The Value of Informativeness for Contracting*

Pierre Chaigneau
HEC Montreal

Alex Edmans
Wharton, NBER, and ECGI

June 6, 2013

Abstract

A key result from contract theory, the informativeness principle, is often violated in practice. We demonstrate an important cost of increasing informativeness: reduced incentives. We study a standard agency model with risk-neutrality and limited liability (as in Innes (1990)), where the optimal contract is a call option. The direct effect of reducing signal volatility – for example, by filtering out luck – is a fall in the value of the option and thus the agent's rents, benefiting the principal. We identify a second indirect effect: if the optimal strike price is not too far in-the-money, the agent can only beat the strike price if he exerts effort and there is a high noise realization. Thus, a fall in volatility weakens effort incentives. As the target performance level rises, the gains from increased precision fall towards zero. Our model offers a potential justification for pay-for-luck and the absence of relative performance evaluation.

KEYWORDS: Executive compensation, pay-for-luck, principal-agent model, relative performance evaluation, options, contract theory, agency theory.
JEL CLASSIFICATION: D86, J33

*pierre.chaigneau@hec.ca, aedmans@wharton.upenn.edu. We thank seminar participants at Wharton for helpful comments and Shiying Dong for excellent research assistance.

A major result in contract theory is the informativeness principle (Holmstrom (1979), Shavell (1979), Gjesdal (1982), Grossman and Hart (1983), Kim (1995)).[1] According to this principle, designers of contracts (such as managers, boards of directors, or regulators) should maximize the precision of the overall vector of performance measures used to evaluate the agent. In turn, maximizing precision involves incorporating all informative signals, and filtering out irrelevant noise such as industry or market shocks that are outside the agent's control. However, numerous violations of the latter have been found in practice. Aggarwal and Samwick (1999) and Murphy (1999) find that CEO pay is determined only by absolute performance, rather than performance relative to industry peers, and thus affected by industry shocks. Jenter and Kanaan (2013) similarly find an absence of relative performance evaluation ("RPE") in CEO firing decisions. Bertrand and Mullainathan (2001) find that CEOs are paid for "luck", i.e. positive exogenous shocks, and that such practices are particularly strong among firms with weak governance, consistent with the view that they are inefficient. Indeed, Bebchuk and Fried (2004) argue that the absence of RPE is a key piece of evidence that CEO compensation is not determined by efficient contracting by shareholders, and instead results from rent extraction by CEOs.

This paper reaches a different conclusion. Traditional arguments in favor of signal precision focus on the fact that a reduction in noise will reduce the risk borne by the agent, and thus the cost of compensation. We analyze an important additional effect that has been largely ignored: an increase in informativeness may weaken the agent's incentives, thus increasing the cost of compensation required to induce effort from the agent and offsetting the first benefit. Taking into account this second effect can lead to quite different conclusions on the benefits of increasing informativeness. It can reduce the net benefit of augmenting precision sufficiently strongly to make it no longer worthwhile for the principal to bear the cost of doing so. In other situations, the increase in informativeness may strengthen incentives and reinforce the first effect, now rendering it optimal to pay the cost of reducing signal volatility.

[1] Shavell (1979) shows that additional information on the agent's effort has positive value. Gjesdal (1982) and Grossman and Hart (1983) show that if the information structure A is sufficient for the information structure B in the sense of Blackwell, then A is associated with a lower agency cost than B. Holmstrom (1979) shows that any signal which is informative about the agent's action will be included in the contract. Lastly, Kim (1995) shows that the information structure A is more efficient than B if the cumulative distribution function of the likelihood ratio under A is a mean-preserving spread of the one under B.

We consider a standard model of moral hazard under risk neutrality and limited liability, similar to Innes (1990). As shown by Innes (1990), the optimal contract involves giving the agent a call option. A fall in the strike price increases the option's delta and thus the agent's effort incentives, but also augments the value of the option and thus his rents. Thus, the strike price is chosen to be the minimum possible to satisfy the agent's incentive constraint.

We analyze the effect of increasing the precision of the signal on which the option is based. Increased precision has several interpretations. First, it can refer to holding constant the choice of signal and improving the monitoring technology. For example, if a supervisor is monitoring a worker through performance reports, an increase in signal precision corresponds to paying the cost of more accurate reports. Second, it can refer to replacing the signal with a more precise one. In Innes (1990), the call option is on the firm's (absolute) profits, but the principal may be able to use other performance measures to evaluate the agent. For example, she can pay a cost to filter out industry shocks from the performance signal, and instead give the agent a call option on industry-adjusted performance. Effectively, such relative performance evaluation increases the informativeness, and reduces the volatility, of the performance signal.

Regardless of the interpretation, the reduction in volatility in turn has two effects. First, it reduces the value of the option and thus the agent's rents. This is the standard benefit of increased informativeness which is traditionally focused on by advocates of RPE. Our main contribution is to analyze a second effect – the reduction in volatility changes the agent's incentives. This effect is often ignored by standard analyses, that do not consider the agent's incentive constraint. Since the agent's effort is unobservable, he receives the call option regardless of whether he exerts effort. However, the distribution of the signal underlying the option – and thus the value of the option to him – depends on his effort decision. His incentives to exert effort stem from the difference in value between two options – the (less valuable) option that he receives when he shirks ("option-when-shirking"), and the (more valuable) option that he receives when he works ("option-when-working"). A fall in volatility increases the value of both options, but to different degrees depending on their vega. Thus, the difference in value between the two options will change, altering his incentives.

We first consider the case in which the cost of effort is high, i.e. the moral hazard problem is severe. Then, if the principal pays the agent according to absolute performance, he will offer an agent an option with a low strike price, so that its delta is sufficiently large to balance the high cost of effort and induce working. Since the strike price is low, the option-when-working will be deeply in-the-money, and the option-when-shirking is closer to at-the-money. The vega of an option is highest when it is at-the-money, and declines when the option moves in or out-of-the-money. Thus, the vega of option-when-shirking is greater, and an increase in informativeness reduces its value faster than the option-when-working. Overall, the fall in volatility increases the agent's incentives. Intuitively, when volatility is high, the agent's effort incentives are weak because, even if he shirked, he would still earn a high wage if he received a very positive shock. Another way to view the intuition is that volatility is valuable to an option holder since he benefits from the asymmetry of its payoff. This asymmetry is strongest when the option is at-the-money; thus, when the strike price is low, the agent benefits most from high volatility when he shirks, as he still receives the full upside of positive shocks but is protected on the downside from sufficiently negative shocks.

In sum, when incentives are strong to begin with, an increase in informativeness further increases incentives, *ceteris paribus*. This reinforces the gains from informativeness stemming from the first effect. Formally, since incentives have strengthened, the principal can increase the strike price of the option without eroding the agent's incentives. This increase further reduces the option value and thus the cost of compensation.

We next consider a low cost of effort, where the optimal contract (without filtering) involves a high strike price. Then, the option-when-shirking will be deeply out-of-the-money, and the option-when-working will be closer to at-the-money. Thus, the vega of the latter option is greater, and an increase in informativeness reduces the value of the option-when-working faster than the option-when-shirking. Now, the fall in volatility *reduces* the agent's incentives. Intuitively, when the strike price is high, the agent will only receive a positive wage if he exerts effort *and* also receives a sufficiently positive shock. When volatility falls, such shocks are less likely, and so the agent may not get paid even if he does work. Thus, his effort incentives decline.

In sum, when incentives are weak to begin with, an increase in informativeness further reduces incentives. This reduction forces the principal to lower the strike price of the option to restore incentives, increasing the cost of compensation and partially offsetting the first effect. Indeed, as incentives weaken and the strike price increases, the net benefit of increased informativeness falls to zero. Thus, for sufficiently weak incentives, it becomes efficient for the principal not to increase signal precision, e.g.

to pay the agent for luck. Note that this result continues to hold even if the cost of filtering out the shock is proportional to the original cost of compensation. Thus, when the moral hazard problem is weak, leading to low incentives and compensation costs, then even though it is cheaper to increase informativeness, the principal still chooses not to do so.

At a general level, our analysis highlights the importance of considering the incentive constraint when assessing the value of increased informativeness, as the total savings in compensation can be substantially higher or substantially lower than the amounts calculated when ignoring this constraint. At a detailed level, our analysis shows that the benefits of increasing informativeness depend on the contracting setting. They are lower for agents with weak incentives, such as rank-and-file employees, but higher for agents with strong incentives, such as CEOs.

In addition to studying whether a firm should endogenously choose to increase informativeness (at a cost), our analysis also investigates the impact of exogenous changes in signal precision. An exogenous increase in the volatility of the signal (see Gormley, Matsa, and Milbourn (2013) and DeAngelis, Grullon, and Michenaud (2013) for natural experiments) will increase the effort incentives of agents with out-of-the-money options, and reduce the effort incentives of firms with in-the-money options. Thus, if firms recontract in response to these exogenous shocks, firms with in-the-money options should increase their CEO's incentives relative to firms with out-of-the-money options, either by granting a greater number of additional options, reducing the strike price at which additional options are granted, or reducing the strike price of existing options.[2]

In a similar vein, since an improvement in informativeness increases (reduces) the optimal strike price if the option was initially in-the-money (out-of-the-money), it overall leads the strike price to converge towards the initial stock price and thus for the option to become closer to at-the-money. Thus, improvements in signal precision (e.g. increases in stock market efficiency over time) lead to at-the-money options being optimal. Bebchuk and Fried (2004) argue that the almost universal practice of granting at-the-money options is suboptimal and that out-of-the-money options would be cheaper for the firm, since the agent only gets paid if performance is very high. Such

[2] Acharya, John, and Sundaram (2000) also study the repricing of stock options theoretically, although in response to changes in the mean rather than volatility of the signal. Brenner, Sundaram, and Yermack (2000) analyze repricing empirically.

an argument ignores the incentive effect: out-of-the-money options have lower deltas and thus may provide the agent with insufficient incentives. Murphy (2002) notes that in-the-money options would provide strongest incentives, but that the tax code discourages such options. One potential interpretation is that the tax code leads to firms choosing at-the-money options when in-the-money options may be more efficient. Our analysis indeed suggests that increases in informativeness lead to options optimally being close to at-the-money.

A recent paper by Dittmann, Maug, and Spalt (2012) also considers the incentive constraint when assessing the benefits of a specific form of increased informativeness – indexing stock and options – and similarly show that indexation may weaken incentives. They use a quite different setting from ours, which reflects the different aims of each paper. Their primary goal is to calibrate real-life contracts, and so their model incorporates risk aversion to allow them to input various risk aversion parameters into the calibration. One disadvantage is that, under risk aversion, it is very difficult to solve for the optimal contract. They therefore restrict the contract to comprising salary, stock, and options, and hold stock constant when changing the contract to restore the agent's incentives upon indexation. This approach may underestimate the savings from indexation, since it prevents the principal from responding optimally to the weakening of incentives. In contrast, our primary goal is theoretical. We incorporate risk neutrality, allowing us to take an optimal contracting approach where the principal changes the contract optimally to restore incentives after an increase in informativeness. In addition, our model allows the analysis of reductions in volatility through other means than indexation, for example investing in a superior monitoring technology.

Other explanations for pay-for-luck have been proposed in the literature, partially reviewed by Edmans and Gabaix (2009). Oyer (2004) shows that pay-for-luck may be optimal if the value of employees' outside options vary with economic conditions and if re-contracting is costly. Raith (2008) shows that it may be preferable to base compensation on measures of output rather than input when the agent has private information on the production technology. Axelson and Baliga (2009) argue that, for contracts to be renegotiation-proof, the manager must have private information that causes him to have a different view from the board on the value of his long-term pay. Industry performance is an example of such information, and so it may be efficient not to filter it out. Gopalan, Milbourn, and Song (2010) show that tying the CEO's pay to industry performance induces him to choose the firm's industry exposure correctly.

This paper proceeds as follows. Section 1 presents the model. Section 2 shows that the optimal contract takes the form of a call option. Section 3 derives the gains from a reduction in the variance of the performance measure. Section 4 concludes. Appendix A contains all proofs not in the main text.

1 The Model

We consider a standard principal-agent model with risk neutrality and limited liability, similar to Innes (1990). The timing is as follows. At time $t = -1$, the principal (firm) offers a compensation contract W to the agent (employee). At $t = 0$, the agent chooses his effort level $e \in \{0, \bar{e}\}$. Effort of $e = 0$ is of zero cost to the agent, and $e = \bar{e}$ costs him C. At $t = 1$, the firm's profit is realized. It is given by $\tilde{\pi} = e + \tilde{v} + \widetilde{\varepsilon}$, where $\tilde{\pi}$ is contractible, but the individual components e, $\tilde{v}$, and $\widetilde{\varepsilon}$ are all unobservable. The random variables $\tilde{v} \sim N\left(0, \underline{\sigma}^2\right)$ and $\widetilde{\varepsilon} \sim N\left(0, \bar{\sigma}^2 - \underline{\sigma}^2\right)$ are independent noise terms. Since we focus on informativeness improvements at the margin, we assume that $\bar{\sigma} - \underline{\sigma}$ is arbitrarily small. This assumption is purely for tractability, since it allows us to consider only first-order effects in a Taylor expansion.

The discount rate is zero. The agent is risk-neutral and so maximizes his expected wage $E[W]$, less the cost of effort. He is protected by limited liability and has a reservation utility of zero. The principal is also risk-neutral and chooses the contract W that maximizes the expected profit $E[\tilde{\pi}]$, less the expected wage. Crucially, the contract need not depend on profits $\tilde{\pi}$, but depends on a contractible performance signal $\widetilde{s}$. By default, the signal $\widetilde{s}$ equals the profit $\tilde{\pi}$, but the principal has access to a technology that allows her to filter out the shock $\widetilde{\varepsilon}$ from the signal $\widetilde{s}$ if she pays a cost κ at $t = -1$. The cost κ can stem from multiple sources. Under the interpretation that $\widetilde{\varepsilon}$ is a market or industry shock that affects a firm's peers, removing the shock $\widetilde{\varepsilon}$ corresponds to RPE, in which case the cost κ stems from two sources. First, it can arise from the literal cost of implementing RPE. While the actual calculation of industry performance, given a peer group, is relatively costless, the determination of the peer group may involve the hiring of compensation consultants. Second, the cost can also represent the loss of the benefits of pay-for-luck highlighted by prior work, e.g. Oyer (2004), Raith (2008), Axelson and Baliga (2009), and Gopalan, Milbourn, and Song (2010). Under the alternative interpretation that $\widetilde{\varepsilon}$ arises from signal imprecision, removing the shock corresponds to an improvement in the monitoring technology, in which case κ refers to the cost of such an improvement. For example, Cornelli, Kominek, and Ljungqvist (2013) show that boards of directors engage in extensive (and thus costly) monitoring to gather soft information on the CEO's competence, strategic choice, and effort.

The signal can be written:

$$\widetilde{s} = e + \tilde{v} + \delta\widetilde{\varepsilon}, \tag{1}$$

where $\delta \in \{0, 1\}$ is a dummy variable. It equals zero if the firm filters the shock $\widetilde{\varepsilon}$ out of the signal $\widetilde{s}$, in which case informativeness is maximized and the signal has variance $\underline{\sigma}^2$. It equals one if the firm does not filter out the shock, in which case the signal $\widetilde{s}$ equals the profit $\tilde{\pi}$ and has variance $\bar{\sigma}^2$. The agent is thus paid for "luck" $\widetilde{\varepsilon}$ (in addition to the second "luck" term $\tilde{v}$ that cannot be filtered out, such as an idiosyncratic shock). Given $\delta \in \{0, 1\}$, let $\sigma^2(\delta)$ denote the variance of $\widetilde{s}$, so that $\sigma^2(\delta) = \underline{\sigma}^2 + \delta^2(\bar{\sigma}^2 - \underline{\sigma}^2)$ and $\sigma(\delta) \in \{\underline{\sigma}, \bar{\sigma}\}$. Choosing $\delta \in \{0, 1\}$ is equivalent to choosing $\sigma \in \{\underline{\sigma}, \bar{\sigma}\}$. To economize on notation, we suppress the dependence of σ on δ.

Following Innes (1990), we make two assumptions on the set of feasible contracts. First, the agent is protected by limited liability, so that $W(s) \geq 0 \ \forall \ s$. Second, the pay-performance sensitivity is capped at 1: $W'(s) \leq 1 \ \forall \ s$. Innes (1990) justifies this constraint on two grounds. First, if it did not hold on some interval, the agent could borrow on his own account to artificially increase the value of s on this interval, thus undoing the contract. Second, the principal would exercise her control rights to burn profits or sabotage the firm along this interval, since for any increase in the signal, payments to the agent increase more than one-for-one.

Given a contract $W(s)$ and a level of effort e, the agent's expected wage is

$$E[W(\widetilde{s})\,|e] = \int_{-\infty}^{\infty} W(s)\,\psi(s|e)\,ds,$$

where $\psi(\cdot)$ is the probability density function of $\widetilde{s}$, for a given σ. The agent's utility is given by his expected wage, less the cost of effort. In the first-best, effort is verifiable. Thus, to induce high effort, the principal simply directs the agent to exert $e = \bar{e}$; there is no incentive constraint. To satisfy the agent's participation constraint, the principal pays him an expected wage $E[W(\widetilde{s})\,|\bar{e}]$ that equals his cost of effort C. Thus, if $E[\widetilde{\pi}|\bar{e}] - E[\widetilde{\pi}|0] > C$, i.e. $\bar{e} > C$, high effort is optimal. We thus assume $\bar{e} > C$ throughout, else even under the first-best, the principal would not want to induce effort. Since the agent's expected wage equals his cost of effort, he earns zero rents.

In the second-best, the agent's effort is unverifiable and so the contract must satisfy an incentive constraint. The agent will exert effort if and only if:

$$E[W(\widetilde{s})|\bar{e}] - E[W(\widetilde{s})|0] \geq C. \tag{2}$$

Following standard arguments, this incentive constraint will bind. In contrast, the participation constraint will typically be slack: since the agent's reservation utility is zero, his expected utility under the contract will represent rents. For a contract in which he optimally chooses $e = \bar{e}$, we define the agency rent as the agent's expected wage minus his effort cost:

$$AR \equiv E\left[W(\widetilde{s})|\bar{e}\right] - C. \tag{3}$$

Since the incentive constraint binds, the agency rent can be rewritten:

$$AR = E\left[W(\widetilde{s})|0\right]. \tag{4}$$

The agency rent will be positive due to limited liability. Even if the agent shirks and noise is negative, his wage cannot fall below zero. If he shirks and noise is sufficiently positive, the signal will be favorable enough that there is a sufficiently high posterior probability that the agent has worked, and so the optimal contract will pay him a strictly positive wage. Thus, the expected wage upon shirking is strictly positive and so the agent receives rents from shirking. To satisfy his incentive constraint, the agent must also be offered rents for working.

We define $X^\star$ implicitly by

$$\int_{X^\star}^{\infty} (s - X^\star)\left(\psi(s|\bar{e}) - \psi(s|0)\right) ds \equiv C. \tag{5}$$

Intuitively, if the agent's contract consists of an option on s with a strike price of X^*, X^* is the strike price that satisfies the agent's incentive constraint with equality.[3] We make the following assumption to ensure that $e = \bar{e}$ is second-best optimal:

$$\bar{e} - \int_{X^\star}^{\infty} (s - X^\star)\psi(s|\bar{e}) ds \geq 0. \tag{6}$$

[3]The assumption $\bar{e} > C$ implies $\int_{-\infty}^{\infty} s\left(\psi(s|\bar{e}) - \psi(s|0)\right) ds > C$, which in turn guarantees the existence of X^*, as shown in the proof of Lemma 1.

The first term is the benefit to the principal of inducing $e = \bar{e}$, and the second term is the cost of the contract required to do so. If (6) did not hold, the principal would allow the agent to shirk, in which case the problem would be trivial and the contract would involve $W(\widetilde{s}) = 0 \; \forall \; s$. Note that (6) implies

$$\bar{e} \geq C + \int_{X^\star}^{\infty} (s - X^\star)\psi(s|0)ds,$$

which is a stronger condition than $\bar{e} \geq C$, which guarantees that effort is optimal in the first-best. The additional term $\int_{X^\star}^{\infty}(s - X^\star)\psi(s|0)ds$ stems from the fact that the agent earns rents from shirking, and thus must be given rents for working to induce him to exert effort in the second-best.

The principal's problem is to choose a contract $W(\widetilde{s})$ and the filtering dummy δ to minimize the sum of the agency rent and the cost κ (if paid)[4], subject to the agent's incentive, participation, and limited liability constraints, plus the upper bound on the slope of the contract. Her problem is thus given by:

$$\min_{W(s), \delta \in \{0,1\}} E\left[W(\widetilde{s})|0\right] + (1 - \delta)\,\kappa \text{ s.t.} \tag{7}$$

$$E\left[W(\widetilde{s})|\bar{e}\right] = E\left[W(\widetilde{s})|0\right] + C \tag{8}$$

$$0 \leq W(s) \;\; \forall \; s \tag{9}$$

$$W'(s) \leq 1 \; \forall \; s. \tag{10}$$

Our setup is similar to the classic model of Innes (1990). He considers a financing model where the agent (entrepreneur) chooses a financing contract to maximize his objective function, subject to the incentive constraint of the agent and the participation constraint of the principal (investor). In contrast, we consider a contracting model where the principal (firm) chooses an employment contract to maximize her objective function subject to the incentive and participation constraints of the agent (employee). As per footnote 2 of Innes (1990), these two optimization problems yield the same optimal contracts. In addition, Innes features a general noise distribution that satisfies

[4] More precisely, the principal seeks to maximize her objective function, which is expected gross profits minus the expected wage and the cost κ. Since expected gross profits are unaffected by the wage and cost κ, if the agent's incentive constraint is satisfied, maximization of the objective function is equivalent to minimization of the wage plus the cost κ. Since the wage equals $E\left[W(\widetilde{s})|\bar{e}\right]$ and C is a constant, minimization of the wage equals minimization of the agency rent (3), which equals (4).

the monotone likelihood ratio property ("MLRP") and a continuous action set. His focus was to derive the form of the optimal contract and thus wishes to do so in the most general setting. Our goal is different: given that the optimal contract is a call option (as shown by Innes)[5], we study how changes in informativeness affect the agent's incentives and thus the strike price of the option.[6] We thus specialize to a normal noise distribution and a binary effort level. A normal distribution satisfies the MLRP and can be parameterized by its volatility σ, allowing us to study the effect of informativeness in a tractable manner by varying σ (changes in this single parameter generate mean-preserving spreads). With a continuous effort level, a change in σ may alter the optimal effort level. It is well known that solving for the optimal effort level in addition to the cheapest contract that induces a given effort level is extremely complex (see, e.g., Grossman and Hart (1983)), and thus many papers focus on the implementation of a given effort level (e.g. Dittmann and Maug (2007), Dittmann, Maug, and Spalt (2008, 2011)).[7] Edmans and Gabaix (2011) show that, if the benefits of effort are multiplicative in firm size and the firm is sufficiently large, it is always optimal for the principal to implement the highest effort level and so the optimal effort level is indeed fixed. We thus consider a binary effort setting where high effort is optimal.

2 The Optimal Contract

This section solves for the optimal contract, taking as given the choice of the filtering dummy $\delta \in \{0, 1\}$ and thus signal precision $\sigma \in \{\underline{\sigma}, \bar{\sigma}\}$. The analysis is similar to Innes (1990); our main results will come in Section 3. The principal's objective function (7) thus becomes $\min_{W(s)} E[W(\widetilde{s})|0]$. We let φ and Φ denote the p.d.f. and c.d.f. of the

[5] Since Innes (1990) studies a financing setting, the optimal contract for the principal is debt. Thus, the agent has equity, which is a call option on the firm's assets.

[6] A consequence of these differences is that deviations from the first-best take different forms. Here, the objective is to minimize the agency rent received by the agent, a problem that exists regardless of whether the effort level is binary or continuous. In Innes (1990), the agent chooses the contract and so the goal is not to minimize the agency rent, but the difference between the first- and second-best levels of effort – hence the importance of allowing for a continuum of effort levels in his paper.

[7] Indeed, Innes (1990) does not solve for the optimal effort level or study how it is affected by the parameters of the setting, but shows that an optimum exists.

standard normal variable, respectively. The likelihood ratio is given by:

$$LR\left(i\right)\equiv\frac{\psi(s=i|\bar{e})-\psi(s=i|0)}{\psi(s=i|0)}=\frac{\varphi\left(\frac{i-\bar{e}}{\sigma}\right)-\varphi\left(\frac{i}{\sigma}\right)}{\varphi\left(\frac{i}{\sigma}\right)}. \tag{11}$$

It is strictly increasing in i, so the MLRP is satisfied. For any given $\sigma\in\{\underline{\sigma},\bar{\sigma}\}$, let $\widehat{X}$ be implicitly defined by

$$\varphi\left(\frac{\widehat{X}-\bar{e}}{\sigma}\right)\equiv\varphi\left(\frac{\widehat{X}}{\sigma}\right), \tag{12}$$

which yields $\widehat{X}=\frac{\bar{e}}{2}$. Note that $\widehat{X}$ is the value of the signal such that the likelihood ratio in (11) is zero when evaluated at $i=\widehat{X}$: $\widehat{X}$ is the least informative performance. Since the likelihood ratio is strictly increasing, $LR\left(i\right)>0$ if and only if $i>\widehat{X}$.

The optimal contract is given in Lemma 1.

Lemma 1 *(Optimal contract.) For a given σ, the optimal contract is characterized by*

$$W\left(s\right)=\max\{0,s-X(\sigma)\},\ i.e. \tag{13}$$

$$\begin{aligned}W\left(s\right)&=0\ \forall\ s\leq X\left(\sigma\right)\\W'\left(s\right)&=1\ \forall\ s>X\left(\sigma\right),\end{aligned}$$

where $X\left(\sigma\right)>0$. The value of $X\left(\sigma\right)$ is chosen so that the incentive constraint binds, i.e.:

$$\int_{X(\sigma)}^{\infty}\left(s-X(\sigma)\right)\left(\psi(s|\bar{e})-\psi(s|0)\right)ds=C. \tag{14}$$

There is a unique $X\left(\sigma\right)$ which satisfies the incentive constraint (14).

The contract (13) is the payoff of a call option on s with strike price $X(\sigma)$. While the signal σ, and thus the principal's filtering decision δ, affects the strike price $X\left(\sigma\right)$, the optimal contract remains a call option. The two forces that drive this result are the constraints on contracting and the MLRP of the normal distribution. The intuition is as in Innes (1990). The absolute value of the likelihood ratio is highest in the tails of the distribution of $\widetilde{s}$, so the signal is most informative about the agent's effort in the tails. The left tail cannot be used for incentive purposes due to the limited liability constraint, so that incentives are concentrated in the right tail. With an upper-bound on the slope, the optimal contract involves call options on $\widetilde{s}$ with the maximum feasible slope, i.e. $W'(s) = 1$. This maximizes the likelihood that positive payments are received by an agent who exerts high effort, which minimizes the agency rent.

The following corollary addresses how the strike price $X(\sigma)$ depends on the cost of effort C.

Corollary 1 *(Effect of cost of effort on strike price.) $X(\sigma)$ is strictly decreasing in C.*

The higher the cost of effort C, the stronger incentives must be, so the lower $X(\sigma)$ is to increase the delta of the option and thus the agent's incentives. This result is important, because the value of $X(\sigma)$ will play an important role in the next section.

3 The Value of Informativeness

3.1 Analytical results

We now determine the gains from increased informativeness. More precisely, we relate the agency rent to the variance σ^2 of the performance measure $\widetilde{s}$.

Since the contract in (13) is optimal for a given $\sigma \in \{\underline{\sigma}, \bar{\sigma}\}$, the principal's problem is to choose σ to minimize the sum of the agency rent and the cost κ if paid. Using (4) and (13), the optimization problem may be written:

$$\min_{\delta \in \{0,1\}} \int_{X(\sigma)}^{\infty} (s - X(\sigma))\, \psi(s|0) ds + (1-\delta)\,\kappa \qquad \text{s.t. (14).} \tag{15}$$

Proposition 1 studies how varying σ forces the strike price of the option, $X(\sigma)$, to change in order to maintain incentive compatibility.

Proposition 1 *(Effect of volatility on strike price.) (i) $\frac{dX(\sigma)}{d\sigma} > 0$ if and only if $X > \widehat{X}$. (ii) $\frac{d^2X}{d\sigma dX} > 0$.*

Proof. (Sketch of proof of part (i)). Implicitly differentiating the binding version of the incentive constraint (2) yields:

$$\frac{\partial}{\partial \sigma}\{E[W(\widetilde{s})|\bar{e}] - E[W(\widetilde{s})|0]\} + \frac{\partial}{\partial X}\{E[W(\widetilde{s})|\bar{e}] - E[W(\widetilde{s})|0]\}\frac{dX(\sigma)}{d\sigma} = 0,$$

which becomes:

$$\frac{dX(\sigma)}{d\sigma} = -\frac{\frac{\partial}{\partial\sigma}\left\{E\left[W(\widetilde{s})|\bar{e}\right] - E\left[W(\widetilde{s})|0\right]\right\}}{\frac{\partial}{\partial X}\left\{E\left[W(\widetilde{s})|\bar{e}\right] - E\left[W(\widetilde{s})|0\right]\right\}}. \tag{16}$$

Appendix A shows that the denominator of the right-hand side ("RHS") is always negative, and that the numerator is positive if and only if $X > \widehat{X}$. Thus, $\frac{dX(\sigma)}{d\sigma} > 0$ if and only if $X > \widehat{X}$. ■

The denominator of the RHS of (16) represents the effect of changes in X on the agent's incentives, $E\left[W(\widetilde{s})|\bar{e}\right] - E\left[W(\widetilde{s})|0\right]$. The agent's compensation is given by a call option on the underlying variable $\widetilde{s}$, and his incentives to work arise because effort increases the mean of $\widetilde{s}$ – working gives him an option worth $E\left[W(\widetilde{s})|\bar{e}\right]$ rather than one worth $E[W(\widetilde{s})|0]$. To highlight the dependence of the option values on the strike price, let $Y(e, X)$ denote (the value of) an option where the mean value of the underlying variable is e and the strike price is X. We thus have

$$\begin{aligned} Y(\bar{e}, X) &= E\left[W(\widetilde{s})|\bar{e}\right] \\ Y(0, X) &= E\left[W(\widetilde{s})|0\right]. \end{aligned}$$

An increase in the strike price X reduces the value of both call options, but particularly for $Y(\bar{e}, X)$ as it is more in the money. Thus, a rise in X reduces incentives, leading to a negative denominator. The sign of $\frac{dX(\sigma)}{d\sigma}$ therefore equals the sign of the numerator of (16).

In turn, the numerator represents the effect of changes in σ on the agent's incentives. This is equal to the vega (the sensitivity with respect to σ) of the option worth $Y(\bar{e}, X)$ minus the vega of the option worth $Y(\bar{e}, 0)$. The vega of an option is always positive, highest for an at-the-money option, and declines when the option moves either in-the-money or out-of-the-money. If the option has a strike price of $\widehat{X} = \frac{\bar{e}}{2}$ (the least informative performance measure), then $Y\left(\bar{e}, \frac{\bar{e}}{2}\right)$ is in-the-money by $\frac{\bar{e}}{2}$, and $Y\left(0, \frac{\bar{e}}{2}\right)$ is out-of-the-money by $\frac{\bar{e}}{2}$. Thus, both options have the same vega, and so increases in σ reduce the values of $Y\left(\bar{e}, \frac{\bar{e}}{2}\right)$ and $Y\left(0, \frac{\bar{e}}{2}\right)$ equally. The incentives to exert effort, $Y\left(\bar{e}, \frac{\bar{e}}{2}\right) - Y\left(0, \frac{\bar{e}}{2}\right)$, are unchanged, and so the strike price X does not need to change. We thus have $\frac{dX(\sigma)}{d\sigma} = 0$ for $X = \widehat{X}$.

Now consider $X < \widehat{X}$. Then, $Y(0, X)$ is closer to being at-the-money than $Y(\bar{e}, X)$, and so it has a higher vega. The intuition is as follows. Volatility increases the value of an option because the option holder benefits from its asymmetric payoff: his downside risk is limited, but he benefits from the upside gain. Since the strike price is low, if the agent works (and receives an option worth $Y(\overline{e}, X)$), the expected signal $\overline{e}$ is very far from the kink X, and thus the agent benefits little from the asymmetry. Thus, when volatility increases, an agent who works benefits from the upside potential but also bears the downside risk, and so $Y(\overline{e}, X)$ rises little with σ. In contrast, if the agent shirks (and receives $Y(0, X)$), the expected signal 0 is close to the strike price X, i.e. close to the kink. Thus, when volatility increases, an agent who works benefits from the upside potential and is protected from the downside risk. Thus, $Y(0, X)$ rises significantly with σ. In sum, an increase in σ reduces the agent's effort incentives, and so a fall in $X(\sigma)$ is needed to restore incentive compatibility (since it increase the value of $Y(\overline{e}, X)$ more than $Y(0, X)$). In simple language, when volatility rises and $X < \widehat{X}$, the agent thinks: "I'm not going to bother working hard, because even if I do, I might be unlucky and so profits will be low. I might as well take it easy, because even if I get unlucky and profits become very low, that doesn't matter, because I can't get paid less than zero no matter how low profits get."

Finally, consider $X > \widehat{X}$. Then, since $Y(\overline{e}, X)$ is closer to being at-the-money than $Y(0, X)$, it has a higher vega (the intuition is analogous to the case of $X < \widehat{X}$). Since $Y(\overline{e}, X)$ is close to the kink, when volatility increases, an agent who works benefits from the upside potential and is protected from the downside risk. Thus, $Y(\overline{e}, X)$ rises significantly with σ. In contrast, if the agent shirks (and receives $Y(0, X)$), the expected signal 0 is well below the kink. Thus, when volatility increases, the agent does not bear the downside risk, but is unlikely to benefit from the upside potential either: even if noise is positive, the option will still be out-of-the-money. Thus, $Y(0, X)$ rises little with σ. In sum, an increase in σ augments the agent's effort incentives, and so a rise in $X(\sigma)$ is possible without violating the incentive constraint. In simple language, when volatility rises and $X > \widehat{X}$, the agent thinks: "If volatility was low, I wouldn't bother working because the target $\widehat{X}$ is so high that I wouldn't meet it, even if I did work. But, now that volatility is high, I will work – because if I do, and I get lucky, I'll meet the target. But if I get lucky and I don't work, luck alone won't be enough to meet the target."

Part (ii) of Proposition 1 states a related result. As the initial strike price X rises, increases in informativeness (i.e. reductions in the variance) have a more negative effect on the strike price. In other words, improvements in informativeness increasingly erode incentives, and require the principal to lower the strike price to maintain incentive compatibility.

Proposition 1 implies that, in all cases, improvements in informativeness draw the strike price X towards $\widehat{X}$, i.e. bring options closer to at-the-money. Thus, improvements in signal precision (e.g. increases in stock market efficiency) lead to at-the-money options being optimal, in contrast to Bebchuk and Fried's (2004) concern that the almost universal practice of granting at-the-money options is inefficient. They argue that out-of-the-money options would be cheaper for the firm, but this view ignores the incentive effect: out-of-the-money options have lower deltas and thus may provide the agent with insufficient incentives.

In addition, the proof and intuition behind Proposition 1 suggest that exogenous changes in σ will have different effects on the incentives of agents depending on the moneyness of their options. In particular, it will increase (reduce) effort the incentives of agents with out-of-the-money (in-the-money) options. Thus, if firms are able to alter the terms of existing stock options in response to these unexpected changes, they will reduce the strike price of agents with in-the-money options. Such repricing is found empirically by Brenner, Sundaram, and Yermack (2000); Acharya, John, and Sundaram (2000) also study the repricing of options theoretically, although in responses to changes in the mean rather than volatility of the signal.

Corollary 2 gives a necessary and sufficient condition for $X > \hat{X}$, i.e. for improvements in informativeness to weaken incentives.

Corollary 2 *(Condition for high strike price.)* $X > \hat{X}$ *if and only if* $\int_{\hat{X}}^{\infty} (s-\hat{X}) \left(\psi(s|\bar{e}) - \psi(s|0)\right) > C$.

Proof. As shown in the proof of Lemma 1, we have $\int_0^{\infty} s \left(\psi(s|\bar{e}) - \psi(s|0)\right) > C$ and $\lim_{X\to\infty} \int_X^{\infty} (s - X) \left(\psi(s|\bar{e}) - \psi(s|0)\right) < C$, and the LHS of the incentive constraint in (14) is strictly decreasing in X. Given that the equilibrium X satisfies (14) as an equality, we have the above result. ■

We now turn from studying the effect of volatility on the exercise price to its effect on the agency rent. There are two effects. The first is the direct effect: as is well-known, a reduction in σ (an increase in informativeness) reduces the value of an option, and thus the agency rent. This benefits the principal. The second is the indirect effect: as shown in Proposition 1, to maintain incentive compatibility, a change in σ forces the strike price to either increase or decrease, which in turn affects the value of the option and thus the agency rent. Lemma 2 compares the partial (direct) effect of volatility on the agency rent, which ignores the change in the strike price, with the total effect, which also includes the indirect effect via the strike price.

Lemma 2 *(Partial and total effects of volatility on agency rent.)*

$$\frac{dAR(\sigma)}{d\sigma} < \frac{\partial AR(\sigma)}{\partial \sigma} \qquad \text{if and only if } X > \widehat{X}. \tag{17}$$

Proof. The total derivative is given by

$$\frac{dAR(\sigma)}{d\sigma} = \frac{\partial AR(\sigma)}{\partial \sigma} + \frac{\partial AR(\sigma)}{\partial X}\frac{dX(\sigma)}{d\sigma}. \tag{18}$$

The partial derivative

$$\frac{\partial AR(\sigma)}{\partial X} = \frac{\partial}{\partial X}E\left[W(\widetilde{s})|0\right] = \frac{\partial}{\partial X}\int_{X(\sigma)}^{\infty}(s - X(\sigma))\,\psi(s|0)ds = -\int_{X(\sigma)}^{\infty}\psi(s|0)ds$$

is always negative: an increase in the strike price reduces the agency rent by making the option less valuable. The derivative $\frac{dX(\sigma)}{d\sigma}$ reflects how the strike price must change with volatility to maintain incentive compatibility. From Proposition 1, it is positive if and only if $X > \widehat{X}$. Thus, $\frac{dAR(\sigma)}{d\sigma} < \frac{\partial AR(\sigma)}{\partial \sigma}$ if and only if $X > \widehat{X}$. ■

Lemma 2 shows that an increase in volatility has two effects on the agency rent. First, $\frac{\partial AR(\sigma)}{\partial \sigma}$ captures the "volatility effect": the direct effect of volatility on the value of a call option, which is positive. Second, the "strike price effect" arises because an increase in volatility changes the strike price $X(\sigma)$ necessary for the incentive constraint to continue to hold: $\frac{dX(\sigma)}{d\sigma}$ may be positive or negative. Since $\frac{\partial AR(\sigma)}{\partial X} < 0$, any increase (decrease) in the strike price lowers (augments) the total agency rent. Thus, the strike price effect can either reinforce or offset the volatility effect. For high incentives ($X < \widehat{X}$), $\frac{dX(\sigma)}{d\sigma} < 0$ (see Lemma 1), so that the volatility effect and the strike price effect both lead to a lower agency rent as σ is reduced. For low incentives, however ($X > \widehat{X}$), then $\frac{dX(\sigma)}{d\sigma} > 0$, so that the strike price effect partly offsets the volatility effect. In this case, the gains from improved informativeness (i.e., from a lower σ) are smaller when considering the incentive constraint, and thus calculating the gains using the total rather than partial derivative.

Proposition 2 below gives the expression for the effect of changes in informativeness on the agency rent.

Proposition 2 *(Effect of volatility on agency rent.) For a given* $\sigma \in \{\underline{\sigma}, \bar{\sigma}\}$, $\frac{dAR(\sigma)}{d\sigma}$ *is given by:*

$$\frac{dAR(\sigma)}{d\sigma} = \varphi\left(\frac{X(\sigma)}{\sigma}\right) - \left[1 - \Phi\left(\frac{X(\sigma)}{\sigma}\right)\right] \frac{\varphi\left(\frac{X(\sigma)-\bar{e}}{\sigma}\right) - \varphi\left(\frac{X(\sigma)}{\sigma}\right)}{\Phi\left(\frac{X(\sigma)}{\sigma}\right) - \Phi\left(\frac{X(\sigma)-\bar{e}}{\sigma}\right)} \tag{19}$$

and $\frac{dAR(\sigma)}{d\sigma} > 0$.

Proof. (Sketch). We start with the definition of $\frac{dAR(\sigma)}{d\sigma}$ in equation (18). We substitute for $\frac{dX(\sigma)}{d\sigma}$ using (16), and for $\frac{\partial AR(\sigma)}{\partial\sigma}$ and $\frac{\partial AR(\sigma)}{\partial X}$ using (4). This yields:

$$\frac{dAR(\sigma)}{d\sigma} = \frac{\partial}{\partial\sigma}\{E[W(\widetilde{s})|0]\} - \frac{\partial}{\partial X}\{E[W(\widetilde{s})|0]\} \frac{\frac{\partial}{\partial\sigma}\{E\left[W(\widetilde{s})|\bar{e}\right] - E\left[W(\widetilde{s})|0\right]\}}{\frac{\partial}{\partial X}\{E\left[W(\widetilde{s})|\bar{e}\right] - E\left[W(\widetilde{s})|0\right]\}}. \tag{20}$$

Calculating the derivatives yields:

$$\frac{dAR(\sigma)}{d\sigma} = \varphi\left(\frac{X(\sigma)}{\sigma}\right) - \left[1 - \Phi\left(\frac{X(\sigma)}{\sigma}\right)\right] \frac{\varphi\left(\frac{X(\sigma)-\bar{e}}{\sigma}\right) - \varphi\left(\frac{X(\sigma)}{\sigma}\right)}{\Phi\left(\frac{X(\sigma)}{\sigma}\right) - \Phi\left(\frac{X(\sigma)-\bar{e}}{\sigma}\right)}. \tag{21}$$

Appendix A shows that $\frac{dAR(\sigma)}{d\sigma} > 0 \; \forall \; X$. ∎

The result that $\frac{dAR(\sigma)}{d\sigma} > 0$ means that the strike price effect never outweighs the volatility effect: even though improvements in informativeness may weaken incentives, they can never make the principal worse off. However, the strike price effect is still important to consider, since it may reduce the benefits of informativeness sufficiently for it to be efficient for the principal not to pay the cost κ, as we will show in Proposition 3.

Note that $\frac{dAR(\sigma)}{d\sigma} > 0$ is equivalent to $\frac{\varphi\left(\frac{X(\sigma)-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X(\sigma)}{\sigma}\right)}{\varphi\left(\frac{X(\sigma)}{\sigma}\right)} < \frac{\Phi\left(\frac{X(\sigma)}{\sigma}\right)-\Phi\left(\frac{X(\sigma)-\bar{e}}{\sigma}\right)}{1-\Phi\left(\frac{X(\sigma)}{\sigma}\right)}$. The left-hand-side ("LHS") is the (ambiguous) effect of changing σ on incentives, divided by its (positive) effect on the agency rent. This ratio turns out to be the likelihood ratio for $i = X(\sigma)$, i.e. where the option is at the money. The RHS is the (negative) effect of changing X on incentives, divided by its (negative) effect on the agency rent.

Corollary 3 shows that the effect of volatility on the agency rent asymptotes to zero as the moral hazard problem becomes weaker.

Corollary 3 *(Effect of volatility on agency rent, limiting case)*

$$\frac{dAR(\sigma)}{d\sigma} \longrightarrow_{X\to\infty} 0 \qquad \textit{and} \qquad \frac{\frac{dAR(\sigma)}{d\sigma}}{\frac{\partial AR(\sigma)}{\partial\sigma}} \longrightarrow_{X\to\infty} 0 \tag{22}$$

$$\frac{dAR(\sigma)}{d\sigma} \longrightarrow_{C\to 0} 0 \qquad \textit{and} \qquad \frac{\frac{dAR(\sigma)}{d\sigma}}{\frac{\partial AR(\sigma)}{\partial\sigma}} \longrightarrow_{C\to 0} 0. \tag{23}$$

The first part of Corollary 3 shows that, as the strike price X approaches infinity, the total effect of informativeness on the agency rent tends to zero, and the gains from improved informativeness are infinitely smaller when assessed with the total derivative than with the partial derivative.

While the strike price X is an endogenous variable, the effort cost C is the exogenous parameter that drives changes in X. When C tends to zero, the moral hazard problem becomes weaker and fewer incentives (a lower option delta) is required; this is achieved by augmenting the strike price X.

We now compare the principal's gains from reducing the agency rent by increasing informativeness with the cost of doing so. She will filter out the shock (choose $\delta = 0$) if and only if:

$$AR(\overline{\sigma}) > AR(\underline{\sigma}) + \kappa.$$

Due to the assumption that $\bar{\sigma} - \underline{\sigma}$ is arbitrarily small, we can do a first-order Taylor expansion to study how changes in informativeness affect the agency rent:

$$AR(\overline{\sigma}) = AR(\underline{\sigma}) + \frac{dAR(\sigma)}{d\sigma}(\overline{\sigma} - \underline{\sigma}).$$

Thus, the principal increases informativeness if and only if

$$\frac{dAR(\sigma)}{d\sigma} > \frac{\kappa}{\bar{\sigma} - \underline{\sigma}}. \tag{24}$$

Proposition 3 *(Low-powered incentives and informativeness.) For any given κ, there exists $\bar{X}(\kappa) \geq \widehat{X}$ such that, if $X > \bar{X}(\kappa)$, the principal sets $\delta = 1$.*

Proof. The principal sets $\delta = 1$ if and only if the cost of informativeness outweighs the gains, that is:

$$\frac{\kappa}{\bar{\sigma}-\underline{\sigma}} > \frac{dAR(\sigma)}{d\sigma} = \frac{\partial AR(\sigma)}{\partial\sigma} + \frac{\partial AR(\sigma)}{\partial X}\frac{dX(\sigma)}{d\sigma} \tag{25}$$

For $X > \hat{X}$, the second-term on the RHS is negative according to Lemma 2. Therefore, (25) is satisfied if:

$$\frac{\kappa}{\bar{\sigma}-\underline{\sigma}} > \frac{\partial AR(\sigma)}{\partial\sigma}. \tag{26}$$

Using (4) and (39) in the proof of Proposition 1, this becomes:

$$\frac{\kappa}{\bar{\sigma}-\underline{\sigma}} > \varphi\left(\frac{X(\sigma)}{\sigma}\right) > 0. \tag{27}$$

We know that φ, the p.d.f. of the standard normal variable, is monotonically decreasing in $X \in [0,\infty)$, and approaches zero as $X \to \infty$. There are two cases to consider.

(i) If $\kappa > \varphi(0)$, (27) is satisfied for any $X > \hat{X}$. In this case, we can set $\bar{X} = \hat{X}$: $X > \hat{X}$ is sufficient (although unnecessary) for the principal to set $\delta = 1$.

(ii) If $\kappa < \varphi(0)$, it follows that, for any given κ, there exists $\bar{X}(\kappa)$ which is defined as

$$\frac{\kappa}{\bar{\sigma}-\underline{\sigma}} \equiv \varphi\left(\frac{\bar{X}(\kappa)}{\sigma}\right) \tag{28}$$

with $\bar{X}(\kappa) > \hat{X} > 0$, and (27) is satisfied if and only if $X > \bar{X}$.

Since (27) is sufficient for (25), it follows that (25) is satisfied if $X > \bar{X}$.

■

For sufficiently low-powered incentives ($X > \bar{X}$), the principal chooses to not filter out the shock. The intuition is as follows. At $X = \hat{X}$, the indirect effect is zero, so the total benefit of informativeness equal the direct effect, which is given by the option's vega. As X rises above $\hat{X}$, the direct effect becomes less positive, as an option's vega falls as it moves out-of-the-money. In addition, the indirect effect becomes negative as a rise in informativeness weakens incentives. Thus, if the total effect $\frac{dAR(\sigma)}{d\sigma}$ is less than $\frac{\kappa}{\bar{\sigma}-\underline{\sigma}}$ at $X = \hat{X}$ (case (i)), we have $\frac{dAR(\sigma)}{d\sigma} < \frac{\kappa}{\bar{\sigma}-\underline{\sigma}}$ for all $X > \hat{X}$ and so we can set $\bar{X} = \hat{X}$. If $\frac{dAR(\sigma)}{d\sigma} > \frac{\kappa}{\bar{\sigma}-\underline{\sigma}}$ at $X = \hat{X}$, (case ii), since the direct effect is monotonically decreasing in X, there will exist an $\bar{X} > \hat{X}$ for which the direct effect equals $\frac{\kappa}{\bar{\sigma}-\underline{\sigma}}$. Since the indirect effect is negative for $X > \hat{X}$, this is sufficient for $\delta = 1$.

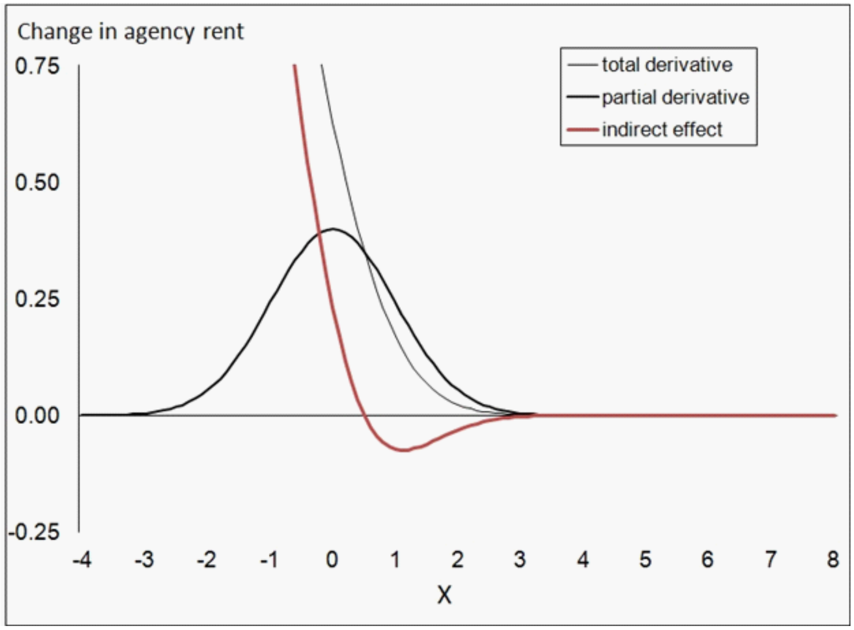

Figure 1: Total and partial derivative of the agency rent with respect to σ for a range of values of X, for $\bar{e} = 1$ and $\sigma = 1$.

Corollary 4 *(Cost of effort and informativeness) For any given κ, there exists a cost of effort $\bar{C}$ such that, for $C < \bar{C}$, the principal sets $\delta = 1$.*

Corollary 4 then shows that, for any $\kappa > 0$, it is suboptimal to set $\delta = 1$ and filter out the shock $\widetilde{\varepsilon}$ as $C \to 0$ (i.e., the cost of effort approaches zero). Thus, pay-for-luck is optimal.

3.2 Graphical illustrations

We now demonstrate graphically the importance of considering the incentive constraint when evaluating the effect of informativeness on incentives, i.e. studying the total rather than partial derivative.

In Figures 1 and 2, we illustrate both the value of the total derivative $\frac{dAR(\sigma)}{d\sigma}$, as calculated in (19), and the value of the partial derivative $\frac{\partial AR(\sigma)}{\partial \sigma}$, as calculated in (39) in the Appendix, for a range of values of X. Figure 1 uses $\sigma = 1$, and Figure 2 uses $\sigma = 3$. Both figures consider $\bar{e} = 1$ (and so $\widehat{X} = 0.5$); numerical simulations show that $\bar{e}$ has little effect on the shapes of the functions displayed in Figures 1 and 2.

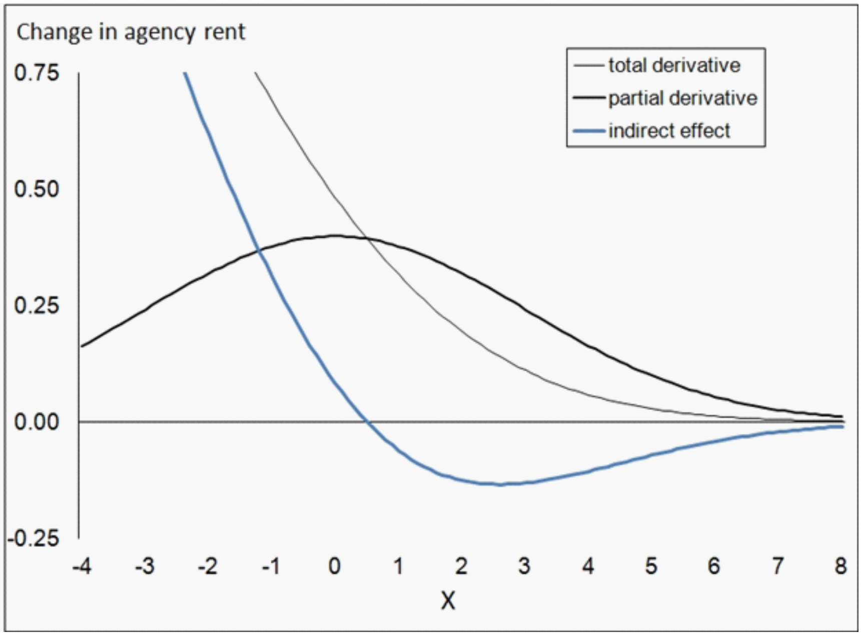

Figure 2: Total and partial derivative of the agency rent with respect to σ for a range of values of X, for $\bar{e} = 1$ and $\sigma = 3$.

To understand the graphs, recall that the total derivative is given by $\frac{dAR(\sigma)}{d\sigma} = \frac{\partial AR(\sigma)}{\partial \sigma} + \frac{\partial AR(\sigma)}{\partial X}\frac{dX(\sigma)}{d\sigma}$. The gains associated with the partial derivative, $\frac{\partial AR(\sigma)}{\partial \sigma}$, tend to zero as the strike price approaches either $-\infty$ or ∞. The vega of an option is greatest when the option is at-the-money, i.e. $X = \bar{e}$. An at-the-money option benefits most from the asymmetry in an option's payoff: a high noise realization leads to a large increase in the option's payoff, but a low noise realization has no effect as the agent will not exercise the option.

The indirect effect, $\frac{\partial AR(\sigma)}{\partial X}\frac{dX(\sigma)}{d\sigma}$, captures the effect of volatility on the agency rent through affecting the agent's incentives. From Proposition 1, $\frac{dX(\sigma)}{d\sigma}$ is positive if and only if $X > \widehat{X}$. Since $\frac{\partial AR(\sigma)}{\partial X} < 0$, the indirect effect in Figures 1 and 2 is negative if and only if $X < \widehat{X} = \frac{1}{2}$. As X decreases below $\widehat{X}$, the incentive benefits of greater informativeness strengthen ($\frac{dX(\sigma)}{d\sigma}$ becomes more negative). Since the option becomes increasingly in the money, $\frac{\partial AR(\sigma)}{\partial X}$ becomes even more negative and falls towards -1. Thus, the indirect effect $\frac{\partial AR(\sigma)}{\partial X}\frac{dX(\sigma)}{d\sigma}$ becomes monotonically more positive as X falls. However, as X rises above $\widehat{X}$, there are two effects working in opposite directions. On the one hand, greater informativeness becomes increasingly detrimental to incentives ($\frac{dX(\sigma)}{d\sigma}$ becomes more positive). On the other hand, the agency rent is less affected by the strike price, and $\frac{\partial AR(\sigma)}{\partial X}$ approaches zero: when the option is deeply out-of-the-money, its value is close to zero and thus it is little affected by changes in the strike price. Thus, the impact of X on the indirect effect is non-monotonic. As X initially rises above $\widehat{X}$, the indirect effect becomes increasingly negative as the option has significant value, so the change in the strike price required to maintain incentives has a large increase on this value. However, as X continues to rise, the option's value asymptotes towards zero and becomes little affected by the strike price, so the indirect effect also asymptotes towards zero and becomes less negative.

The total derivative $\frac{dAR(\sigma)}{d\sigma}$ combines these two effects. While the direct effect is initially increasing in X, this is outweighed by the fact that the indirect effect is initially decreasing in X. Overall, the total gains from increased informativeness are monotonically decreasing in X. The speed of convergence of the total derivative towards zero is striking. For $\sigma = 1$, at $X = 0$ the gains from a marginal change in σ are more than two millions times greater than at $X = 5$. For $\sigma = 3$, however, the gains are only 16.9 times greater at $X = 0$ than at $X = 5$.

In both Figures, consistent with Lemma 2, considering only the partial derivative leads to an underestimation (overestimation) of the total gains from improved informativeness for $X < (>)\, \widehat{X}$. The partial sensitivity measure substantially overestimates the total gains for sufficiently large X. For example, for $\sigma = 1$ and $X = 2$ (which is only one standard deviation away from the expected performance of $\bar{e} = 1$), the gains from a marginal change in σ are 2.4 times larger with the partial sensitivity measure than with the total sensitivity measure. Thus, even for non-extreme parameter values, gains from improved informativeness can be much lower if the strike price effect is taken into account. Comparing across the Figures shows that, for out-of-the-money options, the discrepancy is decreasing in σ: for $X = 2$ and $\sigma = 3$, the gains from a marginal change in σ are 1.6 times larger with the partial sensitivity measure than with the total sensitivity measure. For $X = \bar{e}$ (the benchmark case of an at-the-money option), with $\sigma = 1$ (respectively $\sigma = 3$) the gains are 1.4 (respectively 1.2) times higher with the partial sensitivity measure than with the total sensitivity measure.

Another comparison across the Figures shows that the direct effect of improved informativeness is higher in Figure 2, where initial volatility is greater. This direct effect is given by vega, and so the effect of initial volatility on vega is given by an option's vomma, the second derivative of its value with respect to volatility ($\frac{\partial^2 Y}{\partial \sigma^2}$). For a normally distributed stock price, vomma is always positive:

$$\frac{\partial^2 AR(\sigma)}{\partial \sigma^2} = \frac{\partial}{\partial \sigma}\varphi\left(\frac{X(\sigma)}{\sigma}\right) = \frac{1}{\sqrt{2\pi}}\frac{X^2}{\sigma^3}\exp\left\{-\frac{X^2}{2\sigma^2}\right\} > 0.$$

Thus, a higher initial volatility augments an option's vega, and thus the principal's gain from increased informativeness. That an increase in σ augments the direct effect also explains why the discrepancy between the partial and total sensitivities shrinks with σ: as the direct effect increases, it comprises a larger component of the total.

Both Figures 1 and 2 show that the total gains from increased informativeness are declining in X. One concern is that when X is high, the moral hazard problem is small to begin with, so it may seem unsurprising that the gains from alleviating the moral hazard problem (by increasing informativeness) are small in absolute terms. Thus, it seems natural that, if the cost κ of increasing informativeness is fixed in absolute terms, the principal will only pay the cost if the moral hazard problem is large. We thus now compute the gains from improved informativeness as a *percentage* of the expected cost of compensation, i.e. calculate $\frac{\frac{\partial AR}{\partial \sigma}}{AR+C}$ and $\frac{\frac{dAR}{d\sigma}}{AR+C}$.

These gains are illustrated in Figure 3, for $\sigma = 1$. When the gains are measured in percentage terms, the direction of the difference between the partial and total derivatives naturally remains unchanged, but the magnitude of the difference becomes much greater for a high X, since the expected cost of compensation is decreasing in X. When measured in absolute terms, the partial derivative was maximized at $X = 0$; when measured in relative terms, Figure 3 shows that it is maximized for $X > 0$, because the expected cost of compensation is decreasing in X. When measured in absolute terms, the total derivative was monotonically decreasing in X. When measured in relative terms, there is now an offsetting effect since the cost of compensation is also decreasing in X. However, the second effect is always weaker, so the total derivative in relative terms continues to decrease in X, but just at a slower rate than in absolute terms.[8]

Thus, even if the cost of increasing informativeness κ were a percentage of the cost of compensation, it remains the case that, when the incentive effect is taken into account, the principal will not filter out the shock $\widetilde{\varepsilon}$ when X is sufficiently high. Thus, employees with high-powered incentives (such as CEOs) should be less paid for luck

[8] We intend to prove this result analytically in a future draft.

and more subject to RPE than those with low-powered incentives (such as rank-and-file employees). Oyer and Schaefer (2005), Hochberg and Lindsey (2010), and Kim and Ouimet (2013) indeed find that rank-and-file employees are given significant non-indexed options. Note that an analysis focusing on the partial derivative and ignoring the incentive effect would reach a different opposite conclusion: if the cost of increasing informativeness were in percentage terms, the principal would not filter out the shock not only if X is sufficiently high, but also if it is sufficiently low.

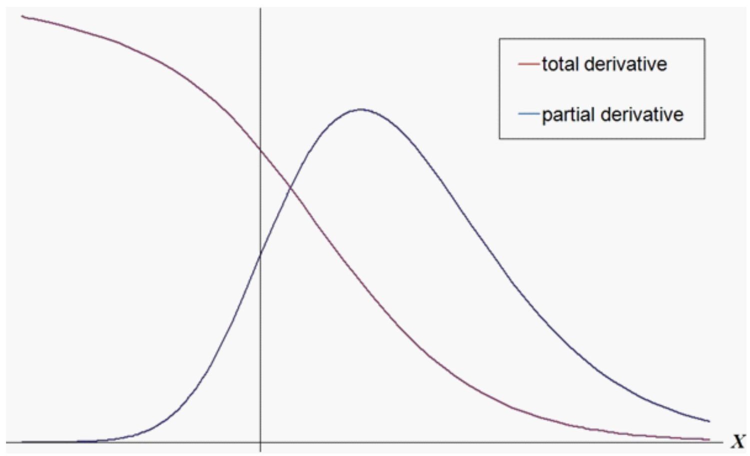

Figure 3: Total and partial derivative of the agency rent with respect to σ for a range of values of X, for $\bar{e} = 1$ and $\sigma = 1$, expressed as a fraction of the expected cost of compensation in equilibrium.

4 Conclusion

This paper studies the principal's benefits from increasing the informativeness of the performance measure used to evaluate the agent. In particular, it shows the criticality of taking into account the effect of increased informativeness on the agent's effort incentives, and that ignoring this effect can lead to incorrect conclusions on whether the principal should incur the cost of filtering out exogenous shocks.

We show that, if incentives are low-powered to begin with (the contract involves an out-of-the-money option), an increase in informativeness reduces the agent's effort incentives. The intuition is that, if the option is sufficiently out-of-the-money when granted, it will only become in-the-money at maturity if the agent both exerts effort and receives a sufficiently positive shock. When informativeness increases, and noise falls, positive shocks are less likely and so the agent may choose not to work. Thus, while the increase in informativeness directly reduces the agent's rents by decreasing the value of his option, this effect is partially offset by the fact that the strike price of the option must fall, increasing the agent's rents, to maintain his effort incentives. If increasing informativeness is costly, it may be optimal for the principal not to filter out exogenous shocks, rationalizing the scarcity of relative performance evaluation in reality. Interestingly, thus result continues to hold when the cost of increasing informativeness is proportional to the cost of compensation and thus the severity of the moral hazard problem.

In contrast, when incentives are high-powered to begin with, an increase in informativeness augments the agent's effort incentives, which provides an additional gain from informativeness over and above the direct effect of reducing volatility traditionally focused upon. Thus, the benefit from filtering out exogenous shocks depends critically on the strength of incentives, and thus the magnitude of the moral hazard problem to begin with.

References

[1] Acharya, Viral V., Kose John, and Rangarajan K. Sundaram (2000): "On the optimality of resetting executive stock options." *Journal of Financial Economics* 57, 65-101.

[2] Aggarwal, Rajesh K. and Andrew A. Samwick (1999): "Executive compensation, relative performance evaluation, and strategic competition: theory and evidence." *Journal of Finance* 54, 1999–2043.

[3] Arnott, Richard and Joseph E. Stiglitz (1988): "Randomization with Asymmetric Information." *RAND Journal of Economics* 19, 344–362.

[4] Axelson, Ulf and Sandeep Baliga (2009): "Liquidity and manipulation of executive compensation schemes." *Review of Financial Studies* 22, 3907–3939.

[5] Bebchuk, Lucian and Jesse Fried (2004): *Pay without performance: the unfulfilled promise of executive compensation.* (Harvard University Press, Cambridge).

[6] Bertrand, Marianne and Sendhil Mullainathan (2001): "Are CEOs rewarded for luck? The ones without principals are." *Quarterly Journal of Economics* 116, 901–932.

[7] Bolton, Patrick and Mathias Dewatripont (2005) *Contract Theory* (MIT Press, Cambridge).

[8] Brenner, Menachem, Rangarajan K. Sundaram, and David Yermack (2000): "Altering the terms of executive stock options." *Journal of Financial Economics 57*, 103–128.

[9] Core, John E. and Wayne R. Guay (2001): "Stock option plans for non-executive employees." *Journal of Financial Economics* 60, 253–267.

[10] Cornelli, Francesca, Zbigniew Kominek, and Alexander Ljungqvist (2013): "Monitoring managers: Does it matter?" *Journal of Finance* 68, 431–481.

[11] Danthine, Jean-Pierre, and John B. Donaldson (2008): "Executive compensation: a general equilibrium perspective." Working Paper, Columbia University.

[12] De Angelis, David, Gustavo Grullon, and Sébastien Michenaud (2013): "Downside risk and the design of CEO incentives: Evidence from a natural experiment." Working Paper, Rice University.

[13] Dewatripont, Mathias, Patrick Legros, and Steven A. Matthews (2003): "Moral hazard and capital structure dynamics." *Journal of the European Economic Association* 1, 890–930.

[14] Dittmann, Ingolf and Ernst Maug (2007): "Lower salaries and no options? On the optimal structure of executive pay." *Journal of Finance* 62, 303–343.

[15] Dittmann, Ingolf, Ernst Maug, and Oliver G. Spalt (2010): "Sticks or carrots? On the optimal structure of executive pay." *Journal of Finance* 65, 2015–2050.

[16] Dittmann, Ingolf, Ernst Maug, and Oliver G. Spalt (2011): "Indexing executive compensation contracts." Working Paper, Erasmus University Rotterdam.

[17] Edmans, Alex and Xavier Gabaix (2009): "Is CEO Pay Really Inefficient? A Survey of New Optimal Contracting Theories." *European Financial Management* 15, 486–496.

[18] Edmans, Alex and Xavier Gabaix (2011): "Tractability in Incentive Contracting." *Review of Financial Studies* 24, 2865–2894.

[19] Frydman, Carola and Dirk Jenter (2010): "CEO compensation." *Annual Review of Financial Economics* 2, 75–102.

[20] Gjesdal, Froystein (1982): "Information and incentives: the agency information problem." *Review of Economic Studies* 49, 373–390.

[21] Gopalan, Radhakrishnan, Todd T. Milbourn and Fenghua Song (2010): "Strategic flexibility and the optimality of pay for sector performance." *Review of Financial Studies* 23, 2060–2098.

[22] Gormley, Todd A., David A. Matsa, and Todd T. Milbourn (2013): "CEO compensation and corporate risk-taking: Evidence from a natural experiment." Working Paper, University of Pennsylvania.

[23] Grossman, Sanford J., and Oliver D. Hart (1983): "An analysis of the principal-agent problem." *Econometrica* 51, 7–45.

[24] Hall, Brian J., Kevin J. Murphy (2002): "Stock options for undiversified executives." *Journal of Accounting and Economics* 33, 3–42.

[25] Hochberg, Yael V. and Laura Lindsey (2010): "Incentives, Targeting, and Firm Performance: An Analysis of Non-executive Stock Options." *Review of Financial Studies* 23, 4148–4186

[26] Holmstrom, Bengt (1979): "Moral hazard and observability." *Bell Journal of Economics* 10, 74–91.

[27] Innes, Robert D. (1990): "Limited liability and incentive contracting with ex-ante action choices." *Journal of Economic Theory* 52, 45–67.

[28] Jenter, Dirk and Fadi Kanaan (2013): "CEO turnover and relative performance evaluation." *Journal of Finance*, forthcoming.

[29] Kim, E. Han and Paige Ouimet (2013): "Broad-based employee stock ownership: Motives and outcomes." Working Paper, University of Michigan.

[30] Kim, Son Ku (1995): "Efficiency of an information system in an agency model." *Econometrica* 63, 89–102.

[31] Murphy, Kevin J. (1999): "Executive compensation" in *Handbook of Labor Economics*, O. Ashenfelter and D. Card, eds. Elsevier/North-Holland.

[32] Murphy, Kevin J. (2002): "Explaining executive compensation: Managerial power versus the perceived cost of stock options." *University of Chicago Law Review* 69, 847–869.

[33] Oyer, Paul (2004): "Why do firms use incentives that have no incentive effects?" *Journal of Finance* 59, 1619–1649.

[34] Oyer, Paul, and Scott Schaefer (2005): "Why do firms give stock options to all employees?: An empirical examination of alternative theories." *Journal of Financial Economics* 76, 99–133.

[35] Raith, Michael (2008): "Specific knowledge and performance measurement." *Rand Journal of Economics* 39, 1059–1079.

[36] Shavell, Steven (1979): "Risk sharing and incentives in the principal and agent relationship." *Bell Journal of Economics* 10, 55–73.

A Appendix

Proof of Lemma 1

We start by arguing that the optimal contract $W(s)$ will be continuous. First, the likelihood ratio is increasing in s for any s, so that the optimal contract is nondecreasing in s for any s, which rules out downward discontinuities in $W(s)$. Second, the constraint that $W'(s) \leq 1$ rules out upward discontinuities.

Denoting the Lagrange multipliers associated with the three constraints (8), (9), and (10) for a given s by μ, $\lambda(s)$, and $\eta(s)$ respectively, the Lagrangian is given by:

$$\begin{aligned} L = & \int_{-\infty}^{\infty} W(s)\,\psi(s|0)ds - \mu\Big[\int_{-\infty}^{\infty} W(s)\,\psi(s|\bar{e})ds - C - \int_{-\infty}^{\infty} W(s)\,\psi(s|0)ds\Big] \\ & - \int_{-\infty}^{\infty} \lambda(s)W(s)\,ds - \int_{-\infty}^{\infty} \eta(s)(1 - W'(s))ds \\ = & \int_{-\infty}^{\infty} \Big((1+\mu)W(s) + \mu C - \mu\frac{\psi(s|\bar{e})}{\psi(s|0)}W(s)\Big)\psi(s|0)ds \\ & - \int_{-\infty}^{\infty} \Big(\lambda(s)W(s) - \eta(s)(1 - W'(s))\Big)ds. \end{aligned} \tag{29}$$

Since the optimal contract $W(s)$ is continuous and $\lim_{s\to-\infty} W(s) = 0$[9], we have

$$W(s) = \int_{-\infty}^{s} W'(x)dx = \int_{-\infty}^{s-ds} W'(x)dx + W'(s)ds. \tag{30}$$

It follows that

$$\frac{dW'(s)}{dW(s)} = \left(\frac{dW(s)}{dW'(s)}\right)^{-1} = \left(\frac{d\big(\int_{-\infty}^{s-ds} W'(x)dx + W'(s)ds\big)}{dW'(s)}\right)^{-1} = \frac{1}{ds}. \tag{31}$$

The first-order necessary condition with respect to $W(s)$ to the constrained opti-

[9]Since the likelihood ratio of the normal distribution is increasing, we know that the optimal contract is nondecreasing in π. Therefore, $W(\pi) \geq \lim_{x\to-\infty} W(x)$, for any π. Consider any contract $W(\pi)$ characterized by $\lim_{x\to-\infty} W(x) = \varepsilon > 0$. Then define $W^{\star}(\pi) \equiv W(\pi) - \varepsilon$ for all π. The contract $W^{\star}(\pi)$ does not violate any constraint (in particular, it satisfies the incentive constraint), and it is characterized by a lower agency rent than $W(\pi)$. This shows that the contract $W(\pi)$ is not optimal.

mization problem is

$$\left((1+\mu)+-\mu\frac{\psi(s|\bar{e})}{\psi(s|0)}\right)\psi(s|0)-\lambda(s)+\eta(s)\frac{1}{ds}=0. \tag{32}$$

We define

$$\phi(s)\equiv\left((1+\mu)+-\mu\frac{\psi(s|\bar{e})}{\psi(s|0)}\right)\psi(s|0) \tag{33}$$

If $\phi(s)>0$ for a given s, then (32) imposes:

$$-\lambda(s)+\eta(s)\frac{1}{ds}<0.$$

Given the non-negativity constraints on the Lagrange multipliers, a necessary condition for this inequality to hold is $\lambda(s)>0$. However, $\lambda(s)>0$ if and only if $W(s)=0$.

If $\phi(s)<0$ for a given s, then (32) imposes that

$$-\lambda(s)+\eta(s)\frac{1}{ds}>0. \tag{34}$$

Given the non-negativity constraints on the Lagrange multipliers, this implies that $\eta(s)>0$.

Since μ is a constant and the likelihood ratio is monotonically increasing in s, we know from (33) that $\phi(s)=0$ on a set of probability zero, and that the optimal contract (which is incentive-compatible) exists if there exists a finite and positive X such that

$$\phi(s)>0 \quad\Leftrightarrow\quad \frac{\psi(s|\bar{e})}{\psi(s|0)}<\frac{1+\mu}{\mu} \quad\Leftrightarrow\quad s<X, \tag{35}$$

$$\phi(s)<0 \quad\Leftrightarrow\quad \frac{\psi(s|\bar{e})}{\psi(s|0)}>\frac{1+\mu}{\mu} \quad\Leftrightarrow\quad s>X. \tag{36}$$

We also know that $W(s)=0$ for $s<X$. We still have to determine the form of the optimal contract for $s>X$. Given the non-negativity constraints on the Lagrange multipliers discussed above, we know that, for $s>X$, (10) is binding since $\eta(s)>0$. The optimal contract is therefore characterized by $W(s)=0\ \forall\ s<X$, and $W'(s)=1$ $\forall\ s>X$.

We now prove that there is a unique positive $X(\sigma)$ that satisfies the incentive constraint (14) with equality. First, for a contract characterized by $W(s)=0\ \forall$

$s < X$, and $W'(s) = 1 \; \forall \; s > X$, the incentive constraint with a given X is

$$\int_X^\infty (s - X)\psi(s|\bar{e})ds - \int_X^\infty (s - X)\psi(s|0)ds = C. \tag{37}$$

The first derivative of the LHS of (37) with respect to X is equal to

$$-\int_X^\infty \psi(s|\bar{e})ds + \int_X^\infty \psi(s|0)ds$$

$$= -Pr(s > X|\bar{e}) + Pr(s > X|0)$$

which is strictly negative for any X, because of first-order stochastic dominance ("FOSD"), which is implied by MLRP.

Second, the LHS of (37) tends to zero as X approaches ∞. Third, the condition for effort to be optimal under the first-best, $\bar{e} > C$, implies $\int_{-\infty}^{\infty} s\left(\psi(s|\bar{e}) - \psi(s|0)\right) ds > C$, which in turn guarantees that the LHS of (37) is strictly larger than C for $X = 0$.

Therefore, there exists one and only one positive and finite X such that the LHS of (37) equals C, i.e., such that the incentive-constraint with the optimal contract is satisfied as an equality.

Proof of Corollary 1

On the one hand, $X(\sigma)$ satisfies the incentive constraint in (37) as an equality, by definition. On the other hand, the LHS of the incentive constraint in (37) is strictly decreasing in X (see Lemma 1). Since the RHS of (37) is equal to C, this implies that $X(\sigma)$ is strictly decreasing in C.

Proof of Proposition 1

We start with part (i). First,

$$\frac{\partial}{\partial\sigma}E[W(\widetilde{s})|e] = \frac{\partial}{\partial\sigma}\int_{X(\sigma)}^\infty (x - X(\sigma))\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-e)^2}{2\sigma^2}\right\}dx \tag{38}$$

$$= \frac{\partial}{\partial\sigma}\int_{X(\sigma)-e}^\infty (x + e - X(\sigma))\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}dx$$

$$= \frac{\partial}{\partial\sigma}\int_{X(\sigma)-e}^\infty x\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}dx - (X(\sigma) - e)\frac{\partial}{\partial\sigma}\int_{X(\sigma)-e}^\infty \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}dx$$

$$= \frac{\partial}{\partial \sigma}\Big\{\frac{\sigma}{\sqrt{2\pi}}\exp\Big\{-\frac{(X(\sigma)-e)^2}{2\sigma^2}\Big\}\Big\} - (X(\sigma)-e)\frac{\partial}{\partial \sigma}\int_{\frac{X(\sigma)-e}{\sigma}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\Big\{-\frac{u^2}{2}\Big\}du$$

$$= \frac{1}{\sqrt{2\pi}}\exp\Big\{-\frac{(X(\sigma)-e)^2}{2\sigma^2}\Big\}\Big[1+\frac{(X(\sigma)-e)^2}{\sigma^2}\Big] - \frac{1}{\sqrt{2\pi}}\frac{(X(\sigma)-e)^2}{\sigma^2}\exp\Big\{-\frac{(X(\sigma)-e)^2}{2\sigma^2}\Big\}$$

$$= \frac{1}{\sqrt{2\pi}}\exp\Big\{-\frac{(X(\sigma)-e)^2}{2\sigma^2}\Big\} = \varphi\Big(\frac{X(\sigma)-e}{\sigma}\Big) \tag{39}$$

which is positive. It follows that

$$\frac{\partial}{\partial \sigma}\left\{E\left[W(\widetilde{s})|\bar{e}\right] - E\left[W(\widetilde{s})|0\right]\right\} = \Big[\varphi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X(\sigma)}{\sigma}\Big)\Big].$$

That is, the RHS of the numerator of (16) is positive if and only if

$$\varphi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big) > \varphi\Big(\frac{X(\sigma)}{\sigma}\Big). \tag{40}$$

Since

$$\varphi\Big(\frac{\widehat{X}-\bar{e}}{\sigma}\Big) \equiv \varphi\Big(\frac{\widehat{X}}{\sigma}\Big)$$

by definition of $\widehat{X}$, and since two p.d.f.s of normal variables with the same variance intersect only once, we know that (40) is satisfied, and $\frac{dX(\sigma)}{d\sigma} > 0$, if and only if $X(\sigma) > \widehat{X}$.

Second,

$$\frac{\partial}{\partial X}E[W(\widetilde{s})|e] = \frac{\partial}{\partial X}\int_{X(\sigma)}^{\infty}(x - X(\sigma))\frac{1}{\sqrt{2\pi}\sigma}\exp\Big\{-\frac{(x-e)^2}{2\sigma^2}\Big\}dx$$

$$= -\int_{X(\sigma)}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\Big\{-\frac{(x-e)^2}{2\sigma^2}\Big\}dx = -\Big(1-\Phi\Big(\frac{X(\sigma)-e}{\sigma}\Big)\Big). \tag{41}$$

It follows that

$$\frac{\partial}{\partial X}\left\{E\left[W(\widetilde{s})|\bar{e}\right] - E\left[W(\widetilde{s})|0\right]\right\} = -\big(1-\Phi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big)\big) + \big(1-\Phi\Big(\frac{X(\sigma)}{\sigma}\Big)\big) = \Phi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big) - \Phi\Big(\frac{X(\sigma)}{\sigma}\Big).$$

MLRP implies FOSD, therefore

$$\Phi\Big(\frac{X(\sigma)}{\sigma}\Big) > \Phi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big).$$

which implies

$$\frac{\partial}{\partial X}\left\{E\left[W(\widetilde{s})|\bar{e}\right]-E\left[W(\widetilde{s})|0\right]\right\}<0.$$

Thus, the RHS of the denominator of (16) is negative.

We now move to part (ii). To simplify notations, define

$$x \equiv \frac{X}{\sigma}, t \equiv \frac{\bar{e}}{\sigma}.$$

We wish to show that $\forall t > 0$

$$f(x,t) \equiv [\varphi(x)-\varphi(x-t)]^2 - [\Phi(x)-\Phi(x-t)][\varphi'(x)-\varphi'(x-t)] > 0, \quad \forall x, \tag{42}$$

where

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

and

$$\Phi(x) = \int_{-\infty}^{x} \varphi(y)\, dy$$

For $t = 0$, $f(x,0)$ is trivially 0.

Since $\varphi(x) = \varphi(-x)$, we have $\Phi(x) - \Phi(x-t) = \Phi(-x+t) - \Phi(-x)$ and $\varphi'(x) - \varphi'(x-t) = \varphi'(-x+t) - \varphi'(-x)$. As a consequence, $f(x,t) = f(-x+t,t)$. We thus only have to study $x \geq \frac{t}{2} > 0$.

We first analyze the term $\varphi'(x) - \varphi'(x-t)$. Since

$$\varphi'(x) = -\frac{x}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$$

$$\varphi'(x) - \varphi'^{-t(x-t/2)} + x - t).$$

When $x \geq t/2$, the function $e^{-t(x-t/2)} - 1 + \frac{t}{x}$ is only equal to zero at one point, since it monotonically decreases from 2 to -1. Let that point be x_0. Then

$$\varphi'(x) - \varphi'(x-t) \begin{cases} < 0 & \frac{t}{2} \leq x < x_0 \\ = 0 & x = x_0 \\ > 0 & x > x_0 \end{cases}.$$

We know that when $x \in [\frac{t}{2}, x_0]$, $f(x,t) > 0$ since $[\varphi(x) - \varphi(x-t)]^2 > 0$ and $\Phi(x) - \Phi(x-t) > 0$ $\forall x$, so that (42) is proven for $x \in [\frac{t}{2}, x_0]$

We now prove (42) for $x > x_0$. Within that range (we omit the argument t in the following)

$$f(x,t) > 0 \iff g(x) \equiv \frac{f(x,t)}{\varphi'(x) - \varphi'(x-t)} > 0.$$

To prove the latter, we first calculate

$$\begin{aligned}
g'(x) =& \frac{2[\varphi(x) - \varphi(x-t)][\varphi'(x) - \varphi'^2 - [\varphi(x) - \varphi(x-t)]^2[\varphi''(x) - \varphi''(x-t)]}{[\varphi'(x) - \varphi'^2} \\
& - [\varphi(x) - \varphi(x-t)] \\
=& \frac{[\varphi(x) - \varphi(x-t)][\varphi'(x) - \varphi'^2 - [\varphi(x) - \varphi(x-t)]^2[\varphi''(x) - \varphi''(x-t)]}{[\varphi'(x) - \varphi'^2} \\
=& \frac{[\varphi(x) - \varphi(x-t)]\varphi(x-t)^2}{[\varphi'(x) - \varphi'^2} \\
& \left\{ \left(-xe^{-t(x-t/2)} + x - t \right)^2 - \left[(x^2-1)e^{-t(x-t/2)} - (x-t)^2 + 1 \right] \left(e^{-t(x-t/2)} - 1 \right) \right\} \\
=& \frac{[\varphi(x) - \varphi(x-t)]\varphi(x-t)^2}{[\varphi'(x) - \varphi'^2} \left[\left(e^{-t(x-t/2)} - 1 \right)^2 + t^2 e^{-t(x-t/2)} \right] \\
& < 0, \quad x \in (x_0, \infty),
\end{aligned}$$

where in the last step we used the fact that $\varphi(x) < \varphi(x-t)$ when $x > t/2$. Therefore,

$$g(x) > 0 \quad \forall x \in (x_0, \infty) \iff \lim_{x \to \infty} g(x) \geq 0.$$

Since

$$\begin{aligned}
g(x) =& \frac{[\varphi(x) - \varphi(x-t)]^2}{\varphi'(x) - \varphi'(x-t)} - \Phi(x) + \Phi(x-t) \\
=& \frac{1}{\sqrt{2\pi}} e^{-(x-t)^2/2} \frac{\left(e^{-t(x-t/2)} - 1 \right)^2}{-xe^{-t(x-t/2)} + x - t} - \Phi(x) + \Phi(x-t)
\end{aligned},$$

it is clear that

$$\lim_{x \to \infty} g(x) = 0.$$

Proof of Proposition 2

We wish to show for $t > 0$,

$$\frac{\Phi(x) - \Phi(x-t)}{1 - \Phi(x)} > \frac{\varphi(x-t) - \varphi(x)}{\varphi(x)}$$

which can be rewritten as

$$\frac{\Phi(t-x)}{\Phi(-x)} > \frac{\varphi(x-t)}{\varphi(x)}, \quad \forall x. \tag{43}$$

or equivalently

$$h(x) \equiv \Phi(t-x)\varphi(-x) - \Phi(-x)\varphi(t-x) > 0, \quad \forall x. \tag{44}$$

First, when $x \leq \frac{t}{2}$, $\varphi(t-x) \leq \varphi(-x)$. Therefore

$$\frac{\Phi(t-x)}{\Phi(-x)} > 1 \geq \frac{\varphi(t-x)}{\varphi(-x)}, \quad \forall x \leq \frac{t}{2}.$$

Thus, (44) is proven for $x \leq \frac{t}{2}$.

Next, we focus on the case when $x \geq t$. Since all terms in (43) are positive, we have to show

$$\left[\frac{\Phi(t-x)}{\varphi(t-x)}\right]^2 > \left[\frac{\Phi(-x)}{\varphi(-x)}\right]^2.$$

Written explicitly, we have

$$\Phi(-x)^2 = \frac{1}{2\pi}\int_{-\infty}^{-x}\int_{-\infty}^{-x} e^{-(a^2+b^2)/2}\,da\,db = \frac{1}{2\pi}\int_{x}^{\infty}\int_{x}^{\infty} e^{-(a^2+b^2)/2}\,da\,db = \frac{1}{\pi}\int_{x}^{\infty}\int_{b}^{\infty} e^{-(a^2+b^2)/2}\,da\,db.$$

Changing to polar coordinates[10] where $a = \rho\cos\theta$ and $b = \rho\sin\theta$, the integration

[10] The polar coordinates ρ(or r used in many cases) and θ(or φ) can be converted to the Cartesian coordinates a and b (or x and y used in most cases) by using the trigonometric functions sine and cosine:

$$a = \rho\cos\theta, \quad b = \rho\sin\theta.$$

The double integral form is

$$da\,db = \rho\,d\rho\,d\theta.$$

In our case, in the integral

$$\Phi(-x)^2 = \frac{1}{\pi}\int_{x}^{\infty}\int_{b}^{\infty} e^{-(a^2+b^2)/2}\,da\,db,$$

bound is $\theta \in [0, \pi/4]$ and $b = \rho \sin\theta \geq x$. Therefore

$$\begin{aligned}
\Phi(-x)^2 =& \frac{1}{\pi}\int_0^{\frac{\pi}{4}}\int_{\frac{x}{\sin\theta}}^{\infty} e^{-\rho^2/2}\rho\, d\rho\, d\theta \\
=& \frac{1}{\pi}\int_0^{\frac{\pi}{4}} \exp\left\{-\frac{x^2}{2\sin^2\theta}\right\} d\theta = \frac{1}{\pi}\int_0^{\frac{\pi}{4}} \exp\left\{-\frac{x^2}{1-\cos(2\theta)}\right\} d\theta \\
=& \frac{1}{2\pi}\int_0^{\frac{\pi}{2}} \exp\left\{-\frac{x^2}{1-\cos\theta}\right\} d\theta.
\end{aligned}$$

Therefore

$$\left[\frac{\Phi(-x)}{\varphi(-x)}\right]^2 = \int_0^{\frac{\pi}{2}} \exp\left\{x^2 - \frac{x^2}{1-\cos\theta}\right\} d\theta = \int_0^{\frac{\pi}{2}} \exp\left\{-\frac{x^2\cos\theta}{1-\cos\theta}\right\} d\theta.$$

Similarly

$$\left[\frac{\Phi(t-x)}{\varphi(t-x)}\right]^2 = \int_0^{\frac{\pi}{2}} \exp\left\{-\frac{(x-t)^2\cos\theta}{1-\cos\theta}\right\} d\theta.$$

Therefore

$$\left[\frac{\Phi(t-x)}{\varphi(t-x)}\right]^2 - \left[\frac{\Phi(-x)}{\varphi(-x)}\right]^2 = \int_0^{\frac{\pi}{2}} \left[\exp\left\{-\frac{(x-t)^2\cos\theta}{1-\cos\theta}\right\} - \exp\left\{-\frac{x^2\cos\theta}{1-\cos\theta}\right\}\right] d\theta > 0, \tag{45}$$

---

both a and b are nonnegative, and so we can take θ to be $\arctan(y/x)$ which takes value in $[0, \pi/2]$. Our integration bound is $x \leq b \leq a$, which means

$$x \leq \rho\sin\theta \leq \rho\cos\theta.$$

The first inequality means

$$\rho \geq \frac{x}{\sin\theta},$$

and the second means(on $[0, \pi/2]$)

$$\tan\theta \leq 1 \Rightarrow 0 \leq \theta \leq \frac{\pi}{4}.$$

Using the fact that $\rho^2 = a^2 + b^2$, we obtain

$$\Phi(-x)^2 = \frac{1}{\pi}\int_x^{\infty}\int_b^{\infty} e^{-(a^2+b^2)/2}\, da\, db. = \frac{1}{\pi}\int_0^{\frac{\pi}{4}}\int_{\frac{x}{\sin\theta}}^{\infty} e^{-\rho^2/2}\rho\, d\rho\, d\theta.$$

The remainder is straightforward integration.

where in the last step we have used the fact that $x^2 > (x-t)^2 \geq 0$, so that

$$-\frac{(x-t)^2\cos\theta}{1-\cos\theta} > -\frac{x^2\cos\theta}{1-\cos\theta}, \quad \forall\theta \in (0, \frac{\pi}{2}).$$

which establishes the inequality in (45). We have shown that $h(x) > 0 \; \forall x \geq t$.

Next we prove (44) for $x \in (t/2, t)$. On this interval, we have

$$h'(x) = -\Phi(t-x)\varphi'(-x) + \Phi(-x)\varphi'(t-x) = -x\Phi(t-x)\varphi(-x) + (x-t)\Phi(-x)\varphi(t-x) < 0.$$

Since $h(t) > 0$ as shown above, $\forall x \in (t/2, t)$, $h(x) > 0$. We have proven (44) on the whole real axis.

Proof of Corollary 3

Using (39) and (41), we can also rewrite (19) as

$$\frac{d}{d\sigma}AR(\sigma) = \varphi\Big(\frac{X(\sigma)}{\sigma}\Big) + \Big(1 - \Phi\Big(\frac{X(\sigma)}{\sigma}\Big)\Big)\frac{\Big[\varphi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X(\sigma)}{\sigma}\Big)\Big]}{-\Big(1 - \Phi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big)\Big) + \Big(1 - \Phi\Big(\frac{X(\sigma)}{\sigma}\Big)\Big)}$$

$$= \varphi\Big(\frac{X(\sigma)}{\sigma}\Big) - \frac{1 - \Phi\Big(\frac{X(\sigma)}{\sigma}\Big)}{\Phi\Big(\frac{X(\sigma)}{\sigma}\Big) - \Phi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big)}\Big[\varphi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X(\sigma)}{\sigma}\Big)\Big]. \tag{46}$$

This expression is strictly positive if and only if

$$\frac{\varphi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X(\sigma)}{\sigma}\Big)}{\varphi\Big(\frac{X(\sigma)}{\sigma}\Big)} < \frac{\Phi\Big(\frac{X(\sigma)}{\sigma}\Big) - \Phi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big)}{1 - \Phi\Big(\frac{X(\sigma)}{\sigma}\Big)}.$$

Define

$$f(X) \equiv \Big(\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X}{\sigma}\Big)\Big)\Big(1 - \Phi\Big(\frac{X}{\sigma}\Big)\Big) \quad \text{and} \quad g(X) \equiv \Big(\Phi\Big(\frac{X}{\sigma}\Big) - \Phi\Big(\frac{X-\bar{e}}{\sigma}\Big)\Big)\varphi\Big(\frac{X}{\sigma}\Big). \tag{47}$$

We need to show that

$$\lim_{X\to\infty} g(X) - f(X) \leq 0. \tag{48}$$

First, if two continuous functions intersect at least twice, then the difference between their first derivatives at the intersection points alternate signs. Thus, if the difference between the first derivatives of two continuous functions at their intersection point(s) always has the same sign, then these two functions only intersect once, i.e., the intersection point is unique. We denote by X_0 any X such that $f(X) = g(X)$. That is,

$$\frac{\varphi\Big(\frac{X_0-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X_0}{\sigma}\Big)}{\varphi\Big(\frac{X_0}{\sigma}\Big)} = \frac{\Phi\Big(\frac{X_0}{\sigma}\Big) - \Phi\Big(\frac{X_0-\bar{e}}{\sigma}\Big)}{1-\Phi\Big(\frac{X_0}{\sigma}\Big)}. \tag{49}$$

Equivalently,

$$\Big(\varphi\Big(\frac{X_0-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X_0}{\sigma}\Big)\Big)\Big(1-\Phi\Big(\frac{X_0}{\sigma}\Big)\Big) = \varphi\Big(\frac{X_0}{\sigma}\Big)\Big(\Phi\Big(\frac{X_0}{\sigma}\Big) - \Phi\Big(\frac{X_0-\bar{e}}{\sigma}\Big)\Big). \tag{50}$$

We now compare the first derivatives of the LHS and RHS of (19), as derived in (53) and (54), respectively. At any given X, (53) is larger than (54) if and only if

$$\Big(-\frac{X-\bar{e}}{\sigma}\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big) + \frac{X}{\sigma}\varphi\Big(\frac{X}{\sigma}\Big)\Big)\Big(1-\Phi\Big(\frac{X}{\sigma}\Big)\Big) > -\frac{X}{\sigma}\varphi\Big(\frac{X}{\sigma}\Big)\Big(\Phi\Big(\frac{X}{\sigma}\Big) - \Phi\Big(\frac{X-\bar{e}}{\sigma}\Big)\Big). \tag{51}$$

At $X = X_0$, (53) is larger than (54) if and only if

$$\frac{\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X}{\sigma}\Big)}{\varphi\Big(\frac{X}{\sigma}\Big)} - \frac{\bar{e}}{X}\frac{\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big)}{\varphi\Big(\frac{X}{\sigma}\Big)} < \frac{\Phi\Big(\frac{X}{\sigma}\Big) - \Phi\Big(\frac{X-\bar{e}}{\sigma}\Big)}{1-\Phi\Big(\frac{X}{\sigma}\Big)} \tag{52}$$

Since $\frac{\bar{e}}{X}\frac{\varphi\left(\frac{X-\bar{e}}{\sigma}\right)}{\varphi\left(\frac{X}{\sigma}\right)} > 0$ for any $X > 0$, and because of (49), we know that (53) > (54) at any X. We have shown that the first derivative of the LHS of (50) is always larger than the first derivative of the RHS of (50) at a given X that satisfies (49), which implies that the function on the LHS of (50) and the one on the RHS intersect only once, so that X_0 is unique.

Second, the first derivative of $f(X)$ with respect to $\frac{X}{\sigma}$ is

$$\begin{aligned}&\Big(\varphi'\Big(\frac{X-\bar{e}}{\sigma}\Big) - \varphi'\Big(\frac{X}{\sigma}\Big)\Big)\Big(1-\Phi\Big(\frac{X}{\sigma}\Big)\Big) - \Big(\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X}{\sigma}\Big)\Big)\varphi\Big(\frac{X}{\sigma}\Big)\\ =&\Big(-\frac{X-\bar{e}}{\sigma}\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big) + \frac{X}{\sigma}\varphi\Big(\frac{X}{\sigma}\Big)\Big)\Big(1-\Phi\Big(\frac{X}{\sigma}\Big)\Big) - \Big(\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X}{\sigma}\Big)\Big)\varphi\Big(\frac{X}{\sigma}\Big).\end{aligned} \tag{53}$$

The first derivative of $g(X)$ with respect to $\frac{X}{\sigma}$ is

$$\varphi'\Big(\frac{X}{\sigma}\Big)\Big(\Phi\Big(\frac{X}{\sigma}\Big)-\Phi\Big(\frac{X-\bar{e}}{\sigma}\Big)\Big)+\varphi\Big(\frac{X}{\sigma}\Big)\Big(\varphi\Big(\frac{X}{\sigma}\Big)-\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big)\Big)$$

$$=-\frac{X}{\sigma}\varphi\Big(\frac{X}{\sigma}\Big)\Big(\Phi\Big(\frac{X}{\sigma}\Big)-\Phi\Big(\frac{X-\bar{e}}{\sigma}\Big)\Big)+\varphi\Big(\frac{X}{\sigma}\Big)\Big(\varphi\Big(\frac{X}{\sigma}\Big)-\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big)\Big). \tag{54}$$

Using (53) and (54), we can write

$$g'(X)-f'(X)=\big(-g(X)+f(X)\big)\frac{X}{\sigma}-\frac{\bar{e}}{\sigma}\varphi\Big(\frac{X-\bar{e}}{\sigma}\Big)\Big(1-\Phi\Big(\frac{X}{\sigma}\Big)\Big). \tag{55}$$

We first show by contradiction that $\lim_{X\to\infty} g(X)-f(X)<0$. Because of the asymptotic properties of the normal distribution's p.d.f. and c.d.f., (55) implies that

$$\lim_{X\to\infty} g'(X)-f'(X)=\big(-g(X)+f(X)\big)\frac{X}{\sigma}. \tag{56}$$

Suppose that $\lim_{X\to\infty} g(X)-f(X)<0$. Using (56), this implies that

$$\lim_{X\to\infty} g'(X)-f'(X)=\infty, \tag{57}$$

which is inconsistent with $\lim_{X\to\infty} g(X)-f(X)<0$.

In the remainder of the proof, we take as given $\lim_{X\to\infty} g(X)-f(X)\geq 0$. For $X>0$,

$$g(X)-f(X)=g(0)-f(0)+\int_0^X \big(g'(a)-f'(a)\big)da. \tag{58}$$

Using (55),

$$g(X)-f(X)=g(0)-f(0)+\int_0^X\Big(\big(-g(a)+f(a)\big)\frac{a}{\sigma}-\frac{\bar{e}}{\sigma}\varphi\Big(\frac{a-\bar{e}}{\sigma}\Big)\Big(1-\Phi\Big(\frac{a}{\sigma}\Big)\Big)\Big)da. \tag{59}$$

In addition, with $\lim_{X\to\infty} g(X)-f(X)\geq 0$ in the case under consideration, the fact that $\lim_{X\to-\infty} g(X)-f(X)>0$ (since f is positive if and only if $X>\widehat{X}$, while g is positive for any X), and the fact that the functions f and g intersect at most once (as proven above) imply that $g(X)-f(X)\geq 0$ for any X. This result and (55) imply that

$g'(X) - f'(X)$ is negative for any $X > 0$, so that, for any given $X > 0$,

$$\arg\max_{a\in[0,X]}\{-g(X) + f(X)\} = X. \tag{60}$$

Using this result and (59), we have:

$$g(X) - f(X) < g(0) - f(0) + \big(-g(X) + f(X)\big)\frac{X}{\sigma} - \int_0^X \frac{\bar{e}}{\sigma}\varphi\Big(\frac{a-\bar{e}}{\sigma}\Big)\Big(1 - \Phi\Big(\frac{a}{\sigma}\Big)\Big)da \tag{61}$$

Since the last term on the RHS of (61) is negative, we obtain

$$g(X) - f(X) < \frac{g(0) - f(0)}{1 + \frac{X}{\sigma}}. \tag{62}$$

Then, for finite $g(0) - f(0)$ and finite σ,

$$\lim_{X\to\infty}\frac{g(0) - f(0)}{1 + \frac{X}{\sigma}} = 0. \tag{63}$$

It follows from (62) that

$$\lim_{X\to\infty} g(X) - f(X) = 0. \tag{64}$$

Then (64) implies that

$$\lim_{X\to\infty}\frac{dAR(\sigma)}{d\sigma} = 0. \tag{65}$$

In addition, (39) implies that

$$\lim_{X\to\infty}\frac{\partial AR(\sigma)}{\partial\sigma} = 0. \tag{66}$$

Using (39) and (46) gives

$$\lim_{X\to\infty}\frac{\frac{dAR(\sigma)}{d\sigma}}{\frac{\partial AR(\sigma)}{\partial\sigma}} = \lim_{X\to\infty}\left(1 - \frac{1-\Phi\Big(\frac{X(\sigma)}{\sigma}\Big)}{\Phi\Big(\frac{X(\sigma)}{\sigma}\Big) - \Phi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big)}\frac{\varphi\Big(\frac{X(\sigma)-\bar{e}}{\sigma}\Big) - \varphi\Big(\frac{X(\sigma)}{\sigma}\Big)}{\varphi\Big(\frac{X(\sigma)}{\sigma}\Big)}\right) = 0, \tag{67}$$

where the last equality arises from (64).

As in the proof of Lemma 1, the incentive constraint for a given X is

$$\int_X^\infty (s - X)\psi(s|\bar{e})ds - \int_X^\infty (s - X)\psi(s|0)ds = C.$$

As $X \to \infty$, this equation is satisfied if and only if $C \to 0$. Applying this result to (65) and (67) yields (23).

Proof of Corollary 4

Proposition 3 states that the principal optimally sets $\delta = 1$ for $X > \bar{X}$. In addition, Corollary 1 states that the optimal X is strictly decreasing in C. Define $\bar{C}$ as the cost of effort associated with $\bar{X}$. It follows that the principal optimally sets $\delta = 1$ for $C < \bar{C}$.