The Value of Information for Contracting*

Pierre Chaigneau
Queen's University

Alex Edmans
LBS, CEPR, and ECGI

Daniel Gottlieb
Washington University in St Louis

July 27, 2016

Abstract

To assess the value of information for contracting, it is necessary to study how contracts change with signal precision. This paper studies the standard setting of risk neutrality and limited liability, which permits an optimal contracting approach. One application is to executive compensation, where the contract is an option. The direct effect of reducing signal volatility is a fall in the option's value, which benefits the principal. The indirect effect is on effort incentives. If the original option is sufficiently out-of-the-money, the agent can only beat the strike price if he works and there is a high noise realization. Thus, a fall in volatility reduces effort incentives, lowering the value of information. In contrast, standard option theory suggests that volatility has greatest effect for at-the-money options. A second application is to financing, where the contract is debt. The model has implications for the value of risk management and a firm's ability to raise financing.

Keywords: executive compensation, limited liability, options, pay-for-luck, relative performance evaluation, risk management.
JEL Classification: D86, G32, G34, J33.

*pierre.chaigneau@queensu.ca, aedmans@london.edu, dgottlieb@wustl.edu. We thank Lin William Cong, Ohad Kadan, seminar participants at Queen's, Wharton, Washington University Conference on Corporate Finance, and Risk Theory Society for helpful comments, and Shiying Dong for excellent research assistance.

Since the seminal contributions of Holmstrom (1979) and Shavell (1979), the moral hazard literature has shown that superior information on agent performance can reduce the principal's cost of implementing a given action. This result has implications for many contracting applications, such as compensation, financing, insurance, and regulation. While information can be valuable, it is also costly. Thus, to determine whether investing in information is efficient for the principal, we need to quantify its benefits so that we can compare it against its costs. This involves solving for how the optimal contract changes under additional information, and calculating the resulting cost savings. Relating these savings to the underlying parameters of the agency problem will identify the situations in which information is most valuable.

As is well-known, solving for the optimal contract in a general setting is highly complex (e.g. Grossman and Hart (1983)). Without doing so, we cannot determine how the contract changes in response to superior information, and thus quantify the cost savings. This paper addresses this open question. We consider the standard framework of risk neutrality and limited liability, originally analyzed by Innes (1990) and widely used in a number of settings (e.g. Biais et al. (2010), Clementi and Hopenhayn (2006), DeMarzo and Fishman (2007a, 2007b), DeMarzo and Sannikov (2006)). This framework allows us to take an optimal contracting approach and leads to contracts that we observe in practice. As shown by Innes (1990), this model has two major applications. The first is an employment setting, where the principal (firm) hires an agent (manager), in which case the optimal contract involves giving the agent a call option.[1] 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. The second is a financing setting where the agent (entrepreneur) raises financing from the principal (investor), in which case the contract is debt, and the strike price represents its face value. We initially present the results using the employment application, since we can discuss the intuition using option theory, but later discuss the implications for financing.

We model the option contract as based on output and precision as affecting output volatility, but the results are unchanged if the contract is instead based on a separate performance signal and precision affects the volatility of this signal. We consider

[1] While options are not the only instruments used in practice, Dittmann and Maug (2007) find that the payoff structure provided by a CEO's overall compensation package resembles an option.

general distributions for the performance measure that have a location parameter, i.e. the location of the distribution (such as its mean) can be changed independently of its shape. Examples include the normal, uniform, logistic, Cauchy, and Laplace distributions. Such distributions are a natural setting to study moral hazard, since effort can be modeled as shifting this location parameter. We show that an increase in precision, in the sense of second order stochastic dominance, has two effects, each with a clear economic interpretation. First, a fall in volatility reduces the value of the option and thus the expected wage: the direct effect. Second, it changes the agent's incentives, thus requiring the principal to change the contract to maintain incentive compatibility. The heart of the paper analyzes this indirect effect. Our optimal contracting approach allows us to determine how the contract changes with precision, ascertain the direction of the incentive effect, and show how it depends on the model's underlying parameters.

The agent's effort incentives are determined by the difference between the value of the option on the signal generated by shirking ("option-when-shirking") and the one generated by working ("option-when-working"). Increases in signal precision affect the values of these two options differentially and thus raise or lower effort incentives. Our main contribution is to derive clean results showing that the direction of the incentive effect depends on the initial strike price of the option. Since the initial strike price depends on the severity of the agency problem (a large agency problem, i.e. a high cost of effort, requires a low strike price to induce effort), we can relate the effect of precision on incentives to the underlying agency problem.

When the initial strike price is below a threshold, i.e. the agency problem is strong, precision increases effort incentives. The intuition is as follows. A decrease in precision (increase in signal volatility) raises the option's value. The magnitude of the gain is increasing in the option's vega, which is highest when the option is at-the-money. When the initial strike price is low, then if the agent shirks, the option is close to at-the-money and has a high vega; if he works, the option is in-the-money and has a low vega. Thus, a fall in precision increases the value of the option-when-shirking more than the option-when-working, and lowers effort incentives. Intuitively, when volatility is high, incentives are weak because, even if the agent shirked, he would still earn a high wage if he received a positive shock. He is not worried about shirking and receiving a negative shock, because his payoff cannot fall below zero due to limited liability.

When the initial strike price is above a second (higher) threshold, i.e. the agency problem is weak, effort and informativeness are substitutes due to the reverse intuition.

The option-when-shirking is deeply out-of-the-money, and the option-when-working is closer to at-the-money. Thus, the vega of the latter option is greater, and its value increases faster with volatility, raising incentives. Intuitively, when the strike price is high, the agent is paid only if he exerts effort *and* receives a sufficiently positive shock. When volatility rises (i.e. precision falls), such shocks are more likely, and so the agent is more likely to be paid. Thus, his effort incentives increase.

For initial strike prices between the two thresholds, precision can either increase or decrease effort incentives. This is because, for general distributions, a decrease in informativeness may not have a consistent effect on the signal distribution: while it shifts mass towards the tails, it could also shift some mass towards the center. Under a simple regularity condition which guarantees that decreases in informativeness consistently shift mass from the center to the tails, the two thresholds now coincide at a single point and there is no ambiguous intermediate region. The effect of informativeness on incentives then depends on whether the initial strike price is above or below this single threshold. Thus, an increase in precision moves the strike price towards this threshold: it reduces it if the strike price is initially above the threshold, and raises it if initially below. A sufficient (although not necessary) condition for regularity is that the signal distribution has a scale as well as a location parameter, as with the normal, uniform, and logistic distributions. Intuitively, when volatility can be characterized by a scale parameter (such as a standard deviation), changes in this parameter consistently move mass towards the tails. Regularity is also automatically satisfied by a mean-preserving spread.

Our results have a number of applications for employment contracts. First, it identifies the settings in which investing in information is optimal for the principal. When incentives are strong (weak) to begin with, e.g. for CEOs (managers), an increase in the precision of the performance measure further increases (reduces) incentives, raising (lowering) the gains from informativeness. Note that it is far from obvious that the value of information is high for severe agency problems. An analysis focusing only on the direct effect of informativeness, and ignoring the incentive constraint, would suggest that the value is highest when the option is at-the-money – i.e. a moderate initial strike price and a moderate agency problem.

One way in which the principal can invest in information is to engage in relative performance evaluation ("RPE"), which is costly as it involves forgoing the benefits of pay-for-luck documented by prior research (e.g. Oyer (2004), Axelson and Baliga

(2009), and Gopalan, Milbourn, and Song (2010)). Such an analysis is also valuable to assess the efficiency of real-life contracts. There is very little evidence that RPE is used for rank-and-file employees, and only modest evidence of its usage for CEOs.[2] Bebchuk and Fried (2004) interpret this rarity as evidence that CEO contracts are inefficient. However, to evaluate this argument, we need to identify the settings in which the value of information is smallest, and compare them to the cases in which RPE is particularly absent in reality. That RPE is more common for CEOs than employees is at least directionally consistent with the idea that precision is more valuable when the agency problem is stronger.

Second, in addition to the gains from precision, our analysis also studies the impact of exogenous changes in precision. An increase in volatility raises (lowers) the incentives of agents with out-of-the-money (in-the-money) options. If firms recontract in response, CEOs with in-the-money options should receive the highest increase in incentives.

Third, the results have implications for how precision affects the probability of firing. The strike price can be thought of as a performance target below which the agent is fired (since he is paid zero). Simple intuition would suggest that more precise monitoring will always increase the firing probability, but this intuition ignores the fact that the target is endogenously chosen. If the target is initially high, precision weakens effort incentives. Thus, to preserve incentive compatibility, the target must be lowered, reducing the probability of firing and in some cases outweighing the first effect.

Fourth, for tractability, the analysis features a binary effort level. In the continuous-effort analog, we show that the threshold for the initial strike price – that determines whether precision increases or decreases effort – is the expected value of the signal. If the initial strike price is above (below) this threshold, increases in precision lower (raise) the strike price towards the threshold, i.e. towards the expected signal value. Thus, such increases (e.g. improvements in stock market efficiency) move the option closer to at-the-money. Bebchuk and Fried (2004) argue that the almost universal use of at-the-money options is suboptimal, versus out-of-the-money options which pay the agent only upon good performance. Our analysis suggests that at-the-money options can be close to optimal if precision is high. This result also suggests that accounting or tax considerations that favor at-the-money options need not induce suboptimal contracting.

[2] While Aggarwal and Samwick (1999) and Murphy (1999) document almost no use, the more recent study of Gong, Li, and Shin (2011) find that 25% of S&P 1500 firms explicitly use RPE. See Edmans and Gabaix (2016) for a review of the evidence on RPE.

Our results also have implications for financing contracts, where the investor has risky debt. Debt contracts are based on output (e.g. cash flow) rather than a separate signal. Thus, our model sheds light on the settings in which the investor's incentive to reduce output volatility is highest, for example through hedging, signing long-term contracts with customers and suppliers, and investing in a less risky production technology. Such risk management has several interpretations: the investor can implement risk management herself since she retains control rights on output; she stipulates in the contract that the entrepreneur implement the above measures; or she has a menu of projects that she can finance and thus can choose project risk.

An analysis based on the direct effect would suggest that risk management is most valuable when the face value of debt is close to expected firm value (i.e. firms at the bankruptcy threshold), as then the value of debt is most sensitive to volatility. This is consistent with standard intuition (e.g. Stulz (1996)) that risk reduction incentives are increasing in loan size (up to the bankruptcy threshold), because the lender has more at stake. This intuition is incomplete due to the incentive effect. When the agency problem is strong and thus the face value of debt is low, risk management raises effort incentives, further increasing its value over and above the direct effect. Thus, surprisingly, risk management may be more valuable for firms that are some distance from bankruptcy, and when the investor has little debt at stake.

This result also has implications for a firm's ability to raise external financing. In standard models (e.g. Innes (1990), Holmstrom and Tirole (1997), Tirole (2006)), a strong agency problem reduces a firm's pledgeable income (i.e. leads to a low face value of debt) since the entrepreneur must retain a sufficient share of output to induce effort. Our results suggest that, when the agency problem is strong, the investor is more likely to reduce risk, increasing effort incentives and thus the firm's pledgeable income and face value of debt. Thus, the endogenous response of risk management mitigates the negative effect of agency problems on pledgeable income.

Dittmann, Maug, and Spalt (2013) also consider the incentive constraint when assessing the benefits of a specific form of increased precision – indexing stock and options – and similarly show that indexation may weaken incentives. They use a quite different setting, 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 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 recontracts optimally. In contrast, our primary goal is theoretical. We incorporate risk neutrality and limited liability, allowing an optimal contracting approach to solve analytically for how the contract changes in response to information. In addition, our theory is somewhat broader and allows the analysis of reductions in volatility through other means than indexation, for example investing in a superior monitoring technology, and for an application to debt contracts. Our main contribution is not to explain the rarity of indexation, but to characterize the contracting settings in which improving signal precision is most valuable, once we take into account the change in the contract necessary to maintain incentive compatibility.

This paper proceeds as follows. Section 1 presents the model, and Section 2 shows that the optimal contract is a call option, as in Innes (1990). Section 3 presents our main results. It derives the gains from increased signal precision, and in particular relates the effect on effort incentives to the underlying parameters of the agency problem. 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, as in Innes (1990). At time $t = -1$, the principal offers a contract to the agent. At $t = 0$, the agent chooses effort $e \in \{0, \overline{e}\}$, where $e = 0$ (low effort, shirking) costs him zero and $e = \overline{e}$ (high effort, working) costs him $C > 0$. At $t = 1$, output q is realized. Since output is a signal of the agent's effort, we sometimes refer to q as the "signal" going forwards. While effort is unobservable, output is contractible. The agent is paid a wage $W(q)$ and the principal receives $R(q) = q - W(q)$.

The distribution of output belongs to a location family with effort e being its location parameter. More precisely, output equals

$$q = e + \varepsilon, \tag{1}$$

where ε is continuously distributed according to a probability density function ("PDF") g_θ, with full support on an interval of the real line. Equation (1) is without loss of generality, since we can always define "noise" ε as the difference between effort and output. In practice, noise can result from a market or industry shock, the contribution of other managers, or measurement error. To belong to a location family, the noise distribution g_θ cannot be a function of e: exerting effort shifts the entire distribution of output rightward without affecting its shape.

Let $G_\theta(\varepsilon)$ denote the cumulative distribution function ("CDF") of ε, and let $f_\theta(q|e) \equiv g_\theta(q-e)$ and $F_\theta(q|e) \equiv G_\theta(q-e)$ denote the PDF and CDF of output conditional on effort e. High output is good news about effort in the sense of the monotone likelihood ratio property ("MLRP"): $\frac{g_\theta(q-\bar{e})}{g_\theta(q)}$ is strictly increasing in q for any fixed θ.[3]

The real-valued parameter θ orders the precision of the signal distribution in the sense of second-order stochastic dominance ("SOSD").[4] It thus captures the informativeness of the signal q for the agent's effort ε. Formally, the mean of ε is independent of θ and

$$\theta \geq \theta' \implies \int_{-\infty}^{t} G_\theta(\varepsilon)d\varepsilon \leq \int_{-\infty}^{t} G_{\theta'}(\varepsilon)d\varepsilon \tag{2}$$

for all t. Thus, increases in θ generate more precise signal distributions. The parameter θ may be chosen by the principal, or result from exogenous changes such as a reduction in economic uncertainty. Our goal is to analyze the value of information, which applies to either setting.

The agent is risk-neutral and so maximizes his expected wage

$$\mathbb{E}\left[W(q)|e\right] = \int_{-\infty}^{\infty} W(q) f_\theta\left(q|e\right) dq,$$

less the cost of effort. His reservation utility is zero and there is no discounting. The principal is also risk-neutral and chooses a contract $W(\cdot)$ and an effort level e to maximize expected output $\mathbb{E}[q]$ less the expected wage $\mathbb{E}[W]$.

Following Innes (1990), we make two assumptions on the set of feasible contracts.

[3]Using the definition of f_θ, we can rewrite this condition as the usual definition: $\frac{f_\theta(q_1|\bar{e})}{f_\theta(q_1|0)} > \frac{f_\theta(q_0|\bar{e})}{f_\theta(q_0|0)}$ for all q_1 and q_0 with $q_1 > q_0$.

[4] Following Innes (1990), the principal contracts on output q. Since signal equals output, changes in the precision of the signal automatically lead to changes in the volatility of output. A previous version of the paper assumed that q was non-contractible and instead that there was a separate signal $s = q+\eta$ on which contracts could be written. Thus, the precision of the signal s could be affected without changing the volatility of output. All results continue to hold (because of risk neutrality, changing the volatility of output has no effect), but the notation is more complex due to the introduction of an additional variable.

First, there is a limited liability constraint ("LL") on the agent: $W(q) \geq 0 \, \forall \, q$. Second, a monotonicity constraint ensures the principal's payoff is non-decreasing in output:

$$\eta \geq W(q+\eta) - W(q) \tag{3}$$

for all $\eta > 0$. Innes (1990) justifies this constraint on two grounds. First, if it did not hold, the principal would have incentives to sabotage output. Second, if it did not hold, the agent would gain more than one-for-one for increases in output. Thus, he would have incentives to borrow on his own account to increase output.

In the first best, effort is verifiable. There is no incentive constraint ("IC") and only an individual rationality constraint ("IR"). As long as $\mathbb{E}\left[q|\overline{e}\right] - \mathbb{E}\left[q|0\right] \geq C$, the principal wishes to induce high effort, in which case the optimal contract is determined by the binding IR: $\mathbb{E}\left[W\left(q\right)|\overline{e}\right] = C$.

In the second best, the agent's effort is unverifiable and so the contract must satisfy an IC. The agent works if and only if:

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

Following standard arguments, this IC will bind, in which case the IR will be slack and can be ignored in the analysis that follows. We define X_θ implicitly by the binding IC:

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

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 q, X_θ is the strike price such that working increases the option's value by an amount equal to the cost of effort, so that the IC is satisfied with equality.

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

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

The left-hand side ("LHS") is the benefit to the principal of inducing $e = \overline{e}$ and the right-hand side ("RHS") is the cost of the contract required to do so.

Given θ, the principal's problem is to choose a contract $W_\theta\left(\cdot\right)$ to minimize the expected wage $\mathbb{E}\left[W_\theta\left(q\right)|\bar{e}\right]$ subject to the IC, LL, and monotonicity constraints:

$$\mathbb{E}\left[W_\theta\left(q\right)|\bar{e}\right] \geq \mathbb{E}\left[W_\theta\left(q\right)|0\right] + C, \tag{7}$$

$$W_\theta\left(q\right) \geq 0 \quad \forall\, q, \quad \text{and} \tag{8}$$

$$\epsilon \geq W_\theta\left(q+\epsilon\right) - W_\theta\left(q\right). \tag{9}$$

To ensure that $\lim_{\varepsilon\searrow-\infty} G_\theta(\varepsilon)$ and $\lim_{\varepsilon\nearrow\infty} G_\theta(\varepsilon)$ are differentiable with respect to θ, we make the technical assumptions that G_θ is differentiable with respect to θ and that the sequences $\{G_\theta(-n)\}_{n\in\mathbb{N}}, \left\{\frac{\partial G_\theta}{\partial\theta}(-n)\right\}_{n\in\mathbb{N}}, \{G_\theta(n)\}_{n\in\mathbb{N}}, \left\{\frac{\partial G_\theta}{\partial\theta}(n)\right\}_{n\in\mathbb{N}}$ are uniformly convergent. These assumptions are automatically satisfied if the noise has bounded support and are also satisfied under standard unbounded distributions (such as normal, uniform, logistic, Cauchy, and Laplace).

One difference from Innes (1990) is that he 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 precision 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 precision 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, such as the related paper by Dittmann, Maug, and Spalt (2013) on indexation. (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.) Edmans and Gabaix (2011) show that, if the benefits of effort are multiplicative in firm size and the firm is sufficiently large, it is always optimal for the principal to implement the highest effort level and so the optimal effort level is indeed fixed. We thus consider a binary effort setting where high effort is optimal. Appendix B.4 presents a continuous effort analog of the core model.[5]

[5]A second difference is that, in Innes, the agent offers the contract and maximizes his utility subject to the principal's participation constraint. Since it is the principal who will typically invest in information, we model her as offering the contract so that she will reap the benefits.

2 The Optimal Contract

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

Let $W_\theta(\cdot)$ and $R_\theta(\cdot)$ denote the optimal payments to the agent and principal for a given θ. Lemma 1 establishes that $W_\theta(\cdot)$ is a call option on q, where the strike price X_θ is chosen to satisfy the binding IC (5). Alternatively, $R_\theta(\cdot)$ can be viewed as risky debt with face value X_θ. This application is relevant for both mature firms, and also young firms since they frequently raise debt and the entrepreneur holds levered equity, as shown by Robb and Robinson (2014). We will initially discuss the former application, as this will allow us to use option theory to explain the intuition.

Lemma 1 *(Optimal Contract) For each θ, there exists an optimal contract with*

$$W_\theta(q) = \max\{0, q - X_\theta\}, \tag{10}$$

$$R_\theta(q) = \min\{q, X_\theta\} \tag{11}$$

where X_θ is the unique solution of (6). Any optimal contract coincides with $W_\theta(q)$ except on a set of outputs with probability zero under effort $\bar{e}$.

As in Innes (1990), the intuition is as follows. The absolute value of the likelihood ratio is highest in the tails of the distribution of q, so output is most informative about effort in the tails. The left tail cannot be used for incentive purposes due to limited liability, 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 q with the maximum feasible slope, i.e. $\frac{\partial}{\partial q}W_\theta(q) = 1$.

Lemma 2 below shows that the strike price falls with the cost of effort, which parametrizes the severity of the agency problem.

Lemma 2 *(Effect of effort cost on strike price): 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.*

3 The Value of Information

This section calculates the value of information to the principal, by studying the effect of changes in signal precision on the optimal contract and its cost to the principal. Section 3.1 provides a condition that relates the effect of precision on effort to the strike price of the option and thus the severity of the agency problem. Section 3.2 graphically illustrates the value of precision for the normal distribution. It also proves analytically that, for this distribution, the value of precision is monotonically increasing in the cost of effort, and thus monotonically decreasing in the initial strike price. Section 3.3 discusses applications of our results to executive compensation, employee compensation, and financing contracts.

3.1 Distribution in the Location Family

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

$$\frac{d}{d\theta}\mathbb{E}\left[W(q)|\overline{e}\right] = \underbrace{\frac{\partial}{\partial\theta}\mathbb{E}\left[W\left(q\right)|\overline{e}\right]}_{\textit{direct effect}} + \underbrace{\frac{\partial}{\partial X_\theta}\mathbb{E}\left[W\left(q\right)|\overline{e}\right]\frac{dX_\theta}{d\theta}}_{\textit{incentive effect}}. \tag{12}$$

The first component is the *direct effect*, $\frac{\partial}{\partial\theta}\mathbb{E}\left[W\left(q\right)|\overline{e}\right]$. 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 pay is the benefit of precision 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, additional information reduces the risk borne by the agent and thus allows the principal to lower the expected wage. In our setting of risk neutrality and limited liability, an increase in precision directly reduces the expected wage by lowering the value of the option.

The second component is the *incentive effect*, $\frac{\partial}{\partial X_\theta}\mathbb{E}\left[W\left(q\right)|\overline{e}\right]\frac{dX_\theta}{d\theta}$, which arises because the increase in precision requires the strike price to rise by $\frac{dX_\theta}{d\theta}$ to maintain incentive compatibility. $\frac{\partial}{\partial X_\theta}\mathbb{E}\left[W\left(q\right)|\overline{e}\right]$ is negative – any increase in the strike price reduces the value of the option – but the sign of $\frac{dX_\theta}{d\theta}$ is unclear and depends on how changes in precision affect the IC (4). The agent's incentives arise because exerting effort increases the option value. If he works, his option is worth $\mathbb{E}\left[W\left(q\right)|\overline{e}\right]$; we refer to this as an "option-when-working." If he shirks, he receives an "option-when-shirking" worth

$\mathbb{E}[W(q)|0]$. His effort incentives are given by the difference, i.e.

$$\mathbb{E}[W(q)|\bar{e}] - \mathbb{E}[W(q)|0]. \tag{13}$$

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 (13). When increasing precision raises the agent's effort incentives,

$$\frac{\partial}{\partial\theta}\{\mathbb{E}[W(q)|\bar{e}] - \mathbb{E}[W(q)|0]\} > 0, \tag{14}$$

we say that precision and effort are complements; when it reduces incentives, they are substitutes.

Even when $\frac{dX_\theta}{d\theta} < 0$ and the incentive effect counteracts the direct effect, it can never outweigh it. The total effect $\frac{d}{d\theta}\mathbb{E}[W(q)|\bar{e}]$ is always weakly negative 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. Appendix B.1 presents an example of the limit case where the incentive effect exactly offsets the direct effect, so that the total value of information equals exactly zero.[6]

Proposition 1 states that whether effort and information are complements or substitutes depends on the initial strike price of the option:

Proposition 1 *(Effect of information on the strike price): There exist $\widehat{X}_1$ and $\widehat{X}_2 \geq \widehat{X}_1$ such that*

(i) If $X_\theta < \widehat{X}_1$, effort and information are complements and so $\frac{dX_\theta}{d\theta} \geq 0$

(ii) If $X_\theta > \widehat{X}_2$, effort and information are substitutes and so $\frac{dX_\theta}{d\theta} \leq 0$.

Effort and information are complements when the initial strike price X_θ is below a threshold $\widehat{X}_1$, substitutes when X_θ exceeds a higher threshold $\widehat{X}_2$, and may be either complements or substitutes for $\widehat{X}_1 \leq X_\theta \leq \widehat{X}_2$. The intuition is as follows. A decrease in precision (from θ to θ') increases both tails of the signal distribution. If the initial strike price of the option is sufficiently low ($X_\theta < \widehat{X}_1$), then the signal distribution upon shirking has significant mass on both sides of X_θ. The agent benefits from high

[6]While we consider the effect of changing the precision of a given signal, Chaigneau, Edmans, and Gottlieb (2016) derive a condition for whether the addition of a new signal has strictly positive value for contracting under risk neutrality and limited liability.

signal realizations ($q > X_\theta$), since the option-when-shirking is in-the-money ("ITM") and so he exercises it, but does not lose from low signals ($q < X_\theta$) as the option is OTM and he does not exercise it. Thus, when precision falls, a shirking agent benefits from the growth in the right tail, but does not lose from the growth in the left tail, and so the value of the option-when-shirking increases significantly.

Working shifts the signal distribution rightwards. Thus, for $X_\theta < \widehat{X}_1$, the signal distribution upon working is mostly to the right of X_θ, and remains this way after precision falls. Since the option usually ends up ITM, the agent usually exercises it. Thus, a working agent benefits from the growth in the right tail *and* loses from the growth in the left tail, and so the value of the option-when-working is little changed.

Put differently, reductions in precision increase the value of an option due to its asymmetric payoff: the agent benefits from $q > X_\theta$, but does not lose from $q < X_\theta$. When X_θ is low and the agent shirks, the mean signal 0 is close to the kink X_θ and the agent benefits from the asymmetry. When the agent works, the mean signal $\overline{e}$ is far above the kink X_θ, and so he enjoys little asymmetry. Overall, a fall in θ raises the value of the option-when-shirking more than the option-when-working and thus reduces effort incentives. In simple language, the agent thinks "I'm not going to bother working, because even if I do, I might be unlucky and so my performance will be low. I might as well shirk, because even if I get unlucky and performance become very low, that doesn't matter, because I can't get paid less than zero."

For a sufficiently high strike price ($X_\theta > \widehat{X}_2$), the signal distribution upon shirking is mostly to the left of X_θ – and remains this way even after precision falls and the right tail expands. Thus, the option-when-shirking usually ends up OTM and its value is little changed. In contrast, if the agent works, this shifts the signal distribution rightward and so decreases in precision now push the right tail above X_θ. Thus, when precision falls, a working agent benefits from the growth in the right tail (since the option may end up ITM and he can now exercise it) but does not lose from the growth in the left tail (since still he does not exercise it). Put differently, when X_θ is high and the agent works, the mean signal $\overline{e}$ is close to the kink X_θ and the agent benefits from the asymmetry. When the agent shirks, the mean signal 0 is far from the kink X_θ, and so he enjoys little asymmetry. Overall, a fall in precision raises the value of the option-when-working more than the option-when-shirking, and thus increases effort incentives. In simple language, the agent thinks "If precision were high, I wouldn't bother working because the target X_θ is so high that I wouldn't meet it, even if I did

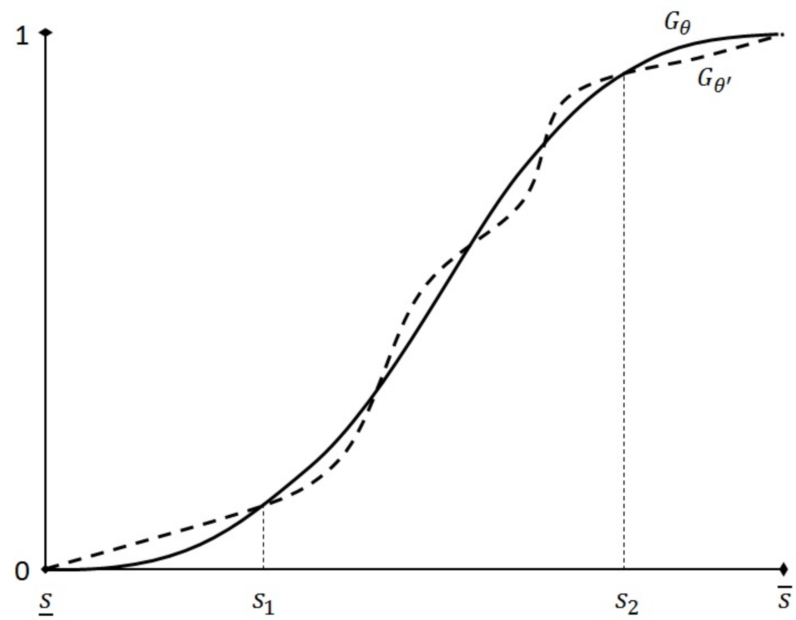

Figure 1: Signal distributions in the location family.

work. But, now that the signal is more noisy, I will work – because if I do, and I get lucky, I'll meet the target."

From Lemma 2, the initial strike price X_θ is decreasing in the cost of effort, and thus the severity of the agency problem. When the agency problem is mild (severe), the initial strike price is high (low); increases in precision reduce (increase) effort incentives, causing the strike price to fall (rise). Thus, precision improves effort incentives if the agency problem was initially severe.

However, for arbitrary distributions, it is unclear how changes in θ affect the distribution between $\widehat{X}_1$ and $\widehat{X}_2$. A fall in θ need not consistently shift mass from the center of the distribution towards the tails. It could shift some mass *towards* the center, as long as it also moves mass towards a more extreme tail point. Figure 1 shows that, while a fall in θ increases $G_\theta(q)$ for low q below a threshold q_1 (i.e. increases the left tail) and increases $1 - G_\theta(q)$ for high q above a threshold $q_2 > q_1$ (i.e. increases the right tail), the effect of θ on $G_\theta(q)$ is unclear for intermediate q. The CDFs G_θ and $G_{\theta'}$ could cross many times between $\widehat{X}_1$ and $\widehat{X}_2$.

Definition 1 below introduces a simple regularity condition that guarantees that falls in precision have a "consistent" effect on the distribution – they shift mass from

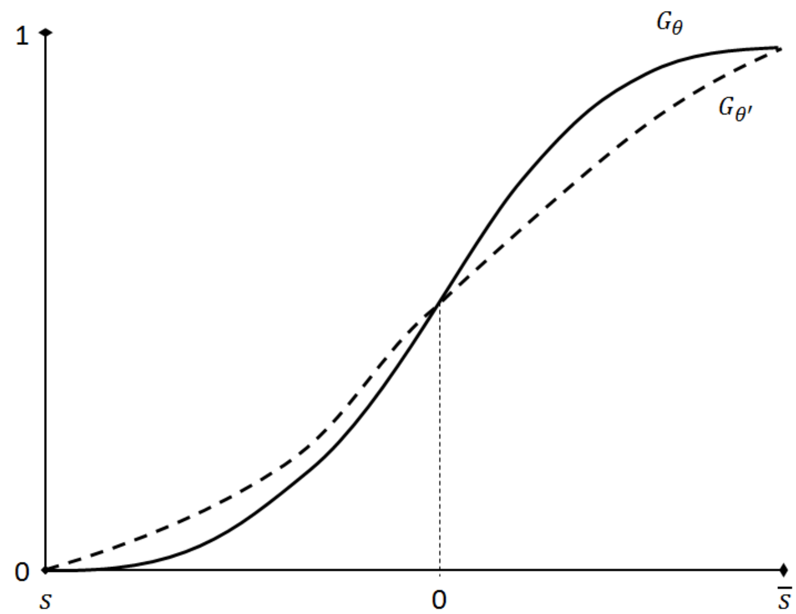

Figure 2: Regular signal distributions.

the center towards the tails. Proposition 2 shows that, for regular distributions, $\widehat{X}_1 = \widehat{X}_2 = \widehat{X}$, and so there is a single threshold below (above) which effort and information are complements (substitutes): the CDFs cross at a single point $\widehat{X}$, as in Figure 2. There is no intermediate range in which changes in precision have an ambiguous effect. Thus, if the initial strike price is above (below) $\widehat{X}$, increases in precision lower (raise) it; in both cases it moves towards $\widehat{X}$.

Definition 1 *The distribution G_θ is regular if there exists $\widehat{X}$ such that*

$$q \left\{ \begin{array}{c} < \\ > \end{array} \right\} \widehat{X} \implies \frac{\partial G_\theta}{\partial \theta}(q) \left\{ \begin{array}{c} \leq \\ \geq \end{array} \right\} 0.$$

Proposition 2 *(Effect of precision with regular distributions): Suppose that the noise distribution G_θ is regular. Then there exists $\widehat{X}$ such that $\frac{dX_\theta}{d\theta} \geq 0$ if $X_\theta < \widehat{X}$, and $\frac{dX_\theta}{d\theta} \leq 0$ if $X_\theta > \widehat{X}$.*

Regularity is not automatically implied by SOSD, but is satisfied by most standard distributions. For example, it is automatically satisfied by a mean-preserving spread as defined in Rothschild and Stiglitz (1970) (see Machina and Pratt (1997)).[7] In addition, Corollary 1 shows that regularity is satisfied by any signal distribution that has a scale parameter (in addition to a location parameter).

Corollary 1 *(Distributions with a scale parameter): If the signal distribution F has a scale parameter, i.e. its CDF can be written as $F_\sigma(q|e) = G\left(\frac{q-e}{\sigma}\right)$, then the noise distribution G is regular and so there exists $\widehat{X}$ such that $\frac{dX_\theta}{d\theta} \geq 0$ if $X_\theta < \widehat{X}$, and $\frac{dX_\theta}{d\theta} \leq 0$ if $X_\theta > \widehat{X}$.*

A distribution with a location and scale parameter can be fully characterized by its mean e and standard deviation σ. It is natural to consider distributions with scale parameters when studying changes in precision, since they can be represented by changes in the scale parameter σ. Since the volatility of a signal is the inverse of its precision, we have $\sigma = \frac{1}{\sqrt{\theta}}$ and so:

$$\frac{\partial}{\partial \sigma} G\left(\frac{q-e}{\sigma}\right) = -\frac{q-e}{\sigma^2} g\left(\frac{q-e}{\sigma}\right) \begin{cases} < 0 & \text{if } q > e \\ > 0 & \text{if } q < e \end{cases} \tag{15}$$

as required by Definition 1. Intuitively, the existence of a scale parameter σ means that precision is characterized by the parameter σ, and so changes in σ have a consistent effect on the shape of the distribution, moving mass towards its tails, thus satisfying the regularity condition.

While regularity guarantees a single cutoff $\widehat{X}$, for general regular distributions we do not know where this cutoff lies. Indeed, Claim 2 in Appendix B.2 shows that, for distributions with a scale parameter, $\widehat{X}$ may lie anywhere between 0 and $\overline{e}$. Proposition 3 shows that, when the distribution is not only regular but also symmetric (as with the normal, uniform, logistic, Cauchy, and Laplace distributions), $\widehat{X}$ lies half-way between 0 and $\overline{e}$, i.e. $\widehat{X} = \frac{\overline{e}}{2}$, as is intuitive. Thus, we can compare the initial strike price, which depends on the underlying parameters of the agency problem (see Lemma 2) to the threshold $\frac{\overline{e}}{2}$. Hence, we can relate whether effort and information are complements or substitutes to model primitives.

[7] A mean-preserving spread, as defined in Rothschild and Stiglitz (1970), is a change in the probability distribution which leaves the mean unchanged, and the probability mass or density is lower in some interval and higher to the left and the right of this interval. Holding the mean constant, SOSD is equivalent to a sequence of mean-preserving spreads (e.g., Gollier (2001, p.44)).

Proposition 3 *(Symmetric regular distributions): Suppose that the noise distribution G_θ is regular and symmetric. Then,* $\frac{dX_\theta}{d\theta} \geq 0$ *if* $X_\theta < \widehat{X}$*, and* $\frac{dX_\theta}{d\theta} \leq 0$ *if* $X_\theta > \widehat{X}$*, where* $\widehat{X} \equiv \frac{\bar{e}}{2}$*.*

In addition to being sufficient for regularity, the presence of a scale parameter also clarifies the intuition because we can fully parametrize changes in precision by changes in volatility σ. We can thus examine how changes in σ affect the values of the two options using the familiar concept of the option "vega": the sensitivity of its value to σ. The vega of each option depends on its strike price, and thus model primitives. With a scale parameter, equation (14) now becomes

$$\frac{\partial}{\partial \sigma} \{\mathbb{E}[W(q)|\bar{e}] - \mathbb{E}[W(q)|0]\} < 0. \tag{16}$$

The LHS of inequality (16) – the vega of the option-when-working minus the vega of the option-when-shirking – represents the effect of changes in σ on incentives. The vega of an option is always positive, highest for an at-the-money ("ATM") option (see Claim 3 in Appendix B.2[8]), and declines when the option moves either ITM or OTM. Thus, the vega of the option-when-working is highest at $X = \bar{e}$, and so if it has a strike price of $\widehat{X} = \frac{\bar{e}}{2}$, it is ITM by $\frac{\bar{e}}{2}$. The vega of the option-when-shirking is highest at $X = 0$, and so if it has a strike price of $\widehat{X} = \frac{\bar{e}}{2}$, it is OTM by $\frac{\bar{e}}{2}$. Overall, at a strike price of $\widehat{X} = \frac{\bar{e}}{2}$, both options are equally away-from-the-money and have the same vega (see Claim 4 in Appendix B.2), and so effort incentives are independent of σ. We thus have $\frac{dX_\sigma}{d\sigma} = 0$. When $X < \widehat{X}$, the option-when-shirking is closer to ATM, and so it has a higher vega. An increase in σ reduces effort incentives, and so $\frac{dX_\sigma}{d\sigma} < 0$. When $X > \widehat{X}$, the option-when-working is closer to ATM than the option-when-shirking. An increase in σ lowers effort incentives, and so $\frac{dX_\sigma}{d\sigma} > 0$.

Note that our analysis takes an optimal contracting approach, so the slope of the contract is the maximum possible without violating the monotonicity constraint (9). ($W'(q) = 1$ for $q \geq X_\theta$). Thus, the principal changes X_θ to ensure that the IC binds. An alternative approach is to restrict the contract to comprising ATM options, e.g. for accounting or tax reasons, and instead meet the IC by varying the slope of the contract. Appendix B.3 demonstrates an analogous result to Proposition 3 for this case. With ATM options, we have $X = \bar{e} \geq \widehat{X} = \frac{\bar{e}}{2}$ and so effort and information are substitutes.

[8] It is well-known that for lognormal distributions, the vega is highest for ATM options. Claim 3 in Appendix B.2 extends this result to all distributions with a location and scale parameter.

An increase in precision requires the number of options granted to increase to maintain incentive compatibility. This augments the expected wage, just like a decrease in the strike price, and so the total effect of precision 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.

3.2 Normal Distribution

We now demonstrate graphically the direct and incentive effects. We need to assume a specific distribution to enable us to calculate the derivatives, and so we consider the common case of a normal distribution (which is symmetric and regular). Let φ and Φ denote the PDF and CDF of the standard normal distribution, respectively. As shown in Appendix A, the total and direct effects are respectively given by:

$$\frac{d\mathbb{E}\left[W\left(q\right)|\overline{e}\right]}{d\sigma} = \varphi\left(\frac{X_\sigma - \overline{e}}{\sigma}\right) - \left[1 - \Phi\left(\frac{X_\sigma - \overline{e}}{\sigma}\right)\right] \frac{\varphi\left(\frac{X_\sigma - \overline{e}}{\sigma}\right) - \varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right) - \Phi\left(\frac{X_\sigma - \overline{e}}{\sigma}\right)}, \text{ and} \tag{17}$$

$$\frac{\partial\mathbb{E}\left[W\left(q\right)|\overline{e}\right]}{\partial\sigma} = \varphi\left(\frac{X_\theta - \overline{e}}{\sigma}\right). \tag{18}$$

Figure 3 illustrates how these effects change with the severity of the moral hazard problem (parametrized by the cost of effort C). As is standard for graphs of option values, the figure contains the strike price X on the x-axis; since X is strictly decreasing in C (Lemma 2), there is a one-to-one mapping between X and C.

To understand Figure 3, recall from (12) that the total effect is given by $\frac{d\mathbb{E}[W(q)|\overline{e}]}{d\sigma} = \frac{\partial\mathbb{E}[W(q)|\overline{e}]}{\partial\sigma} + \frac{\partial\mathbb{E}[W(q)|\overline{e}]}{\partial X_\theta}\frac{dX_\theta}{d\sigma}$. The direct effect, $\frac{\partial\mathbb{E}[W(q)|\overline{e}]}{\partial\sigma}$, is the vega of the option-when-working. It tends to zero as the strike price approaches either $-\infty$ or ∞, and is greatest when the option is ATM, i.e. $X = 1$.

The incentive effect, $\frac{\partial\mathbb{E}[W(q)|\overline{e}]}{\partial X_\sigma}\frac{dX_\sigma}{d\sigma}$, consists of two components. The first is the change in strike price required to maintain incentive compatibility, $\frac{dX_\sigma}{d\sigma}$. From Proposition 3 and using $\sigma = \frac{1}{\sqrt{\theta}}$, $\frac{dX_\sigma}{d\sigma} > 0$ 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 3 below:

Lemma 3 *(Normal distribution, change in strike price): Suppose ε is normally distributed. Then, the effect of volatility on the strike price is decreasing in the cost of*

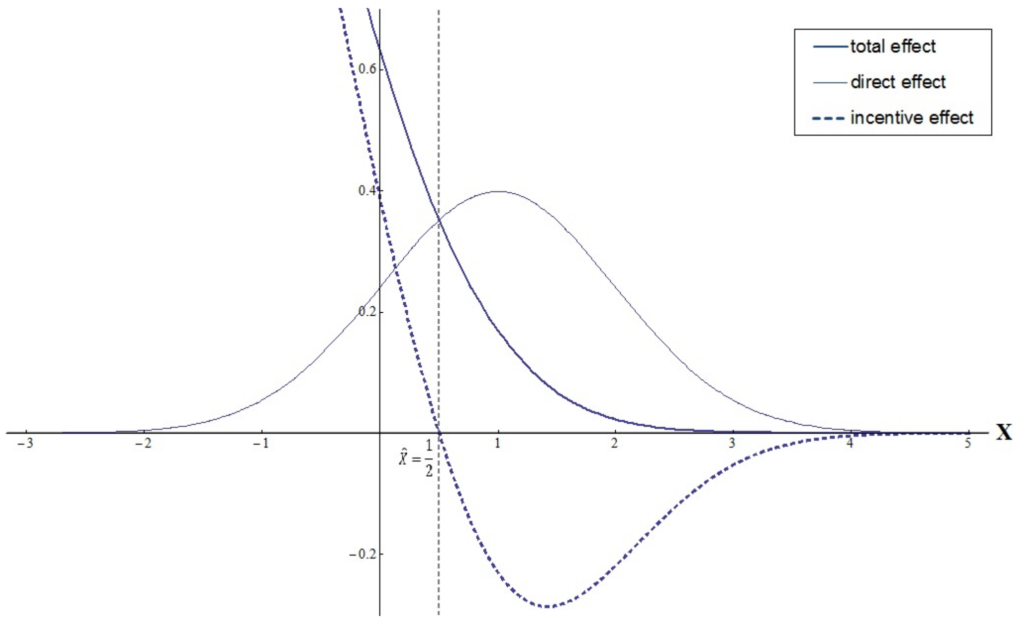

Figure 3: Total and partial derivative of expected pay with respect to σ for a range of values of X, for $\bar{e} = 1$ and $\sigma = 1$.

effort, i.e.

$$\frac{d^2 X_\sigma}{d\sigma dC} < 0. \tag{19}$$

The second component is the change in the value of the option when the strike price increases, $\frac{\partial \mathbb{E}[W(q)|\bar{e}]}{\partial X_\sigma}$. This change is always negative, and its negativity is increasing in the moneyness of the option. Overall, as X falls below $\widehat{X}$ and the option becomes increasingly in the money, both $\frac{dX_\sigma}{d\sigma}$ and $\frac{\partial \mathbb{E}[W(q)|\bar{e}]}{\partial X_\sigma}$ become increasingly negative, and so the overall incentive effect $\frac{\partial \mathbb{E}[W(q)|\bar{e}]}{\partial X_\sigma} \frac{dX_\sigma}{d\sigma}$ becomes monotonically more positive. However, as X rises above $\widehat{X}$, the two components of the incentive effect move in opposite directions. On the one hand, greater precision increasingly worsens incentives ($\frac{dX_\sigma}{d\sigma}$ becomes more positive). On the other hand, $\frac{\partial \mathbb{E}[W(q)|\bar{e}]}{\partial X_\sigma}$ rises towards zero: when the option is deeply OTM, its value is small to begin with and thus little affected by the strike price. Overall, the impact of X on incentives is non-monotonic. As X initially rises above $\widehat{X}$, the incentive effect becomes increasingly negative but subsequently rises to zero.

The total effect $\frac{d\mathbb{E}[W(q)|\bar{e}]}{d\sigma}$ 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 3, the total gains from increased precision are monotonically decreasing in X. Indeed, Proposition 4 proves for the normal distribution that the value of information is monotonically increasing in C (the exogenous parameter that drives X).

Proposition 4 *(Normal distribution, effect of cost of effort on value of information) Suppose ε is normally distributed. Then,* $\frac{d}{dC}\left\{\frac{d\mathbb{E}[W(q)|\bar{e}]}{d\sigma}\right\} > 0$.

An analysis focusing purely on the direct effect would suggest that the value of information is 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 shows that the value of information is monotonically increasing in the severity of the agency problem.

3.3 Applications

We now discuss applications of our results, starting with employment contracts. First, our results highlight the conditions under which employers should invest in increasing the precision with which they monitor their managers' performance. Our analysis suggests that investing in precision is worthwhile where agency problems are strong, and thus incentives are high-powered to begin with. Thus, managers with high-powered incentives (such as CEOs) should be evaluated more precisely than those with low-powered incentives (such as rank-and-file managers). Relatedly, the model has implications for the optimality of pay-for-luck. The results suggest that RPE need not be optimal, as it can reduce effort incentives. This effect is particularly likely where incentives are low-powered to begin with, consistent with RPE being even rarer for rank-and-file employees than for executives.

Second, Proposition 2 suggests that exogenous changes in volatility (see Gormley, Matsa, and Milbourn (2013) and De Angelis, Grullon, and Michenaud (2015) for natural experiments) will have different effects on the incentives of agents depending on the moneyness of their outstanding options. In particular, increases in precision will lower (raise) the incentives of agents with OTM (ITM) options. Thus, firms may wish to reduce the strike prices of OTM 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.[9]

Third, the model has implications for how precision affects the frequency of firing decisions. We can think of X_θ as a performance target below which the manager is fired, and above which he is given a linear contract such as a piece rate. Simple intuition would suggest that more precise monitoring will always increase the probability of firing. However, our results show that this is not the case. If the agency problem is weak to begin with (the target X_θ is initially high), more precise monitoring reduces the agent's effort incentives. To preserve incentive compatibility, the target threshold must be lowered, reducing the probability of firing. Against that, when the threshold is initially high, the agent only avoids being fired upon high performance realizations. A reduction in signal volatility lowers the likelihood of high realizations. Thus, for general distributions, the overall effect on firing probability is ambiguous, contrary to intuition. Moreover, Claim 1 shows that, under certain conditions, firing probability unambiguously *falls* with precision.

Claim 1 *Suppose that the noise distribution G_θ is regular and symmetric. Then, for increases in θ that increase precision in the sense of mean-preserving spreads, we have $\frac{dPr(q \leq X_\theta)}{d\theta} \leq 0$ for $X_\theta \in [\frac{\bar{e}}{2}, \bar{e}]$.*

Fourth, Proposition 3 implies that, for all symmetric regular distributions, improvements in precision draw 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 (see Appendix B.4), where the principal wishes to implement $\bar{e}$, the contract must induce the agent to exert $\bar{e}$ rather than $\bar{e} + \varepsilon$ or $\bar{e} - \varepsilon$, i.e. the IC must be "local". In our discrete model, a local IC resembles the case in which the high effort level ($\bar{e}$) is very close to the low effort level (0). Thus, improvements in precision (e.g. increases in stock market efficiency) will draw X towards $\widehat{X} = \frac{\bar{e}}{2} \simeq 0$. Moreover, since the contract implements effort $\bar{e}$, 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, increases in precision bring the option closer to ATM. Bebchuk and Fried (2004) argue that the almost universal practice of granting ATM options is inefficient and advocate OTM options as they reward the agent only for exceptional performance (see also Rappaport (1999)). However, such views ignore the incentive effect: OTM options have lower deltas and so more would be required to achieve incentive compatibility.

[9] 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.

Murphy (2002) notes that ITM options would provide the strongest incentives, but are discouraged by the tax code. One interpretation is that the tax code leads to firms choosing ATM options when ITM options may be more efficient. Our analysis instead suggests that increases in precision lead to options optimally being close to ATM.[10]

A second application is to financing contract, where the principal (investor) receives debt with face value of X_θ, and the entrepreneur holds equity – a call option on firm value with a strike price equal to the face value of debt. Our results shed light on the investor's incentives to reduce output volatility via risk management.[11] Standard intuition would suggest that these incentives are increasing in the size of her debt claim, and thus her value-at-risk, but this intuition ignores the effect of risk management on effort. If the initial agency problem is strong ($X_\theta < \widehat{X}$), the face value X_θ must be low to induce effort. A fall in output volatility raises effort incentives, and allows the investor to request a higher face value while preserving incentive compatibility.[12] This reinforces the direct effect of risk management, that it increases the value of the investor's risky debt due to its concave payoff structure. If the initial agency problem is weak ($X_\theta > \widehat{X}$), risk management reduces effort incentives, offsetting the direct effect. Indeed, if Proposition 4 holds, the value of information is monotonically increasing in the severity of the agency problem and thus decreasing in the face value of debt, opposite to conventional wisdom.

Our results also have implications for a firm's ability to raise financing. With a strong agency problem, the entrepreneur needs sufficient incentives to exert effort, which requires him to retain a high share of output and thus results in low pledgeable income and a low face value of debt (Innes (1990), Holmstrom and Tirole (1997), and Tirole (2006)). Our results show that there is a second effect of a strong agency problem – it increases the investor's incentives to engage in risk management, which increases

[10]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 and the proportions of his wealth in stock and 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.

[11]Since debt contracts are on output, rather than on a separate signal, for the debt application we take the literal interpretation of changes in θ as affecting the volatility of output (cf. footnote 4.)

[12]This result does not rely on the assumption that the investor has all the bargaining power (see footnote 2 in Innes (1990)). For example, consider the same model except for the assumption that the entrepreneur (rather than investor) has full bargaining power. For each θ, let x_θ be the maximum face value of debt satisfying incentive compatibility (i.e., the entrepreneur's "pledgeable income"). Then $x_\theta = X_\theta$ for any θ, so that it remains the case that, for $x_\theta < \widehat{X}$, higher signal precision increases effort incentives and thus pledgeable income.

pledgeable income and thus the face value of debt. Thus, taking into account the endogenous change in risk management incentives mitigates the effect of the agency problem on the firm's pledgeable income.

4 Conclusion

This paper studies the value of information to the principal in a contracting setting. By taking an optimal contracting approach, we can be specific on how the contract changes in response to increases in precision. This allows us to relate the value of information to the underlying parameters of the agency problem and identify settings in which the value of information is greatest. Under the standard setting of risk neutrality and limited liability, the agent has an option (or, alternatively, the principal holds risky debt). The direct effect of higher signal precision is that it reduces the value of the option and thus expected pay. The focus of the paper is on the indirect effect – we show changes in precision affect the agent's effort incentives and solve precisely for how the contract changes in response.

If effort and information are substitutes, increases in precision weaken incentives. Thus, the principal must reduce the strike price to preserve incentives, increasing the cost of compensation and offsetting the direct effect. Our key result is that we relate whether effort and information are substitutes or complements to the initial strike price of the option, and thus the severity of the agency problem. When the initial strike price is above a threshold, i.e. incentives are weak to begin with, an increase in precision reduces effort incentives. The principal therefore optimally invests less in information. In contrast, if the initial strike price is below a second (lower) threshold, i.e. incentives are strong to begin with, an increase in precision raises effort incentives. This provides an additional gain over and above the direct effect of reducing volatility traditionally focused upon. Thus, the value of information depends on the initial strike price, and thus the severity of the underlying agency problem. For regular signal distributions, such as those with a scale parameter, both thresholds coincide at a single point. Improvements in precision move the strike price towards this point: they lower it if it is initially high, and raise it if initially low.

In an employment setting, our results have implications for the situations in which informativeness is most valuable, for how firms should recontract in response to changes in the informativeness of the performance measure, for how volatility affects the frequency of firing decisions, and for the optimality of at-the-money options. In a financing setting, they have implications for the value of risk management and for how the firm's ability to raise external finance depends on the severity of the agency problem.

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] Aghion, Philippe and Patrick Bolton (1992): "An Incomplete Contracts Approach to Financial Contracting." *Review of Economic Studies* 59, 473–494.

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

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

[6] Biais, Bruno, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve (2010): "Large risks, limited liability and dynamic moral hazard." *Econometrica* 78, 73–118.

[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] Chaigneau, Pierre, Alex Edmans and Daniel Gottlieb (2016): "The informativeness principle under limited liability." NBER Working Paper No. 20456.

[10] Clementi, Gian Luca and Hugo A. Hopenhayn (2006): "A theory of financing constraints and firm dynamics." *Quarterly Journal of Economics* 121, 229–265.

[11] De Angelis, David, Gustavo Grullon, and Sébastien Michenaud (2015): "The effects of short-selling threats on incentive contracts: Evidence from a natural experiment." Working Paper, Rice University.

[12] DeMarzo, Peter M. and Michael J. Fishman (2007a): "Agency and optimal investment dynamics." *Review of Financial Studies* 20, 151–188.

[13] DeMarzo, Peter M. and Michael J. Fishman (2007b): "Optimal long-term financial contracting." *Review of Financial Studies* 20, 2079–2128.

[14] DeMarzo, Peter M. and Yuliy Sannikov (2006): "Optimal security design and dynamic capital structure in a continuous-time agency model." *Journal of Finance* 61, 2681–2726.

[15] Dewatripont, Mathias and Jean Tirole (1994): "A theory of debt and equity: Diversity of securities and manager-shareholder congruence." *Quarterly Journal of Economics* 109, 1027–1054.

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

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

[18] Edmans, Alex and Xavier Gabaix (2016): "Executive compensation: A modern primer." *Journal of Economic Literature*, forthcoming.

[19] Gollier, Christian (2001): *The Economics of risk and time.* (MIT Press, Cambridge).

[20] Gong, Guojin, Laura Yue Li, and Jae Yong Shin (2011): "Relative performance evaluation and related peer groups in executive compensation contracts." *The Accounting Review* 86, 1007–1043.

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

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

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

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

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

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

[27] Jensen, Michael C., and William Meckling (1976): "Theory of the firm: Managerial behavior, agency costs, and capital structure." *Journal of Financial Economics* 3, 305–360.

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

[29] Manso, Gustavo, Bruno Strulovici, and Alexei Tchistyi (2010): "Performance-sensitive debt." *Review of Financial Studies* 23, 1819–1854.

[30] Matthews, Steven A. (2001): "Renegotiating moral hazard contracts under limited liability and monotonicity." *Journal of Economic Theory* 97, 1–29.

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

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

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

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

[35] Robb, Alicia M., and David T. Robinson (2014): "The capital structure decisions of new firms." *Review of Financial Studies* 27, 153-179.

[36] Rothschild, Michael and Joseph E. Stiglitz. (1970): "Increasing risk: I. A definition." *Journal of Economic Theory* 2, 225–243.

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

[38] Stulz, René M. (1996): "Rethinking risk management." *Journal of Applied Corporate Finance* 9, 8–25.

[39] Tirole, Jean (2006): *The theory of corporate finance.* (Princeton University Press, Princeton).

A Proofs

Proof of Lemma 1

The principal's problem is to design a contract that minimizes expected pay:

$$\min_{W_\theta(\cdot)} \int_{-\infty}^{\infty} W_\theta\left(q\right) f_\theta\left(q|\bar{e}\right) dq, \tag{20}$$

subject to the monotonicity constraint, the agent's LL, and the following IC:

$$\int_{-\infty}^{\infty} W_\theta\left(q\right) \left[f_\theta\left(q|\bar{e}\right) - f_\theta\left(q|0\right)\right] dq \geq C. \tag{21}$$

The proof adopts Lemma 1 from Matthews (2001) to a setting with a continuum of signals and general supports. For a given θ, the agent's payoff from an option contract is $W_\theta^*(q) \equiv \max\{q - X, 0\}$, where X is the exercise price. Note that this contract satisfies LL and monotonicity.

Let $\widehat{W}\left(\cdot\right)$ be a contract that satisfies monotonicity, LL, and IC, but differs from an option contract for signals that have positive probability under effort $\bar{e}$, is not an option contract. Without loss of generality, suppose IC holds with equality (otherwise, $\widehat{W}\left(\cdot\right)$ cannot be an optimal contract):

$$\int_{-\infty}^{\infty} \widehat{W}\left(q\right) \left[f_\theta\left(q|\bar{e}\right) - f_\theta\left(q|0\right)\right] dq = C. \tag{22}$$

For any such alternative contract, there exists a unique option contract with the same expected payment, i.e.,

$$\int_{-\infty}^{\infty} W_\theta^*(q) f_\theta\left(q|\bar{e}\right) dq = \int_{-\infty}^{\infty} \widehat{W}(q) f_\theta\left(q|\bar{e}\right) dq. \tag{23}$$

To see this, use the formula for the option contract to write

$$\int_{-\infty}^{\infty} W_\theta^*(q) f_\theta\left(q|\bar{e}\right) dq = \int_{X}^{\infty} \left(q - X\right) f_\theta\left(q|\bar{e}\right) dq,$$

so that (23) can be rewritten as

$$\int_{X}^{\infty} \left(q - X\right) f_\theta\left(q|\bar{e}\right) dq = \int_{-\infty}^{\infty} \widehat{W}(q) f_\theta\left(q|\bar{e}\right) dq. \tag{24}$$

Notice that the derivative of the LHS of (24) with respect to the exercise price is $-\left[1-F_\theta\left(X|\overline{e}\right)\right]$. As $X\nearrow+\infty$, the LHS of (24) converges to $0<C$, so that the LHS is smaller than the RHS because of (22) and $\widehat{W}\left(q\right)\geq 0\ \forall q$ due to LL. As $X\searrow-\infty$, the LHS of (24) converges to $\lim_{X\to-\infty}\int_X^\infty (q-X)f_\theta\left(q|\overline{e}\right)dq$, which is the maximum payment that the agent can receive subject to the monotonicity constraint, and is therefore larger than the RHS of (24). Thus, the Intermediate Value Theorem and the monotonicity of the LHS ensure that a unique solution X to (24) exists.

First, we show that the incentives to exert low effort are higher with the alternative contract than with the option contract:

$$\int_{-\infty}^{\infty}\widehat{W}(q)f_\theta\left(q|0\right)dq>\int_{-\infty}^{\infty}W_\theta^*(q)f_\theta\left(q|0\right)dq. \tag{25}$$

Let $V(q)\equiv\widehat{W}(q)-W_\theta^*(q)$ and note that $V(q)\neq 0$ in a set of positive measure under high effort. By construction, V has mean zero under high effort; thus, $V(q)>0$ for some signals q and $V(q)<0$ for other signals. Let $k\equiv\sup\left\{q:\ V(q)>0\right\}$, so that $V(q)\leq 0$ for all $q>k$. Since W_θ^* is an option contract and $\widehat{W}$ satisfies LL and monotonicity, it follows that $V(q)\geq 0$ for all $q<k$.

Recall that, by MLRP, $\frac{f_\theta(q_1|0)}{f_\theta(q_1|\overline{e})}<\frac{f_\theta(q_0|0)}{f_\theta(q_0|\overline{e})}$ whenever $q_1>q_0$. Then,

$$\begin{aligned}
\int_{-\infty}^{\infty}V(q)f_\theta\left(q|0\right)dq &= \int_{-\infty}^{\infty}V(q)\frac{f_\theta\left(q|0\right)}{f_\theta\left(q|\overline{e}\right)}f_\theta\left(q|\overline{e}\right)dq\\
&= \int_{-\infty}^{k}V(q)\frac{f_\theta\left(q|0\right)}{f_\theta\left(q|\overline{e}\right)}f_\theta\left(q|\overline{e}\right)dq+\int_{k}^{\infty}V(q)\frac{f_\theta\left(q|0\right)}{f_\theta\left(q|\overline{e}\right)}f_\theta\left(q|\overline{e}\right)dq\\
&> \int_{-\infty}^{k}V(q)\frac{f_\theta\left(k|0\right)}{f_\theta\left(k|\overline{e}\right)}f_\theta\left(q|\overline{e}\right)dq+\int_{k}^{\infty}V(q)\frac{f_\theta\left(k|0\right)}{f_\theta\left(k|\overline{e}\right)}f_\theta\left(q|\overline{e}\right)dq\\
&= \frac{f_\theta\left(k|\overline{e}\right)}{f_\theta\left(k|0\right)}\int_{-\infty}^{\infty}V(q)f_\theta\left(q|\overline{e}\right)dq && (26)\\
&= 0, && (27)
\end{aligned}$$

where the first line multiplies and divides by $f_\theta\left(q|\overline{e}\right)$; the second line splits the integral between $q<k$ and $q>k$; the third line uses MLRP, $k>q$, and the fact that $V(q)\leq(\geq)0$ if $q>(<)k$, where each inequality is strict in a set of positive probability measure under high effort; the fourth line takes $\frac{f_\theta(k|\overline{e})}{f_\theta(k|0)}$ outside of the integral; and the last line uses $\frac{f_\theta(k|\overline{e})}{f_\theta(k|0)}>0$ and the fact that, by equation (23), $\int_{-\infty}^{\infty}V(q)f_\theta\left(q|\overline{e}\right)dq=0$.

We have thus established that (25) holds.

Since both contracts pay the same expected amounts under $\bar{e}$ and the option contract pays less under zero effort, it follows that the IC does not bind:

$$\int_{-\infty}^{\infty} W_\theta^*(q) f_\theta(q|\bar{e})\, dq = \int_{-\infty}^{\infty} \widehat{W}(q) f_\theta(q|\bar{e})\, dq \tag{28}$$

$$= \int_{-\infty}^{\infty} \widehat{W}(q) f_\theta(q|0)\, dq + C > \int_{-\infty}^{\infty} W_\theta^*(q) f_\theta(q|0)\, dq + C. \tag{29}$$

Therefore, there exists a small enough increase in the exercise price X such that the new contract, denoted W_θ^{*+}, remains incentive compatible but has a lower expected payment:

$$\int_{-\infty}^{\infty} W_\theta^{*+}(q) f_\theta(q|\bar{e})\, dq < \int_{-\infty}^{\infty} W_\theta^*(q) f_\theta(q|\bar{e})\, dq. \tag{30}$$

Thus, this new option contract W_θ^{*+} satisfies monotonicity, LL, and IC, and has a lower expected cost than the initial non-option contract $\widehat{W}$. Uniqueness of the exercise price follows from the fact that the IC must bind and the existence of a unique exercise price that makes the IC hold as an equality. Thus, the optimal contract is an option contract with an exercise price X_θ that is the unique solution of (6). Moreover, any other optimal contract coincides with this option with probability 1 (under effort $\bar{e}$).

Proof of Lemma 2

Denoting the lower bound of the support of q by $\underline{q}$ and the upper bound by $\overline{q}$, we first show that the IC (5) can also be rewritten as

$$\int_{X_\theta}^{\overline{q}} \left[F_\theta(q|0) - F_\theta(q|\bar{e})\right] dq = C. \tag{31}$$

Opening the expressions inside the brackets in equation (5), we obtain

$$\int_{X_\theta}^{\overline{q}} q f_\theta(q|\bar{e})\, dq - \int_{X_\theta}^{\overline{q}} q f_\theta(q|0)\, dq = \left[F_\theta(X_\theta|0) - F_\theta(X_\theta|\bar{e})\right] X_\theta + C. \tag{32}$$

Integration by parts (for $e \in \{0, \bar{e}\}$) yields:

$$\int_{X_\theta}^{\overline{q}} q f_\theta(q|e)\, dq = \left[q F_\theta(q|e) - \int F_\theta(q|e)\, dq\right]_{X_\theta}^{\overline{q}} = \overline{q} - X_\theta F_\theta(X_\theta|e) - \int_{X_\theta}^{\overline{q}} F_\theta(q|e)\, dq.$$

Substituting into (32) yields:

$$\left[\overline{q} - X_\theta F_\theta \left(X_\theta | \overline{e}\right) - \int_{X_\theta}^{\overline{q}} F_\theta \left(q | \overline{e}\right) dq\right] - \left[\overline{q} - X_\theta F_\theta \left(X_\theta | 0\right) - \int_{X_\theta}^{\overline{q}} F_\theta \left(q | 0\right) dq\right]$$

$$= \left[F_\theta \left(X_\theta | 0\right) - F_\theta \left(X_\theta | \overline{e}\right)\right] X_\theta + C.$$

Canceling terms gives equation (31). Applying the implicit function theorem to (31) yields:

$$\frac{dX_\theta}{dC} = -\frac{1}{F\left(X_\theta | 0\right) - F\left(X_\theta | \overline{e}\right)} < 0. \tag{33}$$

Proof of Proposition 1

It is helpful to start by rewriting the value of the option. Integration by parts yields

$$\mathbb{E}\left[W\left(q\right) | e\right] = \mathbb{E}\left[q | e\right] - X_\theta + \int_{-\infty}^{X_\theta} F_\theta \left(q | e\right) dq. \tag{34}$$

The area under the CDF for signals below X_θ (the third term) is the value of a put option with a strike price of X_θ:

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

where the last equality follows from integration by parts. Therefore, equation (34) can be interpreted as 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.

To study whether precision and effort are complements, we examine each of the three terms on the RHS of (34). While $\mathbb{E}\left[q | e\right]$ depends on e, it is independent of θ since changes in θ do not affect the mean. 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 $\int_{-\infty}^{X_\theta} F_\theta \left(q | e\right) dq$. This observation leads to the following Lemma:

Lemma 4 *Precision and effort are complements (substitutes) if and only if*

$$\frac{\partial}{\partial \theta} \int_{-\infty}^{X_\theta} \left[F_\theta(q|\overline{e}) - F_\theta(q|0)\right] dq \geq (\leq) 0. \tag{35}$$

We can now present the proof of the Proposition.

Since $F_\theta(q|e) = G_\theta(q-e)$, we have

$$\frac{\partial}{\partial\theta}\int_{-\infty}^{X_\theta}\left[F_\theta(q|\bar{e}) - F_\theta(q|0)\right]dq = \frac{\partial}{\partial\theta}\left\{\int_{-\infty}^{X_\theta} G_\theta(q-\bar{e})\,dq - \int_{-\infty}^{X_\theta} G_\theta(q)\,dq\right\}$$

$$= \frac{\partial}{\partial\theta}\left\{\int_{-\infty}^{X_\theta-\bar{e}} G_\theta(q)\,dq - \int_{-\infty}^{X_\theta} G_\theta(q)\,dq\right\} = \frac{\partial}{\partial\theta}\left\{-\int_{X_\theta-\bar{e}}^{X_\theta} G_\theta(q)\,dq\right\}$$

$$= -\int_{X_\theta-\bar{e}}^{X_\theta}\frac{\partial G_\theta}{\partial\theta}(q)\,dq. \tag{36}$$

Therefore, effort and information are complements if and only if (36) > 0.

Let $\xi(\theta) \equiv \lim_{\epsilon\searrow-\infty} G_\theta(\epsilon) = 0$. Since ξ is differentiable at $-\infty$, it follows that $\xi'(\theta) = 0$. Similarly, $\hat{\xi}(\theta) \equiv \lim_{\epsilon\nearrow\infty} G_\theta(\epsilon) = 1$ and the differentiability of $\hat{\xi}$ at ∞ implies that $\hat{\xi}'(\theta) = 0$. Moreover, it is straightforward to show that SOSD implies[13]

$$\int_{-\infty}^{X_\theta}\frac{\partial G_\theta}{\partial\theta}(q-e)\,dq \leq 0 \tag{37}$$

$\forall\ X_\theta$. Thus, $\frac{\partial G_\theta}{\partial\theta} \leq 0$ for q small enough. As a result, there exists $\widehat{X}_1$ such that $\int_{\widehat{X}_1-\bar{e}}^{\widehat{X}_1}\frac{\partial G_\theta}{\partial\theta}(q)\,dq < 0$. Thus, (36) > 0 and so effort and information are complements.

In addition, $\frac{\partial G_\theta}{\partial\theta} = 0$ for $q\to\infty$. Thus, $\frac{\partial G_\theta}{\partial\theta}$ must eventually turn positive: $\frac{\partial G_\theta}{\partial\theta} \geq 0$ for q large enough. As a result, there exists $\widehat{X}_2$ such that $\int_{\widehat{X}_2-\bar{e}}^{\widehat{X}_2}\frac{\partial G_\theta}{\partial\theta}(q)\,dq > 0$. Thus, (36) < 0 and so effort and information are substitutes. In sum, there exists $\widehat{X}_1$ such that $\frac{dX_\theta}{d\theta} \geq 0$ if $X_\theta < \widehat{X}_1$, and $\widehat{X}_2 \geq \widehat{X}_1$ such that $\frac{dX_\theta}{d\theta} \leq 0$ if $X_\theta > \widehat{X}_2$. However, for $\widehat{X}_1 < X_\theta < \widehat{X}_2$, it is possible for $\frac{\partial G_\theta}{\partial\theta}$ to alternate signs several times, and so we cannot sign (36) .

Proof of Proposition 2

From the definition of regular distributions (Definition 1), $\frac{\partial G_\theta}{\partial\theta}$ alternates signs only

[13] Recall that SOSD requires that for all $\theta' \geq \theta$

$$\int_{-\infty}^{X} G_{\theta'}(s-e)\,ds \leq \int_{-\infty}^{X} G_\theta(s-e)\,ds.$$

Taking the limit as $\theta' \searrow \theta$ gives

$$\int_{-\infty}^{X}\frac{\partial G_\theta}{\partial\theta}(s-e)\,ds \leq 0.$$

once. Furthermore, we know from Proposition 1 that $\frac{\partial G_\theta}{\partial \theta} \leq (\geq) \, 0$ for q small (large) enough. Therefore, there exists $\widehat{X}$ such that $-\int_{X_\theta - \overline{e}}^{X_\theta} \frac{\partial G_\theta}{\partial \theta}(q)\,dq$ is nonnegative for $X_\theta < \widehat{X}$, and nonpositive for $X_\theta > \widehat{X}$. It follows from Proposition 1 that $\frac{dX_\theta}{d\theta} \geq 0$ if $X_\theta < \widehat{X}$ and $\frac{dX_\theta}{d\theta} \leq 0$ if $X_\theta > \widehat{X}$.

Proof of Proposition 3

We know from Proposition 1 that $\frac{dX_\theta}{d\theta} \geq (\leq) 0$ if

$$-\int_{X_\theta - \overline{e}}^{X_\theta} \frac{\partial G_\theta}{\partial \theta}(q)\,dq \geq (\leq)\ 0. \tag{38}$$

If G is regular and symmetric for any θ, then

$$\begin{aligned} G(x) &= 1 - G(-x) \\ \frac{\partial G_\theta}{\partial \theta}(x) &= -\frac{\partial G_\theta}{\partial \theta}(-x) \\ \frac{\partial G_\theta}{\partial \theta}(x) &\geq 0 \ \Leftrightarrow \ x \geq 0. \end{aligned}$$

It follows that, for $X_\theta = \overline{e}/2$, the LHS of equation (38) is

$$\int_{-\overline{e}/2}^{\overline{e}/2} -\frac{\partial G_\theta}{\partial \theta}(q)\,dq = 0. \tag{39}$$

For $X_\theta - \overline{e} \geq 0$,

$$\int_{X_\theta - \overline{e}}^{X_\theta} -\frac{\partial G_\theta}{\partial \theta}(q)\,dq \leq 0, \tag{40}$$

and for $X_\theta \leq 0$,

$$\int_{X_\theta - \overline{e}}^{X_\theta} -\frac{\partial G_\theta}{\partial \theta}(q)\,dq \geq 0. \tag{41}$$

Finally, for $X_\theta \in (0, \overline{e})$,

$$\frac{\partial}{\partial X_\theta}\left\{\int_{X_\theta - \overline{e}}^{X_\theta} -\frac{\partial G_\theta}{\partial \theta}(q)\,dq\right\} = \frac{\partial G_\theta}{\partial \theta}(X_\theta - \overline{e}) - \frac{\partial G_\theta}{\partial \theta}(X_\theta) \leq 0 \tag{42}$$

Combining (39)-(42) shows that $\frac{dX_\theta}{d\theta} \geq 0$ if $X_\theta < \frac{\overline{e}}{2}$, and $\frac{dX_\theta}{d\theta} \leq 0$ if $X_\theta > \frac{\overline{e}}{2}$.

Proof of Equations (17) and (18)

First, with σ instead of θ, the decomposition in (12) can be rewritten as

$$\frac{d}{d\sigma}\mathbb{E}\left[W\left(q\right)|\bar{e}\right]=\underbrace{\frac{\partial}{\partial\sigma}\mathbb{E}\left[W\left(q\right)|\bar{e}\right]}_{\textit{direct effect}}+\underbrace{\frac{\partial}{\partial X_{\sigma}}\mathbb{E}\left[W\left(q\right)|\bar{e}\right]\frac{dX_{\sigma}}{d\sigma}}_{\textit{incentive effect}} \tag{43}$$

Second,

$$\begin{aligned}\frac{\partial\mathbb{E}[W\left(q\right)|e]}{\partial\sigma}&=\frac{\partial}{\partial\sigma}\int_{X_{\sigma}}^{\infty}(q-X_{\sigma})\frac{1}{\sigma}\varphi\left(\frac{q-e}{\sigma}\right)dq=\frac{\partial}{\partial\sigma}\int_{X_{\sigma}-e}^{\infty}\frac{q+e-X_{\sigma}}{\sigma}\varphi\left(\frac{q}{\sigma}\right)dq\\&=\frac{\partial}{\partial\sigma}\int_{X_{\sigma}-e}^{\infty}\frac{q}{\sigma}\varphi\left(\frac{q}{\sigma}\right)dq-(X_{\sigma}-e)\frac{\partial}{\partial\sigma}\int_{X_{\sigma}-e}^{\infty}\frac{1}{\sigma}\varphi\left(\frac{q}{\sigma}\right)dq\\&=\frac{\partial}{\partial\sigma}\left\{\left[-\sigma\varphi\left(\frac{q}{\sigma}\right)\right]_{X_{\sigma}-e}^{\infty}\right\}-(X_{\sigma}-e)\frac{\partial}{\partial\sigma}\left\{1-\Phi\left(\frac{X_{\sigma}-e}{\sigma}\right)\right\}\\&=\varphi\left(\frac{X_{\sigma}-e}{\sigma}\right)-\sigma\frac{X_{\sigma}-e}{\sigma^{2}}\varphi'\left(\frac{X_{\sigma}-e}{\sigma}\right)+(X_{\sigma}-e)\left(-\frac{X_{\sigma}-e}{\sigma^{2}}\right)\varphi\left(\frac{X_{\sigma}-e}{\sigma}\right)\\&=\varphi\left(\frac{X_{\sigma}-e}{\sigma}\right)-\frac{X_{\sigma}-e}{\sigma}\varphi'\left(\frac{X_{\sigma}-e}{\sigma}\right)+\frac{X_{\sigma}-e}{\sigma}\varphi'\left(\frac{X_{\sigma}-e}{\sigma}\right)=\varphi\left(\frac{X_{\sigma}-e}{\sigma}\right)\end{aligned} \tag{44}$$

where the fourth and sixth equalities use $\varphi'(x)=-x\varphi(x)$, and the fifth equality uses $\varphi(x)\rightarrow_{x\rightarrow\infty}0$. This establishes (18). In addition, it follows that

$$\frac{\partial}{\partial\sigma}\left\{\mathbb{E}\left[W\left(q\right)|\bar{e}\right]-\mathbb{E}\left[W\left(q\right)|0\right]\right\}=\varphi\left(\frac{X_{\sigma}-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_{\sigma}}{\sigma}\right). \tag{45}$$

Third,

$$\begin{aligned}\frac{\partial\mathbb{E}[W\left(q\right)|e]}{\partial X_{\sigma}}&=\frac{\partial}{\partial X_{\sigma}}\int_{X_{\sigma}}^{\infty}(q-X_{\sigma})\frac{1}{\sigma}\varphi\left(\frac{q-e}{\sigma}\right)dq\\&=\int_{X_{\sigma}}^{\infty}-\frac{1}{\sigma}\varphi\left(\frac{q-e}{\sigma}\right)dq=-\left(1-\Phi\left(\frac{X_{\sigma}-e}{\sigma}\right)\right).\end{aligned} \tag{46}$$

It follows that

$$\begin{aligned}\frac{\partial}{\partial X_{\sigma}}\left\{\mathbb{E}\left[W\left(q\right)|\bar{e}\right]-\mathbb{E}\left[W\left(q\right)|0\right]\right\}&=-\left(1-\Phi\left(\frac{X_{\sigma}-\bar{e}}{\sigma}\right)\right)+\left(1-\Phi\left(\frac{X_{\sigma}}{\sigma}\right)\right)\\&=\Phi\left(\frac{X_{\sigma}-\bar{e}}{\sigma}\right)-\Phi\left(\frac{X_{\sigma}}{\sigma}\right).\end{aligned} \tag{47}$$

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

Fourth, according to Lemma 1, following a change in σ the exercise price X_σ adjusts so that the IC remains satisfied as an equality, so:

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

Rearranging and using the results in equations (45) and (47):

$$\frac{dX_\sigma}{d\sigma}=-\frac{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\Phi\left(\frac{X_\sigma}{\sigma}\right)}. \tag{48}$$

Using the results above, we can rewrite (43) as

$$\frac{d\mathbb{E}[W\left(q\right)|\bar{e}]}{d\sigma}=\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)+\left[1-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right]\frac{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\Phi\left(\frac{X_\sigma}{\sigma}\right)} \tag{49}$$

This establishes (17).

Proof of Lemma 3

As X_σ is strictly decreasing in C (see Lemma 2), inequality (19) holds if and only if $\frac{d\frac{dX_\sigma}{d\sigma}}{dX}>0$. As established in the proof of equations (17) and (18) 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 $\forall t>0$,

$$f(x,t)\equiv[\varphi(x)-\varphi(x-t)]^2-[\Phi(x)-\Phi(x-t)][\varphi'(x)-\varphi'(x-t)]>0,\quad\forall x, \tag{50}$$

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

We now prove (50) for $x > x_0$. In this interval (omitting the argument t):

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

To prove the latter, we first calculate

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

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

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

Since

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

it is clear that

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

Proof of Proposition 4

Using the chain rule,

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

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

Using (17) and $\varphi'(x) = -x\varphi(x)$ for the normal distribution, we have

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

$$= \frac{1}{\sigma}\left(-\frac{X_\sigma-\overline{e}}{\sigma}\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right) + \left[\frac{X_\sigma-\overline{e}}{\sigma}\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right) - \frac{X_\sigma}{\sigma}\varphi\left(\frac{X_\sigma}{\sigma}\right)\right]\frac{1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}\right.$$
$$+\left[\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right]\frac{\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}$$
$$\left.-\frac{1-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)}{\left(\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)\right)^2}\left(\varphi\left(\frac{X_\sigma-\overline{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)\right)^2\right) \tag{51}$$

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 (51) has the same sign as

$$\left[\frac{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)}-\frac{X_\sigma}{\sigma}\right]\left[\varphi\left(\frac{X_\sigma}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right]-\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma}{\sigma}\right)\right]\right]$$
$$-\frac{\bar{e}}{\sigma}\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma}{\sigma}\right)\right]. \tag{52}$$

Since the last term in (52) is always negative, the expression in (52) is negative if the first line in (52) is negative. We now prove the latter.

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

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

Rearranging, we have

$$\varphi\left(\frac{X_\sigma}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\right]-\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)\left[1-\Phi\left(\frac{X_\sigma}{\sigma}\right)\right]>0 \tag{53}$$

Define

$$d(X_\sigma,\bar{e})\equiv\frac{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)}.$$

If $d(X_\sigma,\bar{e})<\frac{X_\sigma}{\sigma}$, then combining with (53) establishes that (52) is negative, as desired. We now show that $d(X_\sigma,\bar{e})<\frac{X_\sigma}{\sigma}$, by proving first that $d(X_\sigma,\bar{e})\longrightarrow_{\bar{e}\to 0}\frac{X_\sigma}{\sigma}$ and second that $d(X_\sigma,\bar{e})$ is decreasing in $\bar{e}$.

First,

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

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

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

Second,

$$\frac{d}{d\bar{e}}\left\{\frac{\varphi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)-\varphi\left(\frac{X_\sigma}{\sigma}\right)}{\Phi\left(\frac{X_\sigma}{\sigma}\right)-\Phi\left(\frac{X_\sigma-\bar{e}}{\sigma}\right)}\right\}=\frac{d}{d\bar{e}}\left\{\frac{\int_{(X_\sigma-\bar{e})/\sigma}^{X_\sigma/\sigma} q\exp\left\{-\frac{q^2}{2}\right\}dq}{\int_{(X_\sigma-\bar{e})/\sigma}^{X_\sigma/\sigma}\exp\left\{-\frac{q^2}{2}\right\}dq}\right\}$$

$$=\frac{1}{\sigma}\frac{\frac{X_\sigma-\bar{e}}{\sigma}\exp\left\{-\frac{(X_\sigma-\bar{e})^2}{2\sigma^2}\right\}\int_{(X_\sigma-\bar{e})/\sigma}^{X_\sigma/\sigma}\exp\left\{-\frac{q^2}{2}\right\}dq-\exp\left\{-\frac{(X_\sigma-\bar{e})^2}{2\sigma^2}\right\}\int_{(X_\sigma-\bar{e})/\sigma}^{X_\sigma/\sigma}q\exp\left\{-\frac{q^2}{2}\right\}dq}{\left(\int_{(X_\sigma-\bar{e})/\sigma}^{X_\sigma/\sigma}\exp\left\{-\frac{q^2}{2}\right\}dq\right)^2}.$$

This expression has the same sign as

$$\frac{X_\sigma-\bar{e}}{\sigma}\int_{(X_\sigma-\bar{e})/\sigma}^{X_\sigma/\sigma}\exp\left\{-\frac{q^2}{2}\right\}dq-\int_{(X_\sigma-\bar{e})/\sigma}^{X_\sigma/\sigma}q\exp\left\{-\frac{q^2}{2}\right\}dq$$

$$=\int_{(X_\sigma-\bar{e})/\sigma}^{X_\sigma/\sigma}\left[\frac{X_\sigma-\bar{e}}{\sigma}-q\right]\exp\left\{-\frac{q^2}{2}\right\}dq<0.$$

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

Proof of Claim 1

First, with a symmetric regular distribution, we know from Proposition 3 that

$$\frac{dX_\theta}{d\theta}\leq 0 \qquad \text{for } X_\theta\geq\frac{\bar{e}}{2}, \qquad \text{and} \qquad \frac{d}{dX_\theta}\int_{-\infty}^{X_\theta}f_\theta(q|\bar{e})dq=f_\theta(X_\theta|\bar{e})\geq 0. \tag{54}$$

Second, for $X_\theta<\bar{e}$, we have

$$\int_{-\infty}^{\bar{e}}f_\theta(q|\bar{e})dq=\frac{1}{2}\qquad\forall\theta\qquad\Rightarrow\qquad\int_{-\infty}^{\bar{e}}\frac{df_\theta(q|\bar{e})}{d\theta}dq=0$$

In addition, for a symmetric regular distribution, and with changes in θ being mean-preserving spreads, there exists z_a such that

$$\frac{df_\theta(q|\bar{e})}{d\theta}\leq 0 \qquad \forall q<z_a, \tag{55}$$

$$\frac{df_\theta(q|\bar{e})}{d\theta}\geq 0 \qquad \forall q\in[z_a,\bar{e}]. \tag{56}$$

This implies that

$$\int_{-\infty}^{X_\theta}\frac{df_\theta(q|\bar{e})}{d\theta}dq\leq 0\qquad\forall X_\theta\leq\bar{e}.$$

In equilibrium,

$$Pr(q \leq X_\theta) = \int_{-\infty}^{X_\theta} f_\theta(q|\bar{e})dq.$$

Furthermore,

$$\frac{d}{d\theta}\int_{-\infty}^{X_\theta} f_\theta(q|\bar{e})dq = \frac{d}{dX_\theta}\int_{-\infty}^{X_\theta} f_\theta(q|\bar{e})dq\frac{dX_\theta}{d\theta} + \int_{-\infty}^{X_\theta} \frac{df_\theta(q|\bar{e})}{d\theta}dq.$$

This gives Claim 1.

B Supplementary Appendix: Not for Publication

B.1 Information Has Zero Value

This section gives an example where the value of information is exactly zero. Denoting the lower bound of the support of q by $\underline{q}$ and the upper bound by $\overline{q}$, from (34) the principal's payoff is

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

where X_θ solves the IC (31):

$$\int_{X_\theta}^{\overline{q}} \left[F_\theta\left(q|0\right) - F_\theta\left(q|\overline{e}\right)\right] dq = C. \tag{57}$$

Let $\underline{q} = 0$ and $\overline{q} = 2$. Suppose that, under low effort, q is uniformly distributed in $[0, 1]$ for any θ:

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

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

Assume $\theta \in \{\theta_L, \theta_H\}$. Under high effort and high precision, q is uniformly distributed in $[0, 2]$:

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

Under high effort and low precision, q has the following density function:

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

Notice that f_L second-order stochastically dominates f_H. Integrating, we obtain the CDF

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

Suppose the parameters are such that $X_\theta \geq 1$. For $q \geq 1$, the CDF are the same under both θ_H and θ_L so that, for $X_\theta \geq 1$, the IC (57) yields:

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

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

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 IC (57) 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 $\int_0^x F_{\theta_H}(q|e)dq = \int_0^x F_{\theta_L}(q|e)dq \ \forall \ x \geq 1.$ [14] Indeed, we can calculate this expression explicitly:

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

Thus, the expected wage is independent of precision: precision has exactly zero value.

B.2 Additional Results for Location-Scale Distributions

Claim 2 states that, if the distribution of q has a location and scale parameter, $\widehat{X} \in (0, \bar{e})$.

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

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

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

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

Claim 2 *Suppose the distribution of* q *belongs to the location-scale family. Then,* $\frac{dX_\theta}{d\theta} \geq 0$ *if* $X_\theta < \widehat{X}$*, and* $\frac{dX_\theta}{d\theta} \leq 0$ *if* $X_\theta > \widehat{X}$*, where* $\widehat{X} \in (0, \overline{e})$*. When* G_θ *is symmetric,* $\widehat{X} \equiv \frac{\overline{e}}{2}$*.*

Proof. Since $F_\theta(q|e) = G\left(\frac{q-e}{\sigma}\right)$, condition (36) from Proposition 1 becomes

$$-\int_{X_\theta - \overline{e}}^{X_\theta} \frac{\partial G\left(\frac{q}{\sigma}\right)}{\partial \theta} dq \geq 0.$$

Using $\sigma = \frac{1}{\sqrt{\theta}}$, this becomes

$$-\int_{X_\theta - \overline{e}}^{X_\theta} \frac{\partial G\left(q\sqrt{\theta}\right)}{\partial \theta} dq \geq 0 \iff -\int_{X_\theta - \overline{e}}^{X_\theta} \frac{q}{2\sqrt{\theta}} g\left(q\sqrt{\theta}\right) dq \geq 0. \tag{58}$$

For a distribution symmetric about its mean of zero, (58) holds if and only if

$$(X_\theta - \overline{e}) + X_\theta \leq 0, \tag{59}$$

that is, if and only if $X_\theta \leq \frac{\overline{e}}{2}$. Since $\frac{dX_\theta}{d\theta} \geq 0$ if (36) > 0 and $\frac{dX_\theta}{d\theta} \leq 0$ if (36) < 0, we conclude that $\frac{dX_\theta}{d\theta} \geq 0$ if $X_\theta \leq \frac{\overline{e}}{2}$, and $\frac{dX_\theta}{d\theta} \leq 0$ if $X_\theta \geq \frac{\overline{e}}{2}$.

Now consider asymmetric distributions. Since $g \geq 0$, (58) holds for $X_\sigma \leq 0$, whereas the LHS of (58) is nonpositive for $X_\sigma \geq \overline{e}$. In addition,

$$\frac{\partial}{\partial X_\theta}\left\{-\int_{X_\theta - \overline{e}}^{X_\theta} \frac{q}{2\sqrt{\theta}} g\left(q\sqrt{\theta}\right) dq\right\} = \frac{X_\theta - \overline{e}}{2\sqrt{\theta}} g\left((X_\theta - \overline{e})\sqrt{\theta}\right) - \frac{X_\theta}{2\sqrt{\theta}} g\left(X_\theta\sqrt{\theta}\right) \tag{60}$$

which is strictly negative for $X_\theta \in (0, \overline{e})$, as both terms on the RHS are negative. We conclude that there exists a unique $\hat{X} \in (0, \overline{e})$ such that condition (36) is satisfied if $X_\theta \leq \hat{X}$, in which case $\frac{dX_\theta}{d\theta} \geq 0$, whereas the LHS of (36) is nonpositive for $X_\theta \geq \overline{e}$, in which case $\frac{dX_\theta}{d\theta} \leq 0$. ■

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 3 shows that this result extends to distributions with location and scale parameters.

Claim 3 *For distributions parametrized with* e *and* σ *such that* $F_\sigma(q|e) = G\left(\frac{q-e}{\sigma}\right)$*, the option vega is highest when* $X_\sigma = e$*.*

Proof. Denoting the lower bound of the support of q by $\underline{q}$ and the upper bound by $\overline{q}$, the agent's expected pay under effort e is given by

$$\mathbb{E}\left[W\left(q\right)|e\right]=\int_{X_{\theta}}^{\overline{q}}\left(q-X_{\theta}\right)f_{\theta}\left(q|e\right)dq. \tag{61}$$

Integration by parts yields:

$$\int_{X_{\theta}}^{\overline{q}}qf_{\theta}\left(q|e\right)dq=q-XF_{\theta}\left(q|e\right)-\int_{X_{\theta}}^{\overline{q}}F_{\theta}\left(q|e\right)dq.$$

Substituting into (61) yields:

$$\begin{aligned}\mathbb{E}\left[W\left(q\right)|e\right]&=\overline{q}-XF_{\theta}\left(q|e\right)-\int_{X_{\theta}}^{\overline{q}}F_{\theta}\left(q|e\right)dq-X\left[1-F_{\theta}\left(q|e\right)\right]\\&=\overline{q}-X-\int_{X_{\theta}}^{\overline{q}}F_{\theta}\left(q|e\right)dq.\end{aligned} \tag{62}$$

The vega of the option is

$$\nu=\frac{\partial}{\partial\sigma}\mathbb{E}\left[W\left(q\right)|e\right]=\frac{\partial}{\partial\sigma}\left\{\overline{q}-X-\int_{X_{\sigma}}^{\overline{q}}F_{\sigma}\left(q|e\right)dq\right\} \tag{63}$$

where we use (62) to derive the second equality. Since $F_{\sigma}\left(q|e\right)=G\left(\frac{q-e}{\sigma}\right)$, we have

$$\nu=\frac{\partial}{\partial\sigma}\left\{-\int_{X_{\sigma}}^{\overline{q}}G\left(\frac{q-e}{\sigma}\right)dq\right\}=\frac{1}{\sigma}\int_{X_{\sigma}}^{\overline{q}}\frac{q-e}{\sigma}g\left(\frac{q-e}{\sigma}\right)dq \tag{64}$$

Using the change of variables $y=\frac{q-e}{\sigma}$ gives

$$\nu=\int_{\frac{X_{\sigma}-e}{\sigma}}^{\frac{\overline{q}-e}{\sigma}}yg(y)dq \tag{65}$$

Since $g(y)>0$, this expression is maximized for $X_{\sigma}=e$, i.e., for an ATM option.[15] ■

Claim 4 shows that, for symmetric distributions with unbounded support, the vegas of the option-when-working and option-when-shirking are equal for $X_{\sigma}=\frac{\overline{e}}{2}$.

[15] 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 $X_{\sigma}=0$.

Claim 4 *For symmetric distributions with unbounded support parametrized by e and σ such that $F_\sigma(q|e) = G\left(\frac{q-e}{\sigma}\right)$, 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 (65) and use the fact that, for a distribution with unbounded support, $\overline{q} = \infty$.

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

$$\nu_{\overline{e}} = \int_{\frac{X_\sigma - \overline{e}}{\sigma}}^{\infty} yg(y)dq = \int_{-\frac{\overline{e}}{2\sigma}}^{\infty} yg(y)dq. \tag{66}$$

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

$$\nu_0 = \int_{\frac{X_\sigma}{\sigma}}^{\infty} yg(y)dq = \int_{\frac{\overline{e}}{2\sigma}}^{\infty} yg(y)dq. \tag{67}$$

In addition,

$$\int_{-\frac{\overline{e}}{2\sigma}}^{\infty} yg(y)dq = \int_{-\frac{\overline{e}}{2\sigma}}^{\frac{\overline{e}}{2\sigma}} yg(y)dq + \int_{\frac{\overline{e}}{2\sigma}}^{\infty} yg(y)dq \tag{68}$$

For a symmetric distribution, we have $\int_{-\frac{\overline{e}}{2\sigma}}^{\frac{\overline{e}}{2\sigma}} yg(y)dq = 0$. Equation (68) then implies that $\nu_{\overline{e}} = \nu_0$. ■

B.3 At-The-Money Options

This Appendix shows that Proposition 3 continues to hold when the principal is restricted to granting ATM options.

As in Proposition 3, we consider symmetric distributions with a location and a scale parameter. However, we now assume that the contract takes the form of ATM options. Considering ATM options requires that we derive 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 $S_0 = e$ ($= \bar{e}$ in equilibrium).

For an ATM option, X is fixed at $S_0 = \bar{e}$, and so the number $n \leq 1$ of ATM options granted adjusts to satisfy the IC.[16] We have the following results:

[16]We only consider the cases such that there exists an incentive compatible contract with ATM

Lemma 5 *(Effect of volatility on number of options) With ATM options,* $\frac{dn}{d\sigma} < 0$*.*

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

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

so that

$$\frac{dn}{d\sigma} = -\frac{\frac{\partial}{\partial\sigma}\left\{\mathbb{E}[W(q)|\bar{e}] - \mathbb{E}[W(q)|0]\right\}}{\frac{\partial}{\partial n}\left\{\mathbb{E}[W(q)|\bar{e}] - \mathbb{E}[W(q)|0]\right\}}. \tag{69}$$

First, if the agent receives n options instead of 1, we have

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

for any given X. With distributions with a location parameter e and scale parameter σ, the numerator of the RHS of (69) is then

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

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

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

where we used the changes of variables $y_L = \frac{q}{\sigma}$ and $y_H = \frac{q-\bar{e}}{\sigma}$. Given the symmetry of g, we have $\int_{\frac{X-\bar{e}}{\sigma}}^{\frac{X}{\sigma}} yg(y)\,dq \geq 0$ if and only if $\frac{X}{\sigma} > -\frac{X-\bar{e}}{\sigma}$, which is always true with ATM options, i.e., with $X = \bar{e}$. We conclude that the numerator of the RHS of (69) is strictly positive.

Second, for an agent who receives n ATM options, the denominator of the RHS of (69) equals

$$\begin{aligned}&\frac{\partial}{\partial n}\left\{\int_X^\infty n(q-X)f(q|\bar{e})dq - \int_X^\infty n(q-X)f(q|0)dq\right\}\\&= \int_X^\infty (q-X)f(q|\bar{e})dq - \int_X^\infty (q-X)f(q|0)dq > 0.\end{aligned} \tag{70}$$

options subject to the constraint $n \leq 1$.

Since both the numerator and the denominator of the RHS of (69) are strictly positive, we have

$$\frac{dn}{d\sigma} < 0. \tag{71}$$

■

With ATM options, $X = \bar{e} > \frac{\bar{e}}{2}$. Thus, an increase in precision (fall in σ) reduces effort incentives, and so n must rise to maintain incentive compatibility.

Corollary 2 compares the partial and total effects of changes in precision on the expected wage.

Corollary 2 *(Partial and total effects of precision on expected wage):*

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

Proof. First,

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

Second,

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

Corollary 2 then follows from this inequality and Lemma 5. ■

With ATM options, the total change in expected pay from a change in precision is smaller than the partial change: while an improvement in precision lowers the value of the agent's options, it also requires that the agent receives more options for incentive compatibility. This incentive effect partially offsets the benefits to the principal.

B.4 Continuous Effort Model

In this section, we present a continuous effort analog of the core model. 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 $\mathbb{E}[W(q)|e] - c\xi(e)$, with $c > 0$, $\xi > 0, \xi' > 0, \xi'' > 0$.

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

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

As in the core model, the principal induces a given level of effort $\bar{e} > 0$. As in Proposition 2, we consider continuously distributed regular distributions with a location parameter, denoted by e. This implies that we can write $q = e + \varepsilon$, where $\mathbb{E}[\varepsilon|e] = 0$.

For a given θ, the principal's problem is to choose a function $W(\cdot)$ to minimize $\mathbb{E}[W(q)|\bar{e}]$ subject to LL, monotonicity, and the following IC:

$$\frac{d}{de}\int_{-\infty}^{\infty} W(q)\, f(q|\bar{e})dq = c\xi'(\bar{e}). \tag{72}$$

Proposition 1 in Innes (1990) implies that, for a given θ, the optimal contract is:

$$W(q) = \max\{0, q - X_\theta\}. \tag{73}$$

As in the core model, there is a unique X_θ that satisfies the IC in (72) with equality. Denoting by $\psi(\cdot|e)$ the PDF of the distribution of q conditional on e, the IC in (72) can be rewritten as

$$\int_{-\infty}^{\infty} W'(q)\psi(q|\bar{e})dq = c\xi'(\bar{e}) \;\Leftrightarrow\; \int_{X}^{\infty} \psi(q|\bar{e})dq = c\xi'(\bar{e}). \tag{74}$$

We consider changes in precision in the sense of a mean-preserving spread ("MPS") of the distribution. Denote by $\bar{\psi}$ the PDF of ε after a decrease in θ, i.e., after a mean-preserving spread of the distribution. By definition of the PDF,

$$\int_{-\infty}^{\infty} \left(\psi(q|\bar{e}) - \bar{\psi}(q|\bar{e})\right) dq = 0. \tag{75}$$

An increase in precision θ in a MPS sense reduces the LHS of the IC in (72) if and only if

$$\int_{X}^{\infty} \left(\psi(q|\bar{e}) - \bar{\psi}(q|\bar{e})\right) dq < 0. \tag{76}$$

Using the definition of a MPS (Rothschild and Stiglitz (1970)), there exists q_a and q_b, with $q_a < q_b$, such that

$$\int_{-\infty}^{q_a} \left(\psi(q|\bar{e}) - \bar{\psi}(q|\bar{e})\right) dq < 0, \tag{77}$$

$$\int_{q_a}^{q_b} \left(\psi(q|\bar{e}) - \bar{\psi}(q|\bar{e})\right) dq > 0, \tag{78}$$

$$\int_{q_b}^{\infty} \left(\psi(q|\bar{e}) - \bar{\psi}(q|\bar{e})\right) dq < 0. \tag{79}$$

These inequalities and (75) imply that there exists $\hat{X} \in (q_a, q_b)$ such that (76) is satisfied if and only if $X > \hat{X}$. In particular, for symmetric distributions, we have $\hat{X} = \bar{e}$ and

$$\int_{\bar{e}}^{\infty} \psi(q|\bar{e})dq = \frac{1}{2} = \int_{\bar{e}}^{\infty} \bar{\psi}(q|\bar{e})dq. \tag{80}$$

As in the core model, the LHS of the IC is strictly decreasing in X. Therefore, for the IC to still be satisfied following a rise in θ with symmetric distributions, we have

$$\frac{dX_\theta}{d\theta} < 0 \qquad \text{if and only if } X_\theta > \bar{e}. \tag{81}$$

Thus, as precision θ increases in a MPS sense, X_θ approaches $\bar{e}$ and the option becomes closer to ATM.

This analysis has held constant the implemented effort level at $\bar{e}$, i.e. solves for the first stage of Grossman and Hart (1983). We thus follow Dittmann, Maug, and Spalt (2013) who study how a specific form of increased precision (indexation) affects the cost of implementing a given effort level; it is well-known that solving also for the optimal effort level is very difficult. Edmans and Gabaix (2011) show that, if the benefits of effort are multiplicative in firm size and the firm is sufficiently large, it is always optimal for the principal to implement the highest effort level and so the optimal effort level is indeed fixed. If this result does not hold, the principal may respond to greater precision by changing the implemented effort level, and so our analysis provides a lower bound for the gains for precision. Where precision has a large (small) effect in reducing the cost of implementing a given effort level, it likely also will have a large (small) effect in changing the optimal effort level. Thus, the situations in which precision has greatest value in implementing a given effort level (the focus of this paper) will also be the situations in which precision has greatest value to the principal overall.