The Value of Informativeness for Contracting*

Pierre Chaigneau
HEC Montreal

Alex Edmans
LBS, Wharton, NBER, CEPR, and ECGI

Daniel Gottlieb
Wharton

June 27, 2014

Abstract

The informativeness principle demonstrates qualitative benefits to increasing signal precision. However, it is difficult to quantify these benefits – and compare them against the costs of monitoring – since we typically cannot solve for the optimal contract and analyze how it changes with precision. We consider 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 is a fall in the value of the option and thus the agent's expected wage, benefiting the principal. The indirect effect is a change in the agent's effort incentives. If the original option is deeply out-of-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 agency problem becomes weaker, the gains from increased precision fall towards zero. These results potentially justify pay-for-luck and the absence of relative performance evaluation. Separately, increases in informativeness lead to at-the-money options being optimal.

Keywords: Contract theory, principal-agent model, executive compensation, limited liability, pay-for-luck, relative performance evaluation, options, informativeness principle.

JEL Classification: D86, J33

*pierre.chaigneau@hec.ca, aedmans@london.edu, dgott@wharton.upenn.edu. We thank seminar participants at Wharton and at the Risk Theory Society 2014 meeting 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] It argues that the principal should maximize the precision of the performance measure used to evaluate the agent. Greater precision allows the principal to use a cheaper contract to implement at least the same effort level. However, in practice, increasing informativeness is costly. Investing in a superior monitoring technology involves direct costs. Engaging in relative performance evaluation ("RPE") involves the indirect costs of forgoing the benefits of pay-for-luck documented by prior research (e.g. Oyer (2004), Raith (2008), Axelson and Baliga (2009), and Gopalan, Milbourn, and Song (2010)). Potentially for this reason, numerous violations of RPE have been found in practice. Aggarwal and Samwick (1999) and Murphy (1999) show that CEO pay is determined by absolute, rather than relative performance. Jenter and Kanaan (2013) similarly find an absence of RPE in CEO firing decisions. Whether these violations are an efficient response to the indirect costs of RPE is unclear. Bertrand and Mullainathan (2001) find that CEOs are paid for positive exogenous shocks, particularly in 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 with shareholders, and instead results from rent extraction by CEOs.

The informativeness principle argues that there are qualitative benefits to increasing signal precision. However, for a principal to decide whether to invest in greater precision, she must quantify these benefits – in particular, relate them to the underlying parameters of the contracting problem – so that she can compare them against the cost of precision. Similarly, to evaluate whether the general absence of RPE is efficient, it is useful to understand under which settings the benefits of informativeness are smallest, and compare them against the cases in which RPE is particularly absent in reality. Such quantification is difficult under the general framework in which the informativeness principle was derived. As is well-known (e.g. Grossman and Hart (1983)), in a

[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. 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.

general setting it is not possible to solve for the optimal contract. We cannot analyze precisely how the contract changes in response to increased informativeness, and thus quantify the cost savings from contract redesign.

This paper addresses this open question. We consider the standard setting of risk neutrality and limited liability, which allows us to take an optimal contracting approach. These restrictions lead to optimal contracts that we observe in practice – as shown by Innes (1990), the agent has 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 expected wage. Thus, the strike price is the minimum possible to satisfy the agent's incentive constraint.

We start by considering general signal distributions. We show that an increase in informativeness (in the sense of second-order stochastic dominance) has two effects, each of which has a clear economic interpretation. First, ignoring the incentive constraint, a fall in volatility directly reduces the value of the option and thus the agent's expected wage. Second, the increase in precision changes the agent's incentives. The heart of the paper analyzes this incentive effect and shows how its direction depends on the model's underlying parameters.

The agent's effort incentives 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 and improves the signal distribution ("option-when-working"). Changes in signal precision affect the values of these options differentially. If the option satisfies *increasing differences*, i.e. effort and precision are *complements* (an increase in precision augments the sensitivity of the option's value to effort), then a rise in informativeness augments effort incentives. The principal can thus increase the strike price of the option, i.e. reduce its delta, without violating the agent's incentive constraint. This strike price increase further reduces the expected wage, and reinforces the first, direct effect. In contrast, if the option satisfies *decreasing differences*, i.e. effort and precision are *substitutes*, an increase in informativeness weakens effort incentives, offsetting the first effect. In the limit, it fully offsets it, rendering the total benefit of precision zero. The key result from the general model is that we derive a simple condition, which holds for any signal distribution and is easy to verify, that governs whether the option satisfies increasing or decreasing differences and thus whether a rise in informativeness raises or lowers effort incentives.

We then focus the model on distributions with a location and scale parameter, such as the Normal and logistic distributions. The existence of a scale parameter – the volatility of the distribution – allows us to fully characterize changes in precision by changes in this volatility parameter. In turn, we can use the concept of the option's vega (sensitivity to volatility) to analyze how changes in precision affect the value of the "option-when-shirking" compared to the "option-when-working", and thus the agent's incentives. The existence of a location parameter allows us to quantify the vega of each option by comparing this location parameter to the option's strike price; the strike price in turn depends on the model's underlying parameters (e.g. the severity of the agency problem). In sum, for distributions with a location and scale parameter, the condition that determines whether informativeness strengthens or weakens effort incentives simplifies to a threshold condition for the strike price.

First, consider the case in which the cost of effort is high, i.e. the moral hazard problem is severe. The option will have a low strike price, so that its delta is sufficiently large to 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. Thus, the vega of the 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 positive shock. The agent is not worried about the fact that, if he shirks and receives a negative shock, the signal will be very low, because his payoff can be no lower than zero due to limited liability.

We next consider a low cost of effort, which leads to the strike price being high. 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 its value falls with informativeness faster than the option-when-shirking, lowering incentives. Intuitively, when the strike price is high, the agent will only receive a positive wage if he exerts effort *and* 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, the effect of informativeness on effort incentives depends on whether the initial strike price of the option is above or below a threshold. Thus, when incentives are strong (weak) to begin with, e.g. for CEOs (rank-and-file workers), an increase in informativeness further increases (reduces) incentives, amplifying (lowering) the gains from informativeness. For the Normal distribution, the gains from informativeness are monotonically increasing in the cost of effort, and thus the severity of the agency problem. In contrast, an analysis focusing only on the direct effect of informativeness on the value of the option and ignoring the incentive constraint, would suggest that the gains from informativeness are highest when the option is at-the-money – i.e. a moderate initial strike price and a moderate agency problem.

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

For tractability, the analysis features a binary effort level. In the continuous-effort analog, in order to implement a given effort level, the contract must ensure that the agent will not deviate to a slightly lower or a slightly higher effort level (i.e. the incentive constraint will be "local"). This situation resembles a binary model in which the low effort level is very close to the (implemented) high effort level. In this case, the threshold for the initial strike price – that determines whether informativeness increases or decreases effort – is the mean value of the signal. If the initial strike price is above (below) this threshold, increases in informativeness lower (raise) the strike price towards the threshold. Thus, improvements in informativeness (e.g. increases in stock market efficiency) move the strike price closer to the mean value of the signal, and thus lead to at-the-money ("ATM") options being optimal. Bebchuk and Fried (2004) argue that the almost universal practice of granting ATM options is suboptimal and that out-of-the-money options are more effective because the agent only gets paid if performance is very high (see also Rappaport (1999)). Such an argument ignores the incentive effect: out-of-the-money options have lower deltas and thus provide fewer incentives. The analysis also suggests that accounting or taxation considerations that

[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.

favor ATM options need not induce suboptimal contracting.

A recent paper by Dittmann, Maug, and Spalt (2013) 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 risk aversion parameters into the calibration. However, 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. They acknowledge that the actual savings from indexation will be different if the principal uses an initially optimal contract and responds optimally to changes in incentives. In contrast, our primary goal is theoretical. We incorporate risk neutrality and limited liability, allowing us to take an optimal contracting approach. 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 workers' 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 (worker). At $t = 0$, the agent chooses his effort level $e \in \{0, \overline{e}\}$. Effort of $e = 0$ is of zero cost to the agent, and $e = \overline{e}$ costs him $C > 0$. We will sometimes refer to $e = \overline{e}$ as "high effort" or "working", and $e = 0$ as "low effort" or "shirking".

At $t = 1$, the agent's contribution to firm value ("output") q is realized. As in the literature on performance measurement (e.g. Baker (1992)), output is generally not contractible, since it is difficult to measure an employee's contribution independently of his colleagues'. Instead, contracts can depend on a performance measure ("signal")

$$s = q + \eta,$$

where η is a mean-zero random variable that is uncorrelated with effort: $\mathbb{E}[\eta|e] = 0$. For example, η may be a market or industry shock, the contribution of other workers, or measurement error. We assume that output q is not contractible and the contract depends on a separate signal s, so that we can change signal precision without affecting output volatility. However, the model allows for the case in which output is contractible, which corresponds to the degenerate distribution concentrated at $\eta = 0$ (i.e. signal is perfectly informative about output).

Conditional on effort e, the signal s is continuously distributed according to the probability density function ("PDF") $f_\theta(s|e)$ with full support on $[\underline{s}, \overline{s}]$, where the bounds $\underline{s}$ and $\overline{s}$ may or may not be finite. Let $F_\theta(s|e)$ denote the cumulative distribution function ("CDF") of s. A high signal is "good news" about effort in the sense of the strict monotone likelihood ratio property (MLRP). Formally, for all θ and for all signals s_1 and s_0 with $s_1 > s_0$,

$$\frac{f_\theta(s_1|\overline{e})}{f_\theta(s_1|0)} > \frac{f_\theta(s_0|\overline{e})}{f_\theta(s_0|0)}.$$

Strict MLRP implies that the distribution of performance is ordered according to strict first-order stochastic dominance ("FOSD"): $F_\theta(s|0) > F_\theta(s|\overline{e})$ for all s and all θ.

The real-valued parameter θ, which lies in an interval Θ, captures the informativeness or precision of the signal, and orders the distributions in terms of second-order stochastic dominance. Formally, the mean of the signal is independent of θ, and

$$\theta \geq \theta' \implies \int_{\underline{s}}^{t} F_{\theta}(s|e)\,ds \leq \int_{\underline{s}}^{t} F_{\theta'}(s|e)\,ds, \tag{1}$$

for all $t \in [\underline{s}, \overline{s}]$, where the bounds may or may not be finite. Thus, increases in θ generate more precise signal distributions in the sense of mean-preserving spreads.

Our analysis solves for the optimal contract for each given level of precision θ. This approach applies to settings in which the principal cannot influence the precision of the signal but it is affected by exogenous forces (such as technological change); the analysis derives empirical predictions on how these changes affect the form of the optimal contract and the expected wage. In addition, our approach also applies to settings in which the principal can choose the level of precision θ at a cost $\kappa(\theta)$, where κ is increasing and convex. Under the interpretation that η arises from measurement error, removing the shock corresponds to an improvement in the monitoring technology, in which case $\kappa(\theta)$ 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. Under the interpretation that η is a market or industry shock, increasing precision θ corresponds to relative performance evaluation (RPE), in which case the cost $\kappa(\theta)$ 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).

The discount rate is normalized to zero. Given a contract $W(\cdot)$ and a level of effort e, the agent's expected wage is

$$\mathbb{E}[W(s)|e] = \int_{\underline{s}}^{\overline{s}} W(s) f_{\theta}(s|e)\,ds.$$

The agent is risk-neutral and so maximizes his expected wage, 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 a contract $W(\cdot)$, an effort level e, and a precision θ that maximizes expected output $\mathbb{E}[q]$, less the expected wage $\mathbb{E}[W]$ and cost of precision $\kappa(\theta)$.

Following Innes (1990), we make two assumptions on the set of feasible contracts. First, the agent is protected by limited liability, so that $W(s) \geq 0$ for all s. Second, pay-performance sensitivity lies between 0 and 1:

$$W(s+\epsilon) \geq W(s) \text{ and } s+\epsilon - W(s+\epsilon) \geq s - W(s)$$

for all s, ϵ. These constraints must be satisfied if the agent can freely borrow to artificially increase output and the principal can freely destroy output. If the constraint on the left did not hold, the agent would artificially increase output, thereby increasing the signal and thus his payoff. If the constraint on the right did not hold, the principal would exercise her control rights to "burn" output, thereby reducing the signal and increasing her payoff. These constraints can be expressed as

$$1 \geq \frac{W(s+\epsilon) - W(s)}{\epsilon} \geq 0$$

for all ϵ. It thus follows that $W(\cdot)$ is Lipschitz continuous and, therefore, differentiable almost everywhere. Hence, with no loss of generality, we can assume that the contract $W(\cdot)$ is a cadlag function satisfying $0 \leq W'(s) \leq 1$ at all points of differentiability.[3]

In the first best, effort is verifiable. There is no incentive constraint and only a participation constraint. If the principal wishes to induce high effort, this constraint is given by:

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

To satisfy (2), the principal pays an expected wage $\mathbb{E}[W(s)|\bar{e}]$ that equals the agent's cost of effort C. Thus, if

$$\mathbb{E}[q|\bar{e}] - \mathbb{E}[q|0] > C, \tag{3}$$

high effort is optimal for the principal. We assume (3) throughout, else even under the first-best, the principal would not want to induce effort.

In the second best, the agent's effort is unverifiable and so the contract must satisfy

[3]A cadlag function is everywhere right-continuous and has left limits everywhere.

an incentive constraint. The agent will exert effort if and only if:

$$\mathbb{E}[W(s)|\overline{e}] - \mathbb{E}\left[W(s)|0\right] \geq C. \tag{4}$$

Following standard arguments, this incentive constraint will bind. In contrast, the participation constraint will be slack if the principal wishes to implement high effort.[4] We thus ignore it in the analysis that follows.

We define X_θ implicitly by

$$\int_{X_\theta}^{\overline{s}} (s - X_\theta)\left[f_\theta(s|\overline{e}) - f_\theta(s|0)\right] ds = C. \tag{5}$$

We will show in Lemma 1 that X_θ exists and is unique. The intuition behind (5) is that, if the agent is given a call option on s, X_θ is the strike price such that working increases the value of the agent's option by an amount equal to the cost of effort, so that the incentive constraint is satisfied with equality.

We make the following assumption to ensure that $e = \overline{e}$ is second-best optimal:

$$\mathbb{E}\left[q|\overline{e}\right] - \mathbb{E}\left[q|0\right] - \int_{X_\theta}^{\overline{s}} (s - X_\theta) f_\theta(s|\overline{e}) ds \geq 0. \tag{6}$$

The first term, $\mathbb{E}\left[q|\overline{e}\right] - \mathbb{E}\left[q|0\right]$, is the benefit to the principal of inducing $e = \overline{e}$. 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 a constant wage. Note that (6) implies

$$\mathbb{E}\left[q|\overline{e}\right] - \mathbb{E}\left[q|0\right] \geq C + \int_{X_\theta}^{\overline{s}} (s - X_\theta) f_\theta(s|0) ds,$$

which is stronger than (3), the condition that guarantees that high effort is optimal in the first-best. The additional term $\int_{X_\theta}^{\overline{s}} (s - X_\theta) f_\theta(s|0) ds$ arises because the agent will earn rents from shirking (as he may enjoy a very positive shock and generate a high signal); thus, he must be offered rents for working to satisfy his incentive constraint.

[4] If the agent shirks, his wage cannot fall below zero no matter how low the signal is, due to limited liability. Thus, the expected wage upon shirking, $\mathbb{E}\left[W(s)|0\right]$, is positive. To satisfy the incentive constraint (4), we must have $\mathbb{E}[W(s)|\overline{e}] - C \geq \mathbb{E}\left[W(s)|0\right] \geq 0$, and so the participation constraint (2) is automatically satisfied.

The principal's problem is to choose a contract $W(\cdot)$ and informativeness θ to minimize the sum of the expected wage and the cost of precision, subject to the agent's incentive and limited liability constraints, plus the constraints on the slope of the contract. She chooses a cadlag function $W(\cdot)$ and an informativeness parameter θ to minimize

$$\mathbb{E}\left[W\left(s\right)|\bar{e}\right]+\kappa\left(\theta\right) \tag{7}$$

subject to

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

$$0 \leq W\left(s\right)\ \forall\ s, \quad \text{and} \tag{9}$$

$$0 \leq W'(s) \leq 1 \text{ at all points of differentiability of } W. \tag{10}$$

Our contracting problem is the dual to the one in Innes (1990). In his model, the agent (entrepreneur) chooses a financing contract subject to his own incentive constraint and the participation constraint of the principal (investor). In our model, the principal (firm) chooses an employment contract subject to the incentive and participation constraints of the agent (worker). In both models, the optimal contract has the same form; the only difference is which party captures the rents. Since Innes 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; here, we will show that the agent receives a call option on the signal.

Another difference is that Innes features a continuous action set. His focus was to derive the form of the optimal contract and thus he wishes to do so in the most general setting. Our goal is different: given that the optimal contract is a call option, we study how changes in informativeness affect the agent's incentives and thus the strike price. We thus specialize to a binary effort level. With a continuous effort level, a change in informativeness θ 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 (2010, 2013)).[5] Edmans and Gabaix (2011) show

[5]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.

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 for a given level of informativeness θ. The analysis is similar to Innes (1990). Our main results will come in Section 3, which analyzes the gains from increasing informativeness θ.

Let $W_\theta(\cdot)$ denote the optimal contract that implements high effort for a given informativeness level θ. Lemma 1 establishes that $W_\theta(\cdot)$ is a call option on s, where the strike price X_θ is chosen to satisfy the incentive constraint (5) with equality:

Lemma 1 *(Optimal contract) For a given θ, there exists a unique optimal contract, characterized by $e = \overline{e}$, and*

$$W_\theta(s) = \max\{0, s - X_\theta\}, \tag{11}$$

where X_θ is determined by the unique solution of (5).

The setting is slightly different from Innes (1990), since the principal is contracting on a signal rather than output. We show that the Innes (1990) result of the optimality of a call option extends to this case, and the intuition is the same. The absolute value of the likelihood ratio is highest in the tails of the distribution of s, so the signal is most informative about effort in the tails. The left tail cannot be used for incentive purposes due to the limited liability constraint, and so incentives are concentrated in the right tail. This maximizes the likelihood that positive payments are received by a working agent. With an upper-bound on the slope, the optimal contract involves call options on s with the maximum feasible slope, i.e. $W'(s) = 1$.

Lemma 2 below shows that the strike price falls with the cost of effort.

Lemma 2 *Let X_θ be the strike price in the optimal contract for a given θ. Then, X_θ is strictly decreasing in the cost of effort C.*

Proof. The Appendix shows that the expression (5), which implicitly determines X_θ, can be rewritten as

$$\int_{X_\theta}^{\overline{s}} \left[F_\theta \left(s|0 \right) - F_\theta \left(s|\overline{e} \right) \right] ds = C. \tag{12}$$

Applying the implicit function theorem yields:

$$\frac{dX_\theta}{dC} = -\frac{1}{F\left(X_\theta|0 \right) - F\left(X_\theta|\overline{e} \right)} < 0. \tag{13}$$

■

By strict FOSD, the denominator in equation (13) is positive – effort improves the distribution of the signal. The higher the cost of effort C, the higher the agent's reward must be for improving the distribution of the signal, to encourage him to induce effort. Lowering the strike price raises the delta of the option (the sensitivity of the option to the value of s) and thus the agent's incentives.

3 The Value of Informativeness

This section calculates the gains from increasing informativeness. Section 3.1 considers general signal distributions and provides a condition under which increases in informativeness raise the agent's effort incentives, which holds for any distribution that satisfies MLRP (so that the optimal contract is a call option). Section 3.2 focuses on symmetric, unbounded distributions with a location and scale parameter and relates this condition – and thus the effect of informativeness on incentives – to the initial strike price and thus the severity of the agency problem. Section 3.3 graphically illustrates the benefits of informativeness for the Normal distribution. It also proves analytically that the benefits from informativeness are monotonically increasing in the cost of effort, and thus monotonically decreasing in the initial strike price.

3.1 General Distributions

The total effect of increasing informativeness on the expected wage can be decomposed as follows:

$$\frac{d}{d\theta}\mathbb{E}\left[W(s)|\overline{e}\right] = \underbrace{\frac{\partial}{\partial\theta}\mathbb{E}\left[W(s)|\overline{e}\right]}_{\textit{direct effect}} + \underbrace{\frac{\partial}{\partial X_\theta}\mathbb{E}\left[W(s)|\overline{e}\right]\frac{dX_\theta}{d\theta}}_{\textit{incentive effect}}. \tag{14}$$

The first component is the *direct effect*, $\frac{\partial}{\partial\theta}\mathbb{E}\left[W(s)|\bar{e}\right]$. Holding constant the strike price, an increase in signal precision changes the value of the option; we will later prove that this effect is negative. This reduction in the cost of compensation is the benefit of informativeness highlighted by Bebchuk and Fried (2004) in their argument that the lack of RPE is inefficient. In the Holmstrom (1979) setting of a risk-averse agent, an increase in informativeness reduces the risk borne by the agent and thus allows the principal to lower the expected wage without violating the agent's participation constraint. In our setting of risk neutrality and limited liability, an increase in precision reduces the value of the option.

The second component is the *incentive effect*, $\frac{\partial}{\partial X_\theta}\mathbb{E}\left[W(s)|\bar{e}\right]\frac{dX_\theta}{d\theta}$, which arises because the increase in precision causes the strike price to rise by $\frac{dX_\theta}{d\theta}$ to maintain incentive compatibility. $\frac{\partial}{\partial X_\theta}\mathbb{E}\left[W(s)|\bar{e}\right]$ is negative – any increase in the strike price reduces the value of the option and thus the cost of compensation – but the sign of $\frac{dX_\theta}{d\theta}$ is unclear. We thus seek to derive conditions under which an increase in precision raises or lowers the optimal strike price. The following definition will be useful:

Definition 1 *Let* $(\Theta, E) \subseteq \mathbb{R}^2$. *A function* $g(\theta, e) : \Theta \times E \to \mathbb{R}$ *satisfies* increasing differences *if, for all* $\theta_L < \theta_H$ *and* $e_L \leq e_H$,

$$g(\theta_H, e_H) - g(\theta_H, e_L) \geq g(\theta_L, e_H) - g(\theta_L, e_L). \tag{15}$$

It satisfies decreasing differences *if* $-g$ *satisfies increasing differences.* θ *and* e *are* complements *if* $g(\theta, e)$ *satisfies increasing differences,* substitutes *if* $g(\theta, e)$ *satisfies decreasing differences, and* neutral *if* $g(\theta, e)$ *satisfies both increasing and decreasing differences.*

The increasing differences condition (15) means that the incremental gain (i.e., increase in the value of g) from effort, $g(\theta, e_H) - g(\theta, e_L)$, is increasing in θ. That is, effort and informativeness are complements in terms of their effect on g. Conversely, decreasing differences means that the incremental gain from effort is decreasing in θ. Thus, effort and informativeness are substitutes. Indeed, increasing differences is the most common definition of complementarity, whereas decreasing differences is the most common definition of substitutability.[6] In our setting, if g is differentiable, increasing

[6]There is a very large literature using these concepts for understanding outcomes of games – e.g. Bulow, Geanakoplos, and Klemperer (1985) and Milgrom and Roberts (1990).

differences is equivalent to the single-crossing condition:

$$\frac{\partial g}{\partial \theta}(\theta, \bar{e}) - \frac{\partial g}{\partial \theta}(\theta, 0)\, ds \geq 0.$$

We are concerned with how changes in precision affect the incentive constraint (4). The agent's incentives stem from the fact that exerting effort increases the value of his option. If he works, his option is worth $\mathbb{E}\left[W(s)\,|\bar{e}\right]$; we refer to this as an "option-when-working." If he shirks, he receives an "option-when-shirking" worth $\mathbb{E}\left[W(s)\,|0\right]$. His effort incentives are given by the difference in the values of these options, i.e.

$$\mathbb{E}\left[W(s)\,|\bar{e}\right] - \mathbb{E}\left[W(s)\,|0\right]. \tag{16}$$

Since a change in precision θ affects the option-when-working and the option-when-shirking to different degrees, it affects the agent's effort incentives (16). When precision and effort are complements, i.e. $\mathbb{E}\left[W(s)|e\right]$ satisfies increasing differences, increases in precision augment the agent's effort incentives:

$$\frac{\partial}{\partial \theta}\left\{\mathbb{E}\left[W(s)\,|\bar{e}\right] - \mathbb{E}\left[W(s)\,|0\right]\right\} > 0, \tag{17}$$

We thus wish to understand the conditions under which $\mathbb{E}\left[W(s)|e\right]$ satisfies increasing differences. We do so by using integration by parts (see Appendix A) to rewrite the agent's expected wage as

$$\mathbb{E}\left[W(s)|e\right] = \mathbb{E}\left[s|e\right] - X_\theta + \int_{\underline{s}}^{X_\theta} F_\theta(s|e)\, ds. \tag{18}$$

The third term, $\int_{\underline{s}}^{X_\theta} F_\theta(s|e)\, ds$, the area under the CDF between $\underline{s}$ and X_θ. It can be interpreted economically as the value of a put option with a strike price of X:

$$\Pr(s < X_\theta|e)\, \mathbb{E}\left[(X_\theta - s)\,|s < X_\theta, e\right] = \int_{\underline{s}}^{X_\theta} -(s - X_\theta)\, f(s|e) ds = \int_{\underline{s}}^{X_\theta} F_\theta(s|e)\, ds,$$

where the last equality follows from integration by parts. Under this interpretation, equation (18) is the put-call parity equation. The agent's call option equals the expected value of the signal, minus the strike price, plus the value of a put option. Let $\pi_X(\theta, e) \equiv \int_{\underline{s}}^{X} F_\theta(s|e)\, ds$ denote the value of a put option with a strike price of X.

By second-order stochastic dominance (equation (1)), the value of the put option is decreasing in the precision of the signal ($\frac{\partial \pi_{X_\theta}}{\partial \theta}(\theta, e) \leq 0$).[7]

To study whether $\mathbb{E}[W(s)|e]$ satisfies increasing differences, we examine each of the three terms on the right-hand side ("RHS") of (18) in turn. While $\mathbb{E}[s|e]$ depends on e, it is independent of θ since changes in θ represent mean-preserving spreads. In addition, X_θ depends on θ but not e. Thus, θ and e are neutral in their effect on both of these terms, and non-neutral only in their effect on the third term $\int_{\underline{s}}^{X_\theta} F_\theta(s|e)\, ds$. This observation leads to the following Lemma:

Lemma 3 *The agent's expected pay $\mathbb{E}[W(s)|e]$ satisfies increasing differences if and only if the area under the CDF, $\int_{\underline{s}}^{X_\theta} F_\theta(s|e)\, ds$, satisfies increasing differences.*

The usefulness of Lemma 3 lies in the fact that, while the value of the call option contains several terms (see (18)), the area under the CDF $\int_{\underline{s}}^{X_\theta} F_\theta(s|e)\, ds$ is a single term, and so it is relatively easy to verify whether it satisfies increasing differences. While it may seem intuitive that the value of the call option satisfies increasing differences if and only if the value of the put option satisfies increasing differences, the value of Lemma 3 is that we can check whether expected pay satisfies increasing differences by studying a single term $\int_{\underline{s}}^{X_\theta} F_\theta(s|e)\, ds$ – not that this term can be interpreted as the value of a put option. The condition in Lemma 3 is simple to check and general: it holds for all signal distributions that satisfy MLRP.

We thus apply the definition of substitutes and complements in Definition 1 to the area under the CDF, $\int_{\underline{s}}^{X_\theta} F_\theta(s|e)\, ds$. This application allows us to determine the effect of informativeness on the strike price X_θ, which is summarized in Proposition 1.

Proposition 1 *The optimal strike price X_θ is increasing in informativeness θ ($\frac{dX_\theta}{d\theta} > 0$) if informativeness and effort are complements at the strike price X_θ, decreasing in informativeness ($\frac{dX_\theta}{d\theta} < 0$) if they are substitutes at X_θ, and constant ($\frac{dX_\theta}{d\theta} = 0$) if they are neutral at X_θ.*

When precision and effort are complements, exerting effort augments the value of the call option by a greater amount when precision is high. As a result, the agent's

[7] In the Black-Scholes model, we have a strict inequality. This is because the Black-Scholes model assumes a lognormal distribution for stock returns, so an increase in precision (which corresponds to a decrease in volatility) affects the whole distribution. However, in our setting with a general distribution, a change in θ may only affect the part of the distribution above the strike price X_θ, where the put option has zero payoff, and so its value does not change.

marginal benefit from effort, $\mathbb{E}[W(s)|\bar{e}] - \mathbb{E}[W(s)|0]$, is increasing in informativeness. This loosens the incentive constraint and allows the principal to increase the strike price (thus reducing the expected wage) while still inducing effort. Thus, in addition to the direct benefit of informativeness (it reduces the expected wage, holding constant the strike price X_θ), the principal further benefits from the incentive effect of greater informativeness (it allows the strike price X_θ to increase, further reducing the expected wage). Proposition 1 in turn leads to Lemma 4 below.

Lemma 4 *(Partial and total effects of informativeness on expected wage):*

$$\left|\frac{d}{d\theta}\mathbb{E}[W(s)|\bar{e}]\right| > \left|\frac{\partial}{\partial\theta}\mathbb{E}[W(s)|\bar{e}]\right| \qquad \text{if and only if } \frac{dX_\theta}{d\theta} > 0. \tag{19}$$

Proof. From equation (14), we have:

$$\begin{aligned}\frac{d}{d\theta}\mathbb{E}[W(s)|\bar{e}] &= \underbrace{\frac{\partial}{\partial\theta}\mathbb{E}[W(s)|\bar{e}]}_{\text{direct effect}} + \underbrace{\frac{\partial}{\partial X_\theta}\mathbb{E}[W(s)|\bar{e}]\frac{dX_\theta}{d\theta}}_{\text{incentive effect}} \\ &= \frac{\partial\pi_X(\theta,e)}{\partial\theta} - \left[1 - \frac{\partial\pi_X(\theta,e)}{\partial X_\theta}\right]\frac{dX_\theta}{d\theta} = \frac{\partial\pi_X(\theta,e)}{\partial\theta} - [1 - F_\theta(X_\theta|e)]\frac{dX_\theta}{d\theta}.\end{aligned}$$

$\frac{d}{d\theta}\mathbb{E}[W(s)|\bar{e}]$ and $\frac{\partial}{\partial\theta}\mathbb{E}[W(s)|\bar{e}]$ are both negative, and the former is more negative (i.e. its absolute value is higher) if and only if $\frac{dX_\theta}{d\theta} > 0$, i.e. effort and precision are complements. ■

The direct effect, $\frac{\partial}{\partial\theta}\mathbb{E}[W(s)|\bar{e}]$, is negative. An increase in precision decreases the value of the put option ($\frac{\partial\pi_{X_\theta}}{\partial\theta} \leq 0$) and thus the expected wage. Turning to the incentive effect, higher precision augments the strike price by $\frac{dX_\theta}{d\theta}$, which in turn requires the principal to pay an additional $\frac{dX_\theta}{d\theta}$ dollars whenever the price exceeds X_θ, which occurs with probability $1 - F_\theta(X_\theta|e)$.

The sign of $\frac{dX_\theta}{d\theta}$ in turn depends on whether informativeness and effort are substitutes or complements. When they are complements, then $\frac{dX_\theta}{d\theta} > 0$. The strike price increases, further reducing the expected wage and reinforcing the direct effect. When they are substitutes, then $\frac{dX_\theta}{d\theta} < 0$ and the effects go in opposite directions. While the value of the option decreases with precision, the agent requires a lower strike price to induce effort, which partially offsets the benefits to the principal.

Even when $\frac{dX_\theta}{d\theta} < 0$ and the incentive effect works in the opposite direction to the direct effect, it can never outweigh it. The total effect $\frac{d}{d\theta}\mathbb{E}\left[W(s)|\overline{e}\right]$ is always weakly negative, i.e. increasing precision weakly reduces the expected wage. This result arises from revealed preference. If reducing precision reduced the expected wage, the principal would have added in randomness to the contract, and so the initial contract would not have been optimal. Even though the incentive effect cannot outweigh the direct effect, it is still important to consider as it affects the optimal level of precision θ that the principal should choose, since increasing precision is costly.[8] Indeed, it is possible that the incentive effect exactly offsets the direct effect, and so that the total gains from informativeness equal exactly zero: see Appendix B for an example.

3.2 Distributions with a Location and Scale Parameter

Section 3.1 shows that, with general distributions, the effect of precision on incentives depends on whether effort and precision are complements or substitutes; it also gives a condition that determines which scenario we are in. We now add more structure to the signal distribution which allows us to relate whether we have complements or substitutes to the underlying parameters of the agency problem. We can thus determine from model primitives whether changes in precision increase or decrease the agent's incentives.

We consider the case in which the signal s has a symmetric distribution with unbounded support and location and scale parameters, i.e., their distribution and density functions can respectively be written as $F_\sigma(s|e) = G\left(\frac{s-e}{\sigma}\right)$ and $f_\sigma(s|e) = \frac{1}{\sigma}g\left(\frac{s-e}{\sigma}\right)$. All such distributions can be fully characterized by their mean e and standard deviation σ. Examples include the Normal, logistic, Cauchy, and Laplace distributions. Since the volatility of a signal is the inverse of its precision, we have $\sigma = \frac{1}{\sqrt{\theta}}$, where θ is the same informativeness / precision parameter we had earlier.

The effect of precision on incentives depends on how changes in precision affect the value of the option-when-working compared to the option-when-shirking. The introduction of a scale parameter σ is useful because we can fully parameterize changes in precision by changes in σ. We can thus examine how changes in volatility affect the values of these two options – and thus the agent's effort incentives – using the familiar

[8]Unfortunately, it is not possible to solve for the optimal level of precision θ^* in closed form. In terms of comparative statics, θ^* is decreasing in the marginal cost of informativeness $\kappa'(\theta)$, but the relationship with the parameters of the agency problem (C) is ambiguous.

concept of the option "vega": the sensitivity of its value to volatility. Specifically, comparing the vegas of the two options will allow us to assess the effect on the agent's overall incentives. The introduction of a location parameter is useful as it allows us to relate the vega of each option to the distance between the option's initial strike price and the location parameter. The initial strike price in turn depends on the underlying parameters of the agency problem. Thus, we will be able to relate the incentive effect to the underlying parameters of the setting. The assumptions of symmetry and unbounded support are useful for tractability. Appendix D extends the results to asymmetric distributions and to distributions with a bounded support (such as the uniform distribution).

With a location and a scale parameter, the signal can now be written as:

$$s = q + \sigma\varepsilon, \tag{20}$$

where ε has a mean of 0 and volatility of 1, and σ parameterizes the volatility of the signal. Since the location of the distribution can be altered without changing its scale, we can set $\mathbb{E}[q] = e$ without loss of generality.

All of the results in Section 3.1 continue to hold, except we will now denote the strike price by X_σ rather than X_θ. Proposition 2 below derives a precise condition under which Lemma 3 and Proposition 1 hold, i.e. the put option satisfies increasing differences, and so the optimal strike price is increasing in informativeness.

Proposition 2 *(Effect of volatility on strike price.)* $\frac{dX_\sigma}{d\sigma} > 0$ *if and only if* $X_\sigma > \hat{X} \equiv \frac{\overline{e}}{2}$*. This in turn holds if and only if*

$$\int_{\hat{X}}^{\overline{s}} (s - \hat{X}) \left(f_\theta(s|\overline{e}) - f_\theta(s|0) \right) > C. \tag{21}$$

Proposition 2 states that the incentive effect is positive, i.e. an increase in precision augments incentives and thus the strike price, if and only if X_σ is below a threshold $\hat{X} \equiv \frac{\overline{e}}{2}$.

To understand the intuition, equation (17) now becomes

$$\frac{\partial}{\partial\sigma} \left\{ \mathbb{E}\left[W(s) \,|\overline{e} \right] - \mathbb{E}\left[W(s) \,|0 \right] \right\} < 0. \tag{22}$$

Since precision can be parameterized by volatility, we are interested in how changes in volatility increase the agent's incentives to exert effort – i.e. whether effort and volatility are complements or substitutes – and how this depends on the initial strike price of the option. 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

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

The left-hand side ("LHS") of inequality (22) represents the effect of changes in σ on incentives. This is equal to the vega of the option worth $Y(\overline{e}, X)$ minus the vega of the option worth $Y(0, X)$. The vega of an option is always positive, highest for an ATM option (see Claim 1 in Appendix C), and declines when the option moves either in-the-money or out-of-the-money. Thus, the vega of the option worth $Y(\overline{e}, X)$ is highest at $X = \overline{e}$, and so if the option has a strike price of $\widehat{X} = \frac{\overline{e}}{2}$, then it is in-the-money by $\frac{\overline{e}}{2}$. The vega of the option worth $Y(0, X)$ is highest at $X = 0$, and so if the option has a strike price of $\widehat{X} = \frac{\overline{e}}{2}$, then it is out-of-the-money by $\frac{\overline{e}}{2}$. Overall, at a strike price of $\widehat{X} = \frac{\overline{e}}{2}$, both options are equally away-from-the-money and have the same vega (see Claim 2 in Appendix C). Thus, increases in σ reduce the values of the option-when-working and option-when-shirking equally. The incentives to exert effort, $Y\left(\overline{e}, \frac{\overline{e}}{2}\right) - Y\left(0, \frac{\overline{e}}{2}\right)$, 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 ATM than $Y(\overline{e}, X)$, 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, a working agent gains from any upside but also loses from any downside, 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 kink X. Thus, when volatility increases, a shirking agent gains from any upside but does not lose from any downside. 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_σ is needed to restore incentive compatibility, since such a fall increases the value of $Y(\bar{e}, X)$ 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 shirk, 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 $Y(\bar{e}, X)$ is closer to being ATM than $Y(0, X)$, it has a higher vega. Since $Y(\bar{e}, X)$ 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, $Y(\bar{e}, X)$ 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_σ is possible without violating the incentive constraint. In simple language, when volatility rises and $X > \widehat{X}$, the agent thinks: "If volatility were 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."

In the language of the general model of Section 3.1, when C is small (the agency problem is weak), then $X > \widehat{X}$ and so we have decreasing differences. Since informativeness and incentives are substitutes, an increase in informativeness reduces incentives and so requires the strike price to fall in response. When C is large (the agency problem is strong), then $X < \widehat{X}$ and so we have increasing differences. An increase in informativeness can be accompanied by an increase in the strike price.

Proposition 2 implies that, in all cases, improvements in informativeness draw the strike price X towards $\widehat{X} = \frac{\bar{e}}{2}$. In the current discrete model, there are two effort levels, $\bar{e}$ and 0. In a continuous-effort analog, where the principal wishes to implement effort of $\bar{e}$, the contract must ensure induce the agent to exert effort of $\bar{e}$ rather than $\bar{e} + \varepsilon$ or $\bar{e} - \varepsilon$, i.e. the incentive constraint must be "local". In our discrete model, a local incentive constraint resembles the case in which the (implemented) high effort level ($\bar{e}$) is very close to the low effort level (0). If $\bar{e} \simeq 0$, then $\widehat{X} \simeq 0$. Moreover, since the contract induces the agent to exert effort of $\bar{e}$, the mean value of the signal is $\bar{e}$ and so an ATM option will have a strike price of $\bar{e} \simeq 0$. Thus, if the initial strike price is higher (lower) than $\widehat{X} \simeq 0$, improvements in informativeness (e.g. increases in stock market efficiency) will lower (raise) the optimal strike price towards 0, i.e. bring the option closer to ATM. (Indeed, Appendix F sketches a continuous effort model which shows that increases in informativeness bring the option closer to ATM.) Bebchuk and Fried (2004) argue that the almost universal practice of granting ATM options is inefficient and that out-of-the-money options would be cheaper for the firm. Similarly, Rappaport (1999) advocates out-of-the-money options as they reward the agent only for exceptional performance. However, such views ignore 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 the strongest incentives, but that the tax code discourages such options. One interpretation is that the tax code leads to firms choosing ATM 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 ATM.[9]

In addition, Proposition 2 suggests that exogenous falls in σ will have different effects on the incentives of agents depending on the moneyness of their options. In particular, it will reduce (increase) the incentives of agents with out-of-the-money (in-the-money) options. Thus, firms may wish to reduce the strike prices of out-of-the-money options to restore incentives. Option repricing is documented empirically by Brenner, Sundaram, and Yermack (2000), although they do not study if it is prompted by falls in volatility. 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.

Finally, note that the above analysis takes an optimal contracting approach, so the slope of the contract is the maximum possible without violating the constraint (10). We thus have $W'(s) = 1$: the agent is the residual "claimant" of any increase in the signal (as long as $s \geq X_\theta$). Thus, the principal changes X_θ to ensure that the incentive constraint binds. An alternative approach is to restrict the contract to comprising ATM options, e.g. for accounting or tax reasons, and instead meet the incentive constraint by

[9]Hall and Murphy (2000) restrict the contract to consist of options, rather than taking an optimal contracting approach, and calibrate the optimal strike price depending on the CEO's risk aversion, the proportion of his wealth in stock, and the proportion of his wealth in options. They show that, in most cases, the range of optimal strike prices includes the current stock price, i.e. corresponds to an ATM option. Dittmann and Yu (2011) feature a risk-taking as well as effort decision, and restrict the contract to consisting of fixed salary, stock, and options. They show that in-the-money options are typically optimal.

varying the slope of the contract. Indeed, Murphy (1999) and Bebchuk and Fried (2004) document that ATM options are almost universally granted. Appendix E demonstrates an analogous result for this case. With ATM options, we have $X = \overline{e} \geq \widehat{X} = \frac{\overline{e}}{2}$ and so effort and precision are substitutes. An increase in informativeness requires the number of options granted to increase, to maintain incentive compatibility. An increase in the number of options granted augments the expected wage, just like a decrease in the strike price. As a result, the total effect of informativeness on expected pay is less than the direct effect. Thus, the results of the core model, where $X > \widehat{X}$, extend to the case of ATM options.

3.3 Normal Distribution

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 direct effect. We need to assume a specific distribution to enable us to calculate these derivatives, and so we consider the case of the Normal distribution. Let φ and Φ denote the PDF and CDF of the standard Normal distribution, respectively. Then, the total effect $\frac{d\mathbb{E}[W(s)|\overline{e}]}{d\sigma}$ and direct effect $\frac{\partial\mathbb{E}[W(s)|\overline{e}]}{\partial\sigma}$ are given analytically by:

$$\frac{d\mathbb{E}\left[W\left(s\right)|\overline{e}\right]}{d\sigma} = \varphi\left(\frac{X_\sigma - \overline{e}}{\sigma}\right) - \left[1 - \Phi\left(\frac{X_\sigma - \overline{e}}{\sigma}\right)\right]\frac{\varphi\left(\frac{X_\sigma - \overline{e}}{\sigma}\right) - \varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right) - \Phi\left(\frac{X_\sigma - \overline{e}}{\sigma}\right)} \tag{23}$$

$$\frac{\partial\mathbb{E}\left[W\left(s\right)|\overline{e}\right]}{\partial\sigma} = \varphi\left(\frac{X_\theta - \overline{e}}{\sigma}\right). \tag{24}$$

These expressions are derived in Appendix A. In Figure 1, we illustrate these effects numerically for a range of values of X. Note that the exogenous parameter that is changing is C, but since X is strictly decreasing in C (equation (13)), there is a one-to-one correspondence between C and X. Thus, we draw the graph with X rather than C on the x-axis, as is standard for graphs of option values.

To understand Figure 1, recall from (14) that the total effect is given by $\frac{d\mathbb{E}[W(s)|\overline{e}]}{d\sigma} = \frac{\partial\mathbb{E}[W(s)|\overline{e}]}{\partial\sigma} + \frac{\partial\mathbb{E}[W(s)|\overline{e}]}{\partial X_\theta}\frac{dX_\theta}{d\sigma}$. The direct effect, $\frac{\partial\mathbb{E}[W(s)|\overline{e}]}{\partial\sigma}$, tends to zero as the strike price approaches either $-\infty$ or ∞. The vega of an option is greatest when the option is ATM, i.e. $X = 1$. An ATM 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.

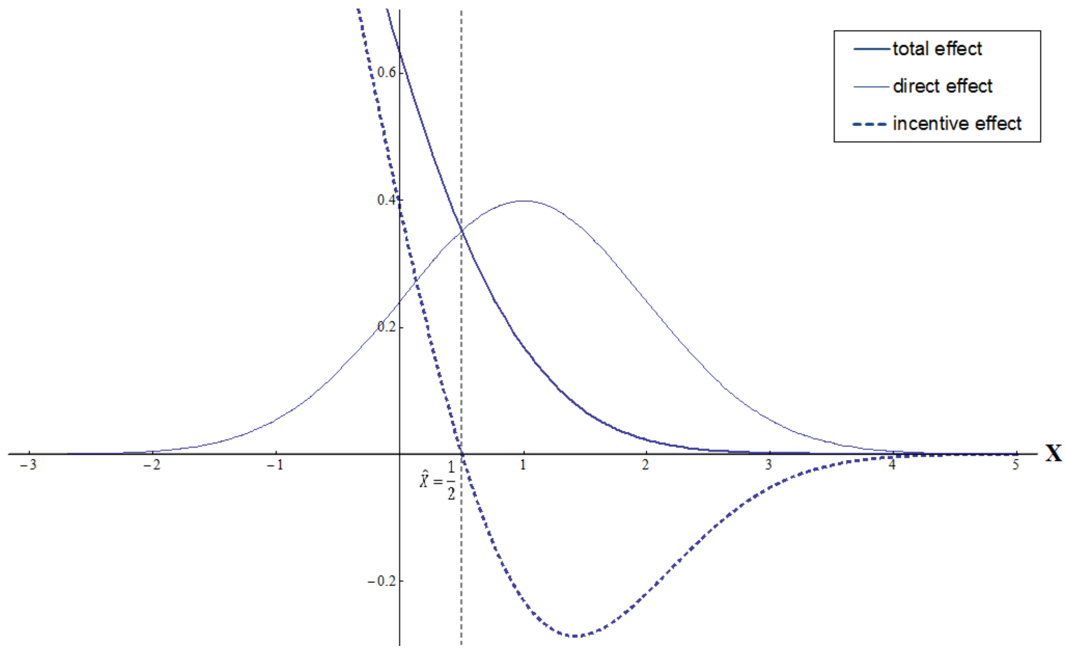

Figure 1: Total and partial derivative of expected pay with respect to σ for a range of values of X, for $\bar{e} = 1$ and $\sigma = 1$.

The incentive effect, $\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial X_\sigma}\frac{dX_\sigma}{d\sigma}$, consists of two components. The first is the change in strike price required to maintain incentive compatibility, $\frac{dX_\sigma}{d\sigma}$. From Proposition 2, $\frac{dX_\sigma}{d\sigma} > 0$ if and only if $X > \widehat{X} = \frac{1}{2}$. Indeed, for the Normal distribution, not only does $\frac{dX_\sigma}{d\sigma}$ turn from negative to positive as X crosses above $\widehat{X}$, but it is also monotonically increasing in X, i.e. monotonically decreasing in the cost of effort. This result is stated in Lemma 5 below:

Lemma 5 *(Normal distribution, change in strike price): The benefits of informativeness are decreasing in the cost of effort, i.e.*

$$\frac{d^2 X_\sigma}{d\sigma dC} < 0. \tag{25}$$

The second is the change in the value of the option when the strike price increases, $\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial X_\sigma}$. This change is always negative, and so the sign of the incentive effect is the opposite of the sign of $\frac{dX_\sigma}{d\sigma} > 0$: indeed, in Figure 1, the incentive effect is positive (negative) for $X < (>)\,\widehat{X}$. In addition, the magnitude of $\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial X_\sigma}$ is increasing in the moneyness of the option. Overall, as X falls below $\widehat{X}$, $\frac{dX_\sigma}{d\sigma}$ becomes increasingly negative (see Lemma 5), and the option becomes increasingly in the money so $\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial X_\sigma}$ also becomes increasingly negative (it falls towards -1). Thus, the overall incentive effect $\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial X_\sigma} \frac{dX_\sigma}{d\sigma}$ becomes monotonically more positive as X falls below $\widehat{X}$. However, as X rises above $\widehat{X}$, the two components of the incentive effect move 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, $\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial X_\sigma}$ falls towards zero: when the option is deeply out-of-the-money, its value is small to begin with and thus little affected by changes in the strike price. Overall, the impact of X on the incentive effect is non-monotonic. As X initially rises above $\widehat{X}$, the incentive effect becomes increasingly negative as the option has significant value, and this value is affected by the change in the strike price required to maintain incentives ($\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial X_\sigma}$ is large). However, as X continues to rise, the option's value falls and so is little affected by the strike price ($\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial X_\sigma}$ is small). Thus, the incentive effect becomes less negative and approaches zero.

The total effect $\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma}$ combines these direct and incentive effects. While the direct effect is initially increasing in X, this is outweighed by the fact that the incentive effect is initially decreasing in X. Thus, in Figure 1, the total gains from increased informativeness are monotonically decreasing in X. Indeed, focusing on the Normal distribution allows us to prove this result analytically: C is the exogenous parameter that drives X, and Lemma 6 shows that the gains from informativeness are monotonically increasing in C.

Lemma 6 *(Normal distribution, effect of cost of effort on gains from informativeness)*

$$\frac{d}{dC}\left\{\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma}\right\} > 0.$$

An analysis focusing purely on the direct effect would suggest that the gains from informativeness are greatest when the initial option is ATM, which in turn corresponds to a moderate strike price and a moderate cost of effort. In contrast, considering the total effect (which incorporates the incentive effect) shows that the gains from informativeness are monotonically increasing in the severity of the agency problem. Thus, workers with high-powered incentives (such as CEOs) should be evaluated more precisely than those with low-powered incentives (such as rank-and-file workers).

Lemma 7 shows that, as the cost of effort goes to zero, the gains from informativeness also approach zero, as does the total gain relative to the direct effect.

Lemma 7 *(Normal distribution, limiting cases). We have the following limiting cases:*

$$\frac{d\mathbb{E}\left[W\left(s\right)|\overline{e}\right]}{d\sigma} \longrightarrow_{C\to 0} 0 \qquad \textit{and} \qquad \frac{\frac{d\mathbb{E}[W(s)|\overline{e}])}{d\sigma}}{\frac{\partial\mathbb{E}[W(s)|\overline{e}]}{\partial\sigma}} \longrightarrow_{C\to 0} 0. \tag{26}$$

As the moral hazard problem becomes weaker, the total effect of informativeness becomes miniscule relative to the direct effect. Thus, ignoring the incentive effect and considering only the direct effect would substantially overestimate the gains from informativeness. Indeed, in Figure 1, the direct effect significantly 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 $\overline{e} = 1$), the gains from a marginal change in σ are 10.8 times larger with the direct effect than with the total effect. Thus, even for non-extreme parameter values, gains from improved informativeness can be much lower if the incentive effect is taken into account. This ratio becomes much greater for higher X, because the total benefits of informativeness fall towards zero.

4 Conclusion

This paper studies the principal's benefits from increasing the informativeness of the signal used to evaluate the agent. The direct effect is that higher signal precision reduces the value of the agent's option and thus expected pay. The core focus of the paper is on the indirect effect – how changes in precision affect the agent's incentives. In particular, by taking an optimal contracting approach, we can be specific on how the contract changes in response to increases in informativeness. With general signal distributions, we show that, if effort and informativeness are substitutes, increases in precision weaken the agent's incentives. Thus, the principal must reduce the strike price of the option to preserve effort incentives, increasing the cost of compensation and offsetting the direct effect. Indeed, we derive a simple condition that governs whether effort and informativeness are substitutes or complements that holds for all distributions that satisfy the monotone likelihood ratio property.

Focusing on signal distributions with a location and scale parameter allows us to derive precise conditions under which substitutes arise. If the initial strike price is high, i.e. incentives are low-powered to begin with, 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, the benefits of informativeness are lower due to the negative effect on incentives. The principal therefore optimally invests less in improving informativeness, potentially rationalizing the scarcity of relative performance evaluation for some agents in reality.

In contrast, if the initial strike price is low, i.e. incentives are high-powered to begin with, an increase in informativeness augments the agent's effort incentives. This 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. When the signal is Normally distributed, the benefits from informativeness are monotonically increasing in the cost of effort, and thus the severity of the agency problem. Finally, if incentive constraints are local, i.e. the implemented effort level is close to other feasible effort levels, then increases in informativeness cause the strike price to optimally move towards the mean signal value, leading to at-the-money options being optimal.

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

[4] Baker, George (1992): "Incentive contracts and performance measurement." *Journal of Political Economy* 100, 598–614.

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

[8] Bulow, Jeremy I., John D. Geanakoplos, and Paul D. Klemperer (1985): "Multimarket oligopoly: Strategic substitutes and complements." *Journal of Political Economy* 93, 488–511.

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

[10] 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.

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

[12] 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.

[13] Dittmann, Ingolf, Ernst Maug, and Oliver G. Spalt (2013): "Indexing executive compensation contracts." *Review of Financial Studies* 26, 3182–3224.

[14] Dittmann, Ingolf and Ko-Chia Yu (2011): "How important are risk-taking incentives in executive compensation?" Working Paper, Erasmus University Rotterdam.

[15] Edmans, Alex and Xavier Gabaix (2009): "Is CEO pay really inefficient? A survey of new optimal contracting theories." *European Financial Management* 15, 486–496.

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

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

[18] 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.

[19] Gormley, Todd A., David A. Matsa, and Todd T. Milbourn (2013): "CEO compensation and corporate risk-taking: Evidence from a natural experiment." *Journal of Accounting and Economics* 56, 79–101.

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

[21] Hall, Brian J. and Kevin J. Murphy (2000): "Optimal exercise prices for executive stock options." *American Economic Review* 90, 209–214.

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

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

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

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

[26] Milgrom, Paul and John Roberts (1990): "Rationalizability, learning, and equilibrium in games with strategic complementarities." *Econometrica* 58, 1255–1277.

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

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

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

[30] Pratt, John and Mark Machina (1997): "Increasing risk: Some direct constructions." *Journal of Risk and Uncertainty* 14, 103–127.

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

[32] Rappaport, Alfred (1999): "New thinking on how to link executive pay with performance." *Harvard Business Review* 72, 91–101

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

[34] Zang, Israel (1980): "A smoothing-out technique for min-max optimization." *Mathematical Programming* 19, 61–77.

A Proofs

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:

$$\begin{aligned}
L = & \int_{-\infty}^{\infty} W(s)\, f_\theta(s|\bar{e}) ds - \mu \Big[\int_{-\infty}^{\infty} W(s)\, f_\theta(s|\bar{e}) ds - C - \int_{-\infty}^{\infty} W(s)\, f_\theta(s|0) ds \Big] \\
& - \int_{-\infty}^{\infty} \lambda(s) W(s)\, ds - \int_{-\infty}^{\infty} \eta(s)(1 - W'(s)) ds \\
= & \int_{-\infty}^{\infty} \Big((1-\mu) W(s) + \mu C + \mu \frac{f_\theta(s|0)}{f_\theta(s|\bar{e})} W(s) \Big) f_\theta(s|\bar{e}) ds \\
& - \int_{-\infty}^{\infty} \Big(\lambda(s) W(s) - \eta(s)(1 - W'(s)) \Big) ds. \qquad (27)
\end{aligned}$$

Since the optimal contract $W(s)$ is continuous and $\lim_{s \to -\infty} W(s) = 0$,[10] we have

$$W(s) = \int_{-\infty}^{s} W'(x) dx = \int_{-\infty}^{s-ds} W'(x) dx + W'(s) ds. \qquad (28)$$

It follows that

$$\frac{dW'(s)}{dW(s)} = \left(\frac{dW(s)}{dW'(s)} \right)^{-1} = \left(\frac{d\Big(\int_{-\infty}^{s-ds} W'(x) dx + W'(s) ds \Big)}{dW'(s)} \right)^{-1} = \frac{1}{ds}. \qquad (29)$$

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

[10] 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

$$\left((1-\mu)+\mu\frac{f_\theta(s|0)}{f_\theta(s|\bar{e})}\right)f_\theta(s|\bar{e})-\lambda(s)+\eta(s)\frac{1}{ds}=0. \tag{30}$$

We define

$$\phi(s)\equiv\left((1-\mu)+\mu\frac{f_\theta(s|0)}{f_\theta(s|\bar{e})}\right)f_\theta(s|\bar{e}) \tag{31}$$

If $\phi(s)>0$ for a given s, then (30) 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 (30) imposes that

$$-\lambda(s)+\eta(s)\frac{1}{ds}>0. \tag{32}$$

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 (31) 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

$$\phi(s)>0 \qquad \Leftrightarrow \qquad \frac{f_\theta(s|\bar{e})}{f_\theta(s|0)}<\frac{\mu}{\mu-1} \qquad \Leftrightarrow \qquad s<X, \tag{33}$$

$$\phi(s)<0 \qquad \Leftrightarrow \qquad \frac{f_\theta(s|\bar{e})}{f_\theta(s|0)}>\frac{\mu}{\mu-1} \qquad \Leftrightarrow \qquad s>X. \tag{34}$$

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 $W'(s)=1$ $\forall\ s>X$.

With this contract, the incentive constraint (4) now becomes (12). In the proof of Lemma 2 we show that (12) may be rewritten as (36). The LHS of (36) is a continuous and strictly increasing function of X (by FOSD). Evaluated at $X = \bar{s}$, it equals $0 < C$. Evaluated at $X = \underline{s}$, it equals

$$\int_{\underline{s}}^{\bar{s}} \left[F_\theta\left(s|0\right) - F_\theta\left(s|\bar{e}\right)\right] ds = \int_{\underline{s}}^{\bar{s}} s f_\theta\left(s|\bar{e}\right) ds - \int_{\underline{s}}^{\bar{s}} s f_\theta\left(s|0\right) ds, \tag{35}$$

using the same integration by parts as in the proof of Lemma 2. Since high effort is first-best efficient, the expression in (35) exceeds C. Thus, the intermediate value theorem and the monotonicity of the LHS ensure that a unique solution to (36) exists.

Proof of Lemma 2

We show that the incentive constraint (5) can also be rewritten as

$$\int_{X_\theta^*}^{\bar{s}} \left[F_\theta\left(s|0\right) - F_\theta\left(s|\bar{e}\right)\right] ds = C. \tag{36}$$

Opening the expressions inside the brackets in equation (4), we obtain:

$$\int_{X_\theta}^{\bar{s}} s f_\theta\left(s|1\right) ds - \left[1 - F_\theta\left(X_\theta|1\right)\right] X_\theta = \int_{X_\theta}^{\bar{s}} s f_\theta\left(s|0\right) ds - \left[1 - F_\theta\left(X_\theta|0\right)\right] X_\theta + C.$$

$$\therefore \int_{X_\theta}^{\bar{s}} s f_\theta\left(s|1\right) ds - \int_{X_\theta}^{\bar{s}} s f_\theta\left(s|0\right) ds = \left[F_\theta\left(X_\theta|0\right) - F_\theta\left(X_\theta|1\right)\right] X_\theta + C.$$

Applying integration by parts (for $e \in \{0,1\}$), gives

$$\int_{X_\theta}^{\bar{s}} s f_\theta\left(s|e\right) ds = \left[s F_\theta\left(s|e\right) - \int F_\theta\left(s|e\right) ds\right]_{X_\theta}^{\bar{s}} = \bar{s} - X_\theta F_\theta\left(X_\theta|e\right) - \int_{X_\theta}^{\bar{s}} F_\theta\left(s|e\right) ds.$$

Plugging back in the previous expression, we obtain

$$\left[\bar{s} - X_\theta F_\theta\left(X_\theta|1\right) - \int_{X_\theta}^{\bar{s}} F_\theta\left(s|1\right) ds\right] - \left[\bar{s} - X_\theta F_\theta\left(X_\theta|0\right) - \int_{X_\theta}^{\bar{s}} F_\theta\left(s|0\right) ds\right]$$
$$= \left[F_\theta\left(X_\theta|0\right) - F_\theta\left(X_\theta|1\right)\right] X_\theta + C.$$

Canceling terms gives equation (36).

Proof of Equation (18)

We start with the case of bounded support. The agent's expected pay in case of effort e is

$$\int_X^{\bar{s}} (s - X) f_\theta (s|e) \, ds = \int_X^{\bar{s}} s f_\theta (s|e) \, ds - X \left[1 - F_\theta (X|e)\right] \tag{37}$$

which yields

$$\begin{aligned}
\int_X^{\bar{s}} (s - X) f_\theta (s|e) \, ds &= \int_{\underline{s}}^{\bar{s}} s f_\theta (s|e) \, ds - \int_{\underline{s}}^{X} s f_\theta (s|e) \, ds - X \left[1 - F_\theta (X|e)\right] \\
&= \mathbb{E} [s|e] - \int_{\underline{s}}^{X} s f_\theta (s|e) \, ds - X \left[1 - F_\theta (X|e)\right].
\end{aligned}$$

A similar integration by parts as before yields

$$\int_{\underline{s}}^{X} s f_\theta (s|e) \, ds = X F_\theta (X|e) - \int_{\underline{s}}^{X} F_\theta (s|e) \, ds.$$

Substituting back in the previous expression, we can rewrite the agent's expected pay as

$$\mathbb{E} [s|e] - X + \int_{\underline{s}}^{X} F_\theta (s|e) \, ds.$$

Moving to the case of unbounded support, equation (37) now becomes:

$$\int_X^{+\infty} (s - X) f_\theta (s|e) \, ds = \int_X^{+\infty} s f_\theta (s|e) \, ds - X \left[1 - F_\theta (X|e)\right],$$

which yields:

$$\int_{\underline{s}}^{+\infty} s f_\theta (s|e) \, ds - X \left[1 - F_\theta (X|e)\right] - \int_{\underline{s}}^{X} s f_\theta (s|e) \, ds.$$

Applying integration by parts, we have:

$$\int_{\underline{s}}^{X} s f_\theta (s|e) \, ds = \left[s F_\theta (s|e) - \int F_\theta (s|e) \, ds \right]_{\underline{s}}^{X} = X F_\theta (X|e) - \int_{\underline{s}}^{X} F_\theta (s|e) \, ds.$$

Substituting this into the previous equation yields:

$$\int_X^{+\infty} (s - X) f_\theta (s|e) \, ds = \int_{\underline{s}}^{+\infty} s f_\theta (s|e) \, ds - X + \int_{\underline{s}}^{X} F_\theta (s|e) \, ds = \mathbb{E} [s|e] - X + \int_{\underline{s}}^{X} F_\theta (s|e) \, ds.$$

Proof of Lemma 3

The agent's expected pay in the case of effort e is given by

$$\mathbb{E}\left[W(s)|e\right] = \int_{X_\theta}^{\bar{s}} (s - X_\theta)\, f_\theta\,(s|e)\, ds. \tag{38}$$

Integration by parts yields:

$$\int_{X_\theta}^{\bar{s}} s f_\theta\,(s|e)\, ds = \bar{s} - X F_\theta\,(s|e) - \int_{X_\theta}^{\bar{s}} F_\theta\,(s|e)\, ds$$

and so (38) can be rewritten:

$$\begin{aligned}\mathbb{E}\left[W(s)|e\right] &= \bar{s} - X F_\theta\,(s|e) - \int_{X_\theta}^{\bar{s}} F_\theta\,(s|e)\, ds - X\left[1 - F_\theta\,(s|e)\right] \\ &= \bar{s} - X - \int_{X_\theta}^{\bar{s}} F_\theta\,(s|e)\, ds. \end{aligned} \tag{39}$$

Since $\bar{s}$ and X are not functions of either θ or e, it follows that the agent's expected pay satisfies increasing differences if and only if $\int_{X_\theta}^{\bar{s}} F_\theta\,(s|e)\, ds$ satisfies decreasing differences (and vice-versa). The following Lemma will be useful for this and future proofs:

Lemma 8 *For any θ, $\int_{\underline{s}}^{\bar{s}} F_\theta\,(s|e)\, ds = \bar{s} - \mathbb{E}\,[s|e]$, which is not a function of θ*

Proof. Applying integration by parts to $\int_{\underline{s}}^{\bar{s}} s f_\theta\,(q|e)\, ds$, we obtain:

$$\int_{\underline{s}}^{\bar{s}} s f_\theta\,(q|e)\, ds = \left[s F_\theta\,(q|e) - \int F_\theta\,(s|e)\, ds\right]_{\underline{s}}^{\bar{s}} = \bar{s} - \int_{\underline{s}}^{\bar{s}} F_\theta\,(s|e)\, ds.$$

Since θ parameterizes mean-preserving spreads, the expression on the LHS, $\mathbb{E}[s|e]$ is not a function of θ. ■

From Lemma 8, $\int_{X_\theta}^{\bar{s}} F_\theta\,(s|e)\, ds$ satisfies decreasing differences if and only if $\int_{\underline{s}}^{X_\theta} F_\theta\,(s|e)\, ds$ satisfies increasing differences (since their sum is independent of θ by second-order stochastic dominance). Thus, the agent's expected pay satisfies increasing differences if and only if $\int_{\underline{s}}^{X_\theta} F_\theta\,(s|e)\, ds$ satisfies increasing differences.

Proof of Proposition 1

Applying the implicit function theorem to equation (12) gives:

$$\frac{dX_\theta}{d\theta} = \frac{\int_X^{\bar{s}} \frac{\partial}{\partial\theta}\left[F_\theta\left(q|0\right) - F_\theta\left(q|\bar{e}\right)\right] dq}{F_\theta\left(X_\theta|0\right) - F_\theta\left(X_\theta|\bar{e}\right)}.$$

By FOSD, the denominator is positive. Lemma 8 yields:

$$\frac{\partial}{\partial\theta}\int_X^{\bar{s}} F_\theta(q|e)dq = -\frac{\partial}{\partial\theta}\int_{\underline{s}}^{X} F_\theta(q|e)dq$$

for all X, θ, e. Plugging back:

$$\frac{dX_\theta}{d\theta} = \frac{\int_{\underline{s}}^{X} \frac{\partial}{\partial\theta}\left[F_\theta\left(q|\bar{e}\right) - F_\theta\left(q|0\right)\right] dq}{F_\theta\left(X_\theta|0\right) - F_\theta\left(X_\theta|\bar{e}\right)}.$$

It is straightforward to show that if $\int_{\underline{s}}^{X} F_\theta\left(q|e\right) dq$ satisfies increasing (decreasing) differences and is differentiable with respect to θ, then

$$\frac{\partial}{\partial\theta}\int_{\underline{s}}^{X} \left[F_\theta\left(q|e_H\right) - F_\theta\left(q|0\right)\right] dq \geq (\leq) 0 \tag{40}$$

for all $e_H > 0$. The denominator is positive by FOSD.

Proof of Proposition 2

From Lemma 3, the agent's expected pay satisfies increasing differences if and only if

$$\pi_X(\theta, e) = \int_{\underline{s}}^{X_\theta} F_\theta(s|e)ds$$

satisfies increasing differences. In turn, this arises if and only if

$$\frac{\partial}{\partial\theta}\int_{\underline{s}}^{X_\theta} \left[F_\theta(s|\bar{e}) - F_\theta(s|0)\right] ds \geq 0, \tag{41}$$

i.e. the single-crossing condition holds. Using $F_\sigma(s|e) = G\left(\frac{s-e}{\sigma}\right)$ and $\sigma = \frac{1}{\sqrt{\theta}}$, inequality (41) yields:

$$\frac{\partial}{\partial\theta}\int_{-\infty}^{X_\sigma} \left[G\left(\frac{s-\bar{e}}{\sigma}\right) - G\left(\frac{s-0}{\sigma}\right)\right] ds \geq 0$$

with $\sigma = \frac{1}{\sqrt{\theta}}$, where the LHS is the difference between the vegas (sensitivities to σ) of the option-when-working and option-when-shirking. This becomes:

$$\frac{\partial}{\partial\theta}\int_{-\infty}^{X_\sigma}\left[G\left(\sqrt{\theta}(s-\overline{e})\right) - G\left(\sqrt{\theta}s\right)\right] ds \geq 0$$

$$\frac{1}{2\theta}\int_{-\infty}^{X_\sigma}\left[\sqrt{\theta}(s-\overline{e})g\left(\sqrt{\theta}(s-\overline{e})\right) - \sqrt{\theta}sg\left(\sqrt{\theta}s\right)\right] ds \geq 0$$

Given that $\theta > 0$, we use the change of variable $y_H \equiv \sqrt{\theta}(s-\overline{e})$ and $y_L \equiv \sqrt{\theta}s$ to rewrite this condition as

$$\int_{-\infty}^{\sqrt{\theta}(X_\sigma-\overline{e})} y_H g\left(y_H\right) dy_H - \int_{-\infty}^{\sqrt{\theta}X_\sigma} y_L g\left(y_L\right) dy_L \geq 0. \tag{42}$$

As $\overline{e} > 0$, this is equivalent to

$$-\int_{\sqrt{\theta}(X_\sigma-\overline{e})}^{\sqrt{\theta}X_\sigma} yg\left(y\right) dy \geq 0. \tag{43}$$

For a distribution symmetric about the mean, this inequality holds if and only if

$$\sqrt{\theta}(X_\sigma - \overline{e}) + \sqrt{\theta}X_\sigma \leq 0, \tag{44}$$

that is, if and only if $X_\sigma \leq \frac{\overline{e}}{2}$. The first part of the proof then follows directly from Proposition 1.

To prove equation (21), as shown in the proof of Lemma 1, we have $\int_0^{\overline{s}} s\left(f_\theta(s|\overline{e}) - f_\theta(s|0)\right) > C$ and $\lim_{X\to\overline{s}}\int_X^{\overline{s}}(s-X)\left(f_\theta(s|\overline{e}) - f_\theta(s|0)\right) < C$. The LHS of the incentive constraint in (12) is strictly decreasing in X. Given that the equilibrium X satisfies (12) as an equality, we have equation (21).

Proof of Equations (23) and (24)

First, with volatility σ instead of precision θ, the decomposition in (14) can be rewritten as

$$\frac{d}{d\sigma}\mathbb{E}\left[W(s)|\overline{e}\right] = \underbrace{\frac{\partial}{\partial\sigma}\mathbb{E}\left[W(s)|\overline{e}\right]}_{\textit{direct effect}} + \underbrace{\frac{\partial}{\partial X_\sigma}\mathbb{E}\left[W(s)|\overline{e}\right]\frac{dX_\sigma}{d\sigma}}_{\textit{incentive effect}} \tag{45}$$

Second,

$$\frac{\partial \mathbb{E}[W(s)|e]}{\partial \sigma} = \frac{\partial}{\partial \sigma}\int_{X_\sigma}^{\infty}(s-X_\sigma)\frac{1}{\sigma}\varphi\left(\frac{s-e}{\sigma}\right)ds = \frac{\partial}{\partial \sigma}\int_{X_\sigma-e}^{\infty}\frac{s+e-X_\sigma}{\sigma}\varphi\left(\frac{s}{\sigma}\right)ds$$

$$= \frac{\partial}{\partial \sigma}\int_{X_\sigma-e}^{\infty}\frac{s}{\sigma}\varphi\left(\frac{s}{\sigma}\right)ds - (X_\sigma - e)\frac{\partial}{\partial \sigma}\int_{X_\sigma-e}^{\infty}\frac{1}{\sigma}\varphi\left(\frac{s}{\sigma}\right)ds$$

$$= \frac{\partial}{\partial \sigma}\left\{\left[-\sigma\varphi\left(\frac{s}{\sigma}\right)\right]_{X_\sigma-e}^{\infty}\right\} - (X_\sigma - e)\frac{\partial}{\partial \sigma}\left\{1-\Phi\left(\frac{X_\sigma-e}{\sigma}\right)\right\}$$

$$= \varphi\left(\frac{X_\sigma-e}{\sigma}\right) - \sigma\frac{X_\sigma-e}{\sigma^2}\varphi'\left(\frac{X_\sigma-e}{\sigma}\right) + (X_\sigma-e)\left(-\frac{X_\sigma-e}{\sigma^2}\right)\varphi\left(\frac{X_\sigma-e}{\sigma}\right)$$

$$= \varphi\left(\frac{X_\sigma-e}{\sigma}\right) - \frac{X_\sigma-e}{\sigma}\varphi'\left(\frac{X_\sigma-e}{\sigma}\right) + \frac{X_\sigma-e}{\sigma}\varphi'\left(\frac{X_\sigma-e}{\sigma}\right) = \varphi\left(\frac{X_\sigma-e}{\sigma}\right) \quad (46)$$

where the fourth and sixth equalities use the property that $\varphi'(x) = -x\varphi(x)$, and the fifth equality uses $\varphi(x) \to_{x\to\infty} 0$. This establishes (24). In addition, it follows that

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

Third,

$$\frac{\partial \mathbb{E}[W\left(s\right)|e]}{\partial X_\theta} = \frac{\partial}{\partial X_\theta}\int_{X_\theta}^{\infty}(s-X_\theta)\frac{1}{\sigma}\varphi\left(\frac{s-e}{\sigma}\right)ds$$

$$= \int_{X_\theta}^{\infty}-\frac{1}{\sigma}\varphi\left(\frac{s-e}{\sigma}\right)ds = -\left(1-\Phi\left(\frac{X_\theta-e}{\sigma}\right)\right). \quad (47)$$

It follows that

$$\frac{\partial}{\partial X_\sigma}\left\{\mathbb{E}\left[W\left(s\right)|\bar{e}\right] - \mathbb{E}\left[W\left(s\right)|0\right]\right\} = -\left(1-\Phi\left(\frac{X_\theta-\bar{e}}{\sigma}\right)\right) + \left(1-\Phi\left(\frac{X_\theta}{\sigma}\right)\right)$$

$$= \Phi\left(\frac{X_\theta-\bar{e}}{\sigma}\right) - \Phi\left(\frac{X_\theta}{\sigma}\right).$$

which is strictly negative because of MLRP, which implies FOSD.

Using the results above, we can rewrite (45) as

$$\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma} = \varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right) + \left[1-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right]\frac{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\Phi\left(\frac{X_\sigma}{\sigma}\right)} \quad (48)$$

This establishes (23).

Proof of Corollary 7

We have $\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma} \geq 0$ for any X_σ: indeed, if $\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma} < 0$ for a given σ, and considering that increasing σ is costless, the given level of σ would not be an equilibrium for any cost of increasing informativeness (including a zero cost).

Using (48), $\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma} \geq 0$ may be rewritten as

$$\frac{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)} \leq \frac{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)}{1-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)}. \tag{49}$$

Define

$$y_N(x_\sigma) \equiv \left(\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)\left(1-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right) \tag{50}$$

$$y_D(x_\sigma) \equiv \left(\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right)\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right). \tag{51}$$

where $x_\sigma \equiv \frac{X_\sigma}{\sigma}$. Differentiating $y_N(X_\sigma)$ and $y_D(x_\sigma)$ with respect to x_σ gives

$$\begin{aligned} y_N'(x_\sigma) &= \left(\varphi'\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi'\left(\frac{X_\sigma}{\sigma}\right)\right)\left(1-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right) \\ &\quad -\left(\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right) \\ &= \left(-\frac{X_\sigma-\bar{e}}{\sigma}\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)+\frac{X_\sigma}{\sigma}\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)\left(1-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right) \\ &\quad -\left(\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right). \end{aligned} \tag{52}$$

$$\begin{aligned} y_D'(x_\sigma) &= \varphi'\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\left(\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right)+\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\left(\varphi\left(\frac{X_\sigma}{\sigma}\right)-\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right) \\ &= -\frac{X_\sigma-\bar{e}}{\sigma}\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\left(\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right)+\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\left(\varphi\left(\frac{X_\sigma}{\sigma}\right)-\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right). \end{aligned} \tag{53}$$

At any given X_σ, (52) is larger than (53) if and only if

$$\left(-\frac{X_\sigma - \bar{e}}{\sigma}\varphi\left(\frac{X_\sigma - \bar{e}}{\sigma}\right) + \frac{X_\sigma}{\sigma}\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)\left(1 - \Phi\left(\frac{X_\sigma - \bar{e}}{\sigma}\right)\right)$$
$$> -\frac{X_\sigma - \bar{e}}{\sigma}\varphi\left(\frac{X_\sigma - \bar{e}}{\sigma}\right)\left(\Phi\left(\frac{X_\sigma}{\sigma}\right) - \Phi\left(\frac{X_\sigma - \bar{e}}{\sigma}\right)\right). \tag{54}$$

Or, for $X_\sigma > \bar{e}$,

$$\frac{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right) - \varphi\left(\frac{X_\sigma}{\sigma}\right)}{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)} - \frac{\bar{e}}{X_\sigma - \bar{e}}\frac{\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)} < \frac{\Phi\left(\frac{X_\sigma}{\sigma}\right) - \Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)}{1 - \Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)} \tag{55}$$

As $\frac{\bar{e}}{X_\sigma-\bar{e}}\frac{\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)} > 0$ for any $X_\sigma > \bar{e}$, and because of (49), we know that (52) > (53) for any $X_\sigma > \bar{e}$, i.e., $y'_N(X_\sigma) > y'_D(X_\sigma)$ for any $X_\sigma > \bar{e}$.

Using (48), we get

$$\frac{\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma}}{\frac{\partial\mathbb{E}[W(s)|\bar{e}]}{\partial\sigma}} = 1 - \frac{\left(1 - \Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right)\left[\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right) - \varphi\left(\frac{X_\sigma}{\sigma}\right)\right]}{\left(\Phi\left(\frac{X_\sigma}{\sigma}\right) - \Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right)\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)}. \tag{56}$$

Note that the numerator and the denominator in the fraction of the right-hand-side of (56) are $y_N(x_\sigma)$ and $y_D(x_\sigma)$, respectively. We have shown above that $y'_N(x_\sigma) > y'_D(x_\sigma)$ for any $x_\sigma > 0$. For any two given b and a such that $b > a \geq 0$, we therefore have

$$y_N(b) - y_N(a) > y_D(b) - y_D(a) \tag{57}$$

Rearranging (57) yields

$$\frac{y_D(b) - y_D(a)}{y_N(b)} < \frac{y_N(b) - y_N(a)}{y_N(b)} \tag{58}$$

for any $b > \frac{1}{2\sigma}$, given that $y_N(x_\sigma) > 0$ for $x_\sigma \geq \frac{1}{2\sigma}$. Setting $a = \frac{1}{2\sigma}$, (58) holds (as $b > a$), and we have $y_N(a) = 0$. In addition, we have $\frac{y_D(a)}{y_N(b)} \longrightarrow_{b\to\infty} 0$, so that

$$\frac{y_D(b) - y_D(a)}{y_N(b)} \longrightarrow_{c\to\infty} \frac{y_D(b)}{y_N(b)} \tag{59}$$

Using (58) with $y_N(a) = 0$, this implies that

$$\lim_{b\to\infty} \frac{y_N(b)}{y_D(b)} \geq 1 \tag{60}$$

In addition, $\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial \sigma} = \varphi\left(\frac{X_\sigma}{\sigma}\right) > 0$, and $\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma} \geq 0$ for any X_σ, as established above. It then follows from (56) that we must have $\frac{y_N(b)}{y_D(b)} \leq 1$. Combining with (60) then implies that

$$\lim_{b\to\infty} \frac{y_N(b)}{y_D(b)} = 1 \tag{61}$$

Given the definitions of $y_N(x_\sigma)$ and $y_D(x_\sigma)$ and (56), this establishes $\frac{\frac{d\mathbb{E}[W(s)|\bar{e}])}{d\sigma}}{\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial \sigma}} \longrightarrow_{X\to 0} 0$. The result that $\frac{d\mathbb{E}[W(s)|\bar{e}])}{d\sigma} \longrightarrow_{X\to 0} 0$ then follows from this result and from $\frac{\partial \mathbb{E}[W(s)|\bar{e}]}{\partial \sigma} = \varphi\left(\frac{X_\sigma - \bar{e}}{\sigma}\right) \longrightarrow_{X_\sigma\to\infty} 0$.

Given the optimal contract derived in Lemma 1, the incentive constraint in (4) for a given X_σ is

$$\int_{X_\sigma}^{\infty} (s - X_\sigma) f(s|\bar{e}) ds - \int_{X_\sigma}^{\infty} (s - X_\sigma) f(s|0) ds = C.$$

As $X_\sigma \to \infty$, this equation is satisfied if and only if $C \to 0$. Combining with the results above yields (26).

Proof of Lemma 5

As X_σ is strictly decreasing in C (see Lemma 2), inequality (25) holds if and only if $\frac{d\frac{dX_\sigma}{d\sigma}}{dX} > 0$. As established in the proof of Equations (23) and (24) above,

$$\frac{dX_\sigma}{d\sigma} = -\frac{\varphi\left(\frac{X_\sigma - \bar{e}}{\sigma}\right) - \varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma - \bar{e}}{\sigma}\right) - \Phi\left(\frac{X_\sigma}{\sigma}\right)}$$

To simplify notation, define

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

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

$$f(x,t) \equiv [\varphi(x) - \varphi(x-t)]^2 - [\Phi(x) - \Phi(x-t)][\varphi'(x) - \varphi'(x-t)] > 0, \quad \forall x, \tag{62}$$

where

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

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

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

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

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

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 (62) is proven for $x \in [\frac{t}{2}, x_0]$

We now prove (62) for $x > x_0$. In this interval (we omit the argument t in what follows):

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

To prove the latter, we first calculate

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

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 Lemma 6

Using the chain rule,

$$
\frac{d}{dC}\left\{\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma}\right\} = \frac{d}{dX_\sigma}\left\{\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma}\right\}\frac{dX_\sigma}{dC}
$$

Given that $\frac{dX_\sigma}{dC} < 0$ (cf. Lemma (2)), we have $\frac{d}{dC}\left\{\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma}\right\} > 0$ if and only if

$\frac{d}{dX_\sigma}\left\{\frac{d\mathbb{E}[W(s)|\overline{e}]}{d\sigma}\right\} < 0$.

Using (23) and $\varphi'(x) = -x\varphi(x)$ for the normal distribution, we have

$$\frac{d}{dX_\sigma}\left\{\frac{d\mathbb{E}[W(s)|\overline{e}]}{d\sigma}\right\} = \frac{d}{dX_\sigma}\left\{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right) - \left[1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right]\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}\right\}$$

$$\begin{aligned}
= \frac{1}{\sigma}\Bigg(&-\frac{X_\sigma-\overline{e}}{\sigma}\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right) + \left[\frac{X_\sigma-\overline{e}}{\sigma}\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right) - \frac{X_\sigma}{\sigma}\varphi\left(\frac{X_\sigma}{\sigma}\right)\right]\frac{1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)} \\
&+\left[\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right]\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)} \\
&-\frac{1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}{\left(\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right)^2}\left(\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)^2\Bigg)
\end{aligned} \tag{63}$$

Multiplying all terms by $\sigma\left(\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right) > 0$, the expression in (63) has the same sign as

$$f(X_\sigma,\overline{e}) \equiv \frac{X_\sigma-\overline{e}}{\sigma}\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma}{\sigma}\right)\right] - \frac{X_\sigma}{\sigma}\varphi\left(\frac{X_\sigma}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right]$$

$$+\left(\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)\left[\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\left[1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right]\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}\right]$$

$$=\varphi\left(\frac{X_\sigma}{\sigma}\right)\left[-\frac{X_\sigma}{\sigma}\left[1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right]+\left(\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)\frac{1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}\right]$$

$$+\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\left[\frac{X_\sigma-\overline{e}}{\sigma}\left[1-\Phi\left(\frac{X_\sigma}{\sigma}\right)\right]+\left(\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)\left(1-\frac{1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}\right)\right]$$

$$\begin{aligned}
&=\varphi\left(\frac{X_\sigma}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right]\left[-\frac{X_\sigma}{\sigma}+\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}\right] \\
&+\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma}{\sigma}\right)\right]\left[\frac{X_\sigma-\overline{e}}{\sigma}-\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}\right]
\end{aligned}$$

$$=\left[\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}-\frac{X_\sigma}{\sigma}\right]\left[\varphi\left(\frac{X_\sigma}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right]-\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma}{\sigma}\right)\right]\right]$$

$$-\frac{\overline{e}}{\sigma}\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma}{\sigma}\right)\right] \tag{64}$$

Given that the last term in (64) is always negative, the expression in (64) is negative if the first line in (64) is negative. We show this in two steps.

To start with, the hazard rate $\varphi(x)/(1-\Phi(x))$ of the normal distribution is increasing, which implies that

$$\frac{\varphi\left(\frac{X_\sigma}{\sigma}\right)}{1-\Phi\left(\frac{X_\sigma}{\sigma}\right)} > \frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}{1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}$$

Rearranging, we have

$$\varphi\left(\frac{X_\sigma}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right]-\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma}{\sigma}\right)\right]>0 \tag{65}$$

Define

$$g(X_\sigma,\overline{e})\equiv\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}$$

If $g(X_\sigma,\overline{e})<\frac{X_\sigma}{\sigma}$, then combining with (65) establishes that the expression in (64) is negative, which is the desired result. We now show that this inequality holds. To this end, we will show in turn that $g(X_\sigma,\overline{e})\longrightarrow_{\overline{e}\to 0}\frac{X_\sigma}{\sigma}$, and that $g(X_\sigma,\overline{e})$ is decreasing in $\overline{e}$.

First,

$$\begin{aligned}\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)&=-\varphi'\left(\frac{X_\sigma}{\sigma}\right)\frac{\overline{e}}{\sigma}+O(\overline{e}^2)\\ \Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)&=\frac{\overline{e}}{\sigma}\varphi\left(\frac{X_\sigma}{\sigma}\right)+O(\overline{e}^2)\end{aligned}$$

Using $\varphi'(x)=-x\varphi(x)$ for the normal distribution, we have

$$\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}\longrightarrow_{\overline{e}\to 0}\frac{\varphi\left(\frac{X_\sigma}{\sigma}\right)\frac{\overline{e}X_\sigma}{\sigma^2}}{\frac{\overline{e}}{\sigma}\varphi\left(\frac{X_\sigma}{\sigma}\right)}=\frac{X_\sigma}{\sigma}$$

Second,

$$\frac{d}{d\overline{e}}\left\{\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}\right\}=\frac{d}{d\overline{e}}\left\{\frac{\int_{(X_\sigma-\overline{e})/\sigma}^{X_\sigma/\sigma}s\exp\left\{-\frac{s^2}{2}\right\}ds}{\int_{(X_\sigma-\overline{e})/\sigma}^{X_\sigma/\sigma}\exp\left\{-\frac{s^2}{2}\right\}ds}\right\}$$

$$= \frac{1}{\sigma} \frac{\frac{X_\sigma - \overline{e}}{\sigma} \exp\left\{-\frac{(X_\sigma-\overline{e})^2}{2\sigma^2}\right\} \int_{(X_\sigma-\overline{e})/\sigma}^{X_\sigma/\sigma} \exp\left\{-\frac{s^2}{2}\right\} ds - \exp\left\{-\frac{(X_\sigma-\overline{e})^2}{2\sigma^2}\right\} \int_{(X_\sigma-\overline{e})/\sigma}^{X_\sigma/\sigma} s \exp\left\{-\frac{s^2}{2}\right\} ds}{\left(\int_{(X_\sigma-\overline{e})/\sigma}^{X_\sigma/\sigma} \exp\left\{-\frac{s^2}{2}\right\} ds\right)^2}$$

This expression has the same sign as

$$\frac{X_\sigma - \overline{e}}{\sigma} \int_{(X_\sigma-\overline{e})/\sigma}^{X_\sigma/\sigma} \exp\left\{-\frac{s^2}{2}\right\} ds - \int_{(X_\sigma-\overline{e})/\sigma}^{X_\sigma/\sigma} s \exp\left\{-\frac{s^2}{2}\right\} ds$$

$$= \int_{(X_\sigma-\overline{e})/\sigma}^{X_\sigma/\sigma} \left[\frac{X_\sigma - \overline{e}}{\sigma} - s\right] \exp\left\{-\frac{s^2}{2}\right\} ds < 0$$

This establishes that $g(X_\sigma, \overline{e})$ is decreasing in $\overline{e}$, which completes the proof.

B Informativeness Has Zero Value

This section provides an example where the value of informativeness is exactly zero. From (7) and (18), the principal's payoff is

$$\mathbb{E}[s|e] - X_\theta + \int_{\underline{s}}^{X_\theta} F_\theta(s|e) ds - \kappa(\theta),$$

where X_θ solves the incentive constraint (12):

$$\int_{X_\theta}^{\bar{s}} \left[F_\theta\left(s|0\right) - F_\theta\left(s|e\right)\right] ds = C. \tag{66}$$

Take $\underline{s} = 0$ and $\bar{s} = 2$. Suppose that, under low effort, s is uniformly distributed in $[0, 1]$ for any level of informativeness θ:

$$F_\theta(s|0) = s \times \mathbf{1}\left(0 \leq s \leq 1\right).$$

This assumption is for concreteness only; the example can be generalized to distributions that, conditional on low effort, are not functions of θ: $F_\theta(s|0) = \zeta(s)$.

Assume two possible informativeness levels: θ_L and θ_H. Under high informativeness, s is uniformly distributed in $[0, 2]$:

$$f_H(x|1) = \frac{1}{2}, \quad F_H(x|1) = \frac{x}{2}.$$

Under low informativeness, s has the following density function:

$$f_L(x|1) = \begin{cases} \frac{1}{4} \text{ if } x \leq .25 \text{ or } .75 \leq x < 1 \\ \frac{3}{4} \text{ if } .25 < x < .75 \\ \frac{1}{2} \text{ if } 1 < x \leq 2 \end{cases}.$$

Notice that θ_H is a mean-preserving spread of θ_L. Integrating, we obtain the CDF

$$F_L(x|1) = \begin{cases} \frac{x}{4} \text{ if } x \leq \frac{1}{4} \\ \frac{1}{16} + \frac{3}{4}\left(x - \frac{1}{4}\right) \text{ if } \frac{1}{4} < x < \frac{3}{4} \\ \frac{7}{16} + \frac{1}{4}\left(x - \frac{3}{4}\right) \text{ if } \frac{3}{4} \leq x < 1 \\ \frac{x}{2} \text{ if } x \geq 1 \end{cases}.$$

Suppose the parameters are such that $X_\theta \geq 1$. For $x \geq 1$, the CDF are the same under both θ_H and θ_L so that, for $X_\theta \geq 1$, the incentive constraint (66) yields:

$$\int_{X_\theta}^{2} \left(1 - \frac{s}{2}\right) ds = C \therefore (2 - X_\theta) - \frac{1}{2}\left[2 - \left(\frac{X_\theta^2}{2}\right)\right] = C$$

$$\therefore \frac{X_\theta^2}{4} - X_\theta + (1 - C) = 0$$

The solution to this quadratic equation is

$$X_\theta = \frac{1 \pm \sqrt{C}}{2}.$$

The relevant root is the smallest one (otherwise, we can relax the incentive constraint (66) by reducing the strike price X_θ):

$$X_\theta = \frac{1 - \sqrt{C}}{2},$$

so the indirect effect is zero (the strike price is the same for both precision levels). The direct effect is also zero since $\int_0^x F_{\theta_H}(s|e)ds = \int_0^x F_{\theta_L}(s|e)ds$ for all $x \geq 1$.[11] Indeed, we

[11]This follows because, since $s|e$ has the same mean under both θ_H and θ_L, integration by parts gives:

$$\int_0^2 F_{\theta_H}(s|e)ds = \int_0^2 F_{\theta_L}(s|e)ds.$$

can calculate this expression explicitly:

$$\int_0^1 F_{\theta_H}(s|e)ds = \int_0^1 F_{\theta_L}(s|e)ds = \frac{1}{4}.$$

Thus, the expected wage is independent of informativeness.

C Option Vega for Distributions with Location and Scale Parameters

Under the Black-Scholes assumption that the stock price is lognormally distributed, the vega of a stock option is highest when the option is ATM. Claim 1 shows that this result extends to distributions with location and scale parameters.

Claim 1 *For distributions parameterized with e and σ such $F_\sigma(s|e) = G\left(\frac{s-e}{\sigma}\right)$, the option vega is highest for an option such that $X_\sigma = e$.*

Proof. By definition, for given e and $X_\sigma \leq \bar{s}$, the vega of the associated option is

$$\nu = \frac{\partial}{\partial \sigma}\mathbb{E}\left[W(s)|e\right] = \frac{\partial}{\partial \sigma}\left\{\bar{s} - X - \int_{X_\sigma}^{\bar{s}} F_\sigma(s|e)\, ds\right\} \tag{67}$$

where we used (39) to derive the second equality. Using $F_\sigma(s|e) = G\left(\frac{s-e}{\sigma}\right)$ for a distribution with location and scale parameters, we have

$$\nu = \frac{\partial}{\partial \sigma}\left\{-\int_{X_\sigma}^{\bar{s}} G\left(\frac{s-e}{\sigma}\right) ds\right\} = \frac{1}{\sigma}\int_{X_\sigma}^{\bar{s}} \frac{s-e}{\sigma} g\left(\frac{s-e}{\sigma}\right) ds \tag{68}$$

Using the change of variable $y = \frac{s-e}{\sigma}$ gives

$$\nu = \int_{\frac{X_\sigma - e}{\sigma}}^{\frac{\bar{s}-e}{\sigma}} yg(y)ds \tag{69}$$

Thus, $\int_1^2 F_{\theta_H}(s|e)ds = \int_1^2 F_{\theta_L}(s|e)ds$ implies that

$$\int_0^1 F_{\theta_H}(s|e)ds = \int_0^1 F_{\theta_L}(s|e)ds.$$

Given that $g(y) > 0$, this expression is maximized for $X_\sigma = e$, i.e., for an ATM option.[12]

■

Claim 2 shows that, for symmetric distributions with unbounded support, the vegas of the option-when-working and option-when-shirking are equal for $X_\sigma = \frac{\overline{e}}{2}$.

Claim 2 *For symmetric distributions with unbounded support parameterized by* e *and* σ *such* $F_\sigma(s|e) = G\left(\frac{s-e}{\sigma}\right)$, *the vegas of the option-when working and the option-when-shirking are equal for* $X_\sigma = \frac{\overline{e}}{2}$.

Proof. We rely on (69) and use the fact that, for a distribution with unbounded support, $\overline{s} = \infty$.

For $X_\sigma = \frac{\overline{e}}{2}$, the vega $\nu_{\overline{e}}$ of the option-when-working ($e = \overline{e}$) is

$$\nu_{\overline{e}} = \int_{\frac{X_\sigma - \overline{e}}{\sigma}}^{\infty} yg(y)ds = \int_{-\frac{\overline{e}}{2\sigma}}^{\infty} yg(y)ds. \tag{70}$$

For $X_\sigma = \frac{\overline{e}}{2}$, the vega ν_0 of the option-when-shirking ($e = 0$) is

$$\nu_0 = \int_{\frac{X_\sigma}{\sigma}}^{\infty} yg(y)ds = \int_{\frac{\overline{e}}{2\sigma}}^{\infty} yg(y)ds. \tag{71}$$

In addition,

$$\int_{-\frac{\overline{e}}{2\sigma}}^{\infty} yg(y)ds = \int_{-\frac{\overline{e}}{2\sigma}}^{\frac{\overline{e}}{2\sigma}} yg(y)ds + \int_{\frac{\overline{e}}{2\sigma}}^{\infty} yg(y)ds \tag{72}$$

For a symmetric distribution, we have $\int_{-\frac{\overline{e}}{2\sigma}}^{\frac{\overline{e}}{2\sigma}} yg(y)ds = 0$. Equation (72) then implies that $\nu_{\overline{e}} = \nu_0$. ■

D Asymmetric and Unbounded Distributions

This section generalizes the results of Section 3.2. We relax the assumption that distributions with a location and scale parameter are symmetric around the mean, and we study distributions with a bounded support.

We start with the case of asymmetric distributions with unbounded support. Proposition 2 now becomes:

[12]With high effort, $e = \overline{e}$, so the "option-when-working" is ATM for $X_\sigma = \overline{e}$. With low effort, $e = 0$, so the "option-when-shirking" is ATM for $X_\sigma = 0$.

Proposition 3 *(Effect of volatility on strike price, asymmetric distributions)* $\frac{dX_\sigma}{d\sigma} > 0$ *if and only if* $X_\sigma > \hat{X} \in (0, \bar{e})$.

Proof. The Proof is similar to the Proof of Proposition 2, and is therefore abbreviated.

Using Lemma 3 and equation (41), the agent's expected pay satisfies increasing differences if and only if

$$\frac{\partial}{\partial \theta} \int_{-\infty}^{X_\theta} \left[G\left(\frac{s - \bar{e}}{\sigma}\right) - G\left(\frac{s - 0}{\sigma}\right) \right] ds \geq 0. \tag{73}$$

Substituting $\sigma = \frac{1}{\sqrt{\theta}}$ and differentiating, this yields:

$$\frac{1}{2\theta} \int_{-\infty}^{X_\sigma} \left[\sqrt{\theta}(s - \bar{e}) g\left(\sqrt{\theta}(s - \bar{e})\right) - \sqrt{\theta} s g\left(\sqrt{\theta} s\right) \right] ds \geq 0$$

Given that $\theta > 0$, we use a changes of variable to rewrite this condition as

$$-\int_{\sqrt{\theta}(X_\sigma - \bar{e})}^{\sqrt{\theta} X_\sigma} y g(y)\, dy \geq 0 \tag{74}$$

Given that $g > 0$, the condition in (74) is satisfied for $X_\sigma \leq 0$, and not satisfied for $X_\sigma \geq \bar{e}$. In addition,

$$\frac{\partial}{\partial X_\sigma} \left\{ -\int_{\sqrt{\theta}(X_\sigma - \bar{e})}^{\sqrt{\theta} X_\sigma} y g(y)\, dy \right\} = -\theta X_\sigma g(\sqrt{\theta} X_\sigma) + \theta (X_\sigma - \bar{e}) g(\sqrt{\theta}(X_\sigma - \bar{e})) \tag{75}$$

which is strictly negative for $X_\sigma \in (0, \bar{e})$, as both terms on the RHS are negative. We conclude that there exists a unique $\hat{X} \in (0, \bar{e})$ such that condition (74) is satisfied if and only if $X_\sigma \leq \hat{X}$. The result then follows directly from Proposition 1. ∎

We now consider the case of symmetric or asymmetric distributions with bounded support, including the uniform distribution.

Proposition 4 *(Effect of volatility on strike price, bounded support).*

(i) For for $X_\sigma \geq \underline{s} + \bar{e}$, $\frac{dX_\sigma}{d\sigma} > 0$ *if and only if*

$$\int_{\sqrt{\theta}(\underline{s} - \bar{e})}^{\sqrt{\theta}\underline{s}} y g(y)\, dy - \int_{\sqrt{\theta}(X_\sigma - \bar{e})}^{\sqrt{\theta} X_\sigma} y g(y)\, dy < 0 \tag{76}$$

(ii) For for $X_\sigma < \underline{s} + \bar{e}$, $\frac{dX_\sigma}{d\sigma} > 0$ if and only if

$$\int_{\sqrt{\theta}(\underline{s}-\bar{e})}^{\sqrt{\theta}(X_\sigma-\bar{e})} yg(y)\,dy - \int_{\sqrt{\theta}\underline{s}}^{\sqrt{\theta}X_\sigma} yg(y)\,dy < 0 \tag{77}$$

(iii) With a uniform distribution, $\frac{dX_\sigma}{d\sigma} > 0$.

Proof. The Proof is similar to the Proof of Proposition 2, and is therefore abbreviated.

Using Lemma 3 and equation (41), the agent's expected pay satisfies increasing differences if and only if

$$\frac{\partial}{\partial\theta}\int_{\underline{s}}^{X_\theta}\left[G\left(\frac{s-\bar{e}}{\sigma}\right) - G\left(\frac{s-0}{\sigma}\right)\right] ds \geq 0. \tag{78}$$

Substituting $\sigma = \frac{1}{\sqrt{\theta}}$ and differentiating, this yields:

$$\frac{1}{2\theta}\int_{\underline{s}}^{X_\sigma}\left[\sqrt{\theta}(s-\bar{e})g\left(\sqrt{\theta}(s-\bar{e})\right) - \sqrt{\theta}sg\left(\sqrt{\theta}s\right)\right] ds \geq 0.$$

Given that $\theta > 0$, and for $X_\sigma \geq \underline{s} + \bar{e}$, we use changes of variables to rewrite this condition as

$$\int_{\sqrt{\theta}(\underline{s}-\bar{e})}^{\sqrt{\theta}(X_\sigma-\bar{e})} yg(y)\,dy - \int_{\sqrt{\theta}\underline{s}}^{\sqrt{\theta}X_\sigma} yg(y)\,dy \geq 0 \tag{79}$$

First, for $X_\sigma \geq \underline{s} + \bar{e}$, (79) rewrites as

$$\int_{\sqrt{\theta}(\underline{s}-\bar{e})}^{\sqrt{\theta}\underline{s}} yg(y)\,dy - \int_{\sqrt{\theta}(X_\sigma-\bar{e})}^{\sqrt{\theta}X_\sigma} yg(y)\,dy \geq 0 \tag{80}$$

Part (i) then follows directly from Proposition 1. For part (iii), with a uniform distribution, the density is equal to a positive constant, so that condition (80) may be rewritten as

$$\int_{\sqrt{\theta}(\underline{s}-\bar{e})}^{\sqrt{\theta}\underline{s}} ydy \geq \int_{\sqrt{\theta}(X_\sigma-\bar{e})}^{\sqrt{\theta}X_\sigma} ydy \tag{81}$$

Given that $X_\sigma - \bar{e} \geq \underline{s}$, this condition does not hold for any X_σ.

Second, for $X_\sigma < \underline{s} + \bar{e}$, part (ii) follows directly from (79) and Proposition 1. For part (iii), with a uniform distribution, the density is equal to a positive constant, so that condition (79) may be rewritten as

$$\int_{\sqrt{\theta}(\underline{s}-\bar{e})}^{\sqrt{\theta}(X_\sigma-\bar{e})} ydy \geq \int_{\sqrt{\theta}\underline{s}}^{\sqrt{\theta}X_\sigma} ydy \tag{82}$$

Given that $X_\sigma - \bar{e} < \underline{s}$, this condition does not hold for any X_σ, which completes the proof. ■

E At-The-Money Options

This Appendix shows that the model's main results continue to hold when the principal is restricted to granting ATM options.

We consider the same problem described in Section 3.1, except that we assume that the contract takes the form of ATM options. Considering ATM options requires that we complement the model by deriving the $t = 0$ stock price. To simplify the exposition, we assume that the firm has a single share outstanding. Denoting the stock price at time 0 by S_0, we have $S_0 = \mathbb{E}[q]$ given the assumptions of a zero discount rate and risk neutrality. In addition, with a symmetric distribution with location parameter e, we have $S_0 = e$.

Since the strike price is fixed (at the stock price), the number $n \leq 1$ of ATM options granted adjusts to satisfy the incentive constraint.[13] It follows that $e = \bar{e}$ in equilibrium, and $S_0 = \bar{e}$. With ATM options, the exercise price is therefore $X = \bar{e}$. Considering the same distributions as in section 3.3, we have the following results:

Lemma 9 *(Effect of volatility on number of options) With ATM options, $\frac{dn}{d\sigma} < 0$.*

Proof. Totally differentiating the LHS of the incentive constraint in (4) with respect to σ yields

$$\frac{d}{d\sigma}\left\{\mathbb{E}[W(s)|\bar{e}] - \mathbb{E}[W(s)|0]\right\} = \frac{\partial}{\partial\sigma}\left\{\mathbb{E}[W(s)|\bar{e}] - \mathbb{E}[W(s)|0]\right\} + \frac{\partial}{\partial n}\left\{\mathbb{E}[W(s)|\bar{e}] - \mathbb{E}[W(s)|0]\right\}\frac{dn}{d\sigma} = 0$$

so that

$$\frac{dn}{d\sigma} = -\frac{\frac{\partial}{\partial\sigma}\left\{\mathbb{E}[W(s)|\bar{e}] - \mathbb{E}[W(s)|0]\right\}}{\frac{\partial}{\partial n}\left\{\mathbb{E}[W(s)|\bar{e}] - \mathbb{E}[W(s)|0]\right\}} \tag{83}$$

[13] We only consider the cases such that there exists an incentive compatible contract with ATM options subject to the constraint $n \leq 1$.

First, if the agent receives n options instead of 1, using the same steps as for (12) yields

$$\mathbb{E}[W(s)|\bar{e}] - \mathbb{E}[W(s)|0] = n\int_X^\infty [F(s|0) - F(s|e^\star)]\, ds$$

for any given X. With distributions with a location parameter e and scale parameter σ, the numerator of the fraction on the RHS of (83) is then

$$\frac{\partial}{\partial\sigma}\left\{\mathbb{E}[W(s)|\bar{e}] - \mathbb{E}[W(s)|0]\right\} = n\frac{\partial}{\partial\sigma}\int_X^\infty \left[G\left(\frac{s}{\sigma}\right) - G\left(\frac{s-\bar{e}}{\sigma}\right)\right] ds$$

$$= n\int_X^\infty \left[-\frac{s}{\sigma^2}g\left(\frac{s}{\sigma}\right) + \frac{s-\bar{e}}{\sigma^2}g\left(\frac{s-\bar{e}}{\sigma}\right)\right] ds$$

$$= n\left[\int_{\frac{X}{\sigma}}^\infty -y_L g\left(y_L\right) ds + \int_{\frac{X-\bar{e}}{\sigma}}^\infty y_H g\left(y_H\right) ds\right] = n\int_{\frac{X-\bar{e}}{\sigma}}^{\frac{X}{\sigma}} y g\left(y\right) ds$$

where we used the changes of variables $y_L = \frac{s}{\sigma}$ and $y_H = \frac{s-\bar{e}}{\sigma}$. Given the symmetry of g, we have $\int_{\frac{X-\bar{e}}{\sigma}}^{\frac{X}{\sigma}} y g\left(y\right) ds \geq 0$ if and only if $\frac{X}{\sigma} > -\frac{X-\bar{e}}{\sigma}$, which is always the case with ATM options, i.e., with $X = \bar{e}$. We conclude that the numerator of the fraction on the RHS of (83) is strictly positive with ATM options.

Second, for an agent who receives n ATM options, the denominator of the fraction on the RHS of (83) is equal to

$$\frac{\partial}{\partial n}\left\{\int_X^\infty n(s-X)f(s|\bar{e})ds - \int_X^\infty n(s-X)f(s|0)ds\right\}$$

$$= \int_X^\infty (s-X)f(s|\bar{e})ds - \int_X^\infty (s-X)f(s|0)ds > 0 \tag{84}$$

As both the numerator and the denominator of the fraction on the RHS of (83) are strictly positive, we have

$$\frac{dn}{d\sigma} < 0 \tag{85}$$

with ATM options. ■

If informativeness is improved (with a lower σ), n must increase to maintain incentive compatibility. This is because $X = \bar{e} > \frac{\bar{e}}{2}$ with ATM options: the exercise price is higher than the threshold $\frac{\bar{e}}{2}$, so that an increase in informativeness reduces effort incentives, ceteris paribus (the intuition is the same as in section 3.3). Incentive compatibility then requires that the agent be given additional options.

Lemma 10 *(Partial and total effects of informativeness on expected pay)*

$$\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma} < \frac{\partial\mathbb{E}[W(s)|\bar{e}]}{\partial\sigma}$$

Proof. First,

$$\frac{d\mathbb{E}[W(s)|\bar{e}]}{d\sigma} = \frac{\partial\mathbb{E}[W(s)|\bar{e}]}{\partial\sigma} + \frac{\partial\mathbb{E}[W(s)|\bar{e}]}{\partial n}\frac{dn}{d\sigma}$$

Second,

$$\frac{\partial\mathbb{E}[W(s)|\bar{e}]}{\partial n} = \frac{\partial}{\partial n}\left\{\int_X^\infty n(s-X)f(s|\bar{e})ds\right\} = \int_X^\infty (s-X)f(s|\bar{e})ds > 0$$

Lemma 10 then follows from this inequality and (85). ■

With ATM options, the total change in expected pay that follows a change in informativeness is of a lesser magnitude than the partial change holding all else constant: while an improvement in informativeness lowers the value of the agent's options, ceteris paribus, it also requires that the agent receives more options for incentive compatibility. This incentive effect partially offsets the benefits to the principal.

F Continuous Effort Model

In this section, we sketch a continuous effort analog of the core model, in which the principal writes a contract to induce a given level of effort $\bar{e} > 0$. The model remains the same, except for the following assumptions:

(A1) The agent chooses effort in $e \in [0, \infty)$.

(A2) The agent's objective function is $\mathbb{E}[W(s)|e] - c\xi(e)$, with $c > 0$, $\xi > 0, \xi' > 0, \xi'' > 0$.

(A3) MLRP: $\frac{d}{ds}\left\{\frac{f_e(s|e)}{f(s|e)}\right\} > 0$, where $f(\pi|e)$ denotes the PDF of s conditional on e, and $f_e(\pi|e)$ denotes its first derivative with respect to e.

(A4) $\mathbb{E}[\max\{s - X_\sigma, 0\}|e] - c\xi(e)$ is concave in e, and $W(s)$ is piecewise smooth with a right derivative, which guarantees that the first-order approach to the effort choice problem applies (cf. footnote 12 in Innes (1990)).

As in section 3.2, we consider continuously distributed symmetric distributions with unbounded support and location and scale parameters, denoted respectively by e and σ. This implies that we can write $s = e + \epsilon$, where ϵ is mean-zero noise that is uncorrelated with effort.

For a given scale parameter σ, the principal's problem is to choose a cadlag function $W(\cdot)$ to minimize $\mathbb{E}[W(s)|\bar{e}]$ subject to the same constraints on contracting as in the core model and the following incentive constraint:

$$\frac{d}{de}\int_{-\infty}^{\infty} W(s)f(s|\bar{e})ds = c\xi'(\bar{e}) \tag{86}$$

Then applying Proposition 1 in Innes (1990) implies that, for a given σ, the optimal contract is characterized by

$$W(s) = \max\{0, s - X_\sigma\} \tag{87}$$

As in the core model, there is a unique X_σ which satisfies the incentive constraint in (86) with equality. Subsequent calculations require that the $W(s)$ function be of class C^2 on the whole domain. This can be achieved by an arbitrarily small change in $W(s)$ on $(X_\sigma - u, X_\sigma + u)$, where $u \to 0$, which smooths out the kink at X_σ (Zang (1980)) which leaving expected pay conditional on any e approximately unchanged.

Denoting by ψ and Ψ the PDF and CDF smoothing-out technique for min-max optimization of ϵ, respectively, the incentive constraint in (86) can be rewritten as

$$\int_{-\infty}^{\infty} W'(\bar{e}+\epsilon)\psi(\epsilon)d\epsilon = c\xi'(\bar{e}). \tag{88}$$

An increase in σ is a mean-preserving spread ("MPS") of the probability distribution of s, or equivalently a MPS to the probability distribution of ϵ. Denote by $\bar{\psi}$ and $\bar{\Psi}$ the PDF and CDF of ϵ after a MPS, respectively. We know that a MPS implies single crossing of the CDFs (e.g., Pratt and Machina (1997)), so that there is a unique $\hat{x} \in (-\infty, \infty)$ such that $\Psi(\hat{x}) = \bar{\Psi}(\hat{x})$. In addition, since we are considering distributions which are symmetrical about the mean, $\hat{x} = 0$, because $\Psi(0) = \bar{\Psi}(0) = 0.5$ and a MPS preserves the mean.

For a given exercise price, an improvement in informativeness or reduction in dispersion in a MPS reduces the LHS of the incentive constraint in (86) if and only if

$$\int_{-\infty}^{\infty} W'(\bar{e}+\epsilon)\left(\psi(\epsilon)-\bar{\psi}(\epsilon)\right)d\epsilon<0. \tag{89}$$

Integrating by parts, this becomes

$$\left[W'(\bar{e}+\epsilon)\left(\Psi(\epsilon)-\bar{\Psi}(\epsilon)\right)\right]_{-\infty}^{\infty}-\int_{-\infty}^{\infty} W''(\bar{e}+\epsilon)\left(\Psi(\epsilon)-\bar{\Psi}(\epsilon)\right)d\epsilon<0. \tag{90}$$

The first term is equal to zero since $\Psi(-\infty)=\bar{\Psi}(-\infty)=0$ and $\Psi(\infty)=\bar{\Psi}(\infty)=1$. In addition, $W''(s)>0$ on $(X_\sigma-u, X_\sigma+u)$, and is equal to zero elsewhere. Since $u\to 0$ and Ψ is continuous, it follows that (90) is satisfied if and only if $\Psi(X_\sigma-\bar{e})>\bar{\Psi}(X_\sigma-\bar{e})$. In turn, because of the single crossing property of the MPS and the symmetry of Ψ, this is satisfied if and only if $X_\sigma>\bar{e}$.

As in the baseline model, the LHS of the incentive constraint is strictly decreasing in X. Therefore, for the incentive constraint to still be satisfied following an improvement in informativeness in a MPS sense (i.e., a lower σ), we have

$$\frac{dX_\sigma}{d\sigma}>0 \qquad \text{if and only if } X_\sigma>\bar{e} \tag{91}$$

Thus, as informativeness increases (σ falls), X_σ approaches $\bar{e}$ and the option becomes closer to ATM.