *The relationship between monitoring intensity and firm value is non-monotonic, and an optimal level of monitoring intensity that maximizes firm value exists.*

In this section, we examine the validity of these hypotheses based on empirical evidence from both the existing literature and data of the US public firms.

4.1 Data

Our main data source is the Institutional Shareholder Services (ISS), which provides individual director-level information – name, age, tenure, committee membership, primary employment, etc – for the universe of S&P 1500 companies at an annual frequency. We combine these data with CEO information from ExecuComp, and the firm's financial and market information from Compustat and CRSP. The final sample spans from 2007 to 2019, containing and 12,677 firm-year observations and 165,388 firm-year-directors observations.

We adopt three measures of corporate monitoring intensity. The first two are director *co-option* and *non-coopted-independence* as defined in Coles, Daniel, and Naveen (2014). Co-option is the fraction of the board comprised of directors appointed after the CEO assumed office; non-co-opted independence measures the fraction of directors who are independent and were appointed before the CEO. Coles, Daniel, and Naveen (2014) show that board monitoring decreases (increases) as co-option (non-co-opted independence) increases, and these two measures have more explanatory power for monitoring effectiveness than the conventional measure of board independence. Our last measure is *busyboard*. Following Fich and Shivdasani (2012), a director is busy if he or she sits on three or more boards, and board is busy if it consists primarily of busy directors. Fich and Shivdasani (2012) find that busy directors place limited attention on overseeing firms' operations and are deemed less effective in monitoring.

In our sample, about half of the board of directors are "co-opted" with the companies' CEOs, which is inline with the evidence in Coles, Daniel, and Naveen (2014). The non-co-opted independence measures is 36.1%, which is lower than traditional independence measure calculated using only financial ties (as in Cai, Xu, and Yang, 2020). On average, a director holds two directorships, 22% of directors in our sample are busy, and 8% of boards consist primarily of busy directors.

4.2 Monitoring Intensity and Managerial Rent

Our **H1** predicts that stronger internal monitoring reduces managerial information rent and lowers the cost of executive compensation. This prediction is shared among a large body of optimal contracting models and has been tested and confirmed repeatedly in the empirical literature. Early studies (Yermack, 1996; Angbazo and Narayanan, 1997; Borokhovich et al., 1997; Hallock, 1997; Core et al., 1999) show that the strength of board monitoring, as reflected by the size, director reputation, and director independence, helps to explain significant fraction of CEO compensation.

The cross-sectional heterogeneity in board size and independence has noticeably declined in recent year, making it difficult to uncover any connection by exploring the relation between such characteristics and managerial compensation.[12] Faced with this challenge, Fich and Shivdasani (2012) and Coles, Daniel, and Naveen (2014) construct new measures based on board members' employment overlapping with the CEO, and their outside committee appointments. Hwang and Kim (2009) and Cai, Xu, and Yang (2020) explore novel dataset containing board members' social ties and corporate charitable donations and find evidence that boards with stronger monitoring incentives are associated with lower managerial compensation. These findings are consistent with those in the prior literature.

In the sections to follow, we will skip the test of **H1**, as we believe the support of this result is already abundant in the literature. We will focus on testing the empirical predictions of **H2** and **H3**, which are specific to our model where internal monitoring can generate an ex-ante screening effect.

4.3 Monitoring Intensity and Managerial Quality

One novel implication of our model is that, weak monitoring by corporate board may have a positive effect on ex ante screening of managerial quality, as predicted by **H2**. In this section, we test this hypothesis by examining the relation between board monitoring and

[12] The boards of large U.S. companies have been overwhelmingly independent following the majority independence rules enacted by major exchanges in 2002.

managerial quality. We gauge managerial quality by firms' innovation activities. Innovation is crucial to a firm's long-run growth, and managerial ability has profound impacts on the success of firm innovation.

We measure a firm's innovation using its R&D expenditure, patent citations, and the market value of patents. We obtain R&D expenditure data from Compustat and the patent citation and market value of patent data following Kogan, Papanikolaou, Seru, and Stoffman (2017).[13] Specifically, we perform our test through the following regression:

$$
\begin{aligned}
Innovation_{i,j,t} = const + \alpha \times Monitoring\ Intensity + \beta \times \mathbf{CEO\ Char_{i,j,t}} \\
+ \gamma \times \mathbf{X_{i,j,t}} + \eta_j + \xi_t + \epsilon_{i,j,t},
\end{aligned}
\tag{22}
$$

where $\mathbf{CEO\ Char_{i,j,t}}$ is the vector of CEO characteristics such as CEO ownership and CEO duality for firm i in industry j and year t. $\mathbf{X_{i,j,t}}$ is the vector of firm characteristics including investment, ROA, leverage, firm size, market-to-book ratio, and sales growth. These control variables are defined in Table 1. We include in the regression the industry fixed effect η_j and the year fixed effect ξ_t. The main coefficient of interest is α.

Table 3 reports the results for R&D expenditure over different horizons. Both busy board and co-option measures are positively associated with a firm's future R&D expenditure in the [t+1, t+2] and [t+3, t+4] year window, while board independence measure is negatively associated with the future R&D investment. If a firm's R&D expenditure reflects its manager's ability in identifying valuable investment opportunities and his commitment to pursuing the innovation projects that spur the firm's long-term growth, our findings are consistent with hypothesis **H2** in the sense that weak monitoring (busier board, more co-opted board members, and low independence) is positively correlated with managerial ability.

We provide further evidence in Table 4 and 5 where we replace the dependent variable R&D expenditure with patent citations and the real market value of patents, and find similar results. These findings lend strong support to hypothesis **H2** because citations and market value of the patents are direct measures of value creation for the firm.

[13] We thank the authors for making their data available online.

4.4 Monitoring Intensity and Firm Value

Finally, we examine the relation between board's monitoring intensity and firm value, we run the following regression:

$$\begin{aligned} \textit{Market-to-book}_{i,j,t} = {} & \alpha_1 \times \textit{Monitoring Intensity} + \alpha_2 \times \textit{Monitoring Intensity}^2 \\ & + \beta \times \mathbf{CEO\ Char_{i,j,t}} + \gamma \times \mathbf{X_{i,j,t}} + \eta_j + \xi_t + \epsilon_{i,j,t}, \end{aligned} \tag{23}$$

We proxy a firm's value using market-to-book ratio. A higher market-to-book ratio indicates that the firm is able to generate higher cash flows out of the current asset in place and hence are deemed more valuable for its investors. At the meantime, high valuation can be achieved through not only the right level of monitoring intensity, but also from the firm's other real and financial decisions. To properly control for such influencing factors, we include a number of firm-level controls. We also include in the regression the board size, the percentage of the firm's common shares owned by executives (as the literature has documented a relationship between share ownership with firm value), and we control for the presence of interlocking directorships between outside directors and the CEO.

The results for the market-to-book regression is reported in Table 6. The marginal effect of monitoring intensity on firm value can be evaluated by $\alpha_1 + \alpha_2 \times \textit{Monitoring Intensity}$, where all of our monitoring intensity measures are bounded naturally with in 0 and 1 by construction. Our results confirms **H3**, which predictions a non-monotonic relationship between the monitoring intensity and the value of the firm. When the monitoring intensity is relatively low, the governance effect dominates, which suggests that in an environment when internal monitoring is relatively weak, an increase in the monitoring intensity can benefit shareholders by reducing excessive managerial rents; on the other hand, in an environment when monitoring is strong and managers are already tightly disciplined, a further increase in the monitoring intensity will make it harder for firms to attract good quality managers, which result in lower firm value. The two offsetting forces implies that there is an interior optimal degree of internal monitoring that lies in between 0 to 1. This optimal screening intensity will maximize the shareholders' value by which balancing the ex-ante screening versus the ex-post governance.

5 Conclusion

There is broad literature on how strong internal monitoring can enhance firm value, especially when firms face the moral hazard problem arising from unobservable managerial effort. In this paper, we provide a complementary view that highlights a novel, positive role of weak monitoring: it facilitates the ex-ante screening of better managers and helps mitigate the adverse selection problem arising from unobservable managerial ability. Using a dynamic agency model with persistent private information, we demonstrate that monitoring has two countervailing effects on firm value. First, strong monitoring reduces the marginal rent that high-ability managers can extract from their compensation contract, which improves firm value once on the manager is employed (i.e., an ex-post disciplining effect). Meanwhile, weak monitoring helps the firm to attract managers with on average higher ability, improving the expected firm value before a manager is hired (i.e., an ex-ante screening effect). The overall impact of monitoring on firm value is determined by the tradeoff between the two effects. Therefore, even in the absence of any explicit monitoring cost, it is not optimal to maximize the monitoring intensity, and the relationship between a firm's value and its monitoring intensity is hump-shaped.

We underpin the model predictions on the screening effect and the overall relationship between firm value and monitoring intensity with empirical evidence from the U.S. public firms. Based on three proxies for monitoring, including director co-option, non-coopted-independence, and board busyness, we confirm that weak monitoring improves the quality of managers being employed, evident by a strong, negative correlation between monitoring intensity and a firm's innovation activities that reflect managerial ability. We further confirm a hump-shaped relation between a firm's market-to-book ratio and its monitoring intensity, lending strong support to the model's predictions on the tradeoff between two opposing effects of monitoring on firm value.

Our study can be expanded in several directions, the most straightforward being perhaps on the methodological side: based on the direct mechanism technique of Bergemann and Strack (2015), we are able to solve a generally difficult dynamic adverse selection problem by converting it into a static mechanism design problem. Moreover, our model allows an extra dimension of control (i.e. the monitoring intensity) than Bergemann and Strack (2015), potentially broadening the applicability of their technique to a wider range of topics in corporate finance that involve agency frictions and private information, such as firm investment (DeMarzo, Fishman, He, and Wang, 2012), liquidity management (Bolton, Chen, and Wang, 2011), resource allocation (Feng and Westerfield, 2020), and risk management (Biais, Mariotti, Rochet, and Villeneuve, 2010). We leave these topics for future studies.

Appendix

A Proofs and Calculations

Proof of Lemma 1

The proof follows Bergemann and Strack (2015) Proposition 1 or Gao and Wong (2017) Proposition 2. Let $\hat{\alpha}$ be the report of the initial information made by the manager with an arbitrary type α. Based on this report, the firm imposes output target $\pi_t^{\mathcal{C}(\hat{\alpha})}$ and recommended effort is $\hat{e}_t \equiv \pi_t^{\mathcal{C}(\hat{\alpha})} - \hat{q}_t$, where $\hat{q}_t$ is the manager's reported project quality. Given the constraint (8), we have

$$\hat{e}_t = e_t + (\hat{q}_t - q_t) \tag{24}$$

Therefore, the payoff from this misreporting is

$$R(\alpha; \hat{\alpha}) = \mathrm{E}^{\alpha}\left[\int_0^T \left(w(\hat{\alpha}, \hat{q}_t) - h(\hat{e}_t)\right) dt\right] \tag{25}$$

Differentiating with respect to α yields

$$\begin{aligned} \frac{\partial}{\partial \alpha} R(\alpha; \hat{\alpha}) &= \mathrm{E}^{\alpha}\left[\int_0^T \left(-\frac{\partial}{\partial \alpha} h(\hat{e}_t)\right) dt\right] && (26)\\ &= \mathrm{E}^{\alpha}\left[\int_0^T \phi_\alpha \hat{e}_t dt\right] && (27) \end{aligned}$$

where the second line utilizes (24) and the fact that $q_t = \phi$ from (3). Letting $\hat{\alpha} = \alpha$ and substituting e_t with $\pi_t - q_t$ implies

$$\frac{\partial}{\partial \alpha} R(\alpha; \hat{\alpha}) = \mathrm{E}^{\alpha}\left[\int_0^T \phi_\alpha \left(\pi_t - q_t\right) dt\right] \tag{28}$$

which is also known as the *dynamic envelop condition*. Integrating (28) up to $\tilde{\alpha}$ yields the information rent (9) for any given type-$\tilde{\alpha}$ manager in the managerial pool $\mathcal{A}$. □

Proof of Proposition 1

The proof follows Bergemann and Strack (2015) Theorem 1 or Gao and Wong (2017) Proposition 3. The definitions of R (Eq. 4) imply that

$$\mathrm{E}\left[\int_0^T w_t dt\right] = R(\alpha) + \mathrm{E}\left[\int_0^T h(e_t) dt\right] \tag{29}$$

Substituting this into the definition of $V(m)$ (Eq. 5) yields

$$V(m) = \max_{\pi_t} \int_{\mathcal{A}(m)} \mathrm{E}\left[\int_0^T \left(\pi_t - \frac{(\pi_t - q_t)^2}{2}\right) dt\right] - \int_{\mathcal{A}(m)} R(\alpha) \tag{30}$$

Applying integration by parts and the fundamental theorem of calculus to the last term yields

$$\int_{\mathcal{A}(m)} R(\alpha) = \int_{\mathcal{A}(m)} R'(\alpha) \left(\frac{1 - F(\alpha)}{f(\alpha)}\right) + R(\underline{\alpha}) \tag{31}$$

Replacing $R'(\alpha)$ with (28) (for $\hat{\alpha} = \alpha$), and substituting back to (30) yields the dynamic virtual surplus (10). Finally, point-wise maximization of (11) with respect to π_t yields the optimal output target π_t^* in (11). □

Proof of Proposition 2

Given (16), $q_t = \alpha\phi_\alpha$. Therefore,

$$\mathrm{E}\left[\int_0^T (\pi_t - q_t)\, \phi_\alpha dt\right] = \mathrm{E}\left[\int_0^T (\pi_t - \alpha\phi_\alpha)\, \phi_\alpha dt\right] \tag{32}$$

The optimal output target (11) implies $\pi_t^* - q_t = 1 - \phi_\alpha g(\alpha)$, thus

$$\mathrm{E}\left[\int_0^T (\pi_t - q_t)\, \phi_\alpha dt\right] = \mathrm{E}\left[\int_0^T (1 - \phi_\alpha g(\alpha))\, \phi_\alpha dt\right] = \eta - \gamma g(\alpha) \tag{33}$$

where $\eta \equiv \mathrm{E}\left[\int_0^T \phi_\alpha dt\right]$ and $\gamma \equiv \mathrm{E}\left[\int_0^T \phi_\alpha^2 dt\right]$, because ϕ_α is *not* a function of α according to (16), and $g(\alpha)$ is *not* a function of t. Substituting $g(\alpha)$ with (18), we have

$$R(\alpha) = \int_{\underline{\alpha}}^{\alpha} [\eta - \gamma g(a)]\, dF(a) + R(\underline{\alpha}) = \eta(\alpha - \underline{\alpha}(m)) + \frac{\lambda^k}{k-2}\gamma \left(\alpha^{2-k} - \underline{\alpha}^{2-k}(m)\right) \tag{34}$$

Moreover, $R'(\alpha) > 0$ implies that

$$\eta - \lambda^k \gamma \alpha^{1-k} > 0 \tag{35}$$

Therefore $\underline{\alpha}(m) = (\gamma\lambda^k/\eta)^{\frac{1}{k-1}}$. □

Calculation of $V(m)$

Under Assumptions 2 and 3, $V(m)$ can be written as

$$V(m) = \int_{\underline{\alpha}}^{+\infty} \left\{\mathrm{E}\left[\int_0^T \left(\pi_t^* - \frac{(\pi_t^* - q_t)^2}{2}\right) dt\right] - R(\alpha)\right\} dF(\alpha) \tag{36}$$

where $R(\alpha)$ is given by (19), and $\underline{\alpha}$ given by (20). $V(m)$ can be calculated with the following steps: first, combining (11), (15), and (18) yields

$$\pi_t^* - \frac{(\pi_t^* - q_t)^2}{2} = \frac{1}{2} + q_t - \frac{\lambda^{2k} q_t^2}{k^2 \alpha^{2k}} \tag{37}$$

Then,

$$\begin{aligned} &\mathrm{E}\left[\int_0^T \left(\pi_t^* - \frac{(\pi_t^* - q_t)^2}{2}\right) dt\right] - R(\alpha) \\ &= \frac{T}{2} + \left(\eta\underline{\alpha} + \frac{\gamma\lambda^k}{k-2}\underline{\alpha}^{2-k}\right) - \gamma\lambda^k \left[\left(\frac{\lambda^k \alpha^{2-2k}}{k^2}\right) + \left(\frac{\alpha^{2-k}}{k-2}\right)\right] \end{aligned} \tag{38}$$

which implies that

$$V(m) = \frac{T}{2} + \left(\eta\underline{\alpha} + \frac{\gamma\lambda^k}{k-2}\underline{\alpha}^{2-k}\right) \tag{39}$$

$$- \gamma\lambda^k \left[\left(\frac{\lambda^{k-1}}{k^2}\right)\Gamma\left(\underline{\alpha}^k, \frac{1}{k}\right) + \frac{1}{\lambda(k-2)}\Gamma\left(\underline{\alpha}^k, \frac{k-1}{k}\right)\right] \tag{40}$$

where $\Gamma(x_1, x_2)$ is the CDF of a gamma distribution with shape x_1 and scale x_2. While (39) cannot be further simplified, the numerical value of a Γ function, and thus $V(m)$, can be easily computed and plotted as shown in Figure 2.

B Tables

Table 1: Variable Definition

This table presents the variable definition.

Variable	Definition
Co-opted Director	=1 if a director joins the board after the CEO assumes office as defined in Coles, Daniel, and Naveen (2014)
Co-option	Number of co-opted directors / Board Size
Tenure-Weighted Co-option	Sum of co-opted directors' tenures divided by the sum of all directors' tenures
Non co-opted Independence	Percentage of directors who are non co-opted and independent as defined in Coles, Daniel, and Naveen (2014)
Busy (outside) Director	=1 if an (outside) director serves on three or more boards as defined in Fich and Shivdasani (2012)
Percentage of Busy Directors	Percentage of busy directors on a board
Board Size	Total number of directors on the board
Board Interlock	=1 if the CEO sits on the board of the outside director
CEO Ownership	Shares held by the CEO / Number of shares outstanding
CEO Tenure	Number of years since an CEO takes office
CEO Duality	= 1 if CEO is also the Chairman and 0 otherwise
Investment	Capital Expenditure / Assets
ROA	Return on Assets = EBITDA/Assets
Leverage	Total Debt / Assets
Firm Size	Log(sales)
Market-to-Book	(Assets – Book equity +Market equity) / Assets
Sales Growth	Sales / Lagged Sales
R&D [1yr, 2yr]	R&D expenditure / Assets in [t+1, t+2] window
R&D [3yr, 4yr]	R&D expenditure / Assets in [t+3, t+4] window
Cites [1yr, 2yr]	Ln(1+citations) in [t+1, t+2], citation is defined as in Kogan et al. (2017)
Cites [3yr, 4yr]	Ln(1+citations) in [t+3, t+4], citation is defined as in Kogan et al. (2017)
Market value [1yr, 2yr]	Ln(1+rv) in [t+1, t+2], rv is the real market value of the patent as in Kogan et al. (2017)
Market value [3yr, 4yr]	Ln(1+rv) in [t+3, t+4], rv is the real market value of the patent as in Kogan et al. (2017)

Table 2: Summary Statistics

In this table, we report summary statistics for our sample. The sample covers S&P 1500 firms over the 2007–2019 period and combines data from the Institutional Shareholder Services (ISS), ExecuComp, Compustat and CRSP. The variables are defined in Table 1. All variables are at an annual frequency. The sample contains 12,204 firm-year observations.

Variable	Mean	std.	P25	P50	P75
Co-option	0.511	0.319	0.250	0.500	0.778
Tenure-Weighted Co-option	0.354	0.344	0.066	0.222	0.574
Non co-opted Independence	0.361	0.291	0.000	0.375	0.600
Percentage of Busy Directors	0.216	0.171	0.100	0.200	0.333
Percentage of Busy Outside Directors	0.228	0.177	0.100	0.200	0.333
Board Size	9.108	2.031	8	9	10
Board Interlock	0.267	0.443	0	0	1
CEO Ownership	2.315	4.844	0.208	0.665	1.896
CEO Tenure	8.519	7.478	3	6	12
CEO Duality	0.590	0.695	0	0	1
Investment	0.043	0.042	0.016	0.030	0.054
ROA	0.053	0.085	0.025	0.057	0.094
Leverage	0.233	0.184	0.078	0.221	0.344
Firm Size	7.798	1.519	6.737	7.690	8.760
Market-to-Book	1.756	1.215	0.962	1.394	2.119
Sales Growth	1.068	0.182	0.986	1.057	1.135
R&D [1yr, 2yr]	0.043	0.090	0.000	0.000	0.044
R&D [3yr, 4yr]	0.042	0.088	0.000	0.000	0.043
Cites [1yr, 2yr]	3.153	2.375	1.099	2.944	4.942
Cites [3yr, 4yr]	2.789	2.289	0.693	2.485	4.471
Market value [1yr, 2yr]	5.907	2.253	4.152	5.831	7.673
Market value [3yr, 4yr]	6.035	2.249	4.351	5.981	7.801

Table 3: Monitoring Intensity and R&D Investments

In this table, we report the estimates of the relation between firms' monitoring intensities and their R&D investments in the following 2- and 4-year horizons. We adopt three measures for board's monitoring intensity: co-option, non-co-opted independence, and percentage of busy directors. These measures are defined in Table 1. Our sample period is from 2007 to 2019. Standard errors for the estimated parameters are reported in parenthesis and are clustered by industry and year.

	R&D [1yr, 2yr]			R&D [3yr, 4yr]		
	(1) Busy Board	(2) Co-option	(3) Independence	(4) Busy Board	(5) Co-option	(6) Independence
Monitoring intensity	-0.042***	-0.012***	-0.009***	-0.043***	-0.012***	-0.007***
	(0.007)	(0.002)	(0.002)	(0.008)	(0.002)	(0.002)
Log board size	-0.009***	-0.007**	-0.008**	-0.007*	-0.005	-0.006
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
CEO ownership	0.000	0.000	0.000	0.000*	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CEO duality	-0.006***	-0.007***	-0.006***	-0.007***	-0.008***	-0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Interlock	0.005*	0.008***	0.007***	0.003	0.006*	0.005
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Firn-level controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	10999	10999	10999	7914	7914	7914
adj. R-sq	0.389	0.386	0.385	0.391	0.388	0.387

Table 4: Monitoring Intensity and Patent Citations

In this table, we report the estimates of the relation between firms' monitoring intensities and their log patent citations in the following 2- and 4-year horizons. We adopt three measures for board's monitoring intensity: co-option, non-co-opted independence, and percentage of busy directors. These measures are defined in Table 1. Our sample period is from 2007 to 2019. Standard errors for the estimated parameters are reported in parenthesis and are clustered by industry and year.

	Cites [1yr, 2yr]			Cites [3yr, 4yr]		
	(1) Busy Board	(2) Co-option	(3) Independence	(4) Busy Board	(5) Co-option	(6) Independence
Monitoring intensity	-0.746***	-0.425***	-0.422***	-0.748***	-0.465***	-0.445***
	(0.187)	(0.077)	(0.082)	(0.206)	(0.100)	(0.123)
Log board size	-0.088	0.013	-0.012	-0.205	-0.117	-0.147
	(0.164)	(0.163)	(0.163)	(0.189)	(0.184)	(0.185)
CEO ownership	0.024***	0.013**	0.014**	0.023***	0.012*	0.013**
	(0.006)	(0.005)	(0.006)	(0.007)	(0.006)	(0.007)
CEO duality	-0.045	-0.077*	-0.063	0.018	-0.018	-0.002
	(0.043)	(0.044)	(0.043)	(0.058)	(0.058)	(0.058)
Interlock	-0.035	-0.002	-0.013	0.093	0.126	0.111
	(0.076)	(0.078)	(0.077)	(0.135)	(0.134)	(0.136)
Firn-level controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	3286	3286	3286	2345	2345	2345
adj. R-sq	0.575	0.576	0.575	0.541	0.542	0.541

Table 5: Monitoring Intensity and Patent Revenue

In this table, we report the estimates of the relation between firms' monitoring intensities and their log revenue from patents in the following 2- and 4-year horizons. We adopt three measures for board's monitoring intensity: co-option, non-co-opted independence, and percentage of busy directors. These measures are defined in Table 1. Our sample period is from 2007 to 2019. Standard errors for the estimated parameters are reported in parenthesis and are clustered by industry and year.

	Market value [1yr, 2yr]			Market value [3yr, 4yr]		
	(1) Busy Board	(2) Co-option	(3) Independence	(4) Busy Board	(5) Co-option	(6) Independence
Monitoring intensity	-0.847***	-0.522***	-0.469***	-0.759***	-0.640***	-0.534***
	(0.162)	(0.087)	(0.082)	(0.184)	(0.116)	(0.112)
Log board size	-0.005	0.115	0.081	-0.034	0.078	0.032
	(0.159)	(0.153)	(0.154)	(0.201)	(0.194)	(0.198)
CEO ownership	-0.009*	-0.022***	-0.020***	-0.007	-0.022***	-0.019***
	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)
CEO duality	-0.040	-0.080**	-0.060	-0.038	-0.092*	-0.063
	(0.038)	(0.040)	(0.039)	(0.045)	(0.048)	(0.046)
Interlock	0.007	0.048	0.032	0.080	0.126	0.102
	(0.083)	(0.084)	(0.083)	(0.124)	(0.124)	(0.124)
Firn-level controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	3286	3286	3286	2345	2345	2345
adj. R-sq	0.689	0.691	0.689	0.654	0.659	0.656

Table 6: Monitoring Intensity and Firm Value

In this table, we report the estimates of the relation between firms' monitoring intensities and their market-to-book ratio. We adopt three measures for board's monitoring intensity: co-option, non-co-opted independence, and percentage of busy directors. These measures are defined in Table 1. Our sample period is from 2007 to 2019. Standard errors for the estimated parameters are reported in parenthesis and are clustered by industry and year.

	(1) Busy Board	(2) Co-option	(3) Independence
Monitoring intensity	0.561***	0.086	0.276**
	(0.219)	(0.114)	(0.121)
Monitoring intensity^2	-0.578**	-0.215**	-0.549***
	(0.285)	(0.107)	(0.146)
Log board size	0.011	0.056	0.039
	(0.045)	(0.046)	(0.046)
CEO ownership	0.003	0.001	0.001
	(0.002)	(0.002)	(0.002)
CEO duality	0.045***	0.028*	0.038**
	(0.015)	(0.016)	(0.015)
Interlock	-0.005	0.018	0.012
	(0.031)	(0.031)	(0.031)
Firm-level controls	YES	YES	YES
Year FE	YES	YES	YES
Industry FE	YES	YES	YES
N	10989	10989	10989
adj. R-sq	0.365	0.364	0.365

References

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

Angbazo, L. and R. Narayanan (1997). Top management compensation and the structure of the board of directors in commercial banks. *Review of Finance 1*(2), 239–259.

Baldenius, T., N. Melumad, and X. Meng (2014). Board composition and CEO power. *Journal of Financial Economics 112*(1), 53–68.

Battaglini, M. (2005). Long-term contracting with markovian consumers. *American Economic Review 95*(3), 637–658.

Bergemann, D. and P. Strack (2015). Dynamic revenue maximization: A continuous time approach. *Journal of Economic Theory 159*, 819–853.

Bergemann, D. and J. Välimäki (2019). Dynamic mechanism design: An introduction. *Journal of Economic Literature 57*(2), 235–74.

Biais, B., T. Mariotti, G. Plantin, and J.-C. Rochet (2007). Dynamic security design: Convergence to continuous time and asset pricing implications. *Review of Economic Studies 74*, 345–390.

Biais, B., T. Mariotti, J.-C. Rochet, and S. Villeneuve (2010). Large risks, limited liability, and dynamic moral hazard. *Econometrica 78*, 73–118.

Boivie, S., M. Bednar, and J. Andrus (2016). Boards aren't the right way to monitor companies by. *Harvard Business Review*. March 10.

Bolton, P., H. Chen, and N. Wang (2011). A unified theory of tobin's q, corporate investment, financing, and risk management. *Journal of Finance 66*(5), 1545–1578.

Borokhovich, K. A., K. R. Brunarski, and R. Parrino (1997). CEO contracting and antitakeover amendments. *Journal of Finance 52*(4), 1495–1517.

Burns, N., S. Kedia, and M. Lipson (2010). Institutional ownership and monitoring: Evidence from financial misreporting. *Journal of Corporate Finance 16*(4), 443–455.

Cai, Y., J. Xu, and J. Yang (2020). Paying by donating: Corporate donations affiliated with independent directors. *Review of Financial Studies*.

Chen, M., P. Sun, and Y. Xiao (2020). Optimal monitoring schedule in dynamic contracts. *Operations Research 68*, 1285–1624.

Chen, Y., E. J. Podolski, and M. Veeraraghavan (2015). Does managerial ability facilitate corporate innovative success? *Journal of Empirical Finance 34*, 313–326.

Coles, J. L., N. D. Daniel, and L. Naveen (2014). Co-opted boards. *Review of Financial Studies 27*(6), 1751–1796.

Core, J. E., R. W. Holthausen, and D. F. Larcker (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics 51*(3), 371–406.

Custódio, C., M. A. Ferreira, and P. Matos (2019). Do general managerial skills spur innovation? *Management Science 65*(2), 459–476.

DeMarzo, P., M. Fishman, Z. He, and N. Wang (2012). Dynamic agency and the q theory of investment. *Journal of Finance 67*.

DeMarzo, P. and Y. Sannikov (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance 61*, 2681–2724.

Eső, P. and B. Szentes (2007). Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies 74*(3), 705–731.

Feeley, J. and M. Bathon (2015). Boards aren't the right way to monitor companies by. *Bloomberg Business*. February 3.

Feng, F. Z. (2020). Financing a black box: Dynamic investment with persistent private information. Working paper. University of Washington.

Feng, F. Z. and M. M. Westerfield (2020). Dynamic resource allocation with hidden volatility. *Journal of Financial Economics*. forthcoming.

Fernandes, A. and C. J. Phelan (2000). A recursive formulation for repeated agency with history dependence. *Journal of Economic Theory 91*(2), 223–247.

Fich, E. M. and A. Shivdasani (2012). *Are busy boards effective monitors?* Springer.

Gao, B. and T.-Y. Wong (2017). Long-term capital budgeting and incentive mechanism. Working paper. Boston University and SUFE.

Garrett, D. F. (2017). Dynamic mechanism design: Dynamic arrivals and changing values. *Games and Economic Behavior 104*, 595–612.

Gershkov, A., B. Moldovanu, and P. Strack (2018). Revenue-maximizing mechanisms with strategic customers and unknown, markovian demand. *Management Science 64*(5), 2031–2046.

Hallock, K. F. (1997). Reciprocally interlocking boards of directors and executive compensation. *Journal of Financial and Quantitative Analysis*, 331–344.

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

Hermalin, B. E. (2005). Trends in corporate governance. *Journal of Finance 60*(5), 2351–2384.

Hwang, B.-H. and S. Kim (2009). It pays to have friends. *Journal of Financial Economics 93*(1), 138–158.

Kapička, M. (2013). Efficient allocations in dynamic private information economies with persistent shocks: A first-order approach. *Review of Economic Studies 80*(3), 1027–1054.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics 132*(2), 665–712.

Krasikov, I. and R. Lamba (2019). On dynamic pricing. Working paper. Tel Aviv University and Penn State University.

Marinovic, I. and F. Varas (2019). CEO horizon, optimal pay duration, and the escalation of short-termism. *Journal of Finance 74*(4), 2011–2053.

Mehran, H. (1992). Executive incentive plans, corporate control, and capital structure. *Journal of Financial and Quantitative analysis*, 539–560.

Orlov, D. (2020). Frequent monitoring in dynamic contracts. *Journal of Economic Theory*. Forthcoming.

Pavan, A., I. Segal, and J. Toikka (2014). Dynamic mechanism design: A myersonian approach. *Econometrica 82*(2), 601–653.

Piskorski, T. and M. M. Westerfield (2016). Optimal dynamic contracts with moral hazard and costly monitoring. *Journal of Economic Theory 166*, 242–281.

Raheja, C. G. (2005). Determinants of board size and composition: A theory of corporate boards. *Journal of Financial and Quantitative Analysis*, 283–306.

Sannikov, Y. (2007). Agency problems, screening and increasing credit lines. Working paper. Stanford University.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *Review of Economic Studies 75*, 957–984.

Tchistyi, A. (2016). Security design with correlated hidden cash flows: The optimality of performance pricing. Working paper. University of Illinois at Urbana-Champaign.

Williams, N. (2011). Persistent private information. *Econometrica 79*(4), 1233–1275.

Williams, N. (2015). A solvable continuous-time dynamic principal–agent model. *Journal of Economic Theory 159*, 989–1015.

Yermack, D. (1996). Higher market valuation of companies with a small board of directors. *Journal of Financial Economics 40*(2), 185–211.

Zhang, Y. (2009). Dynamic contracting with persistent shocks. *Journal of Economic Theory 144*(2), 635–675.

Zhu, J. Y. (2020). Better monitoring...worse productivity? Working paper. University of Kansas.