DISCUSSION PAPER SERIES

DP15625

THE COMPLEMENTARITY BETWEEN SIGNAL INFORMATIVENESS AND MONITORING

Pierre Chaigneau and Nicolas Sahuguet

FINANCIAL ECONOMICS

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

THE COMPLEMENTARITY BETWEEN SIGNAL INFORMATIVENESS AND MONITORING

Pierre Chaigneau and Nicolas Sahuguet

Discussion Paper DP15625
Published 02 January 2021
Submitted 29 December 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

THE COMPLEMENTARITY BETWEEN SIGNAL INFORMATIVENESS AND MONITORING

Abstract

When assessing managerial ability, a firm can rely on two sources of information: a signal of firm value, such as earnings, and monitoring. We show that a more informative signal can surprisingly increase the value of monitoring. This happens if a more informative signal makes some signal realizations more ambiguous indicators of managerial ability, or if the signal leads to negative belief updating on managerial ability yet does not trigger termination. Then, termination decisions will paradoxically rely less on the signal when it is more informative. In private equity owned firms, the model predicts that monitoring intensity is increasing in signal informativeness conditional on a bad performance. These firms can fall into a "bad governance trap" such that a less informative signal is compounded by worse monitoring upon a bad performance.

JEL Classification: N/A

Keywords: board monitoring, corporate governance system, governance complementarity, hard and soft information

Pierre Chaigneau - pierre.chaigneau@queensu.ca
Queen's University

Nicolas Sahuguet - nicolas.sahuguet@hec.ca
HEC Montreal and CEPR

Acknowledgements
*We thank Jeremy Bertomeu, Matthieu Bouvard, Jean-Etienne de Bettignies, Adolfo de Motta, Luke Taylor, Mike Waldman, and seminar participants at Erasmus University Rotterdam, McGill, and the Accounting and Economics Society webinar, for useful comments and suggestions.

The Complementarity between Signal Informativeness and Monitoring*

Pierre Chaigneau† Queen's University

Nicolas Sahuguet‡ HEC Montréal

December 2020

Abstract

When assessing managerial ability, a firm can rely on two sources of information: a signal of firm value, such as earnings, and monitoring. We show that a more informative signal can surprisingly increase the value of monitoring. This happens if a more informative signal makes some signal realizations more ambiguous indicators of managerial ability, or if the signal leads to negative belief updating on managerial ability yet does not trigger termination. Then, termination decisions will paradoxically rely less on the signal when it is more informative. In private equity owned firms, the model predicts that monitoring intensity is increasing in signal informativeness conditional on a bad performance. These firms can fall into a "bad governance trap" such that a less informative signal is compounded by worse monitoring upon a bad performance.

Keywords: corporate governance system, governance complementarity, hard and soft information, informativeness, monitoring.

*We thank Jeremy Bertomeu, Matthieu Bouvard, Jean-Étienne de Bettignies, Adolfo de Motta, Luke Taylor, Mike Waldman, and seminar participants at Erasmus University Rotterdam, McGill, and the Accounting and Economics Society webinar, for useful comments and suggestions.

†Address: Smith School of Business, Goodes Hall, 143 Union Street West, K7L 2P3, Kingston, ON, Canada. Email: pierre.chaigneau@queensu.ca. Tel: 613 533 2312.

‡Address: Applied Economics Department, HEC Montréal, 3000 chemin de la côte Sainte Catherine, H3T 2A7 Montréal, QC, Canada. Email: nicolas.sahuguet@hec.ca. Tel: 514 340 6031.

Assessing the performance of a manager, and using this information to continue or terminate the manager, is a key aspect of corporate governance (Laux (2014), Hermalin and Weisbach (2017)). To this end, investors can rely on accounting measures (e.g. earnings), or market measures (e.g. the stock price). In addition, investors or their agent (the board) can invest resources to acquire additional information, including soft information, about firm performance and management quality. Tirole (2006) refers to the first mechanism as "passive monitoring", and to the second as "active monitoring". Investors can then use this information to intervene, most notably by terminating and replacing the firm's manager.

An important question in the design of corporate governance systems is whether these two types of monitoring are complements or substitutes. According to Tirole (2006), this question "is central to the design of the financial system (...) and yet it has not been investigated in detail in the literature." Leuz and Wysocki (2016) point out that changes in financial reporting standards will likely affect other dimensions of the corporate governance environment, and conclude that "we know relatively little about the nature and importance of such institutional complementarities." Armstrong, Guay, and Weber (2010) likewise argue that "we know relatively little, however, about how firms select among the various mechanisms available to them, or how the various mechanisms interact and serve as complements to and/or substitutes for each other."

This question also has implications for the adequacy of financial reporting standards and corporate governance. When assessing the overall effect of an improvement in the quality of a signal of firm value, it is important to take into consideration the indirect effect of this improvement on the intensity of active monitoring (henceforth "monitoring"), which also provides information about the firm's management. If these sources of information are substitutes, then a lower quality signal of firm value is not highly detrimental for the termination decision because it will result in enhanced monitoring, i.e., the endogenous increase in monitoring counteracts the decrease in signal quality. If these sources of information are complements, however, then a lower quality signal of firm value will lead to less monitoring, i.e., the endogenous decrease in monitoring compounds the information loss due to the decrease in signal quality.

To study this question, we consider a stylized model of corporate governance. Our model can be viewed as an extension of Hermalin (2005) with a signal of firm value. The signal can be interpreted as earnings or the stock price. The firm's manager, whose ability affects the signal distribution, can be terminated or reappointed. Even though the ability of potential managers to manage the firm is unknown, the firm can learn about the ability of the incumbent manager via two channels. First, the firm observes the signal s, which is (imperfectly) informative about managerial ability in a monotone likelihood ratio sense. Second, after observing this signal, the board must decide on the intensity of its monitoring, which may provide additional information about the ability of its manager. Based on all this information, the board has the option to terminate the manager – termination and transition to a new manager can involve a turnover cost. The continuation/termination decision will rely on the signal, which is always observable, and on the information generated by monitoring, if available. As a result, monitoring is useful if and only if it changes the decision that would have been made on the basis of the signal alone. Thus, the value of monitoring is higher if it is more likely that the decision made on the basis of the signal is wrong. We say that the precision of the signal and monitoring are "complements" when an increase in the former increases the latter, and "substitutes" otherwise.

We start by studying a simple model with a symmetric signal distribution which measures managerial ability, albeit imperfectly, and no turnover cost. With these assumptions, which are similar to those used in Hermalin (2005), we establish that an increase in the precision of the signal always reduces the value and therefore the intensity of monitoring. We then relax these assumptions one-by-one to identify sources of complementarity between the quality of the signal and the intensity of monitoring.

First, we assume a positive turnover cost. Absent monitoring, there are thresholds s^T and s^* such that the manager is terminated (continued) for a signal $s < s^T$ ($s \geq s^T$), and the board negatively (positively) updates its beliefs about managerial ability for $s < s^*$ ($s > s^*$). With no turnover cost, the manager is simply dismissed if and only if the signal is "bad news", as in Hermalin (2005), i.e., $s^T = s^*$. With a turnover cost, $s^T < s^*$: the bad news need to be sufficiently bad to trigger termination. There is then a subset of signals (s^T, s^*) of bad news that do not trigger termination. With a more precise signal, a signal $s \in (s^T, s^*)$ is worse news about managerial ability, which raises the value of monitoring. Indeed, more intense monitoring generates additional information on managerial ability, and allows to terminate a low ability manager who would otherwise have remained in place. In short, with a positive turnover cost, there is complementarity between signal precision and monitoring if the signal leads to negative belief updating on managerial ability yet does not trigger termination.

Second, we allow the signal distribution to be asymmetric, by assuming a skew-normal distribution. This distribution generalizes the normal distribution to allow for positive or negative skewness; like the normal distribution, it has location and scale parameters and remains easily interpretable. When the turnover cost is zero, the substitutability result that automatically obtains with a symmetric distribution no longer holds with an asymmetric distribution. There can then be complementarity between signal precision and monitoring. When skewness is positive, this is the case for signals such that higher precision raises the probability that the signal was generated by a high ability manager at a higher rate than for a low ability manager. Then, even though it remains more likely that the signal was generated by a low ability manager, the signal is now paradoxically less informative, which raises the value of monitoring. We establish this result for a subset of intermediate signals.

Third, we allow the signal to be a "composite signal" that measures the value of assets in place as well as the value of managerial ability. In this case, a more precise signal can improve measurement of the former and/or the latter. We derive a condition under which signal precision and monitoring are complements. Complementarity can occur because a more precise signal on one dimension (the value of assets in place) can indirectly make the signal less informative on another dimension (managerial ability).

These results can be viewed as surprising. Since firm value measurement and monitoring both provide information about the manager's ability, we would expect these two sources of information to be substitutes, i.e., the value of one source (monitoring) would diminish when the other (the signal) becomes more informative. However, unless we make strong assumptions, these two sources of information can sometimes be complementary. Moreover, the model suggests that interactions between the quality of the signal and monitoring can also depend on firm performance, as opposed to simply being firm-specific.[1]

The model generates a novel empirical implication for the large literature on CEO turnover and firm performance.[2] When the quality of the signal and monitoring are complements, a more precise signal paradoxically leads to a higher probability of termination after a good signal, and to a lower probability of termination after a bad signal. Indeed, with complementarity, an improvement in signal precision leads to more monitoring, which provides another source of information for termination decisions. The board then relies less on the signal for termination decisions. This result is in contrast with the view that more informative performance measurement should always increase the reliance on firm performance for termination decisions (Engel, Hayes, and Wang (2003)).[3]

We then study an extension of the model where firms can either have a "disciplinarian"

[1] Ward, Brown, and Rodriguez (2009) had already noted that the complementarity between inside monitoring by the board and outside monitoring by investors can depend on firm performance.

[2] Papers in this literature include Murphy and Zimmerman (1993), Huson, Parrino, and Starks (2001), Kaplan and Minton (2012), Eisfeldt and Kuhnen (2013), Jenter and Kanaan (2015), Jenter and Lewellen (2020).

[3] Farrell and Whidbee (2003) had already noted that the CEO turnover decision involves information besides CEO performance. Our contribution is to study how a change in the informativeness of a signal of firm value affects the quality of other available information. This perspective emphasizes a limitation of empirical studies that relate managerial dismissal to observable performance measures (e.g., Kaplan and Minton (2012), Jenter and Lewellen (2020)), whose low explanatory power had already been noted by Brickley (2003). In particular, a low correlation between firm performance and managerial termination does not necessarily reflect a corporate governance failure, but instead can be observed in firms with strong monitoring. This can shed light on the finding of Gao, Harford, and Li (2017) that public firms have greater turnover-performance sensitivity than private firms.

board with more power to terminate the CEO (e.g. empowered independent directors) or an "informed" board with directors better able to assess the CEO's ability. We show that firms with a disciplinarian board will rely more on the signal for termination decisions. This result is consistent with the finding in Faleye, Hoitash, and Hoitash (2011) that the sensitivity of CEO turnover to firm performance is higher in firms with more empowered independent directors. However, it suggests another interpretation for their findings and those of the literature that followed – greater influence for independent directors does not equate to more "monitoring".

Turning to applications, the model generates novel empirical predictions regarding the complementarity of these governance mechanisms. Consider private equity owned firms. Cornelli, Kominek, and Ljungqvist (2013) find that managers in private equity firms do not get fired following a poor performance, but instead only get fired when board monitoring "raise[s] concerns about [their] ability". In our model, this is the case for a very low s^T, which corresponds to a high turnover cost. The latter is consistent with the evidence in Taylor (2010), and by the finding of Gao, Harford, and Li (2017) that private firms have lower turnover rates than public firms, despite the high variability in CEO ability among private firms (Kaplan et al. (2012), Cornelli, Kominek, and Ljungqvist (2013)). Our model predicts that, in these firms, the precision of firm value measurement and the intensity of monitoring are substitutes conditional on a good performance, but complements conditional on a bad performance.

The latter can be either a blessing or a curse. On the one hand, underperforming firms with better firm value measurement will also enjoy better monitoring, and therefore excellent corporate governance. On the other hand, underperforming firms are also prone to fall in a "bad governance trap", such that the information loss due worse firm value measurement will be compounded by a reduction in monitoring. This results in poorly informed continuation and termination decisions, which makes it harder to improve future firm performance.

Our model with unknown managerial ability and monitoring is similar to Hermalin's (2005): we also consider endogenous monitoring of the manager's ability by the board for the purposes of improving the termination decision. In the model of Hermalin (2005), the only source of information on the manager's ability (besides the board's prior) is provided by board monitoring. Likewise, Hermalin and Weisbach (2012) study the consequences of improved information quality for intervention by the firm's owners (such as terminating the manager), but only study one source of information. In contrast, we allow for the existence of a signal of firm value, and we study how an improvement in the information content of this signal will affect monitoring.[4] More generally, our paper contributes to a

[4]The model is also related to von Thadden (1995), where monitoring following a bad performance can improve the termination decision. However, von Thadden focuses on inefficient investment decisions

large recent literature which studies how corporate governance arises endogenously (e.g., Hermalin and Weisbach (1998), Raith (2003), Drimyotes (2007), Kumar and Sivaramakrishnan (2008), Acharya and Volpin (2010), Dicks (2012), Peng and Röell (2014), Misangyi and Acharya (2014), Levit and Malenko (2016), Schroth (2018), Levit (2018), Baldenius, Meng, and Qiu (2019, 2020), Donaldson, Malenko, and Piacentino (2020)).

As opposed to the signal of firm value, which is "hard" or verifiable information, the information gathered by boards via monitoring is typically "soft" or nonverifiable (Cornelli, Kominek, and Ljungqvist (2013)). It is already well-known that hard and soft information can be complementary (Gigler and Hemmer (1998, 2001), Dye and Sridhar (2004), Şabac and Tian (2015), Jiang (2016)). These papers study the complementarity of hard information and soft information when the latter is a report by the manager and the manager can misreport. Christensen, Frimor, and Şabac (2020) study the effects of hard and soft information for managerial incentives. By contrast, we study whether soft information on managerial ability becomes more valuable when a signal whose distribution depends on managerial ability becomes more precise. Since in our paper information is only valuable to assess managerial ability, our results are not driven by incentive provision or its interaction with managerial ability assessment. Several recent papers also study the optimal disclosure of hard and soft information by managers to maximize firm value (Arya, Glover, Mittendorf, and Zhang (2004), Almazan, Banerji, and de Motta (2008), Bertomeu and Marinovic (2016), Cianciaruso and Sridhar (2018)). Thus, we contribute to the growing literature on hard and soft information (Liberti and Petersen (2019)) and our results have implications for the literature on information design (Bergemann and Morris (2019)).

Related papers in the moral hazard literature include Demougin and Fluet (2001) and Piskorski and Westerfield (2016), who study the optimal combination between monitoring and performance-based pay when monitoring provides information about the agent's action rather than his ability. In a model with moral hazard and adverse selection, Banker et al. (2019) study the value of "pre-contract" information about the agent's ability when the agent privately knows his ability – in our model, the principal and agent are symmetrically informed. Crémer (1995) studies the implications of better monitoring for incentive provision. Chaigneau, Edmans, and Gottlieb (2018) study the value of a more precise performance measure in a moral hazard model, but do not study monitoring. Börgers, Hernando-Veciana, and Krähmer (2013) provide general conditions for signals to be complements or substitutes.

due to incentives misalignment, and the value of monitoring in this context, whereas we focus on the implications of a change in the informativeness of firm value measurement for monitoring.

1 The model

We present a principal-agent model with an imperfectly informative signal and monitoring in which the board of a firm must use the information at its disposal to assess a manager with unknown ability. The model can be viewed as an extension of Hermalin (2005) in which the board observes a signal before engaging in monitoring and making the termination decision.

Technology and information

A firm needs a manager to invest in a project that produces output at $t = 2$. At $t = 0$, the firm draws its manager from a pool of managers with unknown ability. A manager from this pool has high ability with probability $p \in (0, 1)$, and low ability with probability $1 - p$.[5] Expected firm output is μ_h with a high ability manager, and μ_l with a low ability manager, with $\mu_h > \mu_l$ and $\Delta \equiv \mu_h - \mu_l$. Ex-ante, the ability of a manager is unknown to all parties: as is common in models of career concerns (Harris and Holmström (1982)), information is symmetric.

At $t = 1$, the board learns about the ability of the manager in place from $t = 0$ to $t = 1$ by observing an imperfectly informative signal $s \in (-\infty, \infty)$. The signal can be interpreted as earnings (a measure of value creation) or as the stock price (a measure of firm value). Let $\mu \in \{\mu_h, \mu_l\}$ denote the manager's ability, and $\pi(s|\mu)$ denote the probability density function (PDF) of the signal s conditional on μ. The quality or "precision" of the signal s is parameterized by σ, with an increase in σ representing a spread in its distribution. We assume that this distribution has full support, and that its PDF is differentiable with respect to σ.[6] Signal realizations are informative about the manager's ability in a monotone likelihood ratio sense: we assume that $\frac{\pi(s|\mu_h)}{\pi(s|\mu_l)}$ is strictly increasing in s (MLRP).

Monitoring

After observing signal s, the board can engage in costly monitoring, by collecting additional soft information to learn more about the ability of the current manager. As in Hermalin (2005), at $t = 1$, the board chooses the intensity of monitoring $m \in [0, 1]$ at cost $c(m)$. We assume that $c(m)$ is differentiable twice, with $c(0) = 0$, $c'(m) > 0$, $c''(m) > 0$, as well as $\lim_{m \to 0} c'(m) = 0$ and $\lim_{m \to 1} c'(m) \to \infty$ to rule out corner solutions. With probability m, monitoring is successful and reveals the ability of the manager to the board; with probability $1 - m$, monitoring is unsuccessful and the board does not learn anything

[5] Managerial ability can also be interpreted as the quality of the match between the manager and the firm's project, i.e., the ability of the manager to manage this particular firm.

[6] The full support assumption allows to keep the model simple. If there were signals s such $\pi(s|\mu_h) > 0$ and $\pi(s|\mu_l) = 0$, or $\pi(s|\mu_h) = 0$ and $\pi(s|\mu_l) > 0$, these signals would be perfectly revealing about the manager's ability so that monitoring would be useless.

new about the manager's ability. This modelization of board monitoring is consistent with the empirical evidence in Cornelli, Kominek, and Ljungqvist (2013) that internal monitoring enables to better assess the manager's ability, and helps avoid terminating a manager for bad luck.

The information generated by monitoring allows the board to make a more informed termination decision. At $t = 1$, after observing the first period signal and the outcome of monitoring, the board can dismiss the manager, and replace him with a new manager (whose ability is high with probability p, as already specified). There is a turnover cost K for the firm, which represents the costs of a transition to a new manager. We assume that $K \in [0, p(\mu_h - \mu_l))$: replacing a manager can be costly, but not so costly as to prevent termination even when the manager in place is known to have low ability.[7] The objective of the board is to maximize expected firm output (including the turnover cost) net of the monitoring cost.

There is an alternative interpretation for the model such that the quality of the firm's project, rather than managerial ability, is unknown. Then, a high quality project would be worth μ_h, and a low quality project would be worth μ_l. As in von Thadden (1995), both the signal and monitoring would then be informative about the quality of the project rather than managerial ability. In this case, the termination decision would refer to the termination of the project rather than managerial dismissal.

2 Governance Mechanisms

The termination decision depends on the monitoring outcome. If monitoring is successful or "conclusive", it reveals the current manager's ability. The board then optimally chooses to retain the manager if ability is high, and to dismiss him if ability is low. If monitoring is unsuccessful, the only information at the board's disposal is the signal s. Section 2.1 studies the termination decision in this case.

2.1 Termination Decision Based on the Signal

When making the termination decision, the board compares the net benefit of replacing the manager against the alternative of retaining the current manager. After observing signal s, the updated probability p_s that the incumbent manager has high ability is:

$$p_s = \frac{p\pi(s|\mu_h)}{p\pi(s|\mu_h) + (1-p)\pi(s|\mu_l)} \tag{1}$$

[7]Indeed, the assumption $K < p(\mu_h - \mu_l)$ can be rewritten as $\mu_l < p\mu_h + (1-p)\mu_l - K$, i.e., the board optimally terminates a manager who has low ability with probability 1. If this assumption were violated, a manager would always be continued, and there would never be any monitoring.

Since by assumption $\frac{\pi(s|\mu_h)}{\pi(s|\mu_l)}$ is strictly increasing in s, p_s is strictly increasing in s. The updated probability p_s is increasing in the proportion of high ability managers in the population, p, and in the likelihood $\frac{\pi(s|\mu_h)}{\pi(s|\mu_l)}$ that a signal s was generated by a high ability manager rather than by a low ability manager.

Let s^* be implicitly defined by $\pi(s^*|\mu_h) \equiv \pi(s^*|\mu_l)$, which is unique by MLRP, and corresponds to the realization of s such that $p_s = p$. For a "negative signal" $s < s^*$, we have $p_s < p$, so that the board negatively updates its belief on the manager's ability after observing the signal. On the contrary, for a "positive signal" $s > s^*$ we have $p_s > p$.

After observing signal s, the expected value of continuing the incumbent manager is: $p_s\mu_h + (1-p_s)\mu_l$, and the expected value of terminating the manager is: $p\mu_h + (1-p)\mu_l - K$. Therefore, based only on the signal s, the manager will be terminated if and only if:

$$\begin{aligned} p_s\mu_h + (1-p_s)\mu_l &< p\mu_h + (1-p)\mu_l - K \\ \Leftrightarrow \qquad p_s &< p - \frac{K}{\mu_h - \mu_l}, \end{aligned} \tag{2}$$

which simply rewrites as $p_s < p$ for $K = 0$: when the turnover cost is nil, the board compares the updated probability that the incumbent manager has high ability, p_s, with the probability that a new manager has high ability, p. The incumbent manager remains in place if $p_s \geq p$, and is replaced otherwise. With a positive turnover cost, the manager remains in place for a larger subset of parameter values. Since p_s is strictly increasing in s, the termination condition in equation (2) can be rewritten as $s < s^T$, for a given s^T.

Lemma 1. *$s^T \leq s^*$, with a strict inequality for $K > 0$, and s^T is strictly decreasing in K.*

Absent conclusive monitoring, the manager is terminated following a "bad signal" ($s < s^T$), and retained following a "good signal" ($s \geq s^T$). For signals $s \in [s^T, s^*]$, the manager is retained even though the board negatively updates its beliefs about his ability: these signals only lead to mild negative belief updating which does not outweigh the turnover cost.

The board can make two types of mistake in its termination decision when monitoring is inconclusive: it sometimes wrongfully terminates a high ability manager following a bad signal and unsuccessful monitoring, and in addition it sometimes wrongfully retains a low ability manager following a good signal.

2.2 Monitoring Decision

We now analyze the monitoring decision of the board at $t = 1$, after it observes the signal s and before it decides on managerial termination.

First, consider the monitoring decision after a "bad signal" ($s < s^T$). Given monitoring intensity m, monitoring is unsuccessful with probability $1 - m$, in which case the board terminates its manager (see section 2.1), and its second period value net of turnover costs is $V^T(p) \equiv p\mu_h + (1-p)\mu_l - K$. Monitoring is successful with probability m, and there are two possible cases. With probability p_s, as defined in equation (1), the incumbent manager has high ability and is retained, in which case firm value is $V^C(1) \equiv \mu_h$. With probability $1 - p_s$, the incumbent manager has low ability and is terminated, in which case firm value is $V^T(p)$. Thus, the choice of monitoring intensity solves the following problem:

$$\begin{aligned} & \max_m \left\{ m \left[p_s V^C(1) + (1-p_s) V^T(p) \right] + (1-m) V^T(p) - c(m) \right\} \\ \Leftrightarrow \quad & \max_m \left\{ m p_s \left(V^C(1) - V^T(p) \right) + V^T(p) - c(m) \right\}. \end{aligned}$$

When the board chooses the monitoring intensity, it compares the benefit of keeping with probability mp_s a high-ability manager who would otherwise have been dismissed, or a net benefit of $V^C(1) - V^T(p)$, against the cost of monitoring, $c(m)$. The necessary and sufficient first-order condition that defines optimal monitoring intensity m_s^* is:

$$c'(m_s^*) = p_s \left(V^C(1) - V^T(p) \right) \quad \Leftrightarrow \quad c'(m_s^*) = p_s \left((1-p)(\mu_h - \mu_l) + K \right). \tag{3}$$

Second, consider the monitoring decision after a "good signal" ($s \geq s^T$). Given monitoring intensity m, monitoring is unsuccessful with probability $1 - m$, in which case the board continues its manager (see section 2.1), and its second period value net of turnover costs is $V^C(p_s) \equiv p_s\mu_h + (1-p_s)\mu_l$. Monitoring is successful and reveals the manager's ability with probability m, with the same outcomes as in the preceding paragraph. Thus, the choice of monitoring intensity solves the following problem:

$$\begin{aligned} & \max_m \left\{ m \left[p_s V^C(1) + (1-p_s) V^T(p) \right] + (1-m) V^C(p_s) - c(m) \right\} \\ \Leftrightarrow \quad & \max_m \left\{ m \left[p_s \left(V^C(1) - V^T(p) \right) + V^T(p) - V^C(p_s) \right] + V^C(p_s) - c(m) \right\}. \end{aligned}$$

Bearing in mind that $K < p(\mu_h - \mu_l)$ by assumption, the first-order condition is:

$$\begin{aligned} c'(m_s^*) &= p_s \left(V^C(1) - V^T(p) \right) + V^T(p) - V^C(p_s) \qquad (4) \\ &= (1-p_s) \left(p(\mu_h - \mu_l) - K \right). \qquad (5) \end{aligned}$$

After a good signal, the "default decision" is to continue the manager. Therefore, the intensity of monitoring is increasing in the probability $1 - p_s$ that the manager has a low ability after a good signal is observed.

Figure 1 summarizes the role of monitoring and the possible outcomes.

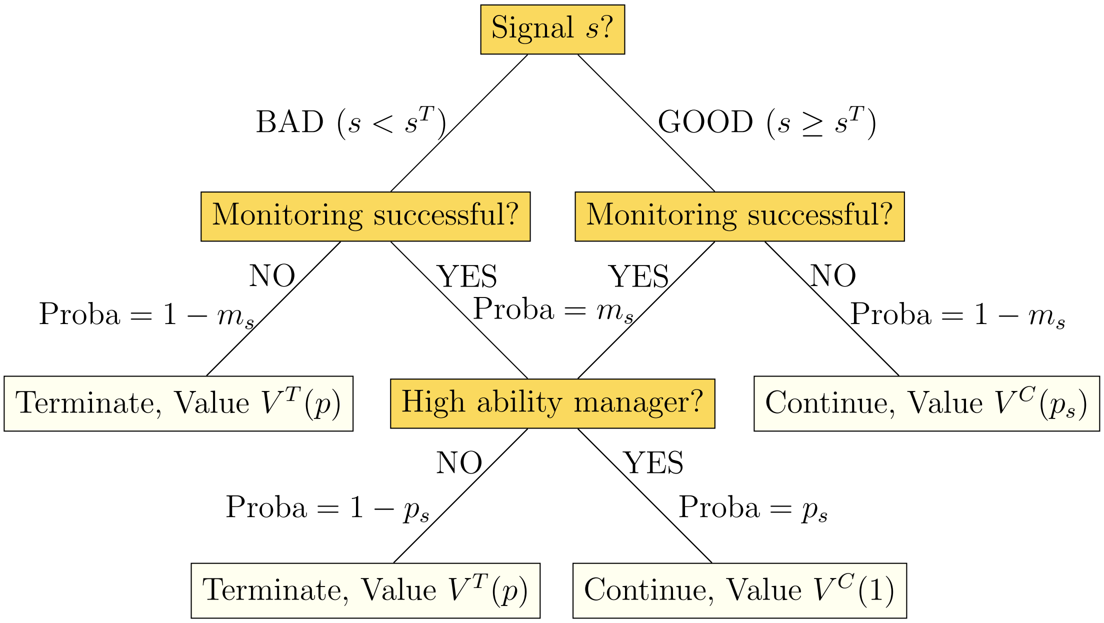

Figure 1: Probability tree.

2.3 A General Condition

This section studies the effect of a change in signal precision on monitoring intensity. If the signal is the stock price, this change could represent an increase in stock market efficiency. If the signal is earnings, this change could be an improvement in the firm's accounting system. In turn, this improvement could be due to a change in listing and disclosure requirements faced by the firm, or a change in accounting regulation and guidelines. In any case, we take this change in the quality of the signal as given, and we study its consequences for monitoring intensity, which is optimally determined by the board.

Definition. *The precision of the signal and monitoring are "complements" when an increase in signal precision increases monitoring, and "substitutes" otherwise.*

Lemma 2 is a key intermediary result which tells us for which signal s the board will choose a higher monitoring intensity when the signal distribution is more precise.

Lemma 2.

- *Following a bad signal ($s < s^T$), an improvement in signal precision (lower σ) increases monitoring intensity m_s^* if and only if it increases the updated probability p_s that the manager has high ability.*

- *Following a good signal ($s \geq s^T$), an improvement in signal precision (lower σ) increases monitoring intensity m_s^* if and only if it decreases the updated probability p_s that the manager has high ability.*

After a bad signal is observed, the default decision is to terminate the manager. Monitoring is then all the more valuable that termination is often wrong, i.e., the updated probability p_s that the manager has high ability is high (in the sense that it is not much lower than p: for a bad signal, $p_s < p$). After a good signal is observed, the default decision is to continue the manager. Monitoring is then all the more valuable that continuing the manager is often wrong, i.e., the updated probability p_s that the manager has high ability is still low.

We now derive a general condition under which the precision of the signal and monitoring are complements or substitutes. Let $\pi'_\sigma(\cdot)$ denote the derivative of $\pi(\cdot)$ with respect to σ.

Proposition 1. *After a bad signal ($s < s^T$), signal precision and monitoring are substitutes if and only if:*

$$\frac{\pi'_\sigma(s|\mu_h)}{\pi(s|\mu_h)} \geq \frac{\pi'_\sigma(s|\mu_l)}{\pi(s|\mu_l)}. \tag{6}$$

After a good signal ($s \geq s^T$), signal precision and monitoring are substitutes if and only if:

$$\frac{\pi'_\sigma(s|\mu_h)}{\pi(s|\mu_h)} \leq \frac{\pi'_\sigma(s|\mu_l)}{\pi(s|\mu_l)}. \tag{7}$$

Proposition 1 tells us for which signal s the intensity of monitoring is lower when the signal distribution is more precise. The conditions in Proposition 1 can be interpreted as follows (remember that an increase in σ reduces precision). After a bad signal, there is "substitutability" if and only if the rate of change in the probability of observing this bad signal as the signal distribution becomes more precise is higher for a manager with low ability than for a manager with high ability. After a good signal, there is "substitutability" if and only if the rate of change in the probability of observing this good signal as the signal distribution becomes more precise is higher for a manager with high ability than for a manager with low ability. These conditions seem natural, which suggests that improvements in the quality or "precision" of the signal will typically decrease monitoring.

Example 1 below studies a commonly used setting. The signal is normally distributed: the mean is the manager's ability, and the variance is the inverse of the precision of the signal.

Example 1. *The signal follows a normal distribution with variance σ^2, and means normalized at $\mu_h = 1$ for high ability managers and $\mu_l = 0$ for low ability managers. There is no turnover cost: $K = 0$. Letting φ be the PDF of the standard normal distribution:*

$$\begin{aligned} \pi\left(s|\mu_h\right) &= \frac{1}{\sigma}\varphi\left(\frac{s-1}{\sigma}\right) \\ \pi\left(s|\mu_l\right) &= \frac{1}{\sigma}\varphi\left(\frac{s}{\sigma}\right) \end{aligned}$$

The probability that the manager has high ability following signal s is:

$$p_s = \frac{p\frac{1}{\sigma}\varphi\left(\frac{s-1}{\sigma}\right)}{p\frac{1}{\sigma}\varphi\left(\frac{s-1}{\sigma}\right) + (1-p)\frac{1}{\sigma}\varphi\left(\frac{s}{\sigma}\right)}.$$

Then we get $s^* = s^T = 0.5$. Figure 2 depicts the relevant functions for the conditions in Proposition 1 in this example.

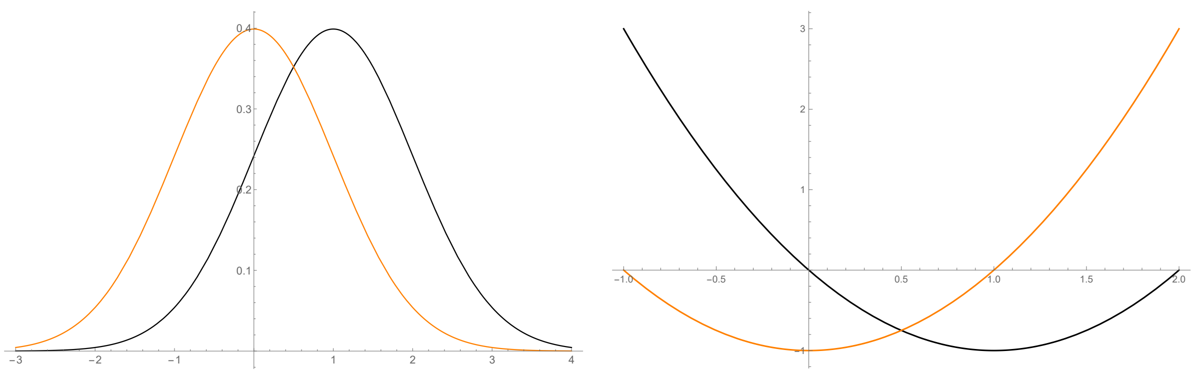

Figure 2: The functions $\frac{\partial\pi(s|\mu_h)/\partial\sigma}{\pi(s|\mu_h)}$ (black) and $\frac{\partial\pi(s|\mu_l)/\partial\sigma}{\pi(s|\mu_l)}$ (orange) as a function of s for a normally distributed signal with $\mu_h = 1$, $\mu_l = 0$, and $\sigma = 1$.

In this example, the conditions in Proposition 1 are always satisfied, i.e. an increase in signal precision always reduces the value and therefore the intensity of monitoring.

In the simple model used in Example 1, the precision of the signal and the intensity of monitoring are substitutes for any signal s.

3 Sources of Complementarity

In this section, we identify three reasons why signal precision and monitoring can be complements.

To this end, we put more structure on the signal distribution, by assuming a single-peaked distribution with location parameter μ and scale parameters σ. This implies that there exists a function $f(\cdot)$ such that the signal distribution can be written as:

$$\pi(s|\mu) = \frac{1}{\sigma} f\left(\frac{s-\mu}{\sigma}\right). \tag{8}$$

An increase in the location parameter represents improvements in the distribution in the sense of first-order stochastic dominance. This captures the effect of managerial ability on the signal distribution. The scale parameter represents changes in the distribution's dispersion in the sense of a spread of the distribution. This captures improvements in signal precision on the distribution. This family of distributions includes the normal, logistic, Cauchy, and Laplace distributions, which are symmetric, but also the skew-normal distribution, which is asymmetric.

3.1 Turnover Cost

We already know from Lemma 1 that, with a turnover cost $K > 0$, "good signals" ($s \geq s^T$) that do not alone lead to managerial termination include a subset of signals with negative belief updating on the manager's ability. Proposition 2 draws the implications for the relation between signal precision and monitoring intensity.

Proposition 2. *Consider symmetric signal distributions with location parameter μ and scale parameter σ. Then, for $K > 0$, we have $s^T < s^* = \frac{\mu_l+\mu_h}{2}$, and signal precision and monitoring are complements if and only if $s \in (s^T, s^*)$. The probability that $s \in (s^T, s^*)$ is zero for $K = 0$ and is increasing in K.*

For $s \in (s^T, s^*)$, even though the board receives a negative signal about the manager's ability ($s < s^*$), this signal alone does not lead to termination due to the cost of managerial termination and replacement ($s > s^T$). Moreover, for negative signals ($s < s^*$), an increase in signal precision increases the rate at which the signal is observed when the manager has low ability more than the rate at which the signal is observed when the manager has high ability. Simply put, a more precise signal distribution increases the information content of a negative signal. This makes monitoring more valuable when the default decision is to retain the manager.

We now revisit Example 1 by introducing a turnover cost $K > 0$. Figure 3 below depicts the probability for the signal to be such that the precision of the signal and monitoring are complements as a function of the turnover cost K for a normally distributed signal with the same parameters as in Figure 2. Proposition 2 already stated that this probability is increasing in K, and Figure 3 depicts the shape of this relation for a normally distributed signal.

For example, with $K = 0.2 \times \Delta$ (the turnover cost is 20% of the firm value gap due to managerial ability), we find that $s^T = -0.35$. As before, $s^* = 0.5$. Then there is complementarity if and only if $s \in (-0.35, 0.50)$, i.e., for signals that lead the board to negatively revise its belief about the manager's ability, but which are such that the manager is not terminated without monitoring. For $p = 0.5$, the signal is in this interval with probability 27.3% (this probability is unconditional: it is the relevant probability for a firm that does not know the type of its manager).

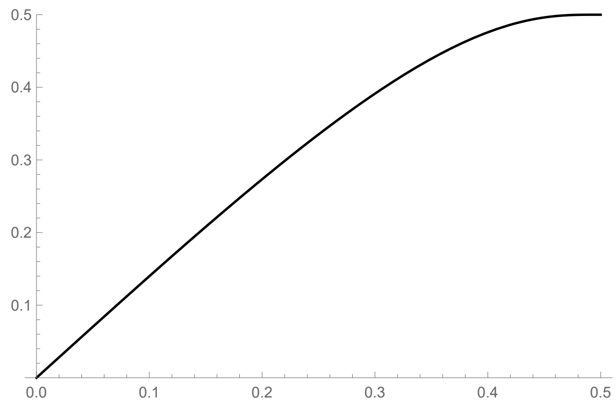

Figure 3: Probability that $s \in \left(s^T, s^*\right)$ as a function of the turnover cost K with a normally distributed signal with $\mu_h = 1$, $\mu_l = 0$, $\sigma = 1$, and $p = 0.5$.

3.2 The signal distribution

This section focuses on the effect of the signal distribution. We now assume no turnover cost ($K = 0$) to identify another potential source of complementarity between the precision of the signal and the value of additional information. With $K = 0$, we know from Lemma 1 that $s^T = s^*$. Let $\hat{s}$ be implicitly defined by:

$$\frac{\pi'_\sigma\left(\hat{s}|\mu_h\right)}{\pi\left(\hat{s}|\mu_h\right)} \equiv \frac{\pi'_\sigma\left(\hat{s}|\mu_l\right)}{\pi\left(\hat{s}|\mu_l\right)} \qquad \Leftrightarrow \qquad \frac{\hat{s}-\mu_l}{\hat{s}-\mu_h} = \frac{-\frac{1}{\sigma}f'\left(\frac{\hat{s}-\mu_h}{\sigma}\right)/f\left(\frac{\hat{s}-\mu_h}{\sigma}\right)}{-\frac{1}{\sigma}f'\left(\frac{\hat{s}-\mu_l}{\sigma}\right)/f\left(\frac{\hat{s}-\mu_l}{\sigma}\right)}.$$

From Proposition 1 with $s^T = s^*$, we know that there is substitutability for all s only if $\hat{s} = s^*$. This means that the cutoff s^* that separates positive from negative signals must be the same as the cutoff $\hat{s}$ such that an improvement in signal precision increases equally the rate at which the signal is generated by managers with high and low ability.

3.2.1 Symmetric distributions

With symmetric distributions, we always have $\hat{s} = s^*$, and the conditions in Proposition 1 always hold:

Proposition 3. *For symmetric distributions and $K = 0$, signal precision and monitoring are substitutes for any s.*

This shows that the result from Example 1 generalizes to any symmetric signal distribution. However, Proposition 3 is but a first step, since with symmetric distributions, the cutoffs $\hat{s}$ and s^* are equal by construction. The next section considers asymmetric distributions.

3.2.2 Asymmetric distributions

We consider the skew-normal distribution, which is an asymmetric distribution with location and scale parameters that satisfies MLRP, as established in the Appendix. The shape parameter α determines its asymmetry: $\alpha > 0$ implies a positive skewness, and $\alpha < 0$ implies a negative skewness – with $\alpha = 0$, skewness is nil and the distribution is normal.

Example 2 shows that, with an asymmetric distribution, the cutoff s^* that separates positive from negative signals is not necessarily the same as the cutoff $\hat{s}$ such that an improvement in signal precision increases equally the rate at which the signal is generated by managers with high and low ability.

Example 2. *The signal follows a skew-normal distribution with scale parameter $\sigma = 1$, shape parameter $\alpha = 4$, and a location parameter $\mu_h = 1$ for high ability managers and $\mu_l = 0$ for low ability managers. Figure 4 depicts the PDFs of the signal distribution and the relevant functions for the condition in Proposition 1 in this example. Numerically, we get $s^* = 1.05$ and $\hat{s} = 0.85$.*

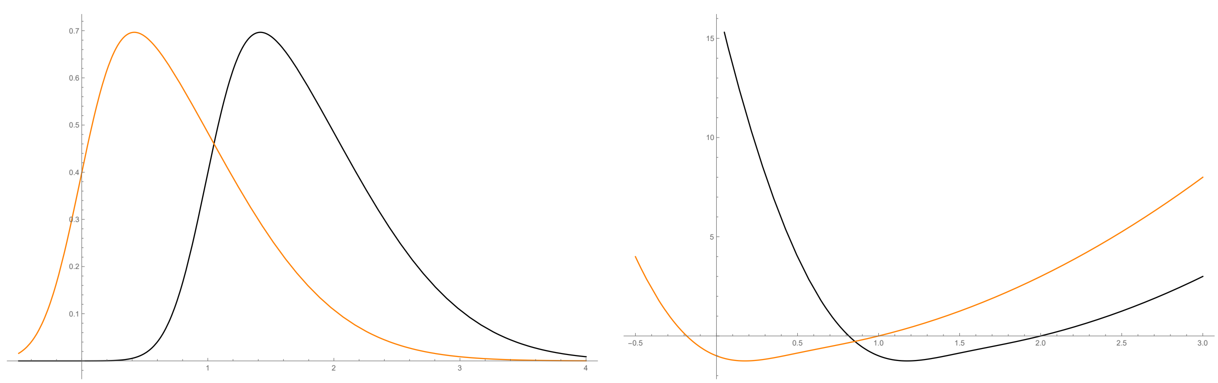

Figure 4: On the left, the PDF of skew-normal distributions with location, scale, and shape parameters $\{1, 1, 4\}$ (black), and $\{0, 1, 4\}$ (orange). On the right, the functions $\frac{\partial \pi(s|\mu_h)/\partial \sigma}{\pi(s|\mu_h)}$ (black) and $\frac{\partial \pi(s|\mu_l)/\partial \sigma}{\pi(s|\mu_l)}$ (orange) as a function of s for $p = 0.5$, and skew-normal distributions with location, scale, and shape parameters $\{1, 1, 4\}$ for $\mu = \mu_h$, and $\{0, 1, 4\}$ for $\mu = \mu_l$.

In this example, the conditions in Proposition 1 are not always satisfied: for $s \in [0.85, 1.05]$, *an increase in signal precision increases the value and therefore the intensity of monitoring. With* $p = 0.5$*, the probability that* $s \in [0.85, 1.05]$ *is 8.2%.*

The intuition for this result is the following. In Example 2, an increase in precision raises monitoring intensity for $s \in [0.85, 1.05]$. Now consider Figure 5, which depicts the two PDFs of skew-normal distributions from Example 2, as well as two PDFs representing a slight decrease in the scale parameter of these two distributions, i.e., an increase in precision. This Figure shows that, for $s \in [0.85, 1.05]$, an increase in precision barely changes the density corresponding to a low ability manager, but it increases the density corresponding to a high ability manager from a low level. That is, even though it remains more likely that the manager has low ability for this subset of signals, the increase in precision increases the likelihood that a signal in this subset was generated by a high ability manager. This in turn increases the probability that the default decision without monitoring, which is to terminate the manager, is the wrong decision. This increases the value of additional information on the manager's ability, and therefore the intensity of monitoring.

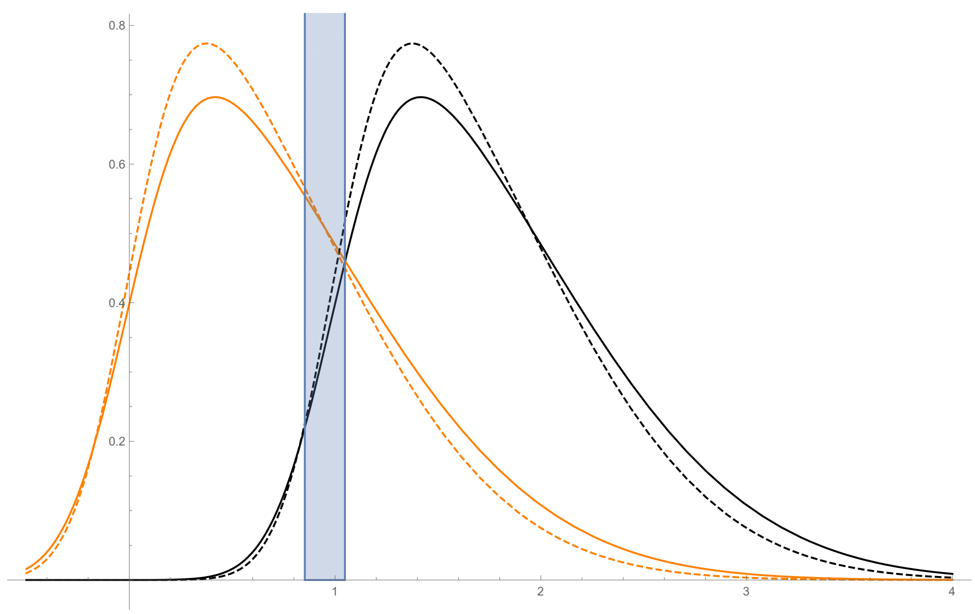

Figure 5: The PDFs of skew-normal distributions with location, scale, and shape parameters $\{1, 1, 4\}$ (black), $\{1, 0.9, 4\}$ (dashed black), $\{0, 1, 4\}$ (orange), and $\{0, 0.9, 4\}$ (dashed orange).

The following Proposition establishes the robustness of the complementarity result from Example 2.

Proposition 4. *For skew-normal distributions with location parameters μ_l and μ_h, and scale and shape parameters σ and α, and $K = 0$, we have:*

(i) For any $\mu_h > \mu_l$, there exists $\bar{\alpha} > 0$ such that for $\alpha > \bar{\alpha}$, $s^* > \hat{s}$, and the signal precision and monitoring are complements for $s \in [\hat{s}, s^*]$; for $\alpha < -\bar{\alpha}$, $s^* < \hat{s}$ and the signal precision and monitoring are complements for $s \in [s^*, \hat{s}]$.

(ii) For any $\alpha > 0$ and any μ_l, there exists $\check{\mu}$ such that when $\mu_h < \mu_l + \check{\mu}$, $s^* > \hat{s}$ and the signal precision and monitoring are complements for $s \in [\hat{s}, s^*]$. Similarly, for any $\alpha < 0$ and any μ_h, there exists $\check{\mu}$ such that when $\mu_l > \mu_h - \check{\mu}$, $s^* < \hat{s}$, and the signal precision and monitoring are complements for $s \in [s^*, \hat{s}]$.

Part (i) of Proposition 4 shows that for any location parameters, if the distribution is sufficiently asymmetric in the sense of a sufficiently large absolute value of its shape parameter α, then there is an interval of signals such that signal precision and monitoring are complements. Part (ii) of Proposition 4 shows that for any asymmetry of the distribution as proxied by its shape parameter α, if managerial ability has a sufficiently important impact on the location of the signal distribution, then there is an interval of signals such that signal precision and monitoring are complements. In short, Proposition 4 establishes that signal precision and monitoring will be complements for some signal realizations when the signal distribution is sufficiently asymmetric, or when it is asymmetric and sufficiently informative about managerial ability.

Results in this section 3.2 depend on two dimensions of the signal distribution: they depend on whether $\hat{s}$, which is by definition such that $\frac{\partial\pi(\hat{s}|\mu_h)/\partial\sigma}{\pi(\hat{s}|\mu_h)} = \frac{\partial\pi(\hat{s}|\mu_l)/\partial\sigma}{\pi(\hat{s}|\mu_l)}$, is higher or lower than s^*, which is by definition such that $\pi(s^*|\mu_h) = \pi(s^*|\mu_l)$. While it is in general possible to normalize distributions without loss of generality (WLOG) if only one dimension of the distribution matters (e.g. in a moral hazard problem, the optimal contract only depends on the likelihood ratio of the output distribution), such a normalization is not necessarily WLOG when more than one dimension of the distribution matters. It can be WLOG in special cases. For example, there is a well-known equivalence between the (asymmetric) lognormal distribution and the (symmetric) normal distribution – if $\tilde{s}$ is lognormally distributed then $\ln(\tilde{s})$ is normally distributed. However, the skew-normal distribution, which is described by three parameters, is irreducible to a normal distribution, which is described by two parameters. That is, the skew-normal distribution is not only asymmetric, but it also fundamentally differs from a symmetric distribution.

3.3 Composite signal

A measure of firm value is often a "composite signal" that aggregates information along several dimensions. For example, the stock price of a publicly listed firm, or the implied valuation in the latest funding round of a private firm, will aggregate information about different components of firm value.

In this context, an improvement in the precision of the signal cannot always be defined straightforwardly. In this section, we model a richer information environment to show how a composite signal can be another source of complementarity between the quality of the signal and monitoring intensity. As in section 3.2, we assume no turnover cost ($K = 0$) to identify another potential source of complementarity between signal precision and monitoring.

We now assume that firm value is the sum of two components. The first component, $\tilde{\mu}$, is as before equal to $\mu = \mu_h$ with a high ability manager, and $\mu = \mu_l$ with a low ability manager. The second component, $\tilde{z}$, is an exogenous variable which is normally distributed with mean $\bar{z}$ and variance $\sigma_z^2 > 0$, and independent from other random variables. Thus, firm value under current management is equal to $\mu + z$, where μ and z denote the realizations of the random variables $\tilde{\mu}$ and $\tilde{z}$. However, the realization of these random variables is not directly observed. A natural interpretation is that μ is the value of the firm's investment opportunities or "growth options" under current management, and z is the value of its assets in place.

3.3.1 A simple signal distribution

We start by assuming a simple signal distribution. The distribution of the signal conditional on the realizations μ and z of the random variables $\tilde{\mu}$ and $\tilde{z}$ is given by:[8]

$$\tilde{s} = \mu + \gamma z + (1 - \gamma)\left(\bar{z} + \tilde{\epsilon}\right), \tag{9}$$

where μ is the ability of the manager in place, and $\tilde{\epsilon}$ is normally distributed with mean 0 and variance $\sigma_\epsilon^2 \geq 0$, and independent from other random variables. The precision of the signal distribution is now parameterized by $\gamma \in [0, 1]$: the signal is imperfectly informative about firm value under current management for $\gamma \in [0, 1)$, and perfectly informative for $\gamma = 1$. The variable $\tilde{\epsilon}$ represents mistakes in asset value measurement when the signal is imperfectly informative ($\gamma < 1$), and its variance σ_ϵ^2 measures the magnitude of these mistakes.

Changes in the precision of the signal, as measured by γ, have two effects.

First, by construction, the signal s reflects more accurately the value z of the firm's assets if γ is higher. For example, with $\gamma = 0$, the conditional signal distribution is $\tilde{s} = \mu + \bar{z} + \tilde{\epsilon}$: the signal is uninformative about the value z of assets in place. With $\gamma = 1$, the conditional signal distribution is $\tilde{s} = \mu + z$: the signal perfectly reveals firm

[8]For comparability, note that the normal distribution of the signal used in Example 1 could have been written as: $\tilde{s} = \mu + \sigma\tilde{\varepsilon}$, where $\tilde{\varepsilon}$ is normally distributed with mean 0 and variance 1.

value under current management. In practice, more stringent accounting regulations, or stricter listing and disclosure rules typically result in a more accurate assessment of the value of firm's assets, i.e., a higher γ.

Second, by construction, changes in the signal's precision do not affect the sensitivity of the signal s to μ, which is equal to 1. In section 3.3.2, we study the more general case when increases in γ can also increase the sensitivity of the signal s to μ.

The following Lemma describes the distribution of composite signal $\tilde{s}$ conditional on managerial ability.

Lemma 3. *The distribution of $\tilde{s}$ conditional on μ is normal with location parameter $\hat{\mu} \equiv \mu + \bar{z}$ and scale parameter $\sigma \equiv \sqrt{\gamma^2\sigma_z^2 + (1-\gamma)^2\sigma_\epsilon^2}$.*

Even though the board does not know managerial ability μ, the variability of the distribution described in Lemma 3 is an inverse measure of the informativeness of the signal with respect to managerial ability. Lemma 3 establishes that the distribution of the signal conditional on managerial ability μ is more variable when $\gamma^2\sigma_z^2 + (1-\gamma)^2\sigma_\epsilon^2$ is higher. The next Proposition draws the implications for the complementarity between signal precision and monitoring.

Proposition 5. *For any signal s, monitoring intensity is decreasing in γ if and only if $\sigma_z^2 < \frac{1-\gamma}{\gamma}\sigma_\epsilon^2$.*

Proposition 5 implies that the precision of firm value measurement (as measured by γ) and the intensity of monitoring are not always substitutes. Indeed, for $\sigma_z^2 > \frac{1-\gamma}{\gamma}\sigma_\epsilon^2$, an increase in γ increases the dispersion of the distribution of the signal conditional on managerial ability, as measured by σ in Lemma 3. This is a decrease in informativeness with respect to managerial ability. A signal which is more informative about firm value (higher γ) can thus paradoxically be less informative about managerial ability, even though managerial ability is one of the components of firm value. In this case, an increase in signal precision (higher γ) will increase the value and therefore the intensity of monitoring. The case with complementarity between the precision of firm value measurement and monitoring is depicted on the left panel of Figure 6, where $\sigma_z^2 > \frac{1-\gamma}{\gamma}\sigma_\epsilon^2$.

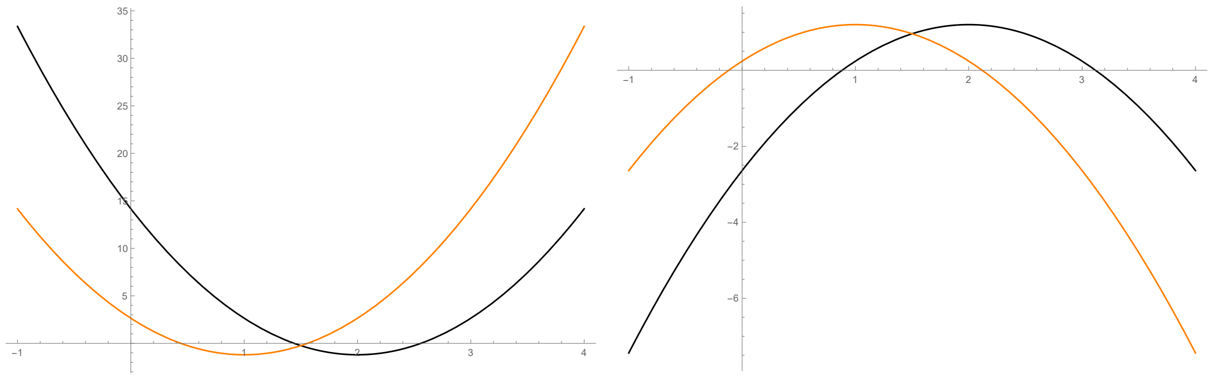

Figure 6: The functions $\frac{\pi'_\gamma(s|\mu_h)}{\pi(s|\mu_h)}$ (black) and $\frac{\pi'_\gamma(s|\mu_l)}{\pi(s|\mu_l)}$ (orange) as a function of s for parameters $\bar{z} = 1$, $\sigma_z = 1$, $\gamma = 0.5$, $\mu_h = 1$, and $\mu_l = 0$, and a normally distributed signal with location and shape parameters as in Lemma 3. We then have $s^T = s^* = 0.5$. On the left, $\sigma_\epsilon = 0.5$. On the right, $\sigma_\epsilon = 2$.

The baseline case presented in Example 1 roughly corresponds to the signal distribution in equation (9) for $\sigma_z = 0$ and $\sigma_\epsilon > 0$. For $\sigma_z = 0$, the signal is only informative about managerial ability; for $\sigma_\epsilon > 0$, an increase in γ makes the signal less noisy. In this case, there is always substitutability between the precision of firm value measurement (higher γ) and monitoring, i.e., the result in Proposition 5 for $\sigma_z = 0$ and $\sigma_\epsilon > 0$ is the same as in Proposition 1.

Proposition 5 is relevant when there is uncertainty about a component of firm value other than managerial ability. The main prediction is then that there is complementarity between improvements in firm value measurement and monitoring if $\sigma_z^2 > \frac{1-\gamma}{\gamma}\sigma_\epsilon^2$, i.e., if components of firm value other than managerial ability, such as the value of assets in place, are sufficiently variable (as measured by σ_z), and if the quality of the signal, as measured by γ and $\frac{1}{\sigma_\epsilon^2}$, is sufficiently high.[9] Intuitively, a higher quality signal is better at measuring the value of assets in place, which makes the signal s highly variable (with a high scale parameter σ in Lemma 3) when the value of assets in place is highly variable (high σ_z).

3.3.2 A more general signal distribution

We now allow changes in γ to also affect the sensitivity of the signal to managerial ability, to verify the robustness of the complementarity result in Proposition 5.

The signal distribution conditional on the realizations of μ and z of $\tilde{\mu}$ and $\tilde{z}$ is now

[9] A higher γ means that the signal better reflects the value of assets in place. A higher $\frac{1}{\sigma_\epsilon^2}$ means that valuation mistakes have a lower magnitude.

given by:

$$\tilde{s} = (1-\xi)\mu + \gamma\left(\xi\mu + z\right) + (1-\gamma)\left(\xi\bar{\mu} + \bar{z} + \tilde{\epsilon}\right), \tag{10}$$

where $\xi \in [0,1]$, $\bar{\mu} \equiv \frac{\mu_h+\mu_l}{2}$, and other variables and parameters are as in section 3.3.1. In this more general model, the sensitivity of s to μ is equal to $1-\xi+\gamma\xi$, i.e., the effect of γ on the sensitivity of s to μ is increasing in ξ. For example, for $\xi = 0$, the sensitivity of s to μ is equal to 1, which is independent of γ (this special case is analyzed in section 3.3.1); for $\xi = 1$, the sensitivity of s to μ is equal to γ.

Lemma 4. *The distribution of $\tilde{s}$ conditional on μ is normal with location parameter $\hat{\mu} \equiv (1-\xi+\gamma\xi)\mu + \bar{z} + (1-\gamma)\xi\bar{\mu}$ and scale parameter $\sigma \equiv \sqrt{\gamma^2\sigma_z^2 + (1-\gamma)^2\sigma_\epsilon^2}$.*

Lemma 4 shows that an increase in signal precision γ now increases the spread between the location parameters respectively associated with managers with ability $\mu = \mu_h$ and $\mu = \mu_l$. Via this new channel, increases in γ make the signal distribution more informative about managerial ability, which favors substitutability between signal precision and monitoring. However, the indirect effect of increases in γ unveiled in section 3.3.1, which can favor complementarity, still exists. Proposition 6 gives a condition for this indirect effect to dominate.

Proposition 6. *For any signal, if $\sigma_z^2 > \frac{1-\gamma}{\gamma}\sigma_\epsilon^2$ and ξ is sufficiently low, monitoring intensity is increasing in γ.*

Proposition 6 shows that the complementarity result from Proposition 5 still holds if $\sigma_z^2 > \frac{1-\gamma}{\gamma}\sigma_\epsilon^2$ (as in Proposition 5) *and* ξ is sufficiently low. This latter condition means that, following an increase in precision γ, improvements in the measurement of the value of assets in place is sufficiently important compared to improvements in the measurement of managerial ability. In this case, the effect of γ on the location parameter of the signal distribution (see Lemma 4) is still outweighed by the effect unveiled in section 3.3.1.

At the end of section 1, we noted that the model admits another interpretation, such that $\tilde{\mu}$ is the quality of a project rather than managerial ability, and both the signal and monitoring provide information about the quality of the firm's project. The results presented in this section can also be interpreted in this light.

4 Empirical Implications

4.1 Firm Performance and Termination Probability

The potential complementarity between the precision of the signal and monitoring has notable implications for the probability that the manager will be terminated following a bad signal:

Proposition 7. *When an increase in the precision of the signal increases monitoring, it also leads to a lower probability of termination after a bad signal ($s < s^T$), and a higher probability of termination after a good signal ($s \geq s^T$).*

This result is driven by the endogenous adjustment of monitoring. When the precision of the signal and monitoring are complements, an increase in signal precision increases monitoring intensity. As a result, the termination decision relies more on the information obtained via monitoring than on the signal realization, which reduces the probability that the manager will be terminated following a bad signal.[10]

4.2 Informed or Disciplinarian Boards

To derive additional empirical implications, we now extend the model by distinguishing between two distinct facets of monitoring:[11] the quality of the board's information on CEO ability, which depends on the cost of monitoring, $c(m)$, and the ability of the board to terminate a manager. So far, we have assumed that the firm's board can freely terminate the manager (at cost K). Now suppose that the board can only terminate a manager with probability ρ. A higher ρ indicates a more powerful board, or equivalently a less entrenched CEO. In practice, firms may face a tradeoff between having "informed" boards that can easily gather and analyze soft information about the CEO's ability on the job (i.e. boards with a lower monitoring cost), or "disciplinarian" boards that are more effective at terminating the CEO (i.e. boards with a higher ρ). To keep this extension simple, suppose that the firm can have two types: type i ("informed") has $\rho < 1$ and the usual monitoring cost function; type d ("disciplinarian") has $\rho = 1$ and $c(m) \to \infty \, \forall m > 0$.

Proposition 8. *For type i firms, Proposition 7 still holds. For type d firms, the probability of termination conditional on any given signal is independent of the precision of the signal. The probability of termination conditional on any given bad signal ($s < s^T$) is higher in type d firms.*

In disciplinarian firms, monitoring intensity does not depend on the signal precision, which is why the result in Proposition 7 does not hold in these firms.

These results allow to revisit the large empirical literature on CEO turnover and firm performance – which can be measured by the stock price or by accounting information such

[10]Consider the hypothetical case without monitoring (which corresponds to $m = 0$ in the model), a manager is automatically dismissed after a bad signal, i.e., there is a perfect correlation between managerial turnover and the signal realization. In the hypothetical case with perfect monitoring ($m = 1$), the board ignores the signal when deciding on manager termination.

[11]Note that this distinction complements the well-known distinction between the monitoring and advisory roles of the board (e.g. Faleye, Hoitash, and Hoitash (2011)).

as earnings (Engel, Hayes, and Wang (2003)). An influential paper (Huson, Parrino, and Starks (2001)) uses the relation between the likelihood of forced CEO turnover and firm performance as a measure of effectiveness of internal monitoring mechanisms. Consistent with this perspective, many recent empirical studies (e.g., Kaplan and Minton (2012), Jenter and Kanaan (2015), Jenter and Lewellen (2020)) focus on the sensitivity of CEO turnover to firm performance. In this perspective, the firm's board is a disciplinary mechanism that relates CEO turnover to firm performance, including by filtering out exogenous shocks to performance. Accordingly, a stronger sensitivity of CEO turnover to performance implies better performing boards. Proposition 7 suggests the opposite: when a board can freely dismiss a CEO, a more informed board will have a lower tendency to dismiss a CEO after a bad performance.

Proposition 8 distinguishes between informed boards and disciplinarian boards. It points out that a stronger sensitivity of CEO turnover to performance following a bad performance can be due to a higher ρ, i.e., a more disciplinarian board. In line with this result, Guo and Masulis (2015) find that firms forced to raise board independence increased the sensitivity of CEO turnover to performance, which they interpret as a positive corporate governance development. Our results suggest that this increased sensitivity can instead be explained by a diminution in monitoring by the board, and consequently worse CEO termination decision. A possibility is that forced increases in board independence result in the appointment of "tough" directors with limited ability for in-depth monitoring (e.g. due to lack of firm-specific or industry knowledge) who instead rely on performance indicators when deciding whether or not to terminate the CEO. This hypothesis is consistent with the empirical findings in Faleye, Hoitash, and Hoitash (2011), who find that firms where independent directors have more power exhibit greater sensitivity of CEO turnover to firm performance. However, it would call for a reinterpretation of their findings because they interpret this as an increase in monitoring quality,[12] whereas the analysis above suggests that this could alternatively reflect a board which is more disciplinarian but less informed.

4.3 Small Private Firms

Small privately owned firms usually face minimal accounting and disclosure requirements, and typically have not accumulated large amounts of assets in place. Instead, their valuation largely depends on their unrealized potential, including their business idea, and their management (Kaplan, Sensoy, and Strömberg (2009), Ewens and Marx (2018)).

In addition, these firms only fire CEOs following monitoring that reveals information

[12] This interpretation is commonplace in the literature, see for example Coles, Daniel, and Naveen (2014).

about the CEO's ability. Indeed, in an empirical study of board monitoring in a sample of private equity owned firms, Cornelli, Kominek, and Ljungqvist (2013) find that "boards in our data fire the CEO only in response to [behavior or decisions that would raise concerns about the CEO's ability] rather than for poor performance that was the result of bad luck or a decision that was wrong ex post but reasonable ex ante. Thus, the boards of the companies we study do not fire their CEOs by mistake." That is, a negative signal is not sufficient for CEO termination, which in our model is the case for $K \to p(\mu_h - \mu_l)$.[13] The empirical evidence is consistent with a high turnover cost (as a fraction of firm value) in these firms. Taylor (2010) finds that boards are more reluctant to fire a CEO in small firms, and Gao, Harford, and Li (2017) find that private firms have lower turnover rates than public firms.

Applying the model from section 3.1 with this high turnover cost yields the following empirical implications:

Empirical Implication 1. *In small private firms, the precision of firm value measurement and the intensity of monitoring are complements following a signal which is bad news about CEO ability ($s < s^*$).*

Empirical Implication 2. *In small private firms, more precise firm value measurement leads to: (i) a higher probability of termination after a signal which is bad news about CEO ability ($s < s^*$); (ii) and a lower probability of termination after a signal which is good news about CEO ability ($s > s^*$).*

Intuitively, an increase in signal precision means that a signal that leads to negative belief updating about the CEO's ability (a "negative signal") is all the more worrying about the manager's ability, which increases the value of monitoring that can lead to a replacement of the CEO. On the contrary, an increase in signal precision means that a positive signal is all the more reassuring about the manager's ability, which reduces the value of monitoring that could lead to CEO turnover. Consequently, private equity firms with more precise value measurement will monitor more intensively and possibly intervene following a negative signal. On the contrary, they will tend not to acquire

[13] Even though our model is stylized and not designed for calibration, it is worth asking whether the magnitude of the turnover cost associated with $K \to p(\mu_h - \mu_l)$ is reasonable in small private firms. To get a rough idea, consider the following simple calibration. First, suppose that a high ability manager improves firm value by 20% relative to a low ability manager – in a sample of buyout firms, Kaplan et al. (2012) find that a one unit change in a CEO's efficiency score increases the success probability by 20.8%. Second, suppose the proportion of high ability managers in the population is 0.2, which might be an overestimate given that the right tail of the ability distribution is very thin. Then $p\frac{\mu_h - \mu_l}{\mu_l} = 4\%$. Thus, we would have $K \to p(\mu_h - \mu_l)$ if the turnover cost as a fraction of firm value μ_l is close to 4%. Remember that the turnover cost is the overall cost of the transition to a new CEO, not just the severance package granted to the departing CEO, and tends to be higher in smaller firms (Taylor (2010)).

much additional information and will consequently generally not intervene following a positive signal (especially with low monitoring, termination following a positive signal is unlikely).[14] Overall, these predictions are in line with the finding in Cornelli, Kominek, and Ljungqvist (2013) that "soft information plays a much larger role in the board's decision to fire the CEO than does hard performance data."

The converse of part (i) of Empirical Implication 2 – that a negative signal generated by a less precise signal distribution leads to reduced monitoring and a lower probability of termination – is the bad corporate governance trap. This makes it all the more difficult to make an informed termination decision upon a bad signal, which in turn makes it harder for the firm to eventually improve its performance.

5 Conclusion

This paper contributes to the emerging research on the optimal combination of corporate governance mechanisms into a corporate governance system by providing a theory of the complementarity between the quality of firm value measurement and the intensity of board monitoring.

We show that, for the purpose of assessing managerial ability, a more informative signal of firm value or value creation can surprisingly increase the value of monitoring. This happens if the signal realization is bad news about managerial ability yet does not trigger termination, which can happen with a turnover cost. This also happens if a more precise signal (paradoxically) makes some signal realizations more ambiguous indicators of managerial ability, which can happen when the signal distribution is asymmetric. In these two cases, signal precision and monitoring are complements for a subset of signals – thus, interactions between the quality of the signal of firm value and monitoring are not only firm-specific but also signal-specific. Complementarity can also arise with a composite signal that measures other aspects of firm value in addition to managerial ability, such as the value of the firm's assets in place. Then, signal precision and monitoring can be complements when the value of assets in place is sufficiently variable.

The contribution of this paper is not only to point out that an improvement in the quality of firm value measurement can paradoxically make the soft information generated by monitoring more valuable, but also to show when this will be the case and what this entails. In small private equity owned firms, the quality of the signal of firm value and monitoring are typically substitutes following a positive signal, but complements following

[14] In the Supplementary Appendix, we show in Empirical Implication 3 that, even when the CEO is not terminated based on the signal s alone, the value of the right to terminate the CEO after a negative signal is higher when the signal is more precise. This is related to the performance-contingent allocation of control rights in many private equity owned firms (Kaplan and Strömberg (2003), Dessein (2005)).

a negative signal. These firms can fall into a bad governance trap: a deterioration in the quality of firm value measurement results in lower monitoring intensity following a negative signal, which compounds the degradation of the information environment in which managerial replacement decisions are made.

Future empirical studies could rely on these predictions, as well as on measures of board monitoring developed in Cornelli, Kominek, and Ljungqvist (2013), to develop a more thorough understanding of corporate governance practices, and their combination into a corporate governance system.

6 References

Acharya, V.V., Volpin, P.F., 2010. Corporate Governance Externalities. *Review of Finance*, 14 (1), 1-33.

Almazan, A., Banerji, S., de Motta, A., 2008. Attracting Attention: Cheap Managerial Talk and Costly Market Monitoring. *Journal of Finance*, 63, 1399-1436.

Armstrong, C.S., Guay, W.R., Weber, J.P., 2010. The role of information and financial reporting in corporate governance and debt contracting. *Journal of Accounting and Economics*, 50 (2-3), 179-234.

Arya, A., Glover, J., Mittendorf, B., Zhang, L., 2004. The Disciplining Role of Accounting in the Long-Run. *Review of Accounting Studies*, 9 (4), 399-417.

Azzalini, A., Capitanio, A.. 2014. *The skew-normal and related families.* Cambridge University Press.

Baldenius, T., Meng, X., Qiu, L., 2019. Biased Boards. *The Accounting Review*, 94 (2), 1-27.

Baldenius, T., Meng, X., Qiu, L., 2020. The Value of Board Commitment. *Review of Accounting Studies*, Forthcoming.

Banker, R.D., Darrough, M., Li, S., Threinen, L., 2019. The Value of Precontract Information About an Agent's Ability in the Presence of Moral Hazard and Adverse Selection. *Journal of Accounting Research*, 57, 1201-1245.

Bergemann, D., Morris, S., 2019. Information Design, *Journal of Economic Literature*, 57 (1), 44-95.

Bertomeu, J., Marinovic, I., 2016. A Theory of Hard and Soft Information. *The Accounting Review*, 91 (1), 1-20.

Börgers, T., Hernando-Veciana, A., Krähmer, D., 2013. When are signals complements or substitutes? *Journal of Economic Theory*, 148 (1), 165-195.

Brickley, J.A., 2003. Empirical research on CEO turnover and firm-performance: a discussion. *Journal of Accounting and Economics*, 36 (1-3), 227-233.

Bushman, R.M., Smith, A.J., 2001. Financial accounting information and corporate governance. *Journal of Accounting and Economics*, 32 (1-3), 237-333.

Bushman, R., Chen, Q., Engel, E., Smith, A., 2004. Financial accounting information, organizational complexity and corporate governance systems. *Journal of Accounting and Economics*, 37 (2), 167-201.

Chaigneau, P., Edmans, A., Gottlieb, D., 2018. Does improved information improve incentives? *Journal of Financial Economics*, 130 (2), 291-307.

Christensen, P.O., Frimor, H., Şabac, F., 2020. Real Incentive Effects of Soft Information. *Contemporary Accounting Research*, 37, 514-541.

Cianciaruso, D., Sridhar, S.S., 2018. Mandatory and Voluntary Disclosures: Dynamic

Interactions. *Journal of Accounting Research*, 56, 1253-1283.

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

Cornelli, F., Kominek, Z., Ljungqvist, A., 2013. Monitoring managers: does it matter? *Journal of Finance*, 68 (2), 431-481.

Crémer, J., 1995, Arm's length relationships. *Quarterly Journal of Economics*, 110, 275-295.

Dechow, P., Ge, W., Schrand, C., 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics*, 50 (2-3), 344-401.

Demougin, D., Fluet, C., 2001. Monitoring versus incentives. *European Economic Review*, 45 (9), 1741-1764.

Dessein, W., 2005. Information and Control in Ventures and Alliances. *Journal of Finance*, 60, 2513-2549.

Dicks, D.L., 2012. Executive Compensation and the Role for Corporate Governance Regulation. *Review of Financial Studies*, 25 (6), 1971-2004.

Drimyotes, G., 2007. The monitoring role of insiders. *Journal of Accounting and Economics*, 44, 359-377.

Donaldson, J.R., Malenko, N., Piacentino, G., 2020. Deadlock on the board. *Review of Financial Studies*, 33 (10), 4445-4488.

Dye, R.A., Sridhar, S.S., 2004. Reliability-Relevance Trade-Offs and the Efficiency of Aggregation. *Journal of Accounting Research*, 42, 51-88.

Eisfeldt, A.L., Kuhnen, C.M., 2013. CEO turnover in a competitive assignment framework, *Journal of Financial Economics*, 109 (2), 351-372.

Engel, E., Hayes, R.M., Wang, X., 2003. CEO turnover and properties of accounting information. *Journal of Accounting and Economics*, 36 (1-3), 197-226.

Ewens, M., Marx, M., 2018. Founder Replacement and Startup Performance. *Review of Financial Studies*, 31(4), 1532-1565.

Faleye, O., Hoitash, R., Hoitash, U., 2011. The costs of intense board monitoring. *Journal of Financial Economics*, 101 (1), 160-181.

Farrell, K.A., Whidbee, D.A., 2003. Impact of firm performance expectations on CEO turnover and replacement decisions. *Journal of Accounting and Economics*, 36 (1-3), 165-196.

Gao, H., Harford, J., Li, K., 2017. CEO Turnover-Performance Sensitivity in Private Firms. *Journal of Financial and Quantitative Analysis*, 52 (2), 583-611.

Gigler, F., Hemmer, T., 1998. On the Frequency, Quality, and Informational Role of Mandatory Financial Reports. *Journal of Accounting Research*, 36, 117-147.

Gigler F.B., Hemmer, T., 2001. Conservatism, Optimal Disclosure Policy, and the Timeliness of Financial Reports. *The Accounting Review*, 76 (4), 471-493.

Guo, L., Masulis, R.W., 2015. Board Structure and Monitoring: New Evidence from CEO Turnovers. *Review of Financial Studies*, 28 (10), 2770-2811.

Harris M., Holmström, B., 1982. A theory of wage dynamics. *Review of Economic Studies*, 49 (3), 315-333.

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

Hermalin, B.E., Weisbach, M., 1998. Endogenously chosen boards of directors and their monitoring of the manager. *American Economic Review*, 88, 96-118.

Hermalin, B.E., Weisbach, M.S., 2012. Information Disclosure and Corporate Governance. *Journal of Finance*, 67, 195-233.

Hermalin, B.E., Weisbach, M., 2017. Chapter 3 - Assessing Managerial Ability: Implications for Corporate Governance. *Handbook of the Economics of Corporate Governance*, 1, 93-176.

Huson, M.R., Parrino, R., Starks, L.T., 2001. Internal Monitoring Mechanisms and CEO Turnover: A Long-Term Perspective. *Journal of Finance*, 56, 2265-2297.

Jenter, D., Kanaan, F., 2015. manager Turnover and Relative Performance Evaluation. *Journal of Finance*, 70, 2155-2184.

Jenter, D., Lewellen, K., 2020. Performance-Induced CEO Turnover. *Review of Financial Studies*, Forthcoming.

Jiang, X., 2016. Biases in Accounting and Nonaccounting Information: Substitutes or Complements? *Journal of Accounting Research*, 54, 1297-1330.

Kaplan, S.N., Klebanov, M.M. and Sorensen, M., 2012. Which CEO Characteristics and Abilities Matter? *Journal of Finance*, 67, 973-1007.

Kaplan, S.N., Minton, B.A. 2012. How Has manager Turnover Changed? *International Review of Finance*, 12: 57-87.

Kaplan, S.N., Sensoy, B.A. and Strömberg, P., 2009. Should Investors Bet on the Jockey or the Horse? Evidence from the Evolution of Firms from Early Business Plans to Public Companies. *Journal of Finance*, 64, 75-115.

Kaplan, S.N., Strömberg, P., 2003. Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts. *Review of Economic Studies* 70 (2), 281-315.

Kumar, P., Sivaramakrishnan, K. 2008. Who Monitors the Monitor? The Effect of Board Independence on Executive Compensation and Firm Value. *Review of Financial Studies*, 21 (3), 1371-1401.

Laux, V., 2014. Corporate Governance, Board Oversight, and manager Turnover. *Foundations and Trends in Accounting*, 8 (1), 1-73.

Leuz, C., Wysocki, P.D., 2016. The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research. *Journal of Accounting Research*, 54, 525-622.

Levit, D., 2018. Expertise, Structure, and Reputation of Corporate Boards. Working paper, Wharton.

Levit, D., Malenko, N., 2016. The Labor Market for Directors and Externalities in Corporate Governance. *Journal of Finance*, 71 (2), 775-808.

Liberti, J.M., Petersen, M.A., 2019. Information: Hard and Soft. *Review of Corporate Finance Studies*, 8 (1), 1-41.

Manso, G., 2011. Motivating Innovation. *Journal of Finance*, 66 (5), 1823-1860.

Misangyi, V.F., Acharya, A.G., 2014. Substitutes or Complements? A Configurational Examination of Corporate Governance Mechanisms. *Academy of Management Journal*, 57, 1681-1705.

Murphy, K.J., Zimmerman, J.L., 1993. Financial performance surrounding CEO turnover. *Journal of Accounting and Economics*, 16 (1-3), 273-315.

Peng, L., Röell, A., 2014. Managerial Incentives and Stock Price Manipulation. *Journal of Finance*, 69 (2), 487-526.

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

Raith, M., 2003. Competition, Risk, and Managerial Incentives. *American Economic Review*, 93 (4), 1425-1436.

Şabac, F., Tian, J., 2015. On the Stewardship Value of Soft Managerial Reports. *The Accounting Review*, 90 (4), 1683-1706.

Schroth, J., 2018. Managerial Compensation and Stock Price Manipulation. *Journal of Accounting Research*, 56, 1335-1381.

Taylor, L.A., 2010. Why Are CEOs Rarely Fired? Evidence from Structural Estimation. *Journal of Finance*, 65, 2051-2087.

Tirole, J., 2006. *The Theory of Corporate Finance.* Princeton University Press, Princeton, NJ.

Von Thadden, E.-L., 1995. Long-term contracts, short-term investment and monitoring. *Review of Economic Studies*, 62 (4), 557-575.

Ward, A.J., Brown, J.A., Rodriguez, D., 2009. Governance Bundles, Firm Performance, and the Substitutability and Complementarity of Governance Mechanisms. *Corporate Governance: An International Review*, 17 646-660.

7 Appendix

Proof of Lemma 1:

From equation (1), $p_s = p$ at $s = s^*$. It follows from the termination condition in equation (2) and from p_s being increasing in s due to MLRP that $s^T \leq s^*$.

In addition, with $\mu_h > \mu_l$, p_s is strictly decreasing in K. It follows from the definition of s^T, the termination condition in equation (2), and p_s being increasing in s due to MLRP that s^T is strictly decreasing in K. ■

Proof of Lemma 2:

When a bad signal ($s < s^T$) is observed, taking the total derivative of equation (3), an equilibrium condition, with respect to the parameter σ, we get:

$$\frac{\partial m_s^*}{\partial \sigma} = \frac{1}{c''(m_s^*)} \frac{\partial p_s}{\partial \sigma} \left[(1-p)\Delta + K\right], \tag{11}$$

where $c''(m_s^*) > 0$ Therefore, the effect of σ on m_s^* has the same sign as the effect of σ on p_s.

When a good signal ($s \geq s^T$) is observed, taking the total derivative of equation (5), an equilibrium condition, with respect to the parameter σ, we get:

$$\frac{\partial m_s^*}{\partial \sigma} = -\frac{1}{c''(m_s^*)} \frac{\partial p_s}{\partial \sigma} \left[p\Delta - K\right], \tag{12}$$

where $c''(m_s^*) > 0$ and $K < p\Delta$ by assumption. Therefore, the effect of σ on m_s^* has the same sign as the effect of σ on $(-p_s)$. ■

Proof of Proposition 1:

Following signal s, from equation (1) we have:

$$\begin{aligned}\frac{\partial p_s}{\partial \sigma} &= \frac{p\pi'_\sigma(s|\mu_h)\left(p\pi(s|\mu_h) + (1-p)\pi(s|\mu_l)\right) - \left(p\pi'_\sigma(s|\mu_h) + (1-p)\pi'_\sigma(s|\mu_l)\right)p\pi(s|\mu_h)}{\left(p\pi(s|\mu_h) + (1-p)\pi(s|\mu_l)\right)^2} \\ &= p(1-p)\frac{\pi'_\sigma(s|\mu_h)\pi(s|\mu_l) - \pi'_\sigma(s|\mu_l)\pi(s|\mu_h)}{\left(p\pi(s|\mu_h) + (1-p)\pi(s|\mu_l)\right)^2}, \end{aligned} \tag{13}$$

where $\pi'_\sigma(\cdot)$ denotes the derivative of $\pi(\cdot)$ with respect to σ.

From equation (13) and Lemma 2, a decrease in σ reduces monitoring intensity following a bad signal ($s < s^T$) if and only if:

$$\pi'_\sigma(s|\mu_h)\pi(s|\mu_l) - \pi'_\sigma(s|\mu_l)\pi(s|\mu_h) \leq 0 \tag{14}$$

From equation (13) and Lemma 2, a decrease in σ reduces monitoring intensity following a good signal ($s \geq s^T$) if and only if:

$$\pi'_\sigma(s|\mu_h)\pi(s|\mu_l) - \pi'_\sigma(s|\mu_l)\pi(s|\mu_h) \geq 0 \tag{15}$$

$\blacksquare$

Proof of Proposition 2:

With a distribution with location and scale parameters:

$$\begin{aligned}\frac{\pi'_\sigma\left(s|\mu\right)}{\pi\left(s|\mu\right)} &= -\frac{\frac{1}{\sigma^2}f\left(\frac{s-\mu}{\sigma}\right)+\frac{s-\mu}{\sigma^3}f'\left(\frac{s-\mu}{\sigma}\right)}{\frac{1}{\sigma}f\left(\frac{s-\mu}{\sigma}\right)} \\ &= -\frac{1}{\sigma}\left[1+\frac{s-\mu}{\sigma}\frac{f'\left(\frac{s-\mu}{\sigma}\right)}{f\left(\frac{s-\mu}{\sigma}\right)}\right] \qquad (16)\end{aligned}$$

For distributions with location and scale parameters as defined in equation (8), the assumed MLRP means:

$$\frac{\partial}{\partial s}\left\{\frac{\pi(s|\mu_h)}{\pi(s|\mu_l)}\right\}=\frac{\partial}{\partial s}\left\{\frac{f\left(\frac{s-\mu_h}{\sigma}\right)}{f\left(\frac{s-\mu_l}{\sigma}\right)}\right\}>0 \;\Leftrightarrow\; \frac{f'\left(\frac{s-\mu_h}{\sigma}\right)f\left(\frac{s-\mu_l}{\sigma}\right)-f\left(\frac{s-\mu_h}{\sigma}\right)f'\left(\frac{s-\mu_l}{\sigma}\right)}{\left(f\left(\frac{s-\mu_l}{\sigma}\right)\right)^2}>0$$

$$\Leftrightarrow\; \frac{f'\left(\frac{s-\mu_h}{\sigma}\right)}{f\left(\frac{s-\mu_l}{\sigma}\right)}>\frac{f\left(\frac{s-\mu_h}{\sigma}\right)}{f\left(\frac{s-\mu_l}{\sigma}\right)}\frac{f'\left(\frac{s-\mu_l}{\sigma}\right)}{f\left(\frac{s-\mu_l}{\sigma}\right)} \;\Leftrightarrow\; \frac{f'\left(\frac{s-\mu_h}{\sigma}\right)}{f\left(\frac{s-\mu_h}{\sigma}\right)}>\frac{f'\left(\frac{s-\mu_l}{\sigma}\right)}{f\left(\frac{s-\mu_l}{\sigma}\right)}$$

Since $\mu_h > \mu_l$, the last inequality implies that we also have, for $s_1 > s_0$:

$$-\frac{f'\left(\frac{s_1-\mu}{\sigma}\right)}{f\left(\frac{s_1-\mu}{\sigma}\right)}>-\frac{f'\left(\frac{s_0-\mu}{\sigma}\right)}{f\left(\frac{s_0-\mu}{\sigma}\right)} \qquad (17)$$

That is, with a distribution that satisfies MLRP, $-\frac{f'\left(\frac{s-\mu}{\sigma}\right)}{f\left(\frac{s-\mu}{\sigma}\right)}$ is strictly increasing in s. For a symmetric single-peaked distribution, it is also equal to 0 at $s=\mu$. Therefore, the expression in equation (16) is strictly decreasing in s for $s<\mu$, and strictly increasing in s for $s>\mu$. Moreover, increasing μ results in a rightward translation of the expression in equation (16). This implies that there exists $s=\hat{s}$ such that for $s<\hat{s}$, we have:

$$\frac{\pi'_\sigma\left(s|\mu_h\right)}{\pi\left(s|\mu_h\right)} \geq \frac{\pi'_\sigma\left(s|\mu_l\right)}{\pi\left(s|\mu_l\right)}, \qquad (18)$$

and that for $s>\hat{s}$, we have:

$$\frac{\pi'_\sigma\left(s|\mu_h\right)}{\pi\left(s|\mu_h\right)} \leq \frac{\pi'_\sigma\left(s|\mu_l\right)}{\pi\left(s|\mu_l\right)}. \qquad (19)$$

With a symmetric distribution, $\hat{s}=\frac{\mu_l+\mu_h}{2}$.

From equation (2) we know that, following unsuccessful monitoring, the manager is terminated if and only if:

$$p_s < p-\frac{K}{\mu_h-\mu_l}\equiv p_{s^T}. \qquad (20)$$

The updated probability p_s, as defined in equation (1), is equal to p if and only if:

$$\frac{p\pi\left(s|\mu_h\right)}{p\pi\left(s|\mu_h\right)+(1-p)\,\pi\left(s|\mu_l\right)}=p \quad \Leftrightarrow \quad \frac{\pi\left(s|\mu_h\right)}{p\pi\left(s|\mu_h\right)+(1-p)\,\pi\left(s|\mu_l\right)}=1$$
$$\Leftrightarrow \quad p+(1-p)\,\frac{\pi\left(s|\mu_l\right)}{\pi\left(s|\mu_h\right)}=1 \quad \Leftrightarrow \quad \pi\left(s|\mu_l\right)=\pi\left(s|\mu_h\right)$$

Thus, we have that $p_s = p$ if and only if $s = s^*$, where s^* is by definition such that $\pi\left(s^*|\mu_l\right) \equiv \pi\left(s^*|\mu_h\right)$ and is unique by MLRP. We already know that p_s is strictly increasing in s (due to MLRP), so that $s^T < s^*$ if and only if $p_{s^T} < p$, i.e., if and only if $K > 0$.

We now establish the second part of the Proposition. From the definition of p_{s^T} in equation (20), p_{s^T} is strictly decreasing in K, all else equal. Moreover, since p_s is strictly increasing in s, we also have that s^T is strictly decreasing in K. The unconditional probability that $s \in [s^T, s^*]$ is:

$$p\left[\int_{s^T}^{s^*}\pi\left(s|\mu_h\right)ds\right]+(1-p)\left[\int_{s^T}^{s^*}\pi\left(s|\mu_l\right)ds\right], \tag{21}$$

where $s^* = 0.5$ for a symmetric distribution. Since K only affects s^T in equation (21), and since s^T is strictly decreasing in K and $\pi\left(s|\mu\right) > 0$ for any s because of the full support assumption, an increase in K strictly increases the expression in equation (21): the probability that $s \in [s^T, s^*]$ is strictly increasing in K. Moreover, at $K = 0$, we have $s^T = s^* = 0.5$, and this probability is zero. As $K \to p(\mu_h - \mu_l)$, from equation (20) we have $p_{s^T} \to 0$. That is, from the definition of p_{s^T}, the manager is never terminated in the limited, i.e., $s^T \to -\infty$. ■

Proof of Proposition 3:

On the one hand, for $K = 0$, by Lemma 1 we have $s^T = s^*$. For a single-peaked symmetric distribution, we have $s < s^*$ if and only if $\pi\left(s|\mu_h\right) < \pi\left(s|\mu_l\right)$, and by definition s^* is such that $\pi\left(s^*|\mu_h\right) = \pi\left(s^*|\mu_l\right)$. For symmetric distributions, this implies $s^T = s^* = \frac{\mu_l+\mu_h}{2}$.

On the other hand, we have:

$$\frac{\pi'_\sigma\left(s|\mu\right)}{\pi\left(s|\mu\right)}=-\frac{\frac{1}{\sigma^2}f\left(\frac{s-\mu}{\sigma}\right)+\frac{s-\mu}{\sigma^3}f'\left(\frac{s-\mu}{\sigma}\right)}{\frac{1}{\sigma}f\left(\frac{s-\mu}{\sigma}\right)}=-\frac{1}{\sigma}\left[1+\frac{s-\mu}{\sigma}\frac{f'\left(\frac{s-\mu}{\sigma}\right)}{f\left(\frac{s-\mu}{\sigma}\right)}\right].$$

For a single-peaked symmetric distribution, $\frac{\pi'_\sigma(s|\mu)}{\pi(s|\mu)} = 0$ if and only if $s = \mu$, and $-f'\left(\frac{s-\mu}{\sigma}\right)/f\left(\frac{s-\mu}{\sigma}\right)$ is strictly increasing in s by equation (17) and equal to zero at $s = \mu$. This implies that $\frac{\pi'_\sigma(s|\mu)}{\pi(s|\mu)}$ is decreasing in s for $s < \mu$ and increasing in s for $s > \mu$. Moreover, for symmetric distributions, $\frac{\pi'_\sigma(s|\mu_h)}{\pi(s|\mu_h)}$ is a rightward translation of $\frac{\pi'_\sigma(s|\mu_l)}{\pi(s|\mu_l)}$ since μ is a location parameter, and $\frac{\pi'_\sigma(s|\mu_h)}{\pi(s|\mu_h)} = \frac{\pi'_\sigma(s|\mu_l)}{\pi(s|\mu_l)}$ at $s = \hat{s} = \frac{\mu_l+\mu_h}{2}$.

In sum:

$$\frac{\pi'_\sigma(s|\mu_h)}{\pi(s|\mu_h)} > \frac{\pi'_\sigma(s|\mu_l)}{\pi(s|\mu_l)} \qquad \text{for } s < s^T,$$
$$\frac{\pi'_\sigma(s|\mu_h)}{\pi(s|\mu_h)} = \frac{\pi'_\sigma(s|\mu_l)}{\pi(s|\mu_l)} \qquad \text{at } s = \hat{s} = s^T,$$
$$\frac{\pi'_\sigma(s|\mu_h)}{\pi(s|\mu_h)} < \frac{\pi'_\sigma(s|\mu_l)}{\pi(s|\mu_l)} \qquad \text{for } s > s^T.$$

Applying Proposition 1 gives the result. ■

Skew-normal distribution and MLRP:

We need to show that:

$$\frac{\partial}{\partial s}\left(\frac{\pi(s|\mu_l)}{\pi(s|\mu_h)}\right) \leq 0 \qquad \Leftrightarrow \qquad \frac{\partial}{\partial s}\left(\frac{\phi\left(\frac{s-\mu_l}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)}{\phi\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right)}\right) \leq 0.$$

We have:

$$\begin{aligned}
&\frac{\partial}{\partial s}\left(\frac{\phi\left(\frac{s-\mu_l}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)}{\phi\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right)}\right) \\
= \; & \frac{\left[-\frac{(s-\mu_l)}{\sigma}\phi\left(\frac{s-\mu_l}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right) + \frac{\alpha}{\sigma}\phi\left(\frac{s-\mu_l}{\sigma}\right)\phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\right)\right]\left(\phi\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right)\right)}{\left(\phi\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right)\right)^2} \\
& - \frac{\left[-\frac{(s-\mu_h)}{\sigma}\phi\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right) + \frac{\alpha}{\sigma}\phi\left(\frac{s-\mu_h}{\sigma}\right)\phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right)\right)\right]\left(\phi\left(\frac{s-\mu_l}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\right)}{\left(\phi\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right)\right)^2} \\
= \; & \frac{\left[\left(\frac{\mu_l-\mu_h}{\sigma}\right)\phi\left(\frac{s-\mu_l}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\phi\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right)\right]}{\left(\phi\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right)\right)^2} \\
& + \frac{\frac{\alpha}{\sigma}\phi\left(\frac{s-\mu_l}{\sigma}\right)\phi\left(\frac{s-\mu_h}{\sigma}\right)\left[\phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right) - \phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\right)\right]}{\left(\phi\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right)\right)^2}
\end{aligned}$$

The denominator is positive. We thus need to determine the sign of the numerator.

When $\alpha = 0$, the first term in the numerator simplifies to $\left[\left(\frac{\mu_l-\mu_h}{\sigma}\right)\phi\left(\frac{s-\mu_l}{\sigma}\right)\phi\left(\frac{s-\mu_h}{\sigma}\right)/4\right]$ and the second term is equal to zero. The sign of the expression depends on the sign $\mu_l - \mu_h < 0$. This corresponds to the MLRP for a normal distribution.

When $\alpha > 0$, the first term is of the sign of $\mu_l - \mu_h < 0$. The second term is of the sign of $\phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right) - \phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\right)$. MLRP of the normal distribution implies that $\phi\left(\frac{s-\mu_l}{\sigma}\right)/\Phi\left(\left(\frac{s-\mu_l}{\sigma}\right)\right) \leq \phi\left(\frac{s-\mu_h}{\sigma}\right)/\Phi\left(\left(\frac{s-\mu_h}{\sigma}\right)\right)$. This in turn, implies that $\phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right) - \phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\right) \leq 0$ and thus that the skew-normal follows MLRP when $\alpha > 0$.

When $\alpha < 0$, $\phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right) = \phi\left(\alpha\left(\frac{\mu_l - s}{\sigma}\right)\right)$. The sign of $\phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\Phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\right) - \phi\left(\alpha\left(\frac{s-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s-\mu_l}{\sigma}\right)\right)\right)$ is now positive but it is mulitplied by $\alpha < 0$. Overall, the numerator is also negative and MLRP still obtains. ■

Proof of Proposition 4:

We prove (i) by showing that for any $\alpha > 0$ $\mu_l < \hat{s} < \mu_h$, while $s^* > \mu_h$ for α large enough.

The PDF of a skew-normal distribution with shape parameter α location parameter μ, and scale parameter σ is given by:

$$\pi(s|\mu) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\alpha\left(\frac{x-\mu}{\sigma}\right)\right),$$

where $\phi(x)$ and $\Phi(x)$ are the PDF and CDF of the normal distribution:

$$\begin{aligned}
\phi(x) &= \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \\
\Phi(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx.
\end{aligned}$$

Condition (*) can be written as:

$$\frac{2}{\sigma}\phi\left(\frac{s^*(\alpha)-\mu_l}{\sigma}\right)\Phi\left(\alpha\left(\frac{s^*(\alpha)-\mu_l}{\sigma}\right)\right) = \frac{2}{\sigma}\phi\left(\frac{s^*(\alpha)-\mu_h}{\sigma}\right)\Phi\left(\alpha\left(\frac{s^*(\alpha)-\mu_h}{\sigma}\right)\right).$$

The PDF of the two distributions are single-peaked and skewed. Further, the PDF with location parameter μ_h is just right transalation of the PDF with location parameter μ_l. The PDFs cross at point s^* which is located in between the peaks (modes) $m(\mu,\alpha)$ of the distribution. Thus $\mu_l < m(\mu_l,\alpha) < s^* < m(\mu_h,\alpha)$. We now show that $s^*(\alpha) > \mu_h$ when α is large. Taking the limit when α goes to infinity, we have

$$\begin{aligned}
\lim_{\alpha\to\infty}\phi\left(\frac{s^*(\alpha)-\mu_l}{\sigma}\right)\Phi\left(\alpha\left(\frac{s^*(\alpha)-\mu_l}{\sigma}\right)\right) &= \lim_{\alpha\to\infty}\phi\left(\frac{s^*(\alpha)-\mu_l}{\sigma}\right) \\
\text{as } \lim_{\alpha\to\infty}\Phi\left(\alpha\left(\frac{s^*(\alpha)-\mu_l}{\sigma}\right)\right) &= \lim_{x\to\infty}\Phi(x) = 1 \text{ given that } s^*(\alpha) > \mu_l.
\end{aligned}$$

Suppose $s^*(\alpha) < \mu_h$, when α is large we have $\Phi\left(\alpha\left(\frac{s^*(\alpha)-\mu_h}{\sigma}\right)\right)$ is close to 0, which leads to an immediate contradiction as condition (*) cannot be satisified in that case.. Thus, $s^*(\alpha) > \mu_h$.

We now turn to $\hat{s}$. We have:

$$\begin{aligned}
\pi_{\sigma}^{'}(s|\mu) &= \frac{\partial \pi (s|\mu)}{\partial \sigma} \\
&= \frac{-2}{\sigma^2}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\alpha\left(\frac{x-\mu}{\sigma}\right)\right) - \frac{2(x-\mu)}{\sigma^3}\phi'\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\alpha\left(\frac{x-\mu}{\sigma}\right)\right) \\
& \quad -\frac{2\alpha (x-\mu)}{\sigma^3}\phi\left(\alpha\frac{x-\mu}{\sigma}\right)\phi\left(\frac{x-\mu}{\sigma}\right) \\
&= \frac{-2}{\sigma^2}\left(\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\alpha\left(\frac{x-\mu}{\sigma}\right)\right) + \frac{(x-\mu)}{\sigma}\phi'\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\alpha\left(\frac{x-\mu}{\sigma}\right)\right)\right. \\
& \quad \left. +\frac{\alpha (x-\mu)}{\sigma}\phi\left(\alpha\frac{x-\mu}{\sigma}\right)\phi\left(\frac{x-\mu}{\sigma}\right)\right)
\end{aligned}$$

Then:

$$\begin{aligned}
\frac{\pi_{\sigma}^{'}(s|\mu)}{\pi (s|\mu)} &= -\frac{1}{\sigma}\left(1+\frac{(x-\mu)}{\sigma}\frac{\phi'\left(\frac{x-\mu}{\sigma}\right)}{\phi\left(\frac{x-\mu}{\sigma}\right)} + \frac{\alpha (x-\mu)}{\sigma}\frac{\phi\left(\alpha\frac{x-\mu}{\sigma}\right)}{\Phi\left(\alpha\left(\frac{x-\mu}{\sigma}\right)\right)}\right) \\
&= -\frac{1}{\sigma}\left(1-\frac{1}{\sigma}\left(\frac{(x-\mu)}{\sigma}\right)^2 + \frac{\alpha (x-\mu)}{\sigma}\frac{\phi\left(\alpha\frac{x-\mu}{\sigma}\right)}{\Phi\left(\alpha\left(\frac{x-\mu}{\sigma}\right)\right)}\right)
\end{aligned}$$

The function $\frac{\pi_{\sigma}^{'}(s|\mu)}{\pi(s|\mu)}$ is U-shaped. The minimum is reached at $s_{\min}(\mu)$:

$$\frac{\partial\left(\frac{\pi_{\sigma}^{'}(s|\mu)}{\pi(s|\mu)}\right)}{\partial s} = \frac{2(s-\mu)}{\sigma^4} - \frac{\alpha}{\sigma^2}\frac{\phi\left(\alpha\frac{s-\mu}{\sigma}\right)}{\Phi\left(\alpha\left(\frac{s-\mu}{\sigma}\right)\right)} - \frac{\alpha (s-\mu)}{\sigma^2}\partial\left(\frac{\phi\left(\alpha\frac{s-\mu}{\sigma}\right)}{\Phi\left(\alpha\left(\frac{s-\mu}{\sigma}\right)\right)}\right)/\partial s = 0.$$

Note that

$$\frac{\partial\left(\frac{\pi_{\sigma}^{'}(s|\mu)}{\pi(s|\mu)}\right)}{\partial \mu} = -\frac{2(s-\mu)}{\sigma^4} + \frac{\alpha}{\sigma^2}\frac{\phi\left(\alpha\frac{s-\mu}{\sigma}\right)}{\Phi\left(\alpha\left(\frac{s-\mu}{\sigma}\right)\right)} + \frac{\alpha (s-\mu)}{\sigma^2}\partial\left(\frac{\phi\left(\alpha\frac{s-\mu}{\sigma}\right)}{\Phi\left(\alpha\left(\frac{s-\mu}{\sigma}\right)\right)}\right)/\partial s$$

This expression is equal to 0 at $s_{\min}(\mu)$ is negative for $s \leq s_{\min}(\mu)$ and is positive for $s \geq s_{\min}(\mu)$. This means that functions $\frac{\pi_{\sigma}^{'}(s|\mu)}{\pi(s|\mu)}$ for two different means are single crossing. We can rewrite the crossing condition as:

$$\frac{(x-\mu_l)}{(x-\mu_h)} = \frac{\alpha\frac{\phi\left(\alpha\frac{x-\mu_h}{\sigma}\right)}{\Phi\left(\alpha\frac{x-\mu_h}{\sigma}\right)} - \frac{(x-\mu_h)}{\sigma^2}}{\alpha\frac{\phi\left(\alpha\frac{x-\mu_l}{\sigma}\right)}{\Phi\left(\alpha\frac{x-\mu_l}{\sigma}\right)} - \frac{(x-\mu_l)}{\sigma^2}}.$$

The left-hand side is decreasing and smaller than 1 for $x < \mu_h$ and is decreasing and larger than 1 for $x > \mu_h$. Similarly, the right hand side is increasing and larger than 1 for $x < y < \mu_l$ with $\alpha\frac{\phi\left(\alpha\frac{y-\mu_l}{\sigma}\right)}{\Phi\left(\alpha\frac{y-\mu_l}{\sigma}\right)} - \frac{(y-\mu_l)}{\sigma^2} = 0$, and increasing and smaller than 1 for $x > y$. The crossing point is thus in $[\mu_l, \mu_h]$.

To prove (ii), we consider the peak (mode) of the PDF of the distribution with location parameter μ_l. We have that $m(\mu_l, \alpha) > \mu_l$ (see Azzalini and Capitanio (2014) pp. 32–33 for an approximation of the value of the mode). As $s^* > m(\mu_l, \alpha)$ and $\hat{s} < \mu_h$, setting $\mu_h < m(\mu_l, \alpha)$ implies that $\hat{s} < s^*$ and complementarity ensues. ■

Proof of Lemma 3:

Conditional on μ, the distribution of the composite signal y is normal. Its location parameter is:

$$\begin{aligned} \mathbb{E}[\tilde{s}|\mu] &= \mathbb{E}\left[\tilde{\mu} + \gamma\tilde{z} + (1-\gamma)(\bar{z} + \tilde{\epsilon})|\mu\right] \\ &= \mathbb{E}\left[\tilde{\mu}|\mu\right] + \gamma\mathbb{E}\left[\tilde{z}|\mu\right] + (1-\gamma)\left(\bar{z} + \mathbb{E}\left[\tilde{\epsilon}|\mu\right]\right) \\ &= \mu + \gamma\bar{z} + (1-\gamma)\bar{z} \\ &= \mu + \bar{z}. \end{aligned}$$

Its variance is:

$$\begin{aligned} var(\tilde{s}|\mu) &= var\left(\tilde{\mu} + \gamma\tilde{z} + (1-\gamma)(\bar{z} + \tilde{\epsilon})|\mu\right) \\ &= var(\tilde{\mu}|\mu) + \gamma^2 var(\tilde{z}|\mu) + (1-\gamma)^2 var(\tilde{\epsilon}|\mu) \\ &= \gamma^2 var(\tilde{z}) + (1-\gamma)^2 var(\tilde{\epsilon}) \\ &= \gamma^2\sigma_z^2 + (1-\gamma)^2\sigma_\epsilon^2. \end{aligned}$$

where we used independence between random variables, and $var(\tilde{\mu}|\mu) = 0$. The scale parameter of a normal distribution is the square root of its variance. ■

Proof of Proposition 5:

Using Bayesian updating, the updated probability p_s that the manager has high ability after the signal s is observed is:

$$p_s = \frac{p\pi(s|\mu_h)}{p\pi(s|\mu_h) + (1-p)\pi(s|\mu_l)} \tag{22}$$

which is increasing in s by MLRP. With $K = 0$, as in equation (2) the manager is terminated if and only if $p_s < p$. Let s^T be such that $p_{s^T} = p$. As in Lemma 1 with $K = 0$, the manager is terminated if and only if $s < s^*$, where s^* is defined by $\pi(s^*|\mu_l) = \pi(s^*|\mu_h)$. Since the variable z is independent of managerial ability, monitoring intensity m_s^* is as in section 2.2 with $K = 0$. The comparative statics of monitoring intensity m_s^* with respect to parameter γ are consequently similar to the proof of Lemma 2. Specifically, when a bad signal ($s < s^T$) is observed:

$$\frac{\partial m_s^*}{\partial\gamma} = \frac{1}{c''(m_s^*)}\frac{\partial p_s}{\partial\gamma}(1-p)\Delta, \tag{23}$$

where $c''(m_s^*) > 0$. Therefore, the effect of γ on m_s^* has the same sign as the effect of γ on p_s. When a good signal ($s \geq s^T$) is observed:

$$\frac{\partial m_s^*}{\partial \gamma} = -\frac{1}{c''(m_s^*)}\frac{\partial p_s}{\partial \gamma} p\Delta, \tag{24}$$

where $c''(m_s^*) > 0$. Therefore, the effect of γ on m_s^* has the same sign as the effect of γ on $(-p_s)$.

We can proceed along the same lines as the proof of Proposition 1: Following signal s, from equation (22) we have:

$$\begin{aligned}
\frac{\partial p_s}{\partial \gamma} &= \frac{p\pi'_\gamma(s|\mu_h)\left(p\pi(s|\mu_h) + (1-p)\pi(s|\mu_l)\right)}{\left(p\pi(s|\mu_h) + (1-p)\pi(s|\mu_l)\right)^2} \\
&\quad - \frac{\left(p\pi'_\gamma(s|\mu_h) + (1-p)\pi'_\gamma(s|\mu_l)\right)p\pi(s|\mu_h)}{\left(p\pi(s|\mu_h) + (1-p)\pi(s|\mu_l)\right)^2} \\
&= p(1-p)\frac{\pi'_\gamma(s|\mu_h)\pi(s|\mu_l) - \pi'_\gamma(s|\mu_l)\pi(s|\mu_h)}{\left(p\pi(s|\mu_h) + (1-p)\pi(s|\mu_l)\right)^2},
\end{aligned} \tag{25}$$

where $\pi'_\gamma(\cdot)$ denotes the derivative of $\pi(\cdot)$ with respect to γ. From equation (25), following a bad signal ($s < s^T$), monitoring intensity is lower if this signal was obtained with a higher γ if and only if:

$$\pi'_\gamma(s|\mu_h)\pi(s|\mu_l) - \pi'_\gamma(s|\mu_l)\pi(s|\mu_h) \leq 0. \tag{26}$$

From equation (25), following a good signal ($s \geq s^T$), monitoring intensity is lower if this signal was obtained with a higher γ if and only if:

$$\pi'_\gamma(s|\mu_h)\pi(s|\mu_l) - \pi'_\gamma(s|\mu_l)\pi(s|\mu_h) \geq 0. \tag{27}$$

Using the notations from Lemma 3:

$$\begin{aligned}
\frac{\pi'_\gamma(s|\mu)}{\pi(s|\mu)} &= \frac{\frac{\partial}{\partial\gamma}\left\{\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}\right\}}{\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}} \\
&= \frac{\frac{\partial}{\partial\gamma}\left\{\frac{1}{\sigma}\right\}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\} + \frac{1}{\sigma}\frac{\partial}{\partial\gamma}\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}}{\frac{1}{\sigma}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}} \\
&= \sigma\frac{\partial}{\partial\gamma}\left\{\frac{1}{\sqrt{\gamma^2\sigma_z^2 + (1-\gamma)^2\sigma_\epsilon^2}}\right\} + \frac{\partial}{\partial\gamma}\left\{-\frac{(s-\hat{\mu})^2}{\gamma^2\sigma_z^2 + (1-\gamma)^2\sigma_\epsilon^2}\right\} \\
&= \sigma\frac{2\gamma\sigma_z^2 - 2(1-\gamma)\sigma_\epsilon^2}{\left(\gamma^2\sigma_z^2 + (1-\gamma)^2\sigma_\epsilon^2\right)^{3/2}} + \frac{\left(2\gamma\sigma_z^2 - 2(1-\gamma)\sigma_\epsilon^2\right)(s-\hat{\mu})^2}{\left(\gamma^2\sigma_z^2 + (1-\gamma)^2\sigma_\epsilon^2\right)^2}
\end{aligned} \tag{28}$$

To establish the sign of $\frac{\pi'_\gamma(s|\mu_h)}{\pi(s|\mu_h)} - \frac{\pi'_\gamma(s|\mu_l)}{\pi(s|\mu_l)}$, only the second term on the RHS of equation (28) matters, and we have:

$$\begin{aligned}\frac{\pi'_\gamma(s|\mu_h)}{\pi(s|\mu_h)} > \frac{\pi'_\gamma(s|\mu_l)}{\pi(s|\mu_l)} \quad &\Leftrightarrow \quad \left(2\gamma\sigma_z^2 - 2(1-\gamma)\sigma_\epsilon^2\right)(s-\mu_h-\bar{z})^2 > \left(2\gamma\sigma_z^2 - 2(1-\gamma)\sigma_\epsilon^2\right)(s-\mu_l-\bar{z})^2 \\ &\Leftrightarrow \quad \left(\gamma\sigma_z^2 - (1-\gamma)\sigma_\epsilon^2\right)\left[(s-\mu_h-\bar{z})^2 - (s-\mu_l-\bar{z})^2\right] > 0 \end{aligned} \tag{29}$$

The term in brackets on the LHS of equation (29) is positive if and only if $s < \bar{z} + \frac{\mu_h+\mu_l}{2} \Leftrightarrow s < s^*$, where $s^* = s^T$ with $K = 0$. The term in parentheses on the LHS of equation (29) is positive if and only if $\frac{\sigma_\epsilon^2}{\sigma_z^2} < \frac{\gamma}{1-\gamma}$. ■

Proof of Lemma 4:

Conditional on μ, the distribution of the composite signal y is normal. Its location parameter is:

$$\begin{aligned}\mathbb{E}[\tilde{s}|\mu] &= \mathbb{E}\left[(\gamma\xi+1-\xi)\tilde{\mu} + \gamma\tilde{z} + (1-\gamma)\left(\xi\bar{\mu} + \bar{z} + \tilde{\epsilon}\right)|\mu\right] \\ &= (\gamma\xi+1-\xi)\mathbb{E}\left[\tilde{\mu}|\mu\right] + \gamma\mathbb{E}\left[\tilde{z}|\mu\right] + (1-\gamma)\left(\xi\bar{\mu} + \bar{z} + \mathbb{E}\left[\tilde{\epsilon}|\mu\right]\right) \\ &= (\gamma\xi+1-\xi)\mu + \gamma\bar{z} + (1-\gamma)\left(\xi\bar{\mu} + \bar{z}\right) \\ &= (\gamma\xi+1-\xi)\mu + \bar{z} + (1-\gamma)\xi\bar{\mu}.\end{aligned}$$

Its variance is:

$$\begin{aligned}var\left(\tilde{s}|\mu\right) &= var\left((\gamma\xi+1-\xi)\tilde{\mu} + \gamma\tilde{z} + (1-\gamma)\left(\xi\bar{\mu} + \bar{z} + \tilde{\epsilon}\right)|\mu\right) \\ &= (\gamma\xi+1-\xi)^2 var\left(\tilde{\mu}|\mu\right) + \gamma^2 var\left(\tilde{z}|\mu\right) + (1-\gamma)^2 var\left(\tilde{\epsilon}|\mu\right) \\ &= \gamma^2 var\left(\tilde{z}\right) + (1-\gamma)^2 var\left(\tilde{\epsilon}\right) \\ &= \gamma^2\sigma_z^2 + (1-\gamma)^2\sigma_\epsilon^2.\end{aligned}$$

where we used independence between random variables, and $var\left(\tilde{\mu}|\mu\right) = 0$. The scale parameter of a normal distribution is the square root of its variance. ■

Proof of Proposition 6:

Some steps are the same as in the proof of Proposition 5 and are omitted. Using the notations from Lemma 3:

$$\begin{aligned}
\frac{\pi'_\gamma(s|\mu)}{\pi(s|\mu)} &= \frac{\frac{\partial}{\partial\gamma}\left\{\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}\right\}}{\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}} \\
&= \frac{\frac{\partial}{\partial\gamma}\left\{\frac{1}{\sigma}\right\}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}+\frac{1}{\sigma}\frac{\partial}{\partial\gamma}\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}}{\frac{1}{\sigma}\exp\left\{-\frac{(s-\hat{\mu})^2}{2\sigma^2}\right\}} \\
&= \sigma\frac{\partial}{\partial\gamma}\left\{\frac{1}{\sqrt{\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2}}\right\}+\frac{\partial}{\partial\gamma}\left\{-\frac{(s-\hat{\mu})^2}{\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2}\right\} \\
&= \sigma\frac{2\gamma\sigma_z^2-2(1-\gamma)\sigma_\epsilon^2}{\left(\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2\right)^{3/2}} \\
&\quad +\frac{2(\xi\mu-\xi\bar{\mu})(s-\hat{\mu})\left(\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2\right)+\left(2\gamma\sigma_z^2-2(1-\gamma)\sigma_\epsilon^2\right)(s-\hat{\mu})^2}{\left(\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2\right)^2}
\end{aligned}$$

To establish the sign of $\frac{\pi'_\gamma(s|\mu_h)}{\pi(s|\mu_h)}-\frac{\pi'_\gamma(s|\mu_l)}{\pi(s|\mu_l)}$, only the second term on the RHS of equation (30) matters, and we have:

$$\begin{aligned}
& \frac{\pi'_\gamma(s|\mu_h)}{\pi(s|\mu_h)} > \frac{\pi'_\gamma(s|\mu_l)}{\pi(s|\mu_l)} \Leftrightarrow \\
& 2\xi(\mu_h-\bar{\mu})\left(s-(\gamma\xi+1-\xi)\mu_h-\bar{z}-(1-\gamma)\xi\bar{\mu}\right)\left(\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2\right) \\
& \quad +\left(2\gamma\sigma_z^2-2(1-\gamma)\sigma_\epsilon^2\right)\left(s-(\gamma\xi+1-\xi)\mu_h-\bar{z}-(1-\gamma)\xi\bar{\mu}\right)^2 \\
& > 2\xi(\mu_l-\bar{\mu})\left(s-(\gamma\xi+1-\xi)\mu_l+\bar{z}+(1-\gamma)\xi\bar{\mu}\right)\left(\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2\right) \\
& \quad +\left(2\gamma\sigma_z^2-2(1-\gamma)\sigma_\epsilon^2\right)\left(s-(\gamma\xi+1-\xi)\mu_l-\bar{z}-(1-\gamma)\xi\bar{\mu}\right)^2 \\
\Leftrightarrow \quad & 2\xi(\underbrace{\mu_h-\bar{\mu}}_{>0})\left(s-(\gamma\xi+1-\xi)\mu_h-\bar{z}-(1-\gamma)\xi\bar{\mu}\right)\underbrace{\left(\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2\right)}_{>0} \\
& -2\xi(\underbrace{\mu_l-\bar{\mu}}_{<0})\left(s-(\gamma\xi+1-\xi)\mu_l+\bar{z}+(1-\gamma)\xi\bar{\mu}\right)\underbrace{\left(\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2\right)}_{>0} \\
& > \left(2\gamma\sigma_z^2-2(1-\gamma)\sigma_\epsilon^2\right)\left(s-(\gamma\xi+1-\xi)\mu_l-\bar{z}-(1-\gamma)\xi\bar{\mu}\right)^2 \\
& -\left(2\gamma\sigma_z^2-2(1-\gamma)\sigma_\epsilon^2\right)\left(s-(\gamma\xi+1-\xi)\mu_h-\bar{z}-(1-\gamma)\xi\bar{\mu}\right)^2
\end{aligned} \tag{30}$$

We study in turn the terms on the LHS of the last inequality in equation (30), and then the terms on the RHS.

The LHS of equation (30) is continuously differentiable in ξ and equal to 0 at $\xi=0$.

It is positive if and only if:

$$
\begin{aligned}
&2\xi(\underbrace{\mu_h-\bar{\mu}}_{>0})(s-(\gamma\xi+1-\xi)\mu_h-\bar{z}-(1-\gamma)\xi\bar{\mu})\underbrace{\left(\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2\right)}_{>0}\\
&>2\xi(\underbrace{\mu_l-\bar{\mu}}_{<0})(s-(\gamma\xi+1-\xi)\mu_l-\bar{z}-(1-\gamma)\xi\bar{\mu})\underbrace{\left(\gamma^2\sigma_z^2+(1-\gamma)^2\sigma_\epsilon^2\right)}_{>0}\\
\Leftrightarrow\quad&(\underbrace{\mu_h-\bar{\mu}}_{>0})(s-(\gamma\xi+1-\xi)\mu_h-\bar{z}-(1-\gamma)\xi\bar{\mu})\\
&>(\underbrace{\mu_l-\bar{\mu}}_{<0})(s-(\gamma\xi+1-\xi)\mu_l-\bar{z}-(1-\gamma)\xi\bar{\mu})
\end{aligned}
\tag{31}
$$

For $s>(\gamma\xi+1-\xi)\mu_h+\bar{z}+(1-\gamma)\xi\bar{\mu}$, the inequality in equation (31) holds. For $s<(\gamma\xi+1-\xi)\mu_l+\bar{z}+(1-\gamma)\xi\bar{\mu}$, the inequality in equation (31) does not hold. That it, the LHS of equation (30) tends to leads to substitutability.

For the terms on the RHS of equation (30):

$$
\begin{aligned}
&\left(2\gamma\sigma_z^2-2(1-\gamma)\sigma_\epsilon^2\right)(s-(\gamma\xi+1-\xi)\mu_h-\bar{z}-(1-\gamma)\xi\bar{\mu})^2\\
&>\left(2\gamma\sigma_z^2-2(1-\gamma)\sigma_\epsilon^2\right)(s-(\gamma\xi+1-\xi)\mu_l-\bar{z}-(1-\gamma)\xi\bar{\mu})^2\\
\Leftrightarrow\quad&\left(\gamma\sigma_z^2-(1-\gamma)\sigma_\epsilon^2\right)\left[(s-(\gamma\xi+1-\xi)\mu_h-\bar{z}-(1-\gamma)\xi\bar{\mu})^2\right.\\
&\left.-(s-(\gamma\xi+1-\xi)\mu_l-\bar{z}-(1-\gamma)\xi\bar{\mu})^2\right]>0
\end{aligned}
\tag{32}
$$

The term in brackets on the LHS of equation (32) is positive if and only if $s<\bar{\mu}+\bar{z}\Leftrightarrow s<s^*$, where $s^*=s^T$ with $K=0$. The term in parentheses on the LHS of equation (32) is positive if and only if $\frac{\sigma_\epsilon^2}{\sigma_z^2}<\frac{\gamma}{1-\gamma}$. Similarly to the case studied in Proposition 5, this leads to complementarity if $\frac{\sigma_\epsilon^2}{\sigma_z^2}<\frac{\gamma}{1-\gamma}$. Moreover, as long as $s\neq\bar{\mu}+\bar{z}$ and $\frac{\sigma_\epsilon^2}{\sigma_z^2}\neq\frac{\gamma}{1-\gamma}$, the RHS of equation (30) is nonzero.

The LHS of equation (30) is continuous in ξ and equal to zero at $\xi=0$, whereas the RHS of equation (30) is generically nonzero regardless of ξ and leads to complementarity for any s if and only if $\frac{\sigma_\epsilon^2}{\sigma_z^2}<\frac{\gamma}{1-\gamma}$. In sum, if $\frac{\sigma_\epsilon^2}{\sigma_z^2}<\frac{\gamma}{1-\gamma}$, we have complementarity if and only if ξ is sufficiently small. ■

Proof of Proposition 7:

For notational convenience we parametrize changes in the precision of the signal by σ, consistent with notations in sections 3.1 and 3.2. For the model studied in section 3.3, replace σ by $1/\gamma$.

For a given monitoring intensity m_s with $s<s^T$, we have:

$$
Pr\left(\text{terminate}|s<s^T\right)=p_s(1-m_s)+(1-p_s)=1-p_s\,m_s. \tag{33}
$$

Indeed, after a signal that leads to termination in the absence of monitoring $(s<s^T)$, with probability p_s the manager has high ability and is only terminated following a monitoring

failure (with probability $1-m_s$), and with probability $1-p_s$ the manager has low ability and is terminated in any case. Then, the relation between the probability of termination after a bad signal and σ is given by:

$$\frac{\partial Pr\left(\text{terminate}|s < s^T\right)}{\partial\sigma} = 1 - p_s\frac{\partial m_s}{\partial\sigma}. \tag{34}$$

That is, the sign of $\frac{\partial Pr\left(\text{terminate}|s<s^T\right)}{\partial\sigma}$ is the opposite of the sign of $\frac{\partial m_s}{\partial\sigma}$. An increase in precision corresponds to a lower σ.

For a given monitoring intensity m_s, we have:

$$Pr\left(\text{terminate}|s \geq s^T\right) = (1-p_s)m_s. \tag{35}$$

Indeed, after a signal that leads to continuation in the absence of monitoring ($s \geq s^T$), with probability p_s the manager has high ability and is never terminated, and with probability $1-p_s$ the manager has low ability and is only terminated if monitoring succeeds (with probability m_s). Then, the relation between the probability of termination after a bad signal and σ is given by:

$$\frac{\partial Pr\left(\text{terminate}|s \geq s^T\right)}{\partial\sigma} = (1-p_s)\frac{\partial m_s}{\partial\sigma}. \tag{36}$$

That is, the sign of $\frac{\partial Pr\left(\text{terminate}|s\geq s^T\right)}{\partial\sigma}$ is the same as the sign of $\frac{\partial m_s}{\partial\sigma}$. An increase in precision corresponds to a lower σ. ■

Proof of Proposition 8:

A type i firm behaves as in the baseline model, but only dismisses its manager when it is optimal to do so with probability ρ. A type d firm does not engage in any monitoring, and therefore only dismisses its manager based on the signal s: the manager is terminated for $s < s^T$, and continued for $s \geq s^T$. ■

Proof of Empirical Implication 2:

For $s \geq s^T$, from equation (12) we have sign $\left\{\frac{\partial m_s^*}{\partial\sigma}\right\} = -\text{sign}\left\{\frac{\partial p_s}{\partial\sigma}\right\}$, which is positive if and only if $s > \hat{s}$, which in the case of a symmetric distribution is such that $\hat{s} = s^* = \frac{\mu_l+\mu_h}{2}$ (see the proof of Proposition 3). Given that $s \geq s^T$ as $K \to p(\mu_h - \mu_l)$, we use equation (36) which give that sign $\left\{\frac{\partial Pr\left(\text{terminate}|s\geq s^T\right)}{\partial\sigma}\right\} = \text{sign}\left\{\frac{\partial m_s^*}{\partial\sigma}\right\}$. Finally, a more precise signal corresponds to a lower σ. ■

Supplementary Appendix

Not for publication

Willingness to Pay for Termination Right

We have assumed that the board always have the control right to make the termination decision. While the optimal allocation of control rights is outside the scope of this paper, the model still has implications for the value of control rights. When as in section 4.3 the manager is only terminated following successful monitoring, the interaction between the informativeness of the signal and monitoring has implications for the value of the control right to terminate the manager:

Empirical Implication 3. *For small private firms, an increase in the quality of the signal increases the value of the right to terminate the manager after a signal which is bad news about CEO ability ($s < s^*$).*

This Empirical Implication is not driven by a change in monitoring: since monitoring is set at the optimal level, a marginal change in monitoring intensity does not affect the value of the right to terminate the manager. Instead, this result is driven by two factors: (i) the default decision (without monitoring) is to retain the manager even after a negative signal; (ii) in this case, when the quality of the signal and monitoring are complements, the probability that the manager has low ability following a negative signal is increasing in the quality of firm value measurement. Due to factors (i) and (ii), the value of the right to terminate the manager after a negative signal is higher with better firm value measurement.

Proof of Empirical Implication 3:

After a signal $s \in (s^T, s^*)$, which is such that $p_s < p$ but the CEO is only terminated following successful monitoring, without a right to terminate the manager, the manager is always continued and expected firm value is $p_s\mu_h + (1 - p_s)\mu_l$. Terminating the manager only happens following successful monitoring (with probability m_s^*) that finds that the incumbent manager has low ability (with probability $1 - p_s$). Without the right to terminate the manager, monitoring is optimally zero and therefore $c(m) = 0$. Therefore, the ex-interim value of the right to terminate the manager (after s is observed but before the monitoring decision) is:

$$\begin{aligned} & m_s^*(1 - p_s)\left[p\mu_h + (1 - p)\mu_l - K - \mu_l\right] - c\left(m_s^*\right) \\ &= m_s^*(1 - p_s)\left[p\Delta - K\right] - c\left(m_s^*\right), \end{aligned} \tag{37}$$

which is strictly positive, otherwise we would have $m_s^* = 0$ (since $c(0) = 0$).

Taking the total derivative of equation (37) with respect to σ:

$$\begin{aligned} & \frac{d}{d\sigma}\left\{m_s^*(1-p_s)\left[p\Delta-K\right]-c\left(m_s^*\right)\right\} \\ = & \frac{dm_s^*}{d\sigma}(1-p_s)\left[p\Delta-K\right]-m_s^*\frac{dp_s}{d\sigma}\left[p\Delta-K\right]-\frac{dm_s^*}{d\sigma}c'\left(m_s^*\right) \\ = & \frac{dm_s^*}{d\sigma}\underbrace{\left[(1-p_s)\left(p\Delta-K\right)-c'\left(m_s^*\right)\right]}_{=0,\text{ see equilibrium condition (5)}}-m_s^*\frac{dp_s}{d\sigma}\underbrace{\left[p\Delta-K\right]}_{>0}, \end{aligned} \tag{38}$$

which has the opposite sign of $\frac{dp_s}{d\sigma}$, which in turn is positive for $s < s^*$. That is, the expression in equation (38) has a negative sign: the value of the right to terminate the manager is decreasing in σ, i.e., the value of the right to terminate the manager after a signal $s \in (s^T, s^*)$ is increasing in the precision of the signal. ■