How Should Performance Signals Affect Contracts?

Finance Working Paper N° 439/2014
August 2020

Pierre Chaigneau
Queen's University

Alex Edmans
London Business School, CEPR and ECGI

Daniel Gottlieb
London School of Economics and Political Science

ECGI Working Paper Series in Finance

How Should Performance Signals Affect Contracts?

Working Paper N° 439/2014
August 2020

Pierre Chaigneau
Alex Edmans
Daniel Gottlieb

We thank three anonymous referees, the Editor (Holger Mueller), Sabrina Buti, Alan Douglas, James Dow, Sergei Glebkin, Barney Hartman-Glaser, Steven Matthews, Sebastian Pfeil, Nicolas Sahuguet, Ming Yang, and conference participants at the AFA, Econometric Society World Congress, HEC-McGill Spring FinanceWorkshop, NFA, Rotterdam Executive Compensation Conference, SFS Cavalcade, Université Paris Dauphine, and the University of Waterloo for valuable comments. Gottlieb thanks the Dorinda and Mark Winkelman Distinguished Scholar Award for research support. This paper was previously circulated under the titles "The Informativeness Principle Under Limited Liability" and "The Value of Performance Signals Under Limited Liability".

Abstract

The informativeness principle demonstrates that a contract should depend on all informative signals. This paper studies how such signals should affect the contract, in a general model with continuous effort, risk aversion, and limited liability. The optimal contract pays zero when output falls below a threshold. Surprisingly, "good" signals of performance need not lower the threshold, because a signal is not only individually informative about effort { it also affects the information output provides about effort. A signal that indicates that the output distribution has shifted to the left (e.g. weak industry performance) lowers the threshold; one that indicates that output is a precise measure of effort (e.g. low volatility) lowers high thresholds and increases low thresholds. We apply our model to performance-based vesting and show that performance measures should affect the threshold (i.e. strike price) rather than the number of vesting options, contrary to common practice.

Keywords: informativeness principle, limited liability, option repricing, pay-for-luck, performance-based vesting, performance-sensitive debt

JEL Classifications: D86, G32, G34, J33.

Pierre Chaigneau
Associate Professor of Finance
Queen's University, Smith School of Business
143 Union Street West
Kingston, ON K7L 2P3, Canada
phone:
e-mail: pierre.chaigneau@queensu.ca

Alex Edmans*
Professor of Finance
London Business School, Institute of Finance and Accounting
Regent's Park
London, NW1 4SA, United Kingdom
phone: +44 20 700 082 58
e-mail: aedmans@london.edu

Daniel Gottlieb
Associate Professor of Managerial Economics and Strategy
Department of Management
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
phone: +44 (0)20 7107 5677
e-mail: d.gottlieb@lse.ac.uk

*Corresponding Author

How Should Performance Signals Affect Contracts?*

Pierre Chaigneau
Queen's University

Alex Edmans
LBS, CEPR, and ECGI

Daniel Gottlieb
LSE

August 14, 2020

Abstract

The informativeness principle demonstrates that a contract should depend on all informative signals. This paper studies *how* such signals should affect the contract, in a general model with continuous effort, risk aversion, and limited liability. The optimal contract pays zero when output falls below a threshold. Surprisingly, "good" signals of performance need not lower the threshold, because a signal is not only individually informative about effort – it also affects the information output provides about effort. A signal that indicates that the output distribution has shifted to the left (e.g. weak industry performance) lowers the threshold; one that indicates that output is a precise measure of effort (e.g. low volatility) lowers high thresholds and increases low thresholds. We apply our model to performance-based vesting and show that performance measures should affect the threshold (i.e. strike price) rather than the number of vesting options, contrary to common practice.

Keywords: informativeness principle, limited liability, option repricing, pay-for-luck, performance-based vesting, performance-sensitive debt

JEL Classification: D86, G32, G34, J33.

*pierre.chaigneau@queensu.ca, aedmans@london.edu, d.gottlieb@lse.ac.uk.

We thank three anonymous referees, the Editor (Holger Mueller), Sabrina Buti, Alan Douglas, James Dow, Sergei Glebkin, Barney Hartman-Glaser, Steven Matthews, Sebastian Pfeil, Nicolas Sahuguet, Ming Yang, and conference participants at the AFA, Econometric Society World Congress, HEC-McGill Spring Finance Workshop, NFA, Rotterdam Executive Compensation Conference, SFS Cavalcade, Université Paris Dauphine, and the University of Waterloo for valuable comments. Gottlieb thanks the Dorinda and Mark Winkelman Distinguished Scholar Award for research support. This paper was previously circulated under the titles "The Informativeness Principle Under Limited Liability" and "The Value of Performance Signals Under Limited Liability".

Executive compensation contracts are typically based on multiple signals of performance. For example, Bettis et al. (2018) find that, in 2012, 70% of large U.S. firms paid their executives with performance-vesting equity, where the number of securities granted depends on performance relative to a threshold. 86% of such grants employ at least one accounting threshold, and so their value depends on factors other than the stock price – the standard "output" measure for executive contracts. Murphy's (2013) survey reports that companies use a variety of financial and non-financial performance measures when determining CEO bonuses. Additional performance signals are also used in financing contracts. Manso, Strulovici, and Tchistyi (2010) document that 40% of loans have performance pricing provisions, where the coupon rate depends on signals such as the firm's credit rating, leverage, and solvency ratios. Thus, the payment to investors depends on factors other than cash flow – the standard "output" measure for financing contracts.

The main theoretical justification for including additional performance measures is Holmström's (1979) informativeness principle. This principle states that any signal should be included in a contract if it provides incremental information about the agent's performance, over and above that already conveyed in output. While the principle has been highly influential, it does not show *how* informative signals should be used in contracts. Two questions are critical for principals to be able to incorporate signals into real-life contracts. First, *which* dimensions of the contract should a signal affect? For example, with performance-vesting options, a signal could affect either the strike price, the number of vesting options, or both. In practice, however, only the number of options and not the strike price depend on performance. Is this practice optimal? Second, *which way* should performance affect the relevant dimensions? In practice, "good" performance measures that indicate high effort always increase the number of vesting options. While intuitive, is this practice actually optimal?

This paper addresses these open questions, in a general optimal contracting model with risk aversion and continuous effort. We study how a contract based on output q (such as the stock price) should be modified when the principal also has access to an additional signal s (such as accounting performance). In order to have practical implications for real-life contracts, we extend the original Holmström (1979) model to incorporate limited liability, because this restriction applies to almost all actual contracting settings. Limited liability of equity applies to contracts between entrepreneurs and investors; the wage paid by a firm to a worker cannot be negative. Explicitly modeling this restriction is necessary, otherwise the model might indicate that a signal should affect the contract for outputs where limited liability binds, and so the contract cannot actually respond to the signal. To do so, we develop a new sufficient condition for the validity of the first-order approach under risk aversion and limited liability.

In the most general model, where the only assumption on the output distribution is the monotone likelihood ratio property, the optimal contract involves a threshold output level. The manager receives zero for outputs below the threshold, and a strictly positive amount above it; this amount is increasing in output but will be typically nonlinear. Indeed, many real-life contracts involve a threshold, such as performance-vesting stock or options, bonuses, or dismissal contracts where the agent is fired if output falls below a certain level.

We use our model to understand how the signal realization optimally affects the threshold. It may seem intuitive that signals that indicate managerial effort – such as high accounting performance – should reduce the threshold and thus increase the payment (the "individual informativeness effect"). However, this is not necessarily the case because the signal also has a second effect. Not only may a signal itself indicate high or low effort, but it can also affect the information the principal infers about effort from the output level (the "output inference effect").

To analyze the output inference effect in more detail, we assume that the output distribution has both a location parameter (such as the mean of the distribution) and also a scale parameter (such as the volatility). Many standard output distributions have both parameters, such as the normal, skew-normal, logistic, Cauchy, and uniform distributions. Doing so allows us to model the signal realization as affecting these parameters, and thus study how the threshold varies with these parameters. We can now decompose the output inference effect into two components. The first is the "location effect" – the signal indicates that the location of the output distribution has shifted, i.e. that the entire distribution has moved to the left or to the right. For example, consider the contract for the manager of an industry incumbent, and let s be the number of competitors in that industry. A low number of competitors is individually a good signal of effort, because it indicates that the incumbent has developed good products that have driven out competition or deterred entry. However, this consideration may be outweighed by the location effect – more competitors cause a leftward shift in the distribution of the firm's output, and so a given output is a more positive signal of effort.

The second component of output inference is the "precision effect". The signal also indicates how precise output is as a measure of effort. This precision, in turn, depends on two parameters – the aforementioned scale parameter, which captures the distribution's volatility, and the impact parameter, which captures the extent to which effort increases the distribution's location or mean. For example, a signal that indicates low volatility, or that effort has a high impact on output, suggests high precision.

Unlike the location effect, where the directional impact on the threshold is unambiguous (signals that indicate a leftward shift in the distribution decrease the threshold), signals that indicate high precision may increase or decrease it. The direction of the effect depends on how high the "original" threshold was. If it is high, the manager is paid only if output is sufficiently high to be good news about effort. If output is a precise measure of effort, even moderately high output is a sufficiently strong indicator of effort to justify the manager being paid – the threshold falls. In contrast, if the original output threshold is low, the agent is paid unless output is sufficiently low to be bad news about effort. If output is a precise measure of effort, even moderately low output is a sufficiently weak indicator of effort to justify the manager not being paid – the threshold rises.

Overall, the precision effect leads to less extreme thresholds: low thresholds rise and high thresholds fall. The output threshold for a given signal realization in turn depends on the exogenous parameters of the model. Moreover, contrary to intuition, we also show that the impact and scale parameters do not always have the opposite effect on the threshold.

We next study how the signal affects the optimal pay-performance sensitivity ("PPS") above the threshold. Starting with the individual informativeness effect, a signal that indicates high effort is associated with greater PPS if and only if risk aversion is low. Low risk aversion means that the contract should be more sensitive to output for high likelihood ratios, and thus be convex in (a linear transformation of) the likelihood ratio. If a signal indicates high effort, the likelihood ratio is high, and so pay should be more sensitive to output. The intuition for the location effect is similar. When the signal indicates that the output distribution has shifted to the left, a given output is a more positive indicator of effort, i.e. the likelihood ratio is higher. Thus, PPS should be higher if the optimal contract is convex above the threshold. Finally, if output is a more precise measure of effort, due to either high impact or low volatility, then pay should generally be more sensitive to output. In sum, incentives are concentrated in states of the world (signal realizations) where output is more informative about effort.

We then further specialize the model to the case in which the optimal contract is linear above the threshold. Doing so allows our model to have implications for performance-vesting options – the contract is now an option, where the threshold is the strike price, and the (linear) slope is the number of vesting options. Despite its popularity, we are unaware of any theories that study under what conditions performance-based vesting is optimal, and what performance signals should be used. We start by deriving the first set of sufficient conditions for options to be the optimal contract when the agent is risk-averse – log utility, normally-distributed output, limited liability on the manager, and a sufficiently convex cost of effort.

We show that the effect of a signal on the number of vesting options depends only on the precision effect, and not the individual informativeness or location effects. This is because the number of vesting options determines the slope of the contract (PPS) and the optimal slope of a linear contract depends only on how precise output is as an indicator of effort. This result suggests that the common practice of making vesting depend on signals such as accounting and stock price performance (either for the firm in question or peer firms) may be suboptimal. Such signals are informative about either the manager's effort and/or the location of the output distribution but are unlikely to affect output precision. Thus, they should optimally affect the overall level of pay, not the sensitivity of pay to performance, and change the strike price rather than the number of vesting options. (If they do also affect output precision, they should affect both dimensions).

In practice, option strike prices are sometimes reset, although on a discretionary basis rather than the resetting being specified in the contract. Such resetting typically involves a lowering of the strike price, and follows poor stock price performance (Brenner, Sundaram, and Yermack (2000)). Acharya, John, and Sundaram (2000) theoretically justify such practices based on the need to restore *future* effort incentives when options fall out of the money. This paper shows that contractual repricing of stock options can be optimal to reward or punish *past* effort. Unlike the number of vesting options, which depends only on the precision effect, the strike price depends also on the individual informativeness and location effects. The location effect means that it may sometimes be optimal to lower the strike price upon a signal that individually conveys bad news about effort (such as a high number of competitors), contrary to conventional wisdom that such repricing necessarily results from rent extraction.

We finally apply to the model to a second commonly-observed contract that is linear above the threshold – debt. Here, the threshold corresponds to the face value of debt, and the manager is the residual claimant, receiving equity. It may seem that debt is simply a special case of the option contract where the slope is always 1, but not all of the results continue to apply. While the individual informativeness and location effects continue to hold, the effect of the impact parameter is different. A higher impact parameter raises the optimal PPS. In the option contract, higher PPS is implemented by more vesting options, and so the effect on the threshold (strike price) is ambiguous. In the debt contract, the slope is always 1, and so higher PPS is instead implemented via a lower threshold (face value of debt).

This paper is related to the theoretical literature on pay-for-performance, surveyed by Holmström (2017). A number of papers extend the original Holmström (1979) informativeness principle and thus study *whether* signals should affect contracts, but not *how* signals should be incorporated into the contract. Examples include Townsend (1979), Gjesdal (1982), Gale and Hellwig (1985), Amershi and Hughes (1989), Allen and Gale (1992), Kim (1995), and Chaigneau, Edmans, and Gottlieb (2019). Similarly, the applied literature on pay-for-performance studies questions such as whether agents should be paid for luck – i.e. *whether* signals of peer performance should affect the contract – such as Oyer (2004), Axelson and Baliga (2009), Gopalan, Milbourn, and Song (2010), Hoffmann and Pfeil (2010), and Hartman-Glaser and Hébert (2019)[1] – rather than *how* signals in general (not just signals of peer performance) should affect contracts.

More closely related is Marinovic and Varas (2019), who show that, if the manager can manipulate output, performance-based vesting is always optimal to deter such manipulation. We show that, even if the manager is unable to manipulate the output measure, performance-based vesting is optimal if the signal affects the precision of output as an effort measure. However, it is not optimal if the signal is informative only along other dimensions – many common performance signals should affect the strike price rather than the number of vesting options. Also related is Chaigneau, Edmans, and Gottlieb (2018), who study how volatility (although not individual informativeness, location, or impact) affects the optimal contract. In their setting, there are no additional signals and the contract is written after volatility is realized, so volatility affects the contract that is offered. In contrast, we study how incentives are allocated across signals – i.e. states of nature – that are associated with different volatility. The contract is written before volatility is realized and the contract is contingent on the signal realization, as is the case for performance-vesting equity or performance-sensitive debt.

1 The Model

1.1 Setup

There are two parties, a principal (firm), and an agent (manager). The manager exerts an unobservable effort $e \in [0, \bar{e}]$ and is protected by limited liability. Effort entails any action that improves output but is costly to the manager, such as working rather than shirking, choosing projects that generate cash flows rather than private benefits, or not extracting rents. The manager's cost of effort $C(\cdot)$ is strictly increasing, strictly convex, twice continuously differentiable in $[0, \bar{e})$, with $C'(0) = 0$ and $\lim_{e \nearrow \bar{e}} C'(e) = +\infty$. His utility over money $u(\cdot)$ is strictly increasing, weakly concave, and twice differentiable. The manager has outside wealth $\bar{W} > 0$ and reservation utility $\overline{u}$.

Effort affects the probability distribution of output q and a signal s, which are both observable and contractible. Output is continuously distributed with full support on $(\underline{q}, +\infty)$, where

[1]Dittmann, Maug, and Spalt (2013) calibrate the cost savings from incorporating peer performance in executive contracts, and Johnson and Tian (2000) compare the incentives provided by indexed and non-indexed options.

$\underline{q}$ is either $-\infty$ or 0. To ensure that an optimal contract exists, we assume that the signal is discrete, $s \in \{s_1, ..., s_S\}$. Note that the signal can have one or multiple dimensions.

The signal is distributed according to the probability mass function $\phi_e^s := \Pr(\tilde{s} = s|\tilde{e} = e)$, which is strictly positive and twice continuously differentiable in e. Output is distributed according to the cumulative distribution function $F(q|e,s)$, which is twice continuously differentiable in q and e and has a strictly positive density $f(q|e,s)$. The joint distribution of output and the signal is $f(q,s|e) := \phi_e^s f(q|e,s)$. The likelihood ratio is defined as:

$$LR_s(q|e) := \frac{\frac{\partial f}{\partial e}(q,s|e)}{f(q,s|e)} = \frac{\partial \phi_e^s/\partial e}{\phi_e^s} + \frac{\frac{\partial f}{\partial e}(q|e,s)}{f(q|e,s)}, \tag{1}$$

which we assume to be strictly increasing in output q ("MLRP"). We call $\frac{\partial \phi_e^s/\partial e}{\phi_e^s}$ the likelihood ratio of the signal, and $\frac{\frac{\partial f}{\partial e}(q|e,s)}{f(q|e,s)}$ the likelihood ratio of output (conditional on the signal). For simplicity, for all s we assume that $\lim_{q \nearrow +\infty} LR_s(q|e) = \infty$, and $\lim_{q \searrow -\infty} LR_s(q|e) = -\infty$ when the support is unbounded below.[2]

The firm has full bargaining power and offers the manager a schedule of payments $\{w_s(q)\}$ conditional on each realization of (q,s). We follow Grossman and Hart (1983) and separate the principal's problem into two stages. The first stage determines the cost of implementing each effort. Given this cost, the second stage determines which effort to implement.

The optimal schedule of payments $\{w_s(q)\}$ that implements effort $\hat{e}$ solves:

$$\min_{\{w_s(q)\}} \sum_s \phi_{\hat{e}}^s \int_{\underline{q}}^{+\infty} w_s(q) f(q|\hat{e},s) dq \tag{2}$$

subject to

$$\sum_s \phi_{\hat{e}}^s \int_{\underline{q}}^{+\infty} u\left(\bar{W} + w_s(q)\right) f(q|\hat{e},s) dq - C(\hat{e}) \geq \overline{u}, \tag{3}$$

$$\hat{e} \in \arg\max_e \sum_s \phi_e^s \int_{\underline{q}}^{+\infty} u\left(\bar{W} + w_s(q)\right) f(q|e,s) dq - C(e), \tag{4}$$

$$w_s(q) \geq 0 \ \ \forall q, s. \tag{5}$$

The firm minimizes the expected payment (2) subject to the manager's individual rationality ("IR") (3), incentive compatibility ("IC") (4), and limited liability constraints ("LL") (5). We say that the manager has "worked" if he exerts effort $\hat{e}$. Otherwise, we say that he has "shirked".

[2] This assumption simplifies expressions by ruling out corner solutions, but is not important for any of our results.

1.2 The Optimal Contract

The first-order approach ("FOA") simplifies the analysis of moral hazard models with a continuum of efforts by allowing a continuum of incentive constraints (equation (4)) to be replaced by the local incentive constraint that prevents local deviations.[3] Although several conditions can be used to justify the FOA (e.g., Rogerson (1985), Jewitt (1988)), using the FOA in models with limited liability is not straightforward.

First, some conditions rely on the optimal contract derived in the absence of contracting constraints (Jewitt (1988), Kim and Jung (2015)), and are thus not applicable under limited liability. Second, Rogerson's (1985) condition on the convexity of the cumulative distribution function, which is generalized to the multi-signal case by Sinclair-Desgagné (1994), does not rely on the contract and thus can be used with contracting constraints (e.g., Kadan and Swinkels (2008), Jewitt, Kadan, and Swinkels (2008)). However, it imposes a strong restriction on the probability distribution of the performance measure, which is not satisfied by probability distributions commonly used to model stock returns, such as the normal, gamma, and lognormal (Hemmer (2004)). Third, an alternative approach, used by Dittmann and Maug (2007) in calibration and Kirkegaard (2017), first derives the optimal contract assuming the FOA is valid, and then imposes conditions based on this contract that guarantee the validity of the FOA. Since the optimal contract generally cannot be characterized in closed form, it is hard to obtain analytical conditions for the validity of the FOA.

We present a new condition for the validity of the FOA under limited liability for utility functions that are bounded from above, which is the case for many common utility functions, such as constant absolute risk aversion and constant relative risk aversion ("CRRA") with a relative risk aversion γ larger than 1. The condition uses limited liability to obtain a lower bound on the manager's utility which, combined with the upper bound, allows us to specify conditions that rule out profitable non-local deviations. Crucially, these conditions do not depend on the (endogenous) contract, nor do they require the strong convexity condition in Rogerson (1985).

Let K_e^+ and K_e^- denote the integral of the positive and negative parts of the second deriva-

[3] See Chaigneau, Edmans, and Gottlieb (2019) for the informativeness principle without the FOA.

tive of the joint distribution $f(q,s|e)$ with respect to effort:

$$K_e^+ \;:=\; \sum_s \int_{\underline{q}}^{+\infty} \max\left\{\frac{\partial^2 f}{\partial e^2}(q,s|e),\ 0\right\} dq, \tag{6}$$

$$K_e^- \;:=\; \sum_s \int_{\underline{q}}^{+\infty} \min\left\{\frac{\partial^2 f}{\partial e^2}(q,s|e),\ 0\right\} dq. \tag{7}$$

Lemma 1 *(First-Order Approach): Suppose that*

$$K_e^- u(\bar{W}) + K_e^+ \lim_{c\nearrow\infty} u(c) < C''(e) \tag{8}$$

for all $e \in (0,\bar{e})$. Then, the FOA is valid.

The intuition for the condition in Lemma 1 is as follows. For the FOA to be valid, we must ensure that the manager's objective, expected utility of payments minus effort costs, is a concave function of effort. In general, the effect of effort on the expected utility of payments may be sufficiently locally convex that it offsets the impact of a strictly convex effort cost. We therefore need to provide an upper bound on the convexity of the (endogenous) payments. With limited liability, $u(\bar{W})$ is a lower bound on the manager's utility, and by monotonicity, $\lim_{c\nearrow\infty} u(c)$ is an upper bound on the manager's utility. Lemma 1 uses these bounds to limit the convexity in the manager's expected utility. Appendix B shows how the condition in Lemma 1 can be applied to specific cases, including those without a signal, which may be of value for future research.

Henceforth, we consider contracts that implement $\hat{e} > 0$ (and thus the IC binds); if the principal wishes to implement $\hat{e} = 0$, she trivially offers a flat wage. Let $\lambda \geq 0$ and $\mu > 0$ denote the Lagrange multipliers associated with the IR (3) and IC (4), respectively. The optimal contract is given by Lemma 2 below.

Lemma 2 *(Contract): Suppose* (8) *holds. The optimal contract to implement $\hat{e} > 0$ satisfies:*

$$w_s(q) = \begin{cases} u'^{-1}\left(1\Big/\left(\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right]\right)\right) - \bar{W} & \text{if } \lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right] \geq \frac{1}{u'(\bar{W})} \\ 0 & \text{if } \lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right] < \frac{1}{u'(\bar{W})} \end{cases}. \tag{9}$$

Note that the IR may not bind in the optimal contract, in which case we have $\lambda = 0$. Due to MLRP, the payment is increasing in the likelihood ratio, but cannot fall below zero due to limited liability. Thus, the manager is only paid when the likelihood ratio exceeds a cutoff. For each signal realization s, this cutoff for the likelihood ratio is associated with a threshold output q_s^* (which may be zero if output has support on $[0, \infty)$). If output exceeds the threshold, it is sufficiently likely that the manager has worked that the firm pays him a strictly positive amount. Again due to MLRP, the payment is monotonically increasing in output, but it will typically be nonlinear.

Since the signal realization cannot affect the contract below the threshold, it can affect the contract in two ways: it can affect the threshold, and it can affect the slope of the contract above the threshold. Section 2 now analyzes these two channels in turn.

2 How Performance Signals Affect Contracts

2.1 Performance Signals and the Threshold

This section studies how the signal realization affects the threshold. Recall from equation (1) that the likelihood ratio can be decomposed into two components: the likelihood ratio of the signal ($\frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s}$) and the likelihood ratio of output ($\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$). A signal can therefore affect the likelihood ratio, and thus the threshold, in two ways. First, it can be individually informative about effort (through the likelihood ratio of the signal) – the "individual informativeness effect". Second, it can affect the information the principal infers about effort from output (through the likelihood ratio of output) – the "output inference effect". For example, even if effort does not affect economic conditions, and so economic conditions are uninformative about effort, these conditions will affect the likelihood ratio if a given output level is more indicative of effort in recessions than booms.

To further analyze these effects, we now parametrize the output distribution. This allows us to model the signal realization as affecting the distribution's parameters, and thus study how the threshold varies with these parameters. Specifically, we consider output distributions with a scale parameter σ_s, which can be interpreted as the distribution's volatility, and a location parameter $h_s(e)$ which, for symmetric distributions such as the normal and logistic, is the mean. We assume $h_s'(e) > 0$ for all e (higher effort shifts the distribution rightward). For distributions with location and scale parameters, there exists a function $g(\cdot)$ such that we can rewrite the density as:

$$f(q|e,s) \equiv \frac{1}{\sigma_s} g\left(\frac{q - h_s(e)}{\sigma_s}\right). \tag{10}$$

Without loss of generality, let $h_s(e) = \xi_s + \zeta_s \Upsilon(e)$ and normalize $\Upsilon(\hat{e}) = 0$ and $\Upsilon'(\hat{e}) = 1$, so that $h_s(\hat{e}) = \xi_s$ and $h'_s(\hat{e}) = \zeta_s > 0$. We refer to ξ_s as the equilibrium location parameter, and to ζ_s as the impact parameter that captures the effect of effort on output. Using equation (10), the likelihood ratio of output can be rewritten:

$$\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} = -\frac{\zeta_s}{\sigma_s}\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)}. \tag{11}$$

We assume that the likelihood ratio of output in equation (11) is continuously differentiable in q, σ_s, ξ_s, and ζ_s. An output distribution with location and scale parameters that satisfies MLRP has a single peak (see the proof of Proposition 1), which we denote by q_s^P. As can be seen in equation (11), the likelihood ratio of output is zero at the peak.[4]

Proposition 1 studies how the signal realization affects the threshold q_s^* above which the manager gets paid. It holds "all else equal across signals", i.e. considers two signals s_i and s_j that vary along only one dimension (e.g. the impact parameter ζ_s); all other dimensions are constant.

Proposition 1 *(Effect of signal on threshold): All else equal across signals:*

(i) If $\frac{\partial\phi_{\hat{e}}^{s_i}/\partial e}{\phi_{\hat{e}}^{s_i}} > \frac{\partial\phi_{\hat{e}}^{s_j}/\partial e}{\phi_{\hat{e}}^{s_j}}$, $q_{s_i}^* \leq q_{s_j}^*$.

(ii) If $\xi_{s_i} < \xi_{s_j}$, $q_{s_i}^* \leq q_{s_j}^*$.

(iii) If $\zeta_{s_i} > \zeta_{s_j}$, $q_{s_i}^* \leq q_{s_j}^*$ *if also* $q_s^* > q_s^P$ *for* $s \in \{s_i, s_j\}$, *and* $q_{s_i}^* \geq q_{s_j}^*$ *if also* $q_s^* < q_s^P$ *for* $s \in \{s_i, s_j\}$.

(iv) If $\sigma_{s_i} > \sigma_{s_j}$ *and* $q_s^* < \min\left\{q_s^P, \xi_s\right\}$ $(q_s^* > \max\left\{q_s^P, \xi_s\right\})$ *for* $s \in \{s_i, s_j\}$, $q_{s_i}^* \leq q_{s_j}^*$ $(q_{s_i}^* \geq q_{s_j}^*)$. *For symmetric distributions, we have* $q_s^P = \xi_s$, *so the conditions simplify to* $q_s^* < (>) q_s^P$.

Part (i) is the individual informativeness effect. If $\frac{\partial\phi_{\hat{e}}^{s_i}/\partial e}{\phi_{\hat{e}}^{s_i}} > \frac{\partial\phi_{\hat{e}}^{s_j}/\partial e}{\phi_{\hat{e}}^{s_j}}$, then signal realization s_i is individually more indicative of high effort than s_j. Thus, the threshold should be lower under s_i than s_j. This is the intuitive effect mentioned previously – if the signal is individually good news about effort, a lower minimum output is required for the manager to be paid. For example, signal s could be accounting profits, with s_i representing higher profits than s_j. High profits indicate high effort and thus increase the payment.

It is important to stress that this is not a "comparative statics" result, but unique to an analysis of how the contract depends on the signal realization. We would not be able to obtain

[4] At the peak of the distribution, we have $f'(q|\hat{e},s) = 0$. From equations (10) and (11), this is where the likelihood ratio is zero; by MLRP, this point is unique.

this result by taking the contract in Section 1.2 and changing the underlying parameters such as $\frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s}$. Changes in the underlying parameters will, in general, change the Lagrange multipliers associated with the optimization program in equations (2)-(5) and thus have an ambiguous effect on the contract. Instead, we are holding constant the contracting setting, and showing how the contract depends on signal realizations that are associated with different parameters.

The output inference effect is given by parts (ii)-(iv) and can be decomposed into two components. The first is the "location effect" and captured by part (ii). If s_i is associated with a lower equilibrium location parameter ξ_{s_i} than s_j, then it indicates that the output distribution has shifted to the left. Due to MLRP, this shift means that achieving any given output level is more indicative of working than shirking. Thus, a lower minimum output is required for the manager to be paid.

For example, let s be industry performance where s_i represents a downturn and s_j an upswing. If achieving an output level is more indicative of effort in a downturn, the threshold should be lower – the intuition behind relative performance evaluation. Note that the location effect may counteract the individual informativeness effect. If industry performance is increasing in effort – for example, the manager's effort improves consumers' perception of the entire sector – then, one might think that high industry performance should be associated with a lower threshold, as it individually indicates high effort. However, this may be outweighed by the location effect: achieving a certain output is easier in an upswing. Thus, we may have $q_{s_i}^* \leq q_{s_j}^*$.

Note that Proposition 1 does not study analytically whether $q_{s_i}^* \lesseqgtr q_{s_j}^*$ when signals s_i and s_j vary across multiple dimensions – it contains "ceteris paribus" results. For example, part (i) assumes that the signal realization s affects the ratio $\frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s}$ but not ξ_s. Section 2.3 imposes more structure on the output distribution and allows us to analytically compare thresholds when signals differ across more than one dimension. We can thus study under what conditions the individual informativeness effect will be outweighed by the location effect and so signal realizations that indicate high effort are not associated with lower thresholds and thus higher payments to the manager. Moreover, even without imposing the assumptions of Section 2.3, we can undertake this comparison numerically. Example 1 illustrates.

Example 1 *Suppose that the manager has CRRA utility with relative risk aversion of 2, outside wealth* $\bar{W} = 1$*, and reservation utility* $\bar{u} = -1$*. The cost of effort is characterized by* $C(\hat{e}) = 0.5$ *and* $C'(\hat{e}) = 0.1$*. The signal* s *is the number of industry competitors, which can be high* ($\tilde{s} = s_H$) *or low* ($\tilde{s} = s_L$)*. A low number of competitors* (s_L) *is individually a better signal of effort than a high number* (s_H)*, because it suggests that the manager has driven out*

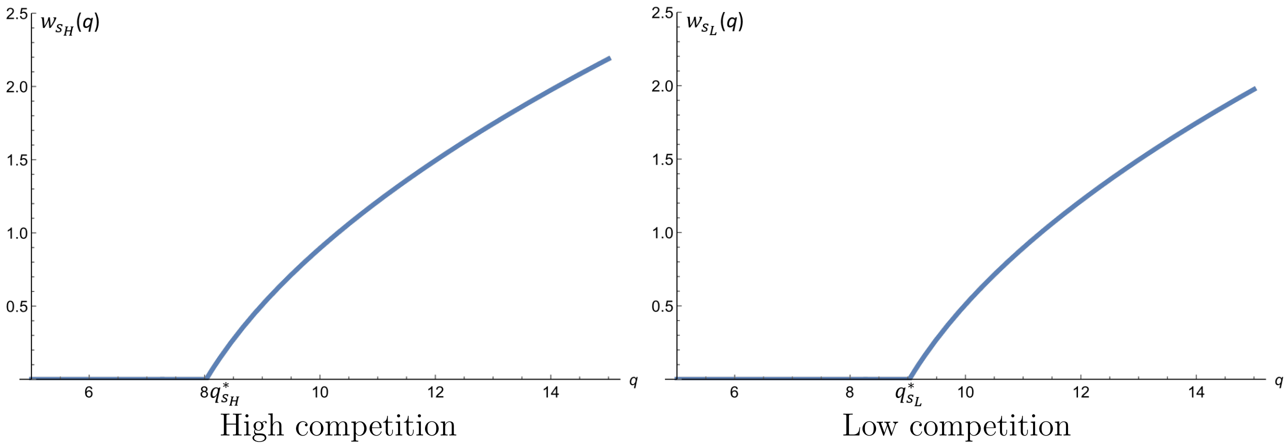

Figure 1: The payoff of the manager as a function of output when the number of competitors is high ($s = s_H$) on the left, and low ($s = s_L$) on the right, in Example 1.

competition or deterred entry. Specifically, the probability of a low (high) number of competitors is $\phi_e^{s_L} = \frac{1+e}{4}$ ($\phi_e^{s_H} = \frac{3-e}{4}$)*, with* $e \in [0,1]$ *and* $\hat{e} = 1$*. This consideration may be outweighed by a second effect – a given level of output is a more positive signal of effort the more competitors there are, since competition causes a downward shift in the location of the output distribution. Specifically, at equilibrium effort* $\hat{e}$*, output is normally distributed with* $\xi_{s_L} = 12$ *and* $\xi_{s_H} = 10$*. For both signals,* $\sigma_s = 1$ *and* $\zeta_s = 1$*.*

The participation constraint is binding, and we find that $q^*_{s_H} \approx 8.03$*,* $q^*_{s_L} \approx 9.03$*. Here, the location effect dominates the individual informativeness effect – even though a low number of competitors individually indicates effort, this is outweighed by the fact that it shifts the distribution of output upward, and so the threshold is higher than with a high number of competitors. The contract is depicted in Figure 1.*

The second component of output inference is the "precision effect". This effect is driven by two parameters of the output distribution: the impact parameter ζ_s (captured by part (iii)) and the scale parameter σ_s (captured by part (iv)). They affect how precise output is as a signal of effort. Unlike the location effect, where the effect of ξ_s is independent of the level of the thresholds, parts (iii) and (iv) show that the precision effect depends critically on the thresholds.

The intuition is as follows. Low σ_s and high ζ_s mean that output is driven more by effort than luck. Thus, they are good news about effort when combined with high output, as they indicate that this high output is likely due to high effort rather than good luck. In contrast, they are bad news when combined with low output, as they indicate that this low output is likely due to low effort rather than bad luck. This contrasts the location effect, where low ξ_s is unambiguously good news about effort.

When the output threshold is high across signal realizations, the manager is paid only if output is sufficiently high to be good news about effort. A high impact of effort on output, or a low volatility of output, mean that output is driven more by effort than luck. Thus, even moderately high output is sufficiently good news about effort to warrant the manager being paid – the threshold falls. Put differently, when the threshold is high, only high output levels are relevant. It does not matter that a low output level is less indicative of effort since, due to the high threshold, the manager receives zero anyway.

When the output threshold is low across signal realizations, the manager is paid unless output is sufficiently low to be bad news about effort. A high impact of effort on output, or a low volatility of output, mean that output is driven more by effort than luck. Thus, even moderately low output is sufficiently bad news about effort to warrant the manager not being paid – the threshold rises. Overall, signals associated with lower σ (such as low market, industry or firm volatility) and/or higher ζ (such as low regulation, high product market fluidity, high industry competition, or high industry disruption)[5] increase the precision of output as a signal of effort and thus lead to less extreme thresholds: low thresholds rise and high thresholds fall. Intuitively, if output is more informative about effort, then less extreme output levels are needed to determine whether the manager should be paid. The output thresholds in turn depends on the exogenous parameters of the model. For example, Appendix C shows that, if the participation constraint is non-binding and so the only goal of the contract is to provide incentives, an increase in the marginal cost of effort lowers the thresholds.

For the impact effect, what determines whether a threshold is "high" or "low" is how it compares with the peak of the distribution. A higher impact parameter corresponds to a steeper likelihood ratio of output: the ratio is more positive above the peak and more negative below the peak; at the peak it is zero and thus unaffected by the impact parameter.[6] Figure 2 illustrates. The solid line corresponds to $\zeta_{s_j} = 1$ and the dotted line to $\zeta_{s_i} = 2$. The left-hand side ("LHS") considers a positive cutoff for the likelihood ratio of output. Thus, the initial threshold $q^*_{s_j}$ is above the peak q^P_s, which is where the ratio is zero and the lines cross the x-axis. A higher impact parameter means that the new likelihood ratio of output at the initial

[5]See RegData's Industry Regulation Index for an example of an industry regulation index, and Hoberg, Phillips, and Prabhala (2013) for an example of a measure of product market fluidity. Merger waves are a potential measure of industry disruption; Harford (2005) finds that they are driven by economic, regulatory, and technological shocks.

[6]Recall that the likelihood ratio of output is given by equation (11). By definition, the (single) peak is such that $g' = 0$, and MLRP ensures that $g' < (>) 0$ below (above) the peak

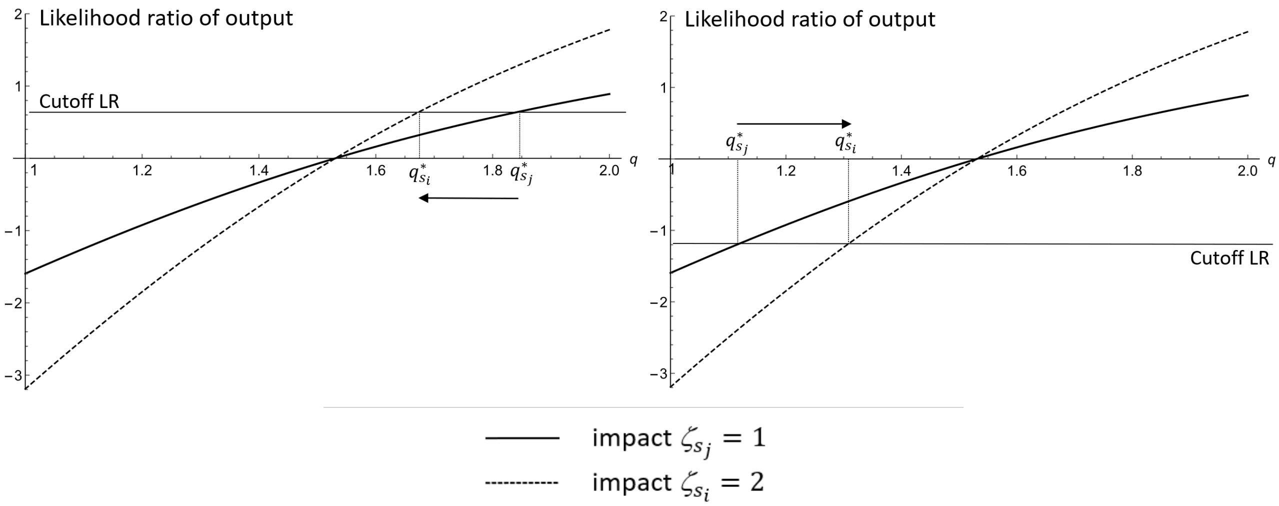

Figure 2: In both figures, the solid line is the likelihood ratio of output for a skew-normal distribution where location, scale, and impact parameters are all 1 and the shape parameter is 2. The peak of the distribution is $q_s^P \approx 1.531$. The dotted line is the likelihood ratio of output of the same distribution but with impact parameter 2.

threshold is even more positive, and so it equals the cutoff at a lower output level – thus the new threshold $q^*_{s_i}$ is lower. The right-hand side ("RHS") considers a negative cutoff, where the initial threshold $q^*_{s_j}$ is below the peak. A higher impact parameter corresponds to a lower likelihood ratio at $q^*_{s_j}$, and so the likelihood ratio equals the cutoff at a higher output level – thus the new threshold $q^*_{s_i}$ is higher. (We abuse language slightly by referring to "initial" and "new" parameters even though we are not conducting comparative statics but comparing parameters under different signal realizations.)

Since the scale parameter has the opposite effect on output precision to the impact parameter, one might think that it always has an opposite effect on the threshold. However, this need not be the case due to an important distinction. While the impact parameter changes the effect of effort on the output distribution, by construction it does not change the *equilibrium* output distribution. Recall that $h_s(\hat{e}) = \xi_s$: the location of the distribution is independent of ζ_s at the equilibrium effort level. Thus, the impact effect depends on the threshold compared only with the distribution's peak, not its location. In contrast, the scale parameter does change the equilibrium output distribution, spreading it out around the equilibrium location parameter – low output below the equilibrium location parameter is a less negative signal of effort and high output is a less positive signal of effort. As a result, the sign of the scale effect depends on two factors – the threshold compared to both the peak q_s^P and the equilibrium location parameter

ξ_s as shown in part (iv). For symmetric distributions, the peak and location are the same, but for asymmetric distributions they are different. If the threshold is above (below) both the peak and location, a lower scale parameter corresponds to a higher (lower) likelihood ratio at the threshold. However, if the threshold is between these two points, the threshold could rise or fall.

Summing up the results of this section, a signal can affect the threshold even if it is not individually informative about effort. If the signal suggests that the entire output distribution has improved, such as good industry performance, then all output levels are less indicative of effort and so the threshold rises. If the signal indicates that output is a more precise measure of effort, such as low industry volatility, then thresholds generally become less extreme – a previously high (low) threshold becomes lower (higher). However, surprisingly, the effects of impact and scale are not always in opposite directions.

Note that the precision effect is quite different from the intuition behind relative performance evaluation, which suggests that individually uninformative signals can only have value if they shift the output distribution (i.e. have a location effect). The precision effect means that, even if a signal neither is individually uninformative nor shifts the output distribution, it may still affect the payment. In practice, thresholds are sometimes changed due to shifts in the location of the distribution. For example, the cash flow target in the bonus of BP CEO Bob Dudley was lowered in 2015 due to the Deepwater Horizon disaster. However, we are not aware of cases in which thresholds are altered due to changes in the precision of output as a measure of effort.

Corollary 1 applies result (iv) of Proposition 1 to the case of a skew-normal output distribution and a non-binding participation constraint, to demonstrate how skewness affects the impact of the scale parameter on the threshold. In addition to location and scale parameters, the skew-normal distribution also has a “shape” parameter, which measures its skewness. Where the shape parameter is zero, the skew-normal distribution becomes a normal distribution.

Corollary 1 *Let the agent have a low reservation utility, $\overline{u} \leq u\left(\bar{W}\right) - C(\hat{e})$, so that the participation constraint is non-binding, and suppose that the output distribution is skew-normal. For a given signal s, the threshold q_s^* is increasing in the scale parameter if the signal is either individually uninformative or bad news about effort ($\frac{\partial \phi_{\hat{e}}^s}{\partial e} \leq 0$) and the output distribution has nonnegative skewness.*

Intuitively, when the only purpose of the contract is to provide incentives, the manager is paid only if q and s are sufficiently good news about effort. When the signal s is bad news, the likelihood ratio of output must be positive at the threshold q_s^*, to offset the bad news of the signal s. Since the likelihood ratio of output is zero at the peak, the threshold exceeds the peak due to MLRP. Moreover, for positively-skewed skew-normal (normal) distributions, the peak is above (equal to) the location parameter, so that $q_s^* \geq \max\left\{q_s^P, \xi_s\right\}$. Then, point (iv) of Proposition 1 implies that the threshold is higher for signals with a higher scale parameter.

This result is consistent with the fact that stock price distributions are typically positively skewed (see, e.g., Albuquerque (2012)), and the finding in Bettis et al. (2010, Table 8) that stock price thresholds in performance-vesting equity are increasing in stock return volatility, even though they do not depend on most other firm-level variables. However, their result studies the impact of volatility ex ante, i.e. before the contract has been signed; we study how the threshold should vary ex post with realized volatility.

This difference also contrasts our analysis with Chaigneau, Edmans, and Gottlieb (2018), who study how the optimal contract is affected by volatility, but not impact, location, or individual informativeness. Moreover, their volatility analysis is quite different. In their setting, there are no additional signals and the contract is written after volatility is realized, so volatility affects the contract that is offered. They show that a change in volatility ex ante affects the contract by altering the manager's incentives. For example, if an increase in volatility reduces incentives, then the threshold is lowered to restore incentives. In contrast, we study how incentives are allocated across signals – i.e. states of nature – that are associated with different output volatility. What matters instead is how the volatility associated with a given signal affects output informativeness. The contract is written before volatility (and other variables affected by the signal) is realized, and the contract is contingent on the signal – as is the case for performance-vesting equity or performance-sensitive debt.

2.2 Performance Signals and Pay-Performance Sensitivity

We now study how the PPS above the threshold varies across signals. To do so, we specialize to CRRA utility. This is because the curvature of the utility function plays an important role in the slope of the contract, and CRRA utility allows us to capture this curvature with a single parameter, γ. Moreover, CRRA is widely used for executive pay, in particular for calibration (see, e.g., Dittmann and Maug (2007) and references cited therein).

We define PPS as $\frac{w_s(q)-w_s(q_0)}{q-q_0}$ for $q > q_0 > \max\{q_s^P, \xi_s, q_s^*\}$ for $s \in \{s_i, s_j\}$. It represents the slope of the contract between any two outputs q and q_0 where the payment is strictly positive (since q_0 is above the threshold). Proposition 2 studies how PPS depends on the signal realization.

Proposition 2 *(Effect of signal on pay-performance sensitivity) Suppose that the FOA holds, and consider outputs q and q_0 such that $q > q_0 > \max\{q_s^P, \xi_s, q_s^*\}$ for $s \in \{s_i, s_j\}$. All else equal across signals:*

(i) If $\frac{\partial \phi_{\hat{e}}^{s_i}/\partial e}{\phi_{\hat{e}}^{s_i}} > \frac{\partial \phi_{\hat{e}}^{s_j}/\partial e}{\phi_{\hat{e}}^{s_j}}$, PPS is higher (lower) under s_i than s_j if $\gamma < (>)1$.

(ii) If $\xi_i > \xi_j$, PPS is higher (lower) under s_i than s_j if the likelihood ratio of output is weakly concave and $\gamma \geq 1$ (weakly convex and $\gamma \leq 1$).

(iii) If $\zeta_i > \zeta_j$, PPS is higher under s_i than s_j if $\gamma \leq 1$.

(iv) If $\sigma_i > \sigma_j$, PPS is lower under s_i than s_j if the likelihood ratio of output is weakly convex and $\gamma \leq 1$.

To understand how a signal realization affects PPS, consider the local PPS at output q and signal s for $w_s(q) > 0$:

$$w_s'(q) = \frac{\mu}{\gamma} \underbrace{\frac{\partial}{\partial q} \left\{ \frac{\frac{\partial f}{\partial e}(q|\hat{e}, s)}{f(q|\hat{e}, s)} \right\}}_{\text{A}} \left(\underbrace{\lambda + \mu \left[\frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e}, s)}{f(q|\hat{e}, s)} \right]}_{\text{B}} \right)^{\frac{1}{\gamma}-1} . \quad (12)$$

A signal can affect PPS by affecting term A or B in equation (12). Term A, which is positive by MLRP, is the slope of the likelihood ratio of output. A signal may affect this slope in two ways. The first is the *translation effect*: the likelihood ratio of output under s_i is a rightward translation of the likelihood ratio under signal s_j. When the likelihood ratio is concave (convex) in output, the translation effect increases (decreases) PPS at a given output level. The second is the *informativeness effect*: the likelihood ratio is steeper under s_i than s_j, i.e. higher for high output and lower for low output. The informativeness effect increases PPS regardless of the curvature of the likelihood ratio – it concentrates incentives in states of the world (signal realizations) where output is more informative about effort. The translation and informativeness effects are analogous to the location and precision effects in Section 2.1; we use different terminology here as this section analyzes the impact of signal realizations on PPS rather than the threshold.

Term B, which is positive (see equation (37) in Appendix A), is a linear transformation of the likelihood ratio. It captures the *contract curvature effect*: if the signal realization affects the payment $w_s(q)$, this in turn affects PPS if the payment above the threshold is nonlinear in term B. When $\gamma < 1$ ($\gamma > 1$), the payment above the threshold is convex (concave) in term B. Intuitively, γ affects not only risk aversion but also prudence – downside risk aversion, which depends on the third derivative of the utility function. When $\gamma < 1$, the effect of prudence (which favors convex contracts as they protect the agent from downside risk) dominates the effect of risk aversion (which favors concave contracts), so that the optimal contract is convex in a linear transformation of the likelihood ratio, i.e. term B.[7] When prudence is high, it is efficient to concentrate rewards on very positive outcomes rather than moderately positive outcomes. If the payment above the threshold is convex (concave) in term B, then if a signal realization increases the payment at a given output, PPS increases (decreases). When $\gamma = 1$, the contract is linear in term B, so changes to the payment do not affect PPS.

The likelihood ratio of the signal affects only term B; the likelihood ratio of output affects both terms. Thus, in part (i) of Proposition 2, which concerns the likelihood ratio of the signal, only the contract curvature effect applies. Signals that are individually indicative of effort increase pay; if $\gamma < 1$, the contract is convex in term B, and so this increase in pay also increases PPS. The effect is reversed if $\gamma > 1$.

In part (ii), which concerns the likelihood ratio of output, both the translation and contract curvature effects apply. Starting with the former, if $\xi_i > \xi_j$, the output distribution under s_i is shifted to the right. If the likelihood ratio is concave in output, a higher location of the output distribution means a greater slope of the likelihood ratio at any given output level, which increases PPS. Moving to the latter, at any given output, the payment is lower under s_i than s_j. When $\gamma \geq 1$ ($\gamma \leq 1$), the payment is concave (convex) in a linear transformation of the likelihood ratio, which means a higher (lower) PPS under s_i.

In parts (iii) and (iv), both the informativeness and contract curvature effects apply. In part (iii), if $\zeta_i > \zeta_j$, output is more informative about effort under s_i and so PPS is higher (the informativeness effect). In addition, above the peak of the distribution, a larger impact parameter leads to a higher likelihood ratio and therefore a higher payment, which weakly increases PPS if and only if $\gamma \leq 1$ (the contract curvature effect). Thus, if $\gamma \leq 1$, both effects go in the same direction and so PPS is unambiguously higher. In part (iv), if $\sigma_i > \sigma_j$, output is less informative about effort under s_i and so PPS is lower (the informativeness effect). In addition, above the peak and location parameter of the distribution, a more volatile output leads to a lower likelihood ratio and therefore a lower payment, which weakly decreases PPS if and only if $\gamma \leq 1$ (the contract curvature effect).[8]

[7] For a formal result, see Proposition 1 and Claim 4 in Chaigneau, Sahuguet, and Sinclair-Desgagné (2017).

[8] There is a subtle distinction between the informativeness effects for the impact and scale parameters, for the same reason that the impact and scale parameters do not always have opposite effects on the threshold in Proposition 1. The impact parameter affects the slope of the likelihood ratio but not the equilibrium output distribution. In contrast, the scale parameter affects the slope of the likelihood ratio and also spreads out the equilibrium output distribution. Thus, a higher scale parameter results in a lower likelihood ratio for high output levels, which further diminishes the sensitivity of the likelihood ratio to output (and thus PPS) if the likelihood ratio of output is convex – reinforcing the effects described in the main text. This explains why part

It is again instructive to highlight that we are comparing pay across signal realizations holding the distribution of signals fixed. Proposition 2 would not arise under a comparative statics analysis that compares contracts under different signal distributions, where the principal observes the parameters of the distribution before writing the contract. Here, the contract is set before the signal is observed, and the agent does not know the signal at the time he makes his effort decision. This is what allows the principal to efficiently concentrate incentives in states of the world where output is more informative. If the agent already knew the state when he took his effort decision, to implement $\hat{e}$ the principal would need to provide the same incentives in all states: see Edmans and Gabaix (2011).

2.3 Option Repricing and Performance-Vesting

Section 2.2 studied how a signal affects the contract's PPS above the threshold. Due to the contract curvature effect, the effect of the signal realization is sometimes ambiguous: for example, in parts (iii) and (iv) of Proposition 2, only if $\gamma \leq 1$ does the contract curvature effect go in the same direction as the informativeness effect, leading to an unambiguous impact on PPS. The analysis is particularly clear if the contract is linear above the threshold, because the contract curvature effect disappears. As well as being theoretically clearer, this case is also practically applicable as many real-life contracts are piecewise linear; in addition, the PPS can be interpreted as the slope of the contract. Indeed, if q is the stock price, the optimal contract is an option, and the PPS above the threshold refers to the number of vesting options. Then, the model provides guidance on performance-based vesting – how the number of vesting options should depend on the performance measure s.

We first derive sufficient conditions under which the optimal contract is an option. Lemma 3 shows this is the case under the standard assumptions of log utility, normally-distributed output and limited liability. Formally, conditional on the signal s, let output q be normally distributed with mean $h_s(e)$ and standard deviation σ_s, where $h_s(0) = 0$. The manager has log utility, $u(w) = \ln w$. The Supplementary Appendix provides a sufficient condition for the validity of the FOA under these assumptions (the condition in Lemma 1 cannot be applied since log utility is unbounded).[9]

(iv) also requires the likelihood ratio of output to be weakly convex whereas part (iii) does not.

[9]The condition for the validity of the FOA in the case without an additional signal is remarkably simple: it is $C''(e) \geq \frac{\bar{e}}{\sigma^2}$ for all $e \in [0, \bar{e}]$. For example, with a quadratic effort cost, $C(e) = \alpha e + \frac{\beta}{2}e^2$, and the condition is $\beta \geq \frac{\bar{e}}{\sigma^2}$.

From equation (1), the likelihood ratio under log utility and normal output is given by:

$$
\begin{aligned}
LR_s(q) := \frac{\frac{\partial f}{\partial e}(q, s|\hat{e})}{f(q, s|\hat{e})} &= \frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e}, s)}{f(q|\hat{e}, s)} \\
&= \frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{1}{\sigma\sqrt{2\pi}} \frac{2h'_s(\hat{e})(q-h_s(\hat{e}))}{2\sigma^2} \exp\left\{-\frac{(q-h_s(\hat{e}))^2}{2\sigma^2}\right\}}{\frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(q-h_s(\hat{e}))^2}{2\sigma^2}\right\}} \\
&= \frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\zeta_s}{\sigma_s^2}\left[q - \xi_s\right]. \qquad (13)
\end{aligned}
$$

The $\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}$ term is the standard individual informativeness effect. In general, the output inference effect $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ comprises the precision and location effects; under the normal distribution (or any distribution with a linear likelihood ratio), $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ can be cleanly decomposed into these two effects, allowing us to see the parameters that drive them. The precision effect is given by the first component of the second term, $\frac{\zeta_s}{\sigma_s^2}$. The signal s increases the precision of output as a measure of effort through increasing the impact of effort on output ζ_s or reducing the volatility of output σ_s. Here, since the normal distribution is symmetric, these parameters do always have opposite effects. The location effect is given by the final component, ξ_s. The signal affects expected output ξ_s and thus changes the location of the output distribution.

Lemma 3 below shows that the optimal contract gives the manager n_s^* options with strike price q_s^*.

Lemma 3 *(Optimal contract, log utility and normal output): The optimal contract under log utility and normally-distributed output consists of* $n_s^* \geq 0$ *options with a strike price of* q_s^**:*

$$
w(q) = n_s^* \max\{q - q_s^*, 0\}. \qquad (14)
$$

Given limited liability, the minimum payment is zero; given MLRP, this minimum payment will be made for all outputs below a threshold. Above the threshold, the payment is positive and determined so that the manager's marginal utility is the inverse of a linear transformation of the likelihood ratio (see Lemma 2). With log utility, marginal utility is the inverse of the payment, and so the payment equals a linear transformation of the likelihood ratio. With normally-distributed output, the likelihood ratio is linear in output, and so the payment is linear in output. Overall, the payment is zero below a threshold and linear in output above the threshold. This corresponds to an option contract, where the strike price is the threshold.

To our knowledge, Lemma 3 and our condition for the validity of the FOA provide the first sufficient conditions for the optimality of options with a risk-averse manager.[10] More generally, the linearity of the contract above the threshold, and thus the optimality of the option contract, holds not only for the normal distribution but for any distribution that has a linear likelihood ratio (for example, the gamma distribution).

Having derived the optimal contract in closed form, Proposition 3 studies how the signal affects each dimension of the contract.

Proposition 3 *(Effect of signal on vesting and strike price):*

(i) The number of options received ex post by the manager n_s^* is proportional to $\frac{\zeta_s}{\sigma_s^2}$.

(ii) The strike price q_s^* is given by:

$$q_s^* = \xi_s + \frac{\sigma_s^2}{\zeta_s}\left(K - \frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s}\right), \tag{15}$$

where $K \in \mathbb{R}$.

The intuition for the number of options is as follows. As in any principal-agent model, pay is increasing in the likelihood ratio. The number of options represents PPS, and is thus increasing in the sensitivity of the likelihood ratio to output, $\frac{dLR_s(q)}{dq} = \frac{\zeta_s}{\sigma_s^2}$. As is standard, the strike price is the level of output that, if not reached, it is sufficiently likely that the agent shirked that it is optimal to pay him zero. In this setting, the strike price can be derived in closed form.

Note that Proposition 3 is different from Propositions 1 and 2 – i.e. it is not simply these propositions applied to the case of log utility and normally-distributed output. Propositions 1 and 2 compared signal realizations that differ only across a single dimension, e.g. either the impact parameter ζ_s or the scale parameter σ_s, but not both. Moving to this setting allows Proposition 3 to compare contracts analytically when signals differ across multiple dimensions. For example, part (i) allows signal realizations to differ in both ζ_s and σ_s. In an economic

[10] Jewitt, Kadan, and Swinkels (2008) show that the contract is "option-like" with risk aversion and agent limited liability, in that incentives are zero for low output and positive for high output, but do not identify conditions under which the increasing portion of the contract is linear. Hemmer, Kim, and Verrecchia (1999) identify a linear likelihood ratio and log utility as leading to the contract having a linear portion, but do not combine them with limited liability to obtain an option contract. In addition, Jewitt, Kadan, and Swinkels (2008) assume the Rogerson (1985) conditions to guarantee the validity of the FOA, but these conditions do not hold under the normal distribution; Hemmer, Kim, and Verrecchia (1999) assume the Jewitt (1988) conditions but they do not hold under limited liability. We derive a general condition for the validity of the FOA that holds in the setting of limited liability, normal output and log utility. The standard model justifying options under moral hazard is Innes (1990), which requires the agent to be risk-neutral. In addition, under risk neutrality, the manager is the residual claimant for $q \geq q^*$, so that the number of options is fixed at 1 and does not depend on the signal realization. Under risk aversion, this need not be the case.

expansion, not only might effort have a greater effect on output ζ_s (e.g. if effort has a multiplicative effect on firm value), but also volatility σ_s might be higher or lower. The ability to compare contracts when signals differ across multiple dimensions will be important for the applications later in this section.

We now discuss the two parts of Proposition 3 in turn.

2.3.1 Performance-based vesting

Part (i) of Proposition 3 studies how a signal realization affects the number of options given to the manager. A signal has value if it affects any component of the likelihood ratio in equation (13): $\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}$ (the individual informativeness effect), $\frac{\zeta_s}{\sigma_s^2}$ (the precision effect), or ξ_s (the location effect). The existence of such a signal will, in general, alter the Lagrange multiplier μ and thus scale up or down the number of options $n_s^* = \mu \frac{\zeta_s}{\sigma_s^2}$ received across all signal realizations. However, part (i) shows that the number of options received will depend on the actual signal realization only via the precision effect and not via the location or individual informativeness effects.

The intuition is as follows. The number of vesting options represents PPS. Pay should be more sensitive to performance, i.e. the number of vesting options should be higher, upon signals where output is a more precise measure of effort. This arises if either effort has a greater effect on output (ζ_s is higher) or output is less volatile (σ_s is lower). In contrast to Proposition 2, Proposition 3 derives PPS in closed form and thus shows how it depends on the *ratio* of impact ζ_s to (the square of) scale σ_s. We can thus understand precisely how PPS varies with parameters of the output distribution under various signal realizations.

To our knowledge, part (i) is the first theoretical justification of why performance-based vesting may be optimal in a standard moral hazard model where the agent only makes an effort decision. The only other justification of performance-based vesting of which we are aware is Marinovic and Varas (2019), where the agent also takes a manipulation action, and the role of performance-based vesting is to deter manipulation. In that model, performance-based vesting is always optimal. Given the "performance-based vesting" terminology, one might think that it will similarly always be optimal in an effort-only model, as long as the performance measure is informative about effort. However, this is not the case, and we provide a framework to understand under what conditions performance-vesting is optimal.

Specifically, it might seem that a signal that is individually indicative of effort (i.e. increases $\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}$) should lead to more vesting, and indeed current performance-vesting practices award more equity after beating performance thresholds. However, Proposition 3 shows that positive signals of effort should increase the level of pay for all output realizations (reduce the strike price) rather than the sensitivity of pay to output (increase the number of vesting options). Similarly, one might think that a signal that indicates that high output is due to luck (i.e. the output distribution has shifted to the right) should lead to less vesting – the intuition behind relative performance evaluation. Indeed, Bettis et al. (2018) find that 48% of firms that use performance-based vesting have at least one performance measure that is calculated relative to peers. However, Proposition 3 shows that the equilibrium location parameter ξ_s (which is affected by peer performance) changes the strike price, not the number of vesting options. In sum, vesting should not depend on signals that do not affect output precision.

We now apply the results of part (i) of Proposition 3 to two types of signal.

Economic conditions

First, let s be a signal of economic conditions, which are outside the manager's control and thus individually uninformative about effort ($\frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s} = 0$). However, part (i) states that individual informativeness is irrelevant; a signal will affect vesting if it affects either ζ_s or σ_s. Starting with the former (ζ_s), if good economic conditions increase the manager's impact on output ζ_s, e.g. if effort has a multiplicative effect on firm value, vesting should be increasing in economic conditions; the intuition behind relative performance evaluation would suggest the opposite.[11] If instead bad economic conditions increase the impact of effort, e.g. if the stakes are higher in bad times, vesting should be decreasing in economic conditions. Moving to the latter (σ_s), if risk σ_s is lower (higher) in good economic conditions, then output is a more (less) precise signal of effort and so vesting should be higher (lower).

The dependence of either ζ_s or σ_s on economic conditions s shows that it may be optimal for vesting to be affected by luck. This contrasts with current performance-vesting practices which assume that vesting should depend on performance measures within the manager's control – and if individually uninformative measures are used, they should only be used to filter out external factors that affect the location of the output distribution.

Accounting performance

Now let s be an accounting performance measure, such as profits or cash flows, which Bettis et al. (2018) show to be common vesting determinants. While accounting performance, unlike economic conditions, is individually informative about effort, this is irrelevant for whether the signal affects vesting. Again, vesting depends on how ζ_s and σ_s vary with the signal. Starting with the latter (σ_s), if volatility is increasing in profits, vesting is decreasing in profits. This

[11]Good economic conditions are also typically associated with a higher ξ_s and thus suggest that a given output level was due to luck rather than effort. However, this location effect will only affect the strike price.

may be the case for a start-up, where the baseline scenario is low profits and a low stock price. High profits increase the variability of the stock price (e.g. because investors speculate as to whether the high profits are sustainable), which makes the stock price less informative about effort. In contrast, if volatility is decreasing in profits, vesting is increasing in profits. This may arise due to the well-known "leverage effect" of Black (1976) and Christie (1982), where higher profitability reduces a firm's leverage and thus equity volatility. It may also be true for a mature firm, where the baseline scenario is high profits and a high stock price. High profits imply "business as usual", where the stock price is less volatile and thus more informative about effort. Low profits likely mean that the business was disrupted, for example by new entrants, so the stock price is more volatile and less informative about effort. Moving to the former (ζ_s), if effort has a multiplicative effect on firm value and is thus more impactful in a more profitable firm, then vesting will be increasing in firm profits.

These two applications are related to studies on the relationship between firm risk and PPS (e.g. Garen (1994), Aggarwal and Samwick (1999)). Those studies estimate this relation cross-sectionally between firms, whereas our prediction is within firm but across states of the world. Aggarwal and Samwick (1999) argue that a negative relation between PPS and stock return volatility is a "key prediction" of the principal-agent model of executive compensation. However, Prendergast's (2002) review of the evidence finds a mixed relationship, and points out that this may arise because risk can be endogenous either to the agent's effort or the job to which he is assigned. Our model yields a related prediction which is not affected by this endogeneity concern: even if the manager could change the overall level of risk, the realized level of risk will still depend on the state of the world. For a given manager, PPS will be higher (via additional vesting) in states where the firm's stock is less volatile because the stock price is a better signal of effort.

2.3.2 Strike price

Part (ii) of Proposition 3 turns to the second dimension of the contract, the strike price. Unlike the number of vesting options, which depends only on the precision effect, the strike price is driven by all three effects. The intuition is as follows. The individual informativeness effect matters because it captures what a signal individually conveys about effort, regardless of the output realization. The strike price affects the payment for all outputs above the strike price. Thus, what the signal individually conveys about effort affects the strike price – signals that are more indicative of effort are associated with lower strike prices. The location effect matters for a similar reason – if the signal indicates that the output distribution has shifted to the left, a given output is more indicative of effort, and so the strike price should be lower.

The intuition for the role of the precision effect is as follows. If strike prices are high across signal realizations, the manager is only paid if output is high. If output is a more precise signal of effort ($\frac{\zeta_s}{\sigma_s^2}$ is higher), even moderately high outputs are positive signals of effort – the strike price falls. Conversely, if strike prices are low across signal realizations, greater precision $\frac{\zeta_s}{\sigma_s^2}$ increases the strike price. Again, this intuition is similar to Proposition 1, but in Proposition 3 we can be even more concrete as to what drives the precision effect: the impact parameter divided by the (squared) scale parameter. In addition, as stated previously, Proposition 3 allows us to compare signals across multiple dimensions, which is important for the examples that we will shortly discuss.

Note that equation (15) shows that the individual informativeness effect is scaled by the inverse of signal precision, $\frac{\sigma_s^2}{\zeta_s}$. The intuition is as follow. A signal that is individually bad news about effort leads to a higher strike price. If output is an imprecise measure of effort, then the strike price needs to be raised a lot, otherwise it would likely be exceeded purely by luck.

Since the strike price is driven by the location and precision effects, as well as the individual informativeness effect, it may be lower even if the signal individually indicates low effort. Consider two signal realizations, L and H, such that $\frac{\partial \phi_{\hat{e}}^L/\partial e}{\phi_{\hat{e}}^L} < \frac{\partial \phi_{\hat{e}}^H/\partial e}{\phi_{\hat{e}}^H}$: L is individually worse news about effort than H. Despite this, the strike price may be lower under L ($q_L^* < q_H^*$). This may occur in two cases. Letting the information output provides about effort be $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} := a_s + b_s q$, we have $a_s = -\frac{\zeta_s}{\sigma_s^2}\xi_s$ (the location effect scaled by the precision effect) and $b_s = \frac{\zeta_s}{\sigma_s^2}$ (the precision effect).

One case is $a_L > a_H$ and $b_L = b_H$, so the signal has a location effect but not a precision effect. Here, any given output q is better news about effort under L than H, since $\frac{\frac{\partial f}{\partial e}(q|\hat{e},L)}{f(q|\hat{e},L)} > \frac{\frac{\partial f}{\partial e}(q|\hat{e},H)}{f(q|\hat{e},H)}$ for all q. Equation (15) shows that the strike price is optimally lower under L if the location effect (the difference between a_L and a_H) outweighs the individual informativeness effect (the difference between $\frac{\partial \phi_{\hat{e}}^L/\partial e}{\phi_{\hat{e}}^L}$ and $\frac{\partial \phi_{\hat{e}}^H/\partial e}{\phi_{\hat{e}}^H}$). This allows us revisit Example 1 analytically. Recall that, in this example, s is the number of industry competitors; $s = L$ reflects few competitors and $s = H$ reflects many. Thus, $s = L$ is individually more indicative of effort, but also shifts the output distribution to the right. Equation (15) shows how much higher the location effect must be relative to the individual informativeness effect for the strike price to be optimally higher when there are few competitors.

A second case is $b_D > b_L$, $a_D = a_L$, and $\frac{\partial \phi_{\hat{e}}^D/\partial e}{\phi_{\hat{e}}^D} < \frac{\partial \phi_{\hat{e}}^L/\partial e}{\phi_{\hat{e}}^L} < 0$. Here the signal has a precision effect but not a location effect. Both signals D ("dire") and L ("low") are individually bad news about effort, with D being worse news. Since $b_D > b_L$, output q is more informative about effort under D than L, and so the manager should be rewarded more for a high output under D. This generates a lower strike price under D, if the precision effect (the difference between b_D and b_L) is sufficiently large to outweigh the individual informativeness effect (the difference between $\frac{\partial \phi_{\hat{e}}^D/\partial e}{\phi_{\hat{e}}^D}$ and $\frac{\partial \phi_{\hat{e}}^L/\partial e}{\phi_{\hat{e}}^L}$).

For example, consider a firm whose credit rating can be downgraded by one notch but remain investment-grade ($s = L$), or downgraded to junk ($s = D$). A downgrade to junk is individually worse news about effort than a one notch downgrade. Such a downgrade also restricts the firm's access to external financing; since it is now financially constrained, its performance may depend more on managerial effort (e.g. to cut costs or reallocate capital across divisions). Thus, output is more informative about effort. As a result, high output following a downgrade to junk can indicate effort more than high output following a one notch downgrade. Even though a downgrade to junk status individually indicates low effort, in combination with high output it indicates high effort, and so can be associated with a lower strike price ($q_D^* < q_L^*$).

Overall, part (ii) provides conditions under which the strike price should depend on additional signals. Liljeblom, Pasternack, and Rosenberg (2011) study the determinants of the option strike price ex ante, i.e., when the contract is set, in Finland where there are no tax and accounting considerations that favor at-the-money options as in the US. However, they do not study how the strike price is reset ex post depending on the signal realization. Brenner, Sundaram, and Yermack (2000) find empirically that repricing nearly always involves a lowering of the strike price, and follows poor stock price performance (both absolute and industry-adjusted). Chance, Kumar, and Todd (2000) also find that repricing follows poor stock returns, but that these are not due to market or industry conditions. Our model suggests that a reduction in the strike price should *generally* be prompted by positive, rather than negative, signals of effort, suggesting that such practices are suboptimal.[12] However, the above examples provide conditions under which such repricing is optimal, contrary to concerns that it is universally inefficient because it rewards failure (e.g. Bebchuk and Fried (2004)).

2.3.3 Summary

Summing up the results of this section, a signal realization should be associated with more vesting options if it is associated with a higher optimal sensitivity of pay to output, due to output being a more precise measure of effort. It should be associated with a lower strike price if it increases the optimal level of pay, by indicating high effort regardless of output. The

[12] Acharya, John, and Sundaram (2000) also study the repricing of options theoretically. In their model, repricing is not undertaken to make use of additional informative signals, but instead to maintain effort incentives when options fall out of the money.

optimal level of pay depends on the individual informativeness, location and precision effects.

These results imply that the common practice of making pay depend on a performance measure only through performance-based vesting may not actually be optimal. This is efficient if the performance measure only has a precision effect, but most measures also have individual informativeness or location effects; thus, they should also affect the strike price. The conventional justification of performance-based vesting is that it rewards the manager for good performance. However, its primary effect is to change the slope of the contract, which is only optimal if the signal realization affects either impact or scale (or both). If the principal wishes to reward the manager for good performance, either because the signal is individually indicative of effort, or because the signal shows that a given output was harder to achieve (the location effect), this is best done by lowering the strike price. Changing the number of vesting options fails to change the payment below the strike price. Thus, the normative implication of this section is that performance-based options should typically involve the strike price, rather than the number of vesting options, depending on performance.

Example 2 illustrates how the precision effect impacts both the strike price and number of vesting options.

Example 2 *Suppose that the manager has log utility, outside wealth $\bar{W} = 2$, and reservation utility $\bar{u} = 0$. The cost of effort is characterized by $C(\hat{e}) = 0.5$ and $C'(\hat{e}) = 0.1$. The signal s is the state of the economy, which can be a recession ($\tilde{s} = s_R$, which occurs with probability 0.25) or expansion ($\tilde{s} = s_E$). The signal is individually uninformative about effort: ($\frac{\partial \phi^s_{\hat{e}}}{\partial e} = 0$). At equilibrium effort $\hat{e}$, output is normally distributed with $\xi_{s_R} = 10$ and $\sigma_{s_R} = 1.7$ in a recession, and $\xi_{s_E} = 11$ and $\sigma_{s_E} = 1.0$ in an expansion. For both signals, $\zeta_s = 1$.*

Since $\ln(\bar{W}) - C(\hat{e}) \geq \bar{u}$, the participation constraint is nonbinding. We find that $q^*_{s_R} \approx 14.27$ and $q^*_{s_E} \approx 12.48$. Indeed, we have $q^*_s = \xi_s + \frac{\sigma_s^2}{\zeta_s}\frac{\bar{W}}{\mu}$: in a recession, a lower equilibrium location parameter ξ_s decreases the strike price, whereas a higher scale parameter σ_s increases it. The latter effect dominates so that the strike price is higher in a recession, which is the opposite of the standard relative performance evaluation effect. Moreover, because output is a more precise measure of effort in an expansion, the number of vesting options is higher in an expansion. The state-contingent contract is depicted in Figure 3.

2.4 Debt Contracts

This section studies a second standard contract that is linear above the threshold – debt. Here, the principal receives debt and the agent receives equity (output minus debt); the threshold thus corresponds to the face value of debt. A debt contract is similar to the option contract

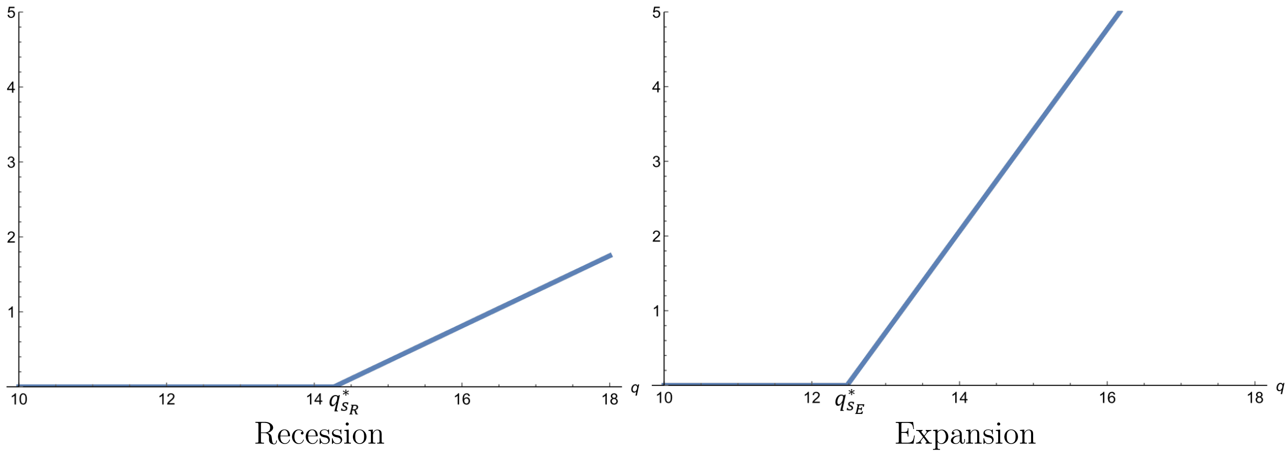

Figure 3: The payoff of the manager as a function of output in a recession ($s = s_R$) on the left, and in an expansion ($s = s_E$) on the right, in Example 2.

of Section 2.3 except that the slope of the contract is always 1. While it might seem that debt is simply a special case of the option contract, we will show that not all of the results of Section 2.3 continue to hold.

To apply our model to debt contracts, we will use the setup of Innes (1990), which is the most standard moral hazard model that generates debt as the optimal contract; otherwise, we retain previous assumptions. In Innes (1990), the agent is risk-neutral, and the principal is also protected by limited liability. Since our setting allows output to be negative, this latter constraint generalizes to $w_s(q) \leq \max\{0, q\}$ – the principal cannot pay more than output.[13] To ensure that an incentive compatible contract exists, we assume:

$$\sum_s \int_0^\infty q \frac{\partial f}{\partial e}(q, s|\hat{e}) dq > C'(\hat{e}). \tag{16}$$

This assumption means that a contract, where the agent receives the entire output subject to limited liability for any signal realization, provides sufficient incentives – thus, the effort level $\hat{e}$ can always be implemented.

With risk aversion, solving the agency problem is always costly to the principal, as doing so requires her to offer a performance-sensitive contract that exposes the agent to risk. With risk neutrality, solving the agency problem is only costly if the IR is slack, and so a performance-

[13]When output can be negative, bilateral limited liability requires a third party – e.g., a creditor, supplier, or the government – to bear the loss.

sensitive contract gives the agent rents. We thus assume $\overline{u} = 0$ and $C(0) = 0$. With a non-binding IR and a risk-neutral manager, we can also assume $\bar{W} = 0$ without loss of generality. The firm solves the following program:

$$\min_{\{w_s(q)\}} \sum_s \phi_{\hat{e}}^s \int_{\underline{q}}^{\infty} w_s(q)\, f(q|\hat{e}, s) dq \tag{17}$$

subject to

$$\sum_s \phi_{\hat{e}}^s \int_{\underline{q}}^{\infty} w_s(q)\, f(q|\hat{e}, s) dq \geq C(\hat{e}), \tag{18}$$

$$\hat{e} \in \arg\max_e \sum_s \phi_e^s \int_{\underline{q}}^{\infty} w_s(q)\, f(q|e, s) dq - C(e), \tag{19}$$

$$w_s(q) \in [0, \max\{0, q\}] \ \ \forall q, s, \tag{20}$$

$$w_s(q + \epsilon) - w_s(q) \leq \epsilon \ \ \forall \epsilon > 0. \tag{21}$$

With $C(0) = 0$, the IC (19) and LL (20) imply that the IR (18) is automatically satisfied, and so we ignore it in the analysis that follows. The monotonicity constraint (21) is the final ingredient of the Innes (1990) model. It means that a dollar increase in output cannot increase the payment to the manager by more than a dollar (else he would inject his own money into the firm to increase output), or equivalently the payoff to the principal cannot decrease in output (else she would exercise her control rights to "burn" output). As Innes (1990) showed, this assumption leads to the optimal contract being debt.

In Lemma 4 below, we present a new condition for the validity of the FOA in the above program (the condition from Lemma 1 is not applicable because it requires utility to be bounded from above).[14] Let K_e be defined as:

$$K_e := \sum_s \int_0^{\infty} q \max\left\{\frac{\partial^2 f}{\partial e^2}(q, s|e), 0\right\} dq.$$

Lemma 4 *(First-order approach, risk-neutral manager and monotonicity constraint) Suppose that $K_e < C''(e)\ \forall e \in (0, \bar{e})$. Then the FOA is valid.*

We henceforth assume that the condition in Lemma 4 holds. The optimal contract is given by Lemma 5 below:

[14] Innes (1990) assumes the FOA and gives examples of sufficient conditions for it to hold, such as Rogerson's (1985) condition on the convexity of the cumulative distribution function. However, this condition is not satisfied by many distributions with location and scale parameters, which our framework uses.

Lemma 5 *(Optimal contract, risk-neutral manager and monotonicity constraint) The optimal contract with a risk-neutral manager and monotonicity constraint is* $w_s(q) = \max\{q - q_s^*, 0\}$*, for some* q_s^**.*

As in Innes (1990), the optimal contract is debt, i.e. an option with a slope of 1. The manager receives zero if output is less than the face value of debt q_s^*, and the residual $q - q_s^*$ otherwise. 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 firm cannot incentivize the manager in the left tail by giving negative payments (due to limited liability), so it incentivizes him in the right tail by giving high payments. Under the monotonicity constraint, the maximum possible incentives involve the manager gaining one-for-one from any increase in output, so he receives the residual. A lower marginal cost of effort $C'(\hat{e})$ corresponds to a higher face value; intuitively, when the moral hazard problem is weak, a low payment is sufficient to induce managerial effort.

The above analysis is standard. We now turn to our key question of how the signal affects the contract. The principal's only degree of freedom is the face value of debt q_s^*. Thus, the signal realization can only affect the contract via changing the required debt repayment, as with performance-sensitive debt. It cannot change the payment for lower output levels ($q < q_s^*$) because the manager is already receiving the minimum possible (zero) due to limited liability. It cannot change the payment for individual higher output levels ($q > q_s^*$) in isolation because the manager is already receiving the maximum possible (the residual $q - q_s^*$) due to monotonicity. This last observation is the critical difference between debt and the option contract; under the latter, the slope can change as it is not always 1.

2.4.1 Performance-Sensitive Debt

Proposition 4 shows how the signal realization affects the face value of debt.

Proposition 4 *(Effect of signal on face value of debt): All else equal across signals:*

(i) If $\frac{\partial \phi_{\hat{e}}^{s_i}/\partial e}{\phi_{\hat{e}}^{s_i}} > \frac{\partial \phi_{\hat{e}}^{s_j}/\partial e}{\phi_{\hat{e}}^{s_j}}$, $q_{s_i}^* \leq q_{s_j}^*$.

(ii) If $\xi_{s_i} < \xi_{s_j}$, $q_{s_i}^* \leq q_{s_j}^*$.

(iii) If $\zeta_{s_i} > \zeta_{s_j}$, $q_{s_i}^* \leq q_{s_j}^*$

(iv) If $\sigma_{s_i} > \sigma_{s_j}$ *and* $q_s^* > \max\{q_s^P, \xi_s\}$, $q_{s_i}^* \geq q_{s_j}^*$.

Parts (i) and (ii), on the individual informativeness effect and the location effect, are standard. They echo the results in the general model of Proposition 1 and also the option contract of Proposition 3. As before, part (ii) may lead to counterintuitive results, since performance measures that indicate low effort (such as low credit ratings) typically increase the required debt repayment. While a low credit rating is indeed a negative individual signal of performance, it may also shift the output distribution to the left as it restricts the firm's access to financing. Thus, achieving a given output is a more positive signal of effort, and so the universal practice of the threshold repayment decreasing in the credit rating may not be optimal.

Part (iii), on the effect of the impact parameter, is different. Recall that, for the option contract, when the impact parameter is high, the principal wishes to provide strong incentives. She does so by increasing the slope of the contract (the number of vesting options); the effect on the strike price is ambiguous as it depends on whether the strike price is high or low across signal realizations. Here, the principal is unable to provide strong incentives by increasing the slope of the contract, as it is always 1. She can only do so by lowering the face value of debt, and so a higher impact parameter ζ_s is always associated with a lower threshold q_s.

Part (iv), on the effect of the scale parameter, is generally the opposite of part (iii). When output volatility is high, the principal wishes to provide weak incentives, which increases the face value of debt. However, recall from Proposition 1 that the impact and scale parameters do not always have opposite effects on the threshold, because the scale parameter changes the equilibrium output distribution but the impact parameter does not. This explains the additional condition, $q_s^* > \max\{q_s^P, \xi_s\}$, in part (iv).

Our rationale for performance-sensitive debt complements existing explanations. Manso, Strulovici, and Tchistyi (2010) model performance-sensitive debt as an optimal signaling device under asymmetric information about the firm's growth rate; there is no effort decision. Bhanot and Mello (2006) and Koziol and Lawrenz (2010) show that performance-sensitive debt deters risk shifting. While none of these papers model an effort decision, Manso et al. (2010, Section 8) conjecture that performance-sensitive debt "could serve as an additional incentive for the firm's manager to exert effort" and Tchistyi (2009) shows that performance-sensitive debt can deter cash flow diversion. This intuition would suggest that the threshold should fall with signals that are individually indicative of effort (part (i) of Proposition 4). Extending this intuition further might also suggest that the threshold should fall with signals that indicate a leftward shift of the output distribution (part (ii)), as this also implies higher effort for a given output level. However, it does not have implications for impact and scale (parts (iii) and (iv)). Innes (1993) derives the optimal contract when profits (which correspond to q in our setting) can be decomposed into output and the output price, i.e. the price is an additional signal that can be used in the contract. He shows that the optimal contract is a price-contingent commodity bond, which has similarities to performance-sensitive debt; however, the only signal that he analyzes is price (i.e. one component of output). In contrast, we consider a broad set of signals, including signals that are informative about the manager's effort, and signals that affect the output distribution in different ways to the price.

Summing up the results of this section, a signal can affect the face value of debt even if it is not individually informative about effort. If the signal indicates that the entire output distribution has improved, such as good industry performance, then all output levels are less indicative of effort and so the face value rises. If the signal indicates that output is a more precise measure of effort, such as a high impact parameter or low industry volatility, then incentives should be increased by lowering the face value of debt.

3 Conclusion

This paper has studied how signal realizations should affect contracts, thus providing guidance on practical contract design. In a general optimal contracting model with limited liability, the optimal contract involves a threshold output level. The signal realization affects the threshold in two main ways. The first is the individual informativeness effect – signals that individually indicate high effort should be associated with lower thresholds, all else equal. The second is the output inference effect – signals affect the information that the principal infers about effort from the output level. Due to this second effect, the impact of a signal realization on the threshold is ambiguous; signals that are individually indicative of effort may end up being associated with higher thresholds. The output inference effect in turn can be decomposed into two components. The location effect arises if the signal indicates that the location of the output distribution has shifted to the left (right), in which case the threshold decreases (increases). The precision effect occurs if the signal indicates that output is riskier and/or more impacted by effort. Greater output precision (higher impact and/or lower risk) leads to high thresholds falling and low thresholds rising.

The signal realization also affects the optimal pay-performance sensitivity above the threshold. A signal that individually indicates high effort, or suggests that the output distribution has shifted to the left, is associated with greater sensitivity if the manager's risk aversion is low. Higher precision is generally associated with greater sensitivity, especially if the manager's risk aversion is low.

We then apply our model to two contracts that are commonly-observed in real life, where the slope of the contract is linear above the threshold. One is an option contract, where the threshold corresponds to the strike price. This application has implications for performance-vesting options. In practice, performance signals only affect the number of vesting options, but our analysis suggests that they should instead affect the strike price. Only a limited set of signals – those that affect output precision – should also affect the number of vesting options. The second application is a debt contract, where the threshold corresponds to the face value of debt. Since the slope of the contract is constrained to be 1, a greater impact of effort on output cannot lead to a steeper contract; instead, it always leads to a lower face value of debt. Thus, debt may optimally be performance-sensitive, even if performance signals are not individually informative about effort.

Our model features a single effort decision by the manager and a single component of output, and studies how to incorporate an additional signal into a contract. In future research, it would be interesting to study a multi-tasking model where there are multiple actions and multiple components of output. For example, in Holmström and Tirole (1993), the manager takes both a short-term and a long-term action, and the principal cares about both short-term earnings and liquidation value. How to incorporate an additional signal, that is informative about one or both actions, into the contract is an open question.

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., Andrew A. Samwick (1999): "The other side of the trade-off: the impact of risk on executive compensation." *Journal of Political Economy* 107, 65–105.

[3] Albuquerque, Rui (2012): "Skewness in stock returns: Reconciling the evidence on firm versus aggregate returns." *Review of Financial Studies* 25, 1630–1673.

[4] Allen, Franklin and Douglas Gale (1992): "Measurement distortion and missing contingencies in optimal contracts." *Economic Theory* 2, 1–26.

[5] Amershi, Amin H. and John S. Hughes (1989): "Multiple signals, statistical sufficiency, and Pareto orderings of best agency contracts." *RAND Journal of Economics* 20, 102–112.

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

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

[8] Bettis, J. Carr, John Bizjak, Jeffrey Coles, and Swaminathan Kalpathy (2018): "Performance-vesting provisions in executive compensation." *Journal of Accounting and Economics* 66, 194–221.

[9] Bhanot, Karan and Antonio S. Mello (2006): "Should corporate debt include a rating trigger?" *Journal of Financial Economics* 79, 69–98.

[10] Black, Fischer (1976): "Studies of stock price volatility changes." *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, American Statistical Association*, 177–181.

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

[12] Chaigneau, Pierre, Alex Edmans and Daniel Gottlieb (2018): "Does improved information improve incentives?" *Journal of Financial Economics* 130, 291–307.

[13] Chaigneau, Pierre, Alex Edmans and Daniel Gottlieb (2019): "The informativeness principle without the first-order approach." *Games and Economic Behavior* 113, 743–755.

[14] Chaigneau, Pierre, Nicolas Sahuguet, and Bernard Sinclair-Desgagné (2017): "Prudence and the convexity of compensation contracts." *Economics Letters* 157, 14–16.

[15] Chance, Don M., Raman Kumar, and Rebecca B. Todd (2000): "The 'repricing' of executive stock options." *Journal of Financial Economics* 57, 129–154.

[16] Christie, Andrew A. (1982): "The stochastic behavior of common stock variances: Value, leverage and interest rate effects." *Journal of Financial Economics* 4, 407–432.

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

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

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

[20] Gale, Douglas and Martin Hellwig (1985): "Incentive-compatible debt contracts: The one-period problem." *Review of Economic Studies* 52, 647–663.

[21] Garen, John E. (1994): "Executive compensation and principal–agent theory." *Journal of Political Economy* 102, 1175–1199.

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

[23] 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.

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

[25] Hartman-Glaser, Barney, and Benjamin Hébert (2020): "The insurance is the lemon: Failing to index contracts." *Journal of Finance* 75, 463–506.

[26] Hemmer (2004): "Lessons lost in linearity: A critical assessment of the general usefulness of LEN models in compensation research." *Journal of Management Accounting Research* 16, 149–162.

[27] Hemmer, Thomas, Oliver Kim and Robert E. Verrecchia (1999): "Introducing convexity into optimal compensation contracts." *Journal of Accounting and Economics* 28, 307–328.

[28] Harford, Jarrad (2005): "What drives merger waves?" *Journal of Financial Economics* 77, 529–560.

[29] Hoberg, Gerard, Gordon Phillips, and Nangpurnanand Prabhala (2013): "Product market threats, payouts, and financial flexibility." *Journal of Finance* 69, 293–324.

[30] Hoffmann, Florian and Sebastian Pfeil (2010): "Reward for luck in a dynamic agency model." *Review of Financial Studies* 23, 3329–3345.

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

[32] Holmström, Bengt and Jean Tirole (1993): "Market liquidity and performance monitoring." *Journal of Political Economy* 101, 678–709.

[33] Holmström, Bengt (2017): "Pay for performance and beyond." *American Economic Review* 107, 1753–1777

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

[35] Innes, Robert D. (1993): "Debt, futures and options: Optimal price-linked financial contracts under moral hazard and limited liability." *International Economic Review* 34, 271–295.

[36] Jewitt, Ian (1988): "Justifying the first-order approach to principal–agent problems." *Econometrica* 56, 1177–1190.

[37] Jewitt, Ian, Ohad Kadan and Jeroen M. Swinkels (2008): "Moral hazard with bounded payments." *Journal of Economic Theory* 143, 59–82.

[38] Johnson, Shane A. and Yisong S. Tian (2000): "Indexed executive stock options." *Journal of Financial Economics* 57, 35–64.

[39] Jung, Jin Yong and Son Ku Kim (2015): "Information space conditions for the first-order approach in agency problems." *Journal of Economic Theory* 160, 243–279.

[40] Kadan, Ohad and Jeroen M. Swinkels (2008): "Stocks or options? Moral hazard, firm viability, and the design of compensation contracts." *Review of Financial Studies* 21, 451–482.

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

[42] Kirkegaard, René (2017): "A unifying approach to incentive compatibility in moral hazard problems. " *Theoretical Economics* 12, 25–51.

[43] Koziol, Christian and Jochen Lawrenz (2010): "Optimal design of rating-triger step-up bonds: Agency conflicts versus asymmetric information." *Journal of Corporate Finance* 16, 182–204.

[44] Liljeblom, Eva, Daniel Pasternack, and Matts Rosenberg (2011): "What determines stock option contract design?" *Journal of Financial Economics* 102, 293–316.

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

[46] Marinovic, Ivan and Felipe Varas (2019): "CEO horizon, optimal pay duration, and the escalation of short-termism." *Journal of Finance* 74, 2011–2053.

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

[48] Murphy, Kevin J. (2013): "Executive compensation: Where we are, and how we got there." *Handbook of the Economics of Finance Volume 2A: Corporate Finance*, edited by George M. Constantinides, Milton Harris, and Rene M. Stulz, 211–356. Elsevier, North-Holland.

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

[50] Prendergast, Canice (2002): "The tenuous trade-off between risk and incentives." *Journal of Political Economy* 110, 1071–1102.

[51] Rogerson, William P. (1985): "The first order approach to principal–agent problems." *Econometrica* 53, 1357–1368.

[52] Sinclair-Desgagné, Bernard (1994): "The first-order approach to multi-signal principal-agent problems." *Econometrica* 62, 459–466.

[53] Tchistyi, Alexei (2009): "Security design with correlated hidden cash flows: The optimality of performance pricing." Working paper, University of Illinois.

[54] Townsend, Robert M. (1979): "Optimal contracts and competitive markets with costly state verification." *Journal of Economic Theory* 21, 265–293.

A Proofs

Proof of Lemma 1: For a given contract $w_s(q)$, the effort choice problem of the agent can be written successively as:

$$\max_e \mathbb{E}\left[u\left(\bar{W}+w_s(q)\right)\right]-C(e) \;\Leftrightarrow\; \max_e \sum_s \int_{\underline{q}}^{+\infty} u\left(\bar{W}+w_s(q)\right) f(q,s|e)dq - C(e)$$

The second derivative of the agent's objective function with respect to e is negative for any e if and only if:

$$\sum_s \int_{\underline{q}}^{+\infty} u\left(\bar{W}+w_s(q)\right) \frac{\partial^2 f(q,s|e)}{\partial e^2} dq < C''(e) \quad \forall e \in (0,\bar{e}). \tag{22}$$

Assume that u is bounded from above, with $\lim_{w\to\infty} u(w) \equiv u^+$. In addition, with limited liability, the minimum payment is $w_s(q)=0$; with an increasing utility function, this implies that the minimum value of u is $u(\bar{W})$. Therefore, for any $\{q,s\}$:

$$u\left(\bar{W}+w_s(q)\right) \in \left[u(\bar{W}),u^+\right]$$

Using notations K_e^+ and K_e^- defined in equations (6) and (7), the expression on the LHS of equation (22) can then be rewritten as:

$$\sum_s \int_{\underline{q}}^{+\infty} u\left(\bar{W}+w_s(q)\right) \min\left\{\frac{\partial^2 f(q,s|e)}{\partial e^2},0\right\}dq + \sum_s \int_{\underline{q}}^{+\infty} u\left(\bar{W}+w_s(q)\right) \max\left\{\frac{\partial^2 f(q,s|e)}{\partial e^2},0\right\}dq \tag{23}$$

As established above, we have $u\left(\bar{W}+w_s(q)\right) \geq u(\bar{W})$ for any q,s, and $u\left(\bar{W}+w_s(q)\right) \leq u^+$ for any q,s. Therefore, for any q,s such that $\frac{\partial^2 f(q,s|e)}{\partial e^2} \leq 0$ we have $u\left(\bar{W}+w_s(q)\right)\frac{\partial^2 f(q,s|e)}{\partial e^2} \leq u(\bar{W})\frac{\partial^2 f(q,s|e)}{\partial e^2}$; and for any q,s such that $\frac{\partial^2 f(q,s|e)}{\partial e^2} \geq 0$ we have $u\left(\bar{W}+w_s(q)\right)\frac{\partial^2 f(q,s|e)}{\partial e^2} \leq u^+\frac{\partial^2 f(q,s|e)}{\partial e^2}$. Integrating and summing over q and s, this implies that the expression in equation (23) is less than $K_e^- u(\bar{W}) + K_e^+ u^+$, which completes the proof. ■

Proof of Lemma 2: For now we ignore the LL (5). By Lemma 1, when (8) holds, we can replace the IC in (4) by the first-order condition ("FOC"):

$$\sum_s \left[\frac{\partial \phi_{\hat{e}}^s}{\partial e}\int_0^{+\infty} u\left(\bar{W}+w_s\left(q\right)\right) f(q|\hat{e},s)dq + \phi_{\hat{e}}^s \int_0^{+\infty} u\left(\bar{W}+w_s\left(q\right)\right)\frac{\partial f}{\partial e}(q|\hat{e},s)dq\right] = C'(\hat{e}). \tag{24}$$

The FOC with respect to $w_s(q)$ in the program in (2), (3), and (4) is:

$$\begin{aligned} & \phi_{\hat{e}}^s f(q|\hat{e},s) - \lambda\phi_{\hat{e}}^s u'(\bar{W}+w_s(q)) f(q|\hat{e},s) - \mu u'(\bar{W}+w_s(q)) \left[\frac{\partial\phi_{\hat{e}}^s}{\partial e} f(q|\hat{e},s) + \phi_{\hat{e}}^s \frac{\partial f}{\partial e}(q|\hat{e},s)\right] = 0 \\ \Leftrightarrow \quad & \frac{1}{u'(\bar{W}+w_s(q))} = \lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right]. \end{aligned} \tag{25}$$

Due to LL, we have $\underline{m}(q) = \bar{W}$ and $\overline{m}(q) = \infty$, using the notations in Jewitt, Kadan, and Swinkels (2008). Using the FOC in (25), the same reasoning as in Proposition 1 in Jewitt, Kadan, and Swinkels (2008) applies for any given signal realization s, so that the optimal contract for a given s is defined implicitly by:

$$\frac{1}{u'(\bar{W}+w_s(q))} = \begin{cases} \lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right] & \text{if } \lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right] \geq \frac{1}{u'(\bar{W})}, \\ \frac{1}{u'(\bar{W})} & \text{if } \lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right] < \frac{1}{u'(\bar{W})}, \end{cases} \tag{26}$$

with $\lambda \geq 0$ and $\mu > 0$. This can be rewritten as equation (9). ■

Proof of Proposition 1: From equation (26), we have $w_s(q) > 0$ if and only if:

$$\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right] > \frac{1}{u'(\bar{W})}.$$

If the support of the output distribution is $[0,\infty)$ and if, for a given s,

$$\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(0|\hat{e},s)}{f(0|\hat{e},s)}\right] > \frac{1}{u'(\bar{W})}, \tag{27}$$

then $q_s^* = 0$. Otherwise (if the support is $[0,\infty)$ and equation (27) does not hold, or if the support is instead $(-\infty,\infty)$), since $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ is increasing in q by MLRP, for $u'(\bar{W}) > 0$ the signal-contingent performance threshold q_s^* is implicitly defined by:

$$\begin{aligned} \lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q_s^*|\hat{e},s)}{f(q_s^*|\hat{e},s)}\right] &= \frac{1}{u'(\bar{W})} \\ \Leftrightarrow \qquad \frac{\frac{\partial f}{\partial e}(q_s^*|\hat{e},s)}{f(q_s^*|\hat{e},s)} &= \frac{1}{\mu}\left[\frac{1}{u'(\bar{W})} - \lambda\right] - \frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} \end{aligned} \tag{28}$$

where the RHS is independent of q, and $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ is increasing in q by MLRP.

Using equation (11) that defines the likelihood ratio of output, we have:

$$\frac{d}{dq}\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}=-\frac{\zeta_s}{\sigma_s^2}\frac{g''\left(\frac{q-\xi_s}{\sigma_s}\right)g\left(\frac{q-\xi_s}{\sigma_s}\right)-\left(g'\left(\frac{q-\xi_s}{\sigma_s}\right)\right)^2}{\left(g\left(\frac{q-\xi_s}{\sigma_s}\right)\right)^2}. \tag{29}$$

To ease notation, let

$$G(q):=-\frac{g''\left(\frac{q-\xi_s}{\sigma_s}\right)g\left(\frac{q-\xi_s}{\sigma_s}\right)-\left(g'\left(\frac{q-\xi_s}{\sigma_s}\right)\right)^2}{\left(g\left(\frac{q-\xi_s}{\sigma_s}\right)\right)^2}. \tag{30}$$

Due to MLRP, equations (29) and (30) imply that $G(q)>0\ \forall q$. From equation (11), since $\zeta_s>0$, $\sigma_s>0$, and $g(\cdot)>0$, a distribution with location and scale parameters that satisfies MLRP is such that $g'(\cdot)>0$ if q is lower than a threshold, and $g'(\cdot)<0$ if q is higher than this threshold, i.e., the probability density function ("PDF") g is single peaked – the output corresponding to the peak is denoted by q_s^P.

For part (i), there are four cases to consider:

- If the support of the output distribution is $[0,\infty)$ and equation (27) holds for s_i and s_j, then $q_{s_i}^*=q_{s_j}^*=0$.
- If the support is $[0,\infty)$ and (27) holds for s_i but not s_j, then $q_{s_i}^*=0<q_{s_j}^*$.
- If the support is $[0,\infty)$, it is impossible for (27) to hold for s_j but not s_i since $LR_{s_i}(q)>LR_{s_j}(q)\,\forall q$.
- If the support is $[0,\infty)$ and (27) neither holds for s_i nor for s_j, or if the support is instead $(-\infty,\infty)$, $q_{s_i}^*$ and $q_{s_j}^*$ are both described by equation (28). Since $\mu>0$ and $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ is increasing in q by MLRP, it follows from equation (28) that $\frac{\partial\phi_{\hat{e}}^{s_i}/\partial e}{\phi_{\hat{e}}^{s_i}}>\frac{\partial\phi_{\hat{e}}^{s_j}/\partial e}{\phi_{\hat{e}}^{s_j}}$ is associated with $q_{s_i}^*<q_{s_j}^*$, all else equal across signals.

For part (ii), there are the same four cases to consider. The first three cases are exactly the same as part (i). For the fourth case, $q_{s_i}^*$ and $q_{s_j}^*$ are both described by equation (28). For $\xi_{s_i}<\xi_{s_j}$, all else equal across signals we have $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s_i)}{f(q|\hat{e},s_i)}>\frac{\frac{\partial f}{\partial e}(q|\hat{e},s_j)}{f(q|\hat{e},s_j)}$ for any q. Then by MLRP and $\mu>0$, it follows from equation (28) that $q_{s_i}^*<q_{s_j}^*$, all else equal across signals.

For part (iii), using equation (11) we have:

$$\frac{\partial \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}}{\partial \zeta_s} = \underbrace{-\frac{1}{\sigma_s}}_{\leq 0} \underbrace{\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)}}_{\geq 0 \text{ for } q \leq q_s^P} . \tag{31}$$

Changes in ζ_s do not change the peak of the distribution (such that $f'(q|\hat{e},s) = 0$), which is the same for s_i and s_j, and denoted by q_s^P. There are four cases to consider:

- If the support is $[0, \infty)$ and (27) holds for s_i and s_j, then $q^*_{s_i} = q^*_{s_j} = 0$.
- If the support is $[0, \infty)$ and (27) holds for s_i but not s_j, then $q^*_{s_i} = 0 < q^*_{s_j}$. Given MLRP and equation (31), this is possible for $q_s^* > q_s^P$ but not for $q_s^* < q_s^P$ $(s \in \{s_i, s_j\})$.
- If the support is $[0, \infty)$ and (27) holds for s_j but not s_i, then $q^*_{s_j} = 0 < q^*_{s_i}$. Given MLRP and equation (31), this is possible for $q_s^* < q_s^P$ but not for $q_s^* > q_s^P$ $(s \in \{s_i, s_j\})$.
- If the support is $[0, \infty)$ and (27) does not hold, or if the support is instead $(-\infty, \infty)$, $q^*_{s_i}$ and $q^*_{s_j}$ are both described by equation (28). Again, since $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ is increasing in q by MLRP, a higher likelihood ratio of output implies a lower threshold q_s^*, as implicitly defined in equation (28). For $q_s^* \leq q_s^P$, the RHS of equation (31) is negative at $q = q_s^*$, so that a higher ζ_s is associated with a higher threshold q_s^*. For $q_s^* \geq q_s^P$, the RHS of equation (31) is positive at $q = q_s^*$, so that a higher ζ_s is associated with a lower threshold q_s^*.

For part (iv), using equation (11) we have:

$$\frac{\partial \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}}{\partial \sigma_s} = \frac{\zeta_s}{\sigma_s^2} \underbrace{\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)}}_{> 0 \text{ for } q < q_s^P} + \frac{\zeta_s}{\sigma_s} \underbrace{G(q)}_{\geq 0} \underbrace{\left(\frac{\xi_s - q}{\sigma_s^2}\right)}_{> 0 \text{ for } q < \xi_s} .$$

We also have $\zeta_s > 0$ and $\sigma_s > 0$, so that:

$$\text{sign}\left(\frac{\partial \frac{\frac{\partial f}{\partial e}(q_s^*|\hat{e},s)}{f(q_s^*|\hat{e},s)}}{\partial \sigma_s}\right) \begin{cases} > 0 & \text{if } q_s^* < q_s^P \text{ and } q_s^* < \xi_s \\ < 0 & \text{if } q_s^* > q_s^P \text{ and } q_s^* > \xi_s \end{cases} . \tag{32}$$

There are four cases to consider:

- If the support is $[0,\infty)$ and (27) holds for s_i and s_j, then $q^*_{s_i} = q^*_{s_j} = 0$.
- If the support is $[0,\infty)$ and (27) holds for s_i but not s_j, then $q^*_{s_i} = 0 < q^*_{s_j}$. Given MLRP and equation (32), this is possible for $q^*_s < q^P_s$ and $q^*_s < \xi_s$ but not for $q^*_s > q^P_s$ and $q^*_s > \xi_s$ ($s \in \{s_i, s_j\}$).
- If the support is $[0,\infty)$ and (27) holds for s_j but not s_i, then $q^*_{s_j} = 0 < q^*_{s_i}$. Given MLRP and equation (32), this is possible for $q^*_s > q^P_s$ and $q^*_s > \xi_s$ but not for $q^*_s < q^P_s$ and $q^*_s < \xi_s$ ($s \in \{s_i, s_j\}$).
- If the support is $[0,\infty)$ and (27) does not hold, or if the support is instead $(-\infty,\infty)$, $q^*_{s_i}$ and $q^*_{s_j}$ are both described by equation (28). Since $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ is increasing in q by MLRP, a higher likelihood ratio of output implies a lower threshold q^*_s, as implicitly defined in equation (28); the relation between signal volatility and the threshold is then given by equation (32). ∎

Proof of Corollary 1: The threshold q^*_s is implicitly defined in equation (28). With $\mu > 0$, $u' > 0$, a non-binding IR ($\lambda = 0$) and a signal realization which is either individually uninformative or bad news about effort ($\partial \phi^s_{\hat{e}}/\partial e \leq 0$), the threshold is such that:

$$\frac{\frac{\partial f}{\partial e}(q^*_s|\hat{e},s)}{f(q^*_s|\hat{e},s)} = \frac{1}{\mu}\frac{1}{u'(\bar{W})} - \frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} > 0$$

Since $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ is increasing in q by MLRP, an increase in the likelihood ratio of output implies a lower threshold q^*_s. Using equation (32) from the proof of Proposition 1, we know that:

$$\text{sign}\left(\frac{dq^*_s}{d\sigma_s}\right) \begin{cases} \leq 0 & \text{if } q^*_s \leq q^P_s \text{ and } q^*_s \leq \xi_s \\ \geq 0 & \text{if } q^*_s \geq q^P_s \text{ and } q^*_s \geq \xi_s \end{cases} \tag{33}$$

For a distribution with location and scale parameters (as defined in equation (10)), $\lambda = 0$, and $\partial \phi^s_{\hat{e}}/\partial e \leq 0$, equation (28) rewrites as:

$$-\frac{\zeta_s}{\sigma_s}\frac{g'\left(\frac{q^*_s-\xi_s}{\sigma_s}\right)}{g\left(\frac{q^*_s-\xi_s}{\sigma_s}\right)} = \frac{1}{\mu}\frac{1}{u'(\bar{W})} - \frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} > 0 \tag{34}$$

The output q_s^P corresponding to the peak of the output distribution is defined implicitly by:

$$\frac{1}{\sigma_s^2} g'\left(\frac{q_s^P - \xi_s}{\sigma_s}\right) = 0. \tag{35}$$

For a single-peaked distribution, equations (34) and (35) and MLRP imply that $q_s^* > q_s^P$. In this case, we know from equation (33) that q_s^* is increasing in σ_s if $q_s^* \geq \xi_s$.

Under a skew-normal distribution for q with location parameter $h_s(e)$, scale parameter σ_s, and shape parameter α_s, the likelihood ratio of output is:

$$\frac{\frac{\partial f}{\partial e}(q|\hat{e}, s)}{f(q|\hat{e}, s)} = \frac{1}{\sigma_s}\left[\frac{q - \xi_s}{\sigma_s} - \alpha_s \frac{\varphi\left(\alpha_s \frac{q-\xi_s}{\sigma_s}\right)}{\Phi\left(\alpha_s \frac{q-\xi_s}{\sigma_s}\right)}\right], \tag{36}$$

In addition, we know from equations (11) and (35) that the likelihood ratio of output is zero at the peak q_s^P of the distribution under signal s. Using equation (36) with $\alpha_s \geq 0$, this implies that $q_s^P \geq \xi_s$, with an equality for $\alpha_s = 0$. In sum, with a skew-normal distribution, we have $q_s^P \geq \xi_s$ if and only if the distribution has nonnegative skewness. In this case, $q_s^* > q_s^P$ implies $q_s^* > \xi_s$.

In sum, when the participation constraint is non-binding and the signal realization is such that $\partial \phi_{\hat{e}}^s / \partial e \leq 0$, for a skew-normal distribution with nonnegative skewness we have $\text{sign}\left(\frac{dq_s^*}{d\sigma_s}\right) \geq 0$. ■

Proof of Proposition 2: From equation (11), since $\zeta_s > 0$, $\sigma_s > 0$, and $g(\cdot) > 0$, a distribution with location and scale parameters that satisfies MLRP is such that $g'(\cdot) > 0$ if q is lower than a threshold, and $g'(\cdot) < 0$ if q is higher than this threshold, i.e., the PDF g is single peaked – the output corresponding to the peak is denoted by q_s^P. For symmetric distributions, the single peak of the distribution, which is such that $g'\left(\frac{q_s^P - \xi_s}{\sigma_s}\right) = 0$, is at $q_s^P = \xi_s$ for $s \in \{s_i, s_j\}$. With CRRA utility, characterized by $u'(w) = w^{-\gamma}$ and $u'^{-1}(w) = w^{-\frac{1}{\gamma}}$, provided that the FOA holds we have:

$$w_s(q) = \begin{cases} \left(\lambda + \mu\left[\frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right]\right)^{\frac{1}{\gamma}} - \bar{W} & \text{if } \lambda + \mu\left[\frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right] \geq \frac{1}{u'(\bar{W})} \\ 0 & \text{if } \lambda + \mu\left[\frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right] < \frac{1}{u'(\bar{W})} \end{cases}. \tag{37}$$

For $\gamma > 1$, we use the condition for the FOA in Lemma 1. For $\gamma \leq 1$, we derive a condition for FOA to hold in this setting. The FOA holds if:

$$\sum_s \int_{\underline{q}}^{+\infty} u\left(\bar{W} + w_s(q)\right) \frac{\partial^2 f(q,s|e)}{\partial e^2} dq < C''(e) \quad \forall e \in (0, \bar{e}), \tag{38}$$

where $u(w) = \frac{w^{1-\gamma}}{1-\gamma}$ if $\gamma < 1$ and $\ln(w)$ if $\gamma = 1$, and $w_s\left(q\right)$ is defined by equation (37).

Part (i): Suppose that signals s_i and s_j differ only in their individual informativeness: $\frac{\partial \phi_{\hat{e}}^{s_i}/\partial e}{\phi_{\hat{e}}^{s_i}} > \frac{\partial \phi_{\hat{e}}^{s_j}/\partial e}{\phi_{\hat{e}}^{s_j}}$ (s_i and s_j are associated with the same output distribution). For notational convenience, let $\tilde{\phi}_s \equiv \frac{\partial \phi_{\hat{e}}^{s}/\partial e}{\phi_{\hat{e}}^{s}}$ and $w_s(q) \equiv W_q(\tilde{\phi}_s)$, which is a continuous function of $\tilde{\phi}_s$. For a given q, we can write:

$$w_{s_i}(q) - w_{s_j}(q) = W_q(\tilde{\phi}_i) - W_q(\tilde{\phi}_j) = \int_{\tilde{\phi}_j}^{\tilde{\phi}_i} \frac{\partial W_q(\tilde{\phi})}{\partial \tilde{\phi}} d\tilde{\phi}.$$

Holding all else constant including Lagrange multipliers (we are comparing two signal realizations, i.e. we do not change parameters of the contracting environment), for given q and s such that $w_s(q) > 0$:

$$\begin{aligned}
\frac{\partial W_q(\tilde{\phi})}{\partial \tilde{\phi}} &= \frac{\partial}{\partial \tilde{\phi}} \left\{ \left(\lambda + \mu \left[\frac{\partial \phi_{\hat{e}}^{s}/\partial e}{\phi_{\hat{e}}^{s}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}} - \bar{W} \right\} \\
&= \underbrace{\frac{\mu}{\gamma}}_{>0} \underbrace{\left(\lambda + \mu \left[\frac{\partial \phi_{\hat{e}}^{s}/\partial e}{\phi_{\hat{e}}^{s}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1}}_{\geq 0 \text{ since } w_s(q) \geq 0},
\end{aligned} \tag{39}$$

For given q and q_0, we have

$$\frac{w_{s_i}(q) - w_{s_i}(q_0)}{q - q_0} \geq \frac{w_{s_j}(q) - w_{s_j}(q_0)}{q - q_0} \quad \Leftrightarrow \quad w_{s_i}(q) - w_{s_j}(q) - \left[w_{s_i}(q_0) - w_{s_j}(q_0)\right] \geq 0. \tag{40}$$

Thus, for given q and q_0 such that $q > q_0$ and $w_s(q_0) > 0$, we have:

$$\begin{aligned}
& w_{s_i}(q) - w_{s_j}(q) - \left[w_{s_i}(q_0) - w_{s_j}(q_0)\right] \\
= \; & \frac{\mu}{\gamma} \int_{\tilde{\phi}_j}^{\tilde{\phi}_i} \left[\left(\lambda + \mu \left[\frac{\partial \phi_{\hat{e}}^{s}/\partial e}{\phi_{\hat{e}}^{s}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} - \left(\lambda + \mu \left[\frac{\partial \phi_{\hat{e}}^{s}/\partial e}{\phi_{\hat{e}}^{s}} + \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} \right] d\tilde{\phi}.
\end{aligned}$$

We have: $q > q_0$ which implies $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} > \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)}$ by MLRP, so that, for a given $\tilde{\phi}_s$:

$$\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \;>\; \lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)} \right]$$

$$\left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right]\right)^{\frac{1}{\gamma}-1} \;>\; \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)} \right]\right)^{\frac{1}{\gamma}-1} \quad \text{iff } \gamma < 1. \tag{41}$$

If $\gamma < 1$, the condition in (41) holds, and so (40) also holds. If $\gamma < 1$, (41) does not hold, and so (40) does not either. If $\gamma = 1$, the equation in (41) is satisfied as an equality so that the expression on the right in (40) is equal to zero.

Part (ii): Suppose that signals s_i and s_j differ only in their equilibrium location parameter, with $\xi_i > \xi_j$. Let $w_s(q) \equiv W_q(\xi_s)$, which is a continuous function of ξ_s, since $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ is by assumption continuously differentiable in the equilibrium location parameter ξ_s. For a given q, we have:

$$w_{s_i}(q) - w_{s_j}(q) = W_q(\xi_i) - W_q(\xi_j) = \int_{\xi_j}^{\xi_i} \frac{\partial W_q(\xi)}{\partial \xi} d\xi.$$

Holding all else constant including Lagrange multipliers, for given q and s such that $w_s(q) > 0$, we have:

$$\begin{aligned}
\frac{\partial W_q(\xi)}{\partial \xi} &= \frac{\partial}{\partial \xi}\left\{ \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right]\right)^{\frac{1}{\gamma}} - \bar{W} \right\} \\
&= \underbrace{\frac{\mu}{\gamma}}_{>0} \frac{\partial}{\partial \xi}\left\{ \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right\} \underbrace{\left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right]\right)^{\frac{1}{\gamma}-1}}_{\geq 0 \text{ since } w_s(q) \geq 0},
\end{aligned} \tag{42}$$

where:

$$\begin{aligned}
\frac{\partial}{\partial \xi}\left\{ \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right\} &= \frac{\partial}{\partial \xi}\left\{ -\frac{\zeta_s}{\sigma_s} \frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)} \right\} \\
&= -\frac{\zeta_s}{\sigma_s^2} \underbrace{G(q)}_{\geq 0},
\end{aligned} \tag{43}$$

where G is defined in equation (30). For given q and q_0, we have:

$$\frac{w_{s_i}(q) - w_{s_i}(q_0)}{q - q_0} \geq \frac{w_{s_j}(q) - w_{s_j}(q_0)}{q - q_0} \quad \Leftrightarrow \quad w_{s_i}(q) - w_{s_j}(q) - \big[w_{s_i}(q_0) - w_{s_j}(q_0)\big] \geq 0. \quad (44)$$

Thus, for given q and q_0 such that $q > q_0$ and $w_s(q_0) > 0$, we have:

$$\begin{aligned} & w_{s_i}(q) - w_{s_j}(q) - \big[w_{s_i}(q_0) - w_{s_j}(q_0)\big] \\ = \; & \frac{\mu}{\gamma} \frac{\zeta_s}{\sigma_s^2} \int_{\xi_j}^{\xi_i} \Bigg[- G(q) \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}} / \partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} \\ & + G(q_0) \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}} / \partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} \Bigg] d\xi. \end{aligned} \quad (45)$$

From equation (29) and the definition of $G(q)$ in equation (30), the likelihood ratio of output is weakly concave in q if and only if $G'(q) \leq 0$ in which case $0 \leq G(q) \leq G(q_0)$ since $q > q_0$. In addition, $q > q_0$ implies $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} > \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)}$ by MLRP, so that, for a weakly concave likelihood ratio and $\gamma \geq 1$:

$$\begin{aligned} \lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}} / \partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] &> \lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}} / \partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)} \right] \\ \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}} / \partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} &\leq \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}} / \partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} \\ G(q) \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}} / \partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} &\leq G(q_0) \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}} / \partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} \end{aligned} \quad (46)$$

We conclude that if the likelihood ratio is nonconvex and $\gamma \geq 1$, then the condition in (46) holds, and, using equation (45), condition (44) holds too. Symmetrically, if the likelihood ratio is weakly convex (so that $G'(q) \geq 0$) and $\gamma \leq 1$, then the inequality in condition in (46) is reversed, so that it is reversed in condition (44) too. Finally, if the likelihood ratio is linear (so that $G'(q) = 0$) and $\gamma = 1$, then the condition in (46) holds as an equality, and so does condition (44).

Part (iii): Suppose that signals s_i and s_j differ only in their impact parameter, with $\zeta_i > \zeta_j$. Let $w_s(q) \equiv W_q(\zeta_s)$, which is a continuous function of ζ_s, since $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ is by assumption continuously differentiable in the parameter ζ_s. For a given q, we have:

$$w_{s_i}(q) - w_{s_j}(q) = W_q(\zeta_i) - W_q(\zeta_j) = \int_{\zeta_j}^{\zeta_i} \frac{\partial W_q(\zeta)}{\partial \zeta} d\zeta. \tag{47}$$

As above, holding all else constant including Lagrange multipliers, for given q and s such that $w_s(q) > 0$:

$$\begin{aligned} \frac{\partial W_q(\zeta)}{\partial \zeta} &= \frac{\partial}{\partial \zeta}\left\{ \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}} - \bar{W} \right\} \\ &= \frac{\mu}{\gamma}\frac{\partial}{\partial \zeta}\left\{ \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right\} \underbrace{\left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1}}_{>0 \text{ since } w_s(q)>0}, \end{aligned} \tag{48}$$

where

$$\frac{\partial}{\partial \zeta}\left\{ \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right\} = \frac{\partial}{\partial \zeta}\left\{ -\frac{\zeta_s}{\sigma_s}\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)} \right\} = \underbrace{-\frac{1}{\sigma_s}}_{<0} \underbrace{\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)}}_{<0 \text{ for } q>q^P_s} . \tag{49}$$

Thus, at output q, we have:

$$\int_{\zeta_j}^{\zeta_i} \frac{\partial W_q(\zeta)}{\partial \zeta} d\zeta = \int_{\zeta_j}^{\zeta_i} \frac{\mu}{\gamma}\frac{1}{\sigma_s} \underbrace{\left(-\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)} \right)}_{>0 \text{ for } q>q^P_s} \underbrace{\left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1}}_{>0 \text{ since } w_s(q)>0} d\zeta.$$

In sum, with $q > q_0 > \max\{q^P_s, q^*_s, \xi_s\}$, we have:

$$\begin{aligned} w_{s_i}(q) - w_{s_j}(q) - \left[w_{s_i}(q_0) - w_{s_j}(q_0) \right] &= W_q(\zeta_i) - W_q(\zeta_j) - \left[W_{q_0}(\zeta_i) - W_{q_0}(\zeta_j) \right] \\ = \frac{\mu}{\gamma}\frac{1}{\sigma_s}\int_{\zeta_j}^{\zeta_i} \Bigg[&\left(-\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)} \right) \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} \\ &- \left(-\frac{g'\left(\frac{q_0-\xi_s}{\sigma_s}\right)}{g\left(\frac{q_0-\xi_s}{\sigma_s}\right)} \right) \left(\lambda + \mu \left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)} \right] \right)^{\frac{1}{\gamma}-1} \Bigg] d\zeta. \end{aligned} \tag{50}$$

For $q > \max\{q_s^P, q_s^*, \xi_s\}$, both $-\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)}$ and $\left(\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right]\right)^{\frac{1}{\gamma}-1}$ are positive and weakly increasing in q (by MLRP) if $\gamma \leq 1$. Therefore, if $\gamma \leq 1$, expression (50) is positive. Using equation (44), this means that, with $\zeta_i > \zeta_j$, the PPS measure $\frac{w_s(q)-w_s(q_0)}{q-q_0}$ is higher under s_i than under s_j.

Part (iv): Suppose that signals s_i and s_j differ only in their scale parameter, with $\sigma_i > \sigma_j$. Let $w_s(q) \equiv W_q(\sigma_s)$, which is a continuous function of σ_s, since $\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ is by assumption continuously differentiable in the scale parameter σ_s. For a given q, we have:

$$w_{s_i}(q) - w_{s_j}(q) = W_q(\sigma_i) - W_q(\sigma_j) = \int_{\sigma_j}^{\sigma_i} \frac{\partial W_q(\sigma)}{\partial \sigma} d\sigma. \tag{51}$$

Holding all else constant including Lagrange multipliers, for given q and s such that $w_s(q) > 0$:

$$\begin{aligned}
\frac{\partial W_q(\sigma)}{\partial \sigma} &= \frac{\partial}{\partial \sigma}\left\{\left(\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right]\right)^{\frac{1}{\gamma}} - \bar{W}\right\} \\
&= \underbrace{\frac{\mu}{\gamma}}_{>0} \frac{\partial}{\partial \sigma}\left\{\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right\} \underbrace{\left(\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right]\right)^{\frac{1}{\gamma}-1}}_{>0 \text{ since } w_s(q)>0},
\end{aligned}$$

where:

$$\begin{aligned}
\frac{\partial}{\partial \sigma}\left\{\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right\} &= \frac{\partial}{\partial \sigma}\left\{-\frac{\zeta_s}{\sigma_s}\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)}\right\} \\
&= \underbrace{\frac{\zeta_s}{\sigma_s^2}}_{>0} \underbrace{\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)}}_{<0 \text{ for } q>q_s^P} + \underbrace{\frac{\zeta_s}{\sigma_s}}_{>0} \underbrace{G(q)}_{\geq 0} \underbrace{\frac{\xi_s - q}{\sigma_s^2}}_{<0 \text{ for } q>\xi_s}.
\end{aligned} \tag{52}$$

So, at output q:

$$\int_{\sigma_j}^{\sigma_i} \frac{\partial W_q(\sigma)}{\partial \sigma} d\sigma = \int_{\sigma_j}^{\sigma_i} \frac{\mu}{\gamma}\frac{\zeta_s}{\sigma_s^2}\left(\underbrace{\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)} + G(q)\frac{\xi_s - q}{\sigma_s}}_{<0 \text{ for } q>\max\{\xi_s, q_s^P\}}\right)\underbrace{\left(\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right]\right)^{\frac{1}{\gamma}-1}}_{>0 \text{ since } w_s(q)>0} d\sigma$$

In sum, with $q > q_0 > \max\{q_s^P, q_s^*, \xi_s\}$:

$$\begin{aligned} w_{s_i}(q) - w_{s_j}(q) - \left[w_{s_i}(q_0) - w_{s_j}(q_0)\right] &= W_q(\sigma_i) - W_q(\sigma_j) - [W_{q_0}(\sigma_i) - W_{q_0}(\sigma_j)] \\ = \frac{\mu}{\gamma}\frac{\zeta_s}{\sigma_s^2}\int_{\sigma_j}^{\sigma_i} \Bigg[& \left(\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)} + G(q)\frac{\xi_s - q}{\sigma_s}\right)\left(\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right]\right)^{\frac{1}{\gamma}-1} \\ & - \left(\frac{g'\left(\frac{q_0-\xi_s}{\sigma_s}\right)}{g\left(\frac{q_0-\xi_s}{\sigma_s}\right)} + G(q)\frac{\xi_s - q_0}{\sigma_s}\right)\left(\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q_0|\hat{e},s)}{f(q_0|\hat{e},s)}\right]\right)^{\frac{1}{\gamma}-1}\Bigg]d\sigma \end{aligned} \tag{53}$$

For $q > \max\{q_s^P, q_s^*, \xi_s\}$, $\left(\frac{g'\left(\frac{q-\xi_s}{\sigma_s}\right)}{g\left(\frac{q-\xi_s}{\sigma_s}\right)} + G(q)\frac{\xi_s-q}{\sigma_s}\right)$ is negative and weakly decreasing in q if the likelihood ratio of output is weakly convex (by MLRP and since then $G'(q) \geq 0$, as per point (ii) above), while $\left(\lambda + \mu\left[\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}\right]\right)^{\frac{1}{\gamma}-1}$ is positive and weakly increasing in q (by MLRP) if $\gamma \leq 1$. Therefore, if $\gamma \leq 1$ and the likelihood ratio of output is weakly convex, expression (53) is negative. Using inequality (44), this means that, with $\sigma_i > \sigma_j$, the PPS measure $\frac{w_s(q)-w_s(q_0)}{q-q_0}$ is lower under s_i than under s_j if the likelihood ratio of output is weakly convex and $\gamma \leq 1$. ■

Proof of Lemma 3: We start by characterizing the optimal contract that induces effort $\hat{e}$. For effort e, we have $q \sim \mathcal{N}(h_s(e), \sigma_s^2)$, with $h_s'(e) > 0$ and $h_s(0) = 0$. Let φ_s be the PDF of the normal distribution with mean zero and standard deviation σ_s.

Letting $W_s(q) = w_s(q) + \bar{W}$ to simplify notation, the IC is:

$$\hat{e} \in \arg\max_{e \in E} \sum_s \phi_e^s \int \ln\left[W_s(q)\right]\varphi_s\left(q - h_s(e)\right)dq - c(e).$$

The IR and LL are, respectively:

$$\sum_s \phi_{\hat{e}}^s \int \ln\left[W_s(q)\right]\varphi_s\left(q - h_s(\hat{e})\right)dq - c(\hat{e}) \geq 0,$$

and

$$W_s(q) \geq \bar{W}\ \forall q, s.$$

To simplify the analysis, we will work with the manager's indirect utility:

$$u_s(q) := \ln\left[W_s(q)\right],$$

so that $W_s(q) = \exp[u_s(q)]$. This part is without loss of generality. The next step, which relies on the FOA, is to replace the IC by its FOC:

$$\sum_s \frac{\partial \phi^s_{\hat{e}}}{\partial e} \int u_s(q)\, \varphi_s(q - h_s(\hat{e}))\, dq + \sum_s \phi^s_{\hat{e}} \int u_s(q)\, \zeta_s \frac{q - h_s(\hat{e})}{\sigma_s^2} \varphi_s(q - \hat{e})\, dq - c'(\hat{e}) \begin{cases} \geq 0 & \text{if } \hat{e} = \bar{e} \\ = 0 & \text{if } \hat{e} \in (0, \bar{e}) \end{cases}, \tag{54}$$

This FOC rewrites as:

$$\sum_s \int u_s(q)\, \varphi_s(q - h_s(\hat{e})) \left[\frac{d\phi^s_{\hat{e}}}{de} + \phi^s_{\hat{e}} \zeta_s \frac{q - h_s(\hat{e})}{\sigma_s^2} \right] dq - c'(\hat{e}) \begin{cases} \geq 0 & \text{if } \hat{e} = \bar{e} \\ = 0 & \text{if } \hat{e} \in (0, \bar{e}) \end{cases}. \tag{55}$$

The principal's relaxed program is:

$$\max_{u_s(\cdot)} \sum_s \phi^s_{\hat{e}} \int \{q - \exp[u_s(q)]\}\, \varphi_s(q - h_s(\hat{e}))\, dq$$

subject to (54),

$$\sum_s \phi^s_{\hat{e}} \int u_s(q)\, \varphi_s(q - h_s(\hat{e}))\, dq - c(\hat{e}) \geq 0, \tag{56}$$

and

$$u_s(q) \geq \ln(\bar{W}) \ \ \forall q, s.$$

The next Lemma solves this optimization problem.

Lemma 6 *The solution of the relaxed program that implements effort $\hat{e} > 0$ satisfies:*

$$W_s(q) = \begin{cases} \bar{W} & \textit{for } q \leq h_s(\hat{e}) + \frac{\sigma_s^2}{\zeta_s}\left[\frac{\bar{W}-\lambda}{\mu} - \frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right] \\ \lambda + \mu\left[\frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s \frac{q - h_s(\hat{e})}{\sigma_s^2}\right] & \textit{for } q > h_s(\hat{e}) + \frac{\sigma_s^2}{\zeta_s}\left[\frac{\bar{W}-\lambda}{\mu} - \frac{\partial \phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right] \end{cases},$$

where $\lambda \geq 0$ and $\mu > 0$.

Proof. The relaxed program maximizes a strictly concave function subject to linear constraints, so the FOC below, the complementary slackness conditions and the constraints are necessary and sufficient. Pointwise optimization gives:

$$\begin{aligned} &-\exp[u_s(q)]\, \phi^s_{\hat{e}} \varphi_s(q - h_s(\hat{e})) + \mu \left[\frac{d\phi^s_{\hat{e}}}{de} + \phi^s_{\hat{e}} \zeta_s \frac{q - h_s(\hat{e})}{\sigma_s^2} \right] \varphi_s(q - h_s(\hat{e})) \\ &+ \lambda \phi^s_{\hat{e}} \varphi_s(q - h_s(\hat{e})) + \lambda_{LL}(q, s) = 0, \end{aligned}$$

where $\lambda_{LL}(q,s)$ are the multipliers associated with the LL. Letting $\tilde{\lambda}_{LL}(q,s) \equiv \frac{\lambda_{LL}(q,s)}{\phi^s_{\hat{e}}\varphi_s(q-h_s(\hat{e}))} \geq 0$, we can rewrite the FOC as:

$$W_s(q) = \lambda + \tilde{\lambda}_{LL}(q,s) + \mu\left[\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q-h_s(\hat{e})}{\sigma^2_s}\right]. \tag{57}$$

There are two cases to consider. First, if $\lambda = 0$ in the optimal contract, it can be verified that the following solves the necessary and sufficient optimality conditions:

$$W_s(q) = \begin{cases} \bar{W} & \text{for } q \leq h_s(\hat{e}) + \frac{\sigma^2_s}{\zeta_s}\left[\frac{\bar{W}}{\mu} - \frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right] \\ \mu\left[\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q-h_s(\hat{e})}{\sigma^2_s}\right] & \text{for } q > h_s(\hat{e}) + \frac{\sigma^2_s}{\zeta_s}\left[\frac{\bar{W}}{\mu} - \frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right] \end{cases}, \tag{58}$$

where μ is chosen so that the IC holds (it can be shown that such $\mu > 0$ exists and is unique). To see this, note that when the LL binds, we have $W_s(q) = \bar{W}$ and $\tilde{\lambda}_{LL}(q,s) \geq 0$. Then, the FOC becomes:

$$\tilde{\lambda}_{LL}(q,s) = \bar{W} - \mu\left[\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q-h_s(\hat{e})}{\sigma^2_s}\right] - \lambda = \bar{W} - \mu\left[\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q-h_s(\hat{e})}{\sigma^2_s}\right] \geq 0,$$

which is positive because $q \leq h_s(\hat{e}) + \frac{\sigma^2_s}{\zeta_s}\left[\frac{\bar{W}}{\mu} - \frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right]$. When the LL does not bind ($W_s(q) > \bar{W}$), we have $\tilde{\lambda}_{LL}(q,s) = 0$, so that the FOC becomes:

$$W_s(q) = \mu\left[\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q-h_s(\hat{e})}{\sigma^2_s}\right],$$

where μ is chosen so that the IC holds. If the resulting contract satisfies the IR in (56), then indeed $\lambda = 0$ at the optimal contract, which is described in (58). This establishes that an option with state-contingent strike price $h_s(\hat{e}) + \frac{\sigma^2_s}{\zeta_s}\left[\frac{\bar{W}}{\mu} - \frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right]$ and impact $n^*_s := \mu\frac{\zeta_s}{\sigma^2_s}$ solves the relaxed program when we are implementing $\hat{e} > 0$.

Now suppose that the resulting contract does not satisfy IR in (56). Then we have $\lambda > 0$. It can be verified that the following solves the necessary and sufficient optimality conditions:

$$W_s(q) = \begin{cases} \bar{W} & \text{for } q \leq h_s(\hat{e}) + \frac{\sigma^2_s}{\zeta_s}\left[\frac{\bar{W}-\lambda}{\mu} - \frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right] \\ \lambda + \mu\left[\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q-h_s(\hat{e})}{\sigma^2_s}\right] & \text{for } q > h_s(\hat{e}) + \frac{\sigma^2_s}{\zeta_s}\left[\frac{\bar{W}-\lambda}{\mu} - \frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right] \end{cases},$$

where λ and μ are chosen so that the IR and IC hold. To see this, note that when the LL binds, we have $W_s(q) = \bar{W}$ and $\tilde{\lambda}_{LL}(q, s) \geq 0$. Then, the FOC becomes:

$$\tilde{\lambda}_{LL}(q, s) = \bar{W} - \mu\left[\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q - h_s(\hat{e})}{\sigma^2_s}\right] - \lambda = \bar{W} - \mu\left[\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q - h_s(\hat{e})}{\sigma^2_s}\right] - \lambda \geq 0,$$

which is positive because $q \leq h_s(\hat{e}) + \frac{\sigma^2_s}{\zeta_s}\left[\frac{\bar{W}-\lambda}{\mu} - \frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right]$. When the LL does not bind ($W_s(q) > \bar{W}$), we have $\tilde{\lambda}_{LL}(q, s) = 0$, so that the FOC becomes:

$$W_s(q) = \lambda + \mu\left[\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q - h_s(\hat{e})}{\sigma^2_s}\right],$$

This establishes that an option with state-contingent strike price $h_s(\hat{e}) + \frac{\sigma^2_s}{\zeta_s}\left[\frac{\bar{W}-\lambda}{\mu} - \frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right]$ and slope $n^*_s := \mu\frac{\zeta_s}{\sigma^2_s}$ solves the relaxed program when we are implementing $\hat{e} > 0$. Let $K := \frac{\bar{W}-\lambda}{\mu}$, which is independent from q and s. ■

■

Proof of Proposition 3: We rely on the proof of Lemma 3.

For part (i), the PPS of the option contract is $n^*_s := \mu\frac{\zeta_s}{\sigma^2_s}$.

For part (ii), we can write the optimal contract as:

$$w_s(q) = \max\left\{\lambda + \mu\left(\frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}} + \zeta_s\frac{q - h_s(\hat{e})}{\sigma^2_s}\right) - \bar{W}, 0\right\}.$$

The strike price q^*_s is such that $w_s(q) > 0$ if and only if $q \geq q^*_s$, as derived in the proof of Lemma 3. With $K = \frac{\bar{W}-\lambda}{\mu}$, we have $q^*_s = \xi_s + \frac{\sigma^2_s}{\zeta_s}\left[K - \frac{\partial\phi^s_{\hat{e}}/\partial e}{\phi^s_{\hat{e}}}\right]$. ■

Proof of Lemma 4: The FOA is valid if the following objective function is concave in e:

$$\sum_s \phi^s_e \int_{\underline{q}}^{\infty} w_s(q) f(q|e, s) dq - C(e).$$

A sufficient condition is:

$$\sum_s \int_{\underline{q}}^{\infty} w_s(q)\frac{\partial^2 f}{\partial e^2}(q, s|e) dq < C''(e) \qquad \forall e. \tag{59}$$

From equation (20), $w_s(q) \in [0, \max\{0, q\}]$ for all q, s, so that $w_s(q) = 0$ for $q \leq 0$. Then, a sufficient condition for equation (59) is:

$$\sum_s \int_0^\infty \max\left\{q\frac{\partial^2 f}{\partial e^2}(q,s|e),0\right\}dq = \sum_s \int_0^\infty q\max\left\{\frac{\partial^2 f}{\partial e^2}(q,s|e),0\right\}dq < C''(e) \quad \forall e. \blacksquare$$

Proof of Lemma 5: Let

$$\overline{LR}_s(q) := \frac{\partial \phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s} + \frac{\int_q^\infty \frac{\partial f}{\partial e}(z|\hat{e},s)dz}{\int_q^\infty f(z|\hat{e},s)dz} \tag{60}$$

denote the likelihood ratio associated with the event $(\tilde{q} \geq q, \tilde{s} = s)$, which is strictly increasing by MLRP (see proof of Proposition 4). The proof is divided into two parts:

Step 1. Conditional on each signal realization, the optimal contract is debt.

This part adapts the proof of Lemma 1 from Chaigneau, Edmans, and Gottlieb (2018) to show that the optimal contract is debt, such that the payoff for the manager is $\max\{q - q_s, 0\}$ for some face value q_s. Let $w_s(q)$ be a contract satisfying bilateral LL, monotonicity, and the IC. Note that there exists a unique debt contract with the same expected payment conditional on each signal realization. In other words, for each s, there exists a unique q_s that solves

$$\int_{\underline{q}}^\infty \max\{q - q_s, 0\} f(q|\hat{e},s)dq = \int_{\underline{q}}^\infty w_s(q) f(q|\hat{e},s)dq. \tag{61}$$

Among the set of contracts that satisfy bilateral LL and monotonicity, a contract with a payment of 0 at $q = 0$ and a slope of 1 for $q > 0$ maximizes the expected payment to the agent for any given e, s. Therefore, for equation (61) to hold, a debt contract must be such that $q_s \geq 0$. Then, following the same steps as the first part of the proof of Lemma 1 from Chaigneau, Edmans, and Gottlieb (2018), using notations $W^O(q) := \max\{q - q_s, 0\}$ and $W^*(q) := w_s(q)$, the equilibrium effort associated with contract $\max\{q - q_s, 0\}$ is larger than with contract $w_s(q)$. We have therefore shown that substituting a non-debt contract with debt allows the firm to relax the IC. Since the IC must bind at the optimum, this establishes that the original contract cannot be optimal.

Step 2. Determining the optimal face value of debt.

Since any debt contract satisfies bilateral LL and monotonicity, and since we assumed that the condition for the FOA in Lemma 4 holds, the firm's program becomes:

$$\min_{\{q_s\}_{s=1,\ldots,S}} \sum_s \int_{q_s}^\infty (q - q_s)\, f(q,s|\hat{e})dq. \tag{62}$$

subject to

$$\sum_s \int_{q_s}^{\infty} (q - q_s) \frac{\partial f}{\partial e}(q, s|\hat{e}) dq = C'(\hat{e}), \tag{63}$$

where $\frac{\partial f}{\partial e}(q, s|\hat{e}) = \frac{\partial \phi^s_{\hat{e}}}{\partial e} f(q|\hat{e}, s) + \phi^s_{\hat{e}} \frac{\partial f}{\partial e}(q|\hat{e}, s)$. The likelihood ratio can be rewritten as follows:

$$\begin{aligned}
\overline{LR}_s(q) &= \frac{\int_q^{\infty} \left[\frac{\partial \phi^s_{\hat{e}}}{\partial e} f(z|\hat{e}, s) + \phi^s_{\hat{e}} \frac{\partial f}{\partial e}(z|\hat{e}, s) \right] dz}{\int_q^{\infty} \phi^s_{\hat{e}} f(z|\hat{e}, s) dz} \\
&= \frac{\int_q^{\infty} \frac{\partial \phi^s_{\hat{e}}}{\partial e} f(z|\hat{e}, s) dz}{\int_q^{\infty} \phi^s_{\hat{e}} f(z|\hat{e}, s) dz} + \frac{\int_q^{\infty} \phi^s_{\hat{e}} \frac{\partial f}{\partial e}(z|\hat{e}, s) dz}{\int_q^{\infty} \phi^s_{\hat{e}} f(z|\hat{e}, s) dz} \\
&= \frac{\partial \phi^s_{\hat{e}} / \partial e}{\phi^s_{\hat{e}}} + \frac{\int_q^{\infty} \frac{\partial f}{\partial e}(z|\hat{e}, s) dz}{\int_q^{\infty} f(z|\hat{e}, s) dz}
\end{aligned}$$

For each fixed κ and signal realization s, construct the threshold $q^*_s(\kappa)$ as follows:

$$q^*_s(\kappa) := \begin{cases} 0 & \text{if } \overline{LR}_s(0) > \kappa \\ \overline{LR}_s^{-1}(\kappa) & \text{if } \overline{LR}_s(0) \leq \kappa \end{cases}. \tag{64}$$

The cutoff κ is implicitly determined by the binding IC:

$$\sum_s \int_{q^*_s(\kappa)}^{\infty} (q - q^*_s(\kappa)) \frac{\partial f}{\partial e}(q, s|\hat{e}) dq = C'(\hat{e}). \tag{65}$$

The necessary first-order conditions associated with the program in equations (62) and (63) are equation (64) and the binding IC:

$$\sum_s \int_{q^*_s(\kappa)}^{\infty} (q - q^*_s(\kappa)) \frac{\partial f}{\partial e}(q, s|\hat{e}) dq = C'(\hat{e}). \tag{66}$$

where $\kappa := \frac{1}{\mu}$ and μ is the Lagrange multiplier associated with the IC.

Each κ determines $q^*_s(\kappa)$ according to equation (64). From the Intermediate Value Theorem, there exists κ that solves equation (66): as $\kappa \searrow -\infty$, the LHS of (66) exceeds $C'(\hat{e})$ since then $q^*_s(\kappa) = 0 \; \forall s$ and

$$\sum_s \int_0^{\infty} q \frac{\partial f}{\partial e}(q, s|\hat{e}) dq \geq C'(\hat{e})$$

by the assumption in equation (16), and it converges to $0 < C'(\hat{e})$ as $\kappa \nearrow +\infty$. Moreover, κ must be unique since our conditions for the validity of the FOA ensure that the agent's program has a unique solution. ■

Proof of Proposition 4: We first prove that the likelihood ratio $\overline{LR}_s(q)$ in equation (60) is increasing in q:

$$\frac{d}{dq}\left\{\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s}+\frac{\int_q^\infty \frac{\partial f}{\partial e}(z|\hat{e},s)dz}{\int_q^\infty f(z|\hat{e},s)dz}\right\}=\frac{-\frac{\partial f}{\partial e}(q|\hat{e},s)\int_q^\infty f(z|\hat{e},s)dz+f(q|\hat{e},s)\int_q^\infty \frac{\partial f}{\partial e}(z|\hat{e},s)dz}{\left(\int_q^\infty f(z|\hat{e},s)dz\right)^2}. \tag{67}$$

For $\frac{\partial f}{\partial e}(q|\hat{e},s)\leq 0$, we have $-\frac{\partial f}{\partial e}(q|\hat{e},s)\int_q^\infty f(z|\hat{e},s)dz\geq 0$. Moreover, $f(q|\hat{e},s)\int_q^\infty \frac{\partial f}{\partial e}(z|\hat{e},s)dz>0$ because of MLRP and $\int_{\underline{q}}^\infty \frac{\partial f}{\partial e}(z|\hat{e},s)dz=0$. In sum, the RHS of equation (67) is positive. For $\frac{\partial f}{\partial e}(q|\hat{e},s)>0$, the RHS of equation (67) is positive if and only if:

$$\begin{aligned}
& f(q|\hat{e},s)\int_q^\infty \frac{\partial f}{\partial e}(z|\hat{e},s)dz \geq \frac{\partial f}{\partial e}(q|\hat{e},s)\int_q^\infty f(z|\hat{e},s)dz \\
\Leftrightarrow\quad & \int_q^\infty \frac{\frac{\partial f}{\partial e}(z|\hat{e},s)}{\frac{\partial f}{\partial e}(q|\hat{e},s)}dz \geq \int_q^\infty \frac{f(z|\hat{e},s)}{f(q|\hat{e},s)}dz \\
\Leftrightarrow\quad & \int_q^\infty \left[\frac{\frac{\partial f}{\partial e}(z|\hat{e},s)}{\frac{\partial f}{\partial e}(q|\hat{e},s)}-\frac{f(z|\hat{e},s)}{f(q|\hat{e},s)}\right]dz\geq 0,
\end{aligned}$$

which holds because by MLRP we have $\frac{\frac{\partial f}{\partial e}(z|\hat{e},s)}{f(z|\hat{e},s)}\geq\frac{\frac{\partial f}{\partial e}(q|\hat{e},s)}{f(q|\hat{e},s)}$ for any $q\geq z$.

For distributions with location and scale parameters, the PDF of output can be written as in equation (10). The likelihood ratio in equation (60) can then be written as:

$$\overline{LR}_s(q)=\frac{\partial\phi_{\hat{e}}^s/\partial e}{\phi_{\hat{e}}^s}-\frac{\zeta_s}{\sigma_s}\frac{\int_q^\infty g'\left(\frac{z-\xi_s}{\sigma_s}\right)dz}{\int_q^\infty g\left(\frac{z-\xi_s}{\sigma_s}\right)dz}.$$

For part (i), suppose that signals s_i and s_j differ only in that $\frac{\partial\phi_{\hat{e}}^{s_i}/\partial e}{\phi_{\hat{e}}^{s_i}}\geq\frac{\partial\phi_{\hat{e}}^{s_j}/\partial e}{\phi_{\hat{e}}^{s_j}}$. Since the likelihood ratio $\overline{LR}_s(q)$ is increasing in q as shown above, and since the face value of debt q_s^* is given by equation (64), with all else equal across signals we have $q_{s_i}^*\leq q_{s_j}^*$.

For part (ii), when $\overline{LR}_{s_i}(q)\geq\overline{LR}_{s_j}(q)$ for any q, since $\overline{LR}_s(q)$ is increasing in q as shown above and the face value q_s^* is given by (64), we have $q_{s_i}^*\leq q_{s_j}^*$. This condition on the likelihood ratios is satisfied for two signals $\{s_i,s_j\}$ such that $\xi_{s_i}\leq\xi_{s_j}$, all else equal across signals.

For part (iii), for single-peaked distributions, there exists $\overline{z}$ such that $g'(z)>0$ for $z<\overline{z}$

and $g'(z) < 0$ for $z > \overline{z}$, and $\int_{\underline{q}}^{\infty} g'(z)\, dz = 0$. Therefore, all else equal:

$$\frac{\partial \overline{LR}_s(q)}{\partial \zeta_s} = -\frac{1}{\sigma_s} \frac{\int_q^{\infty} g'\left(\frac{z-h_s(\hat{e})}{\sigma_s}\right) dz}{\int_q^{\infty} g\left(\frac{z-h_s(\hat{e})}{\sigma_s}\right) dz} > 0 \tag{68}$$

Consider two signals $\{s_i, s_j\}$ such that $\zeta_{s_i} \geq \zeta_{s_j}$. Then, because of equation (68), we have $\overline{LR}_{s_i}(q) \geq \overline{LR}_{s_j}(q)$ for any q. Since the face value q_s^* is given by (64), with all else equal across signals we have $q_{s_i}^* \leq q_{s_j}^*$.

For part (iv), for a given s, use the change of variables $y = \frac{z-h_s(\hat{e})}{\sigma_s}$ to rewrite the likelihood ratio as:

$$\overline{LR}_s(q) = \frac{\partial \phi_{\hat{e}}^s / \partial e}{\phi_{\hat{e}}^s} - \frac{\zeta_s}{\sigma_s} \frac{\int_{\frac{q-h_s(\hat{e})}{\sigma_s}}^{\infty} g'(y)\, dy}{\int_{\frac{q-h_s(\hat{e})}{\sigma_s}}^{\infty} g(y)\, dy}.$$

Then:

$$\begin{aligned} \frac{\partial \overline{LR}_s(q)}{\partial \sigma_s} \quad &= \quad \frac{\zeta_s}{\sigma_s^2} \frac{\int_{\frac{q-h_s(\hat{e})}{\sigma_s}}^{\infty} g'(y)\, dy}{\int_{\frac{q-h_s(\hat{e})}{\sigma_s}}^{\infty} g(y)\, dy} \\ & \quad -\frac{\zeta_s}{\sigma_s} \frac{g'\left(\frac{q-h_s(\hat{e})}{\sigma_s}\right) \frac{q-h_s(\hat{e})}{\sigma_s^2} \int_{\frac{q-h_s(\hat{e})}{\sigma_s}}^{\infty} g(y)\, dy - g\left(\frac{q-h_s(\hat{e})}{\sigma_s}\right) \frac{q-h_s(\hat{e})}{\sigma_s^2} \int_{\frac{q-h_s(\hat{e})}{\sigma_s}}^{\infty} g'(y)\, dy}{\left(\int_{\frac{q-h_s(\hat{e})}{\sigma_s}}^{\infty} g(y)\, dy\right)^2} \end{aligned}$$

The first term on the RHS is negative, for the same reason as in part (iii) above. We now study the sign of the second term on the RHS. Let $\underline{y} \equiv \frac{q-h_s(\hat{e})}{\sigma_s}$. For $q > h_s(\hat{e})$ and $q > q_s^P$ (which implies $g'(y) < 0\ \forall y \geq \underline{y}$), the numerator of the second fraction of the second term on the RHS is positive if and only if:

$$\begin{aligned} & g'\left(\frac{q-h_s(\hat{e})}{\sigma_s}\right) \frac{q-h_s(\hat{e})}{\sigma_s^2} \int_{\frac{q-h_s(\hat{e})}{\sigma_s}}^{\infty} g(y)\, dy - g\left(\frac{q-h_s(\hat{e})}{\sigma_s}\right) \frac{q-h_s(\hat{e})}{\sigma_s^2} \int_{\frac{q-h_s(\hat{e})}{\sigma_s}}^{\infty} g'(y)\, dy > 0 \\ \Leftrightarrow \quad & g'\left(\underline{y}\right) \int_{\underline{y}}^{\infty} g(y)\, dy > g\left(\underline{y}\right) \int_{\underline{y}}^{\infty} g'(y)\, dy \qquad \Leftrightarrow \qquad \int_{\underline{y}}^{\infty} \frac{g(y)}{g\left(\underline{y}\right)} dy < \int_{\underline{y}}^{\infty} \frac{g'(y)}{g'\left(\underline{y}\right)} dy \\ \Leftrightarrow \quad & \int_{\underline{y}}^{\infty} \left[\frac{g(y)}{g\left(\underline{y}\right)} - \frac{g'(y)}{g'\left(\underline{y}\right)}\right] dy < 0. \end{aligned} \tag{69}$$

Since the distribution g is characterized by MLRP, we have $\frac{g'(\underline{y})}{g(\underline{y})} \geq \frac{g'(y)}{g(y)}\ \forall y \geq \underline{y}$ so that $\frac{g(y)}{g(\underline{y})} \leq \frac{g'(y)}{g'(\underline{y})}\ \forall y \geq \underline{y}$. That is, the term in brackets on the same line of equation (69) is negative for all $y \geq \underline{y}$, so that the integral is negative, and the inequality in equation (69) holds. In sum, if $\sigma_{s_i} > \sigma_{s_j}$, all else equal across signals, then for $q > \max\{q_s^P, h_s(\hat{e})\}$, $\overline{LR}_{s_i}(q) < \overline{LR}_{s_j}(q)$. Since the face value q_s^* is given by (64), with all else equal across signals we have $q_{s_i}^* \geq q_{s_j}^*$. ■

B Application of Lemma 1

This section gives an example of how researchers can apply the condition for the FOA in Lemma 1 to specific cases. Example 3 below considers a quadratic effort cost, $C(e) = \alpha e + \frac{\beta}{2}e^2$.

Example 3 *Let* $\bar{W} = 10$, $u(w) = \frac{w^{-2}}{-2}$ *(i.e., the manager has CRRA utility* $\gamma = 3$*), the cost of effort is* $C(e) = \alpha e + \frac{\beta}{2}e^2$, *with* $\alpha \geq 0$ *and* $\beta > 0$, *the set of possible effort levels is* $e \in [0, 10]$, *output follows a normal distribution with mean* e *and standard deviation 2,* $S = 1$, *and the manager is protected by limited liability. Then we have:* $\int_{\underline{q}}^{\infty} \min\left\{\frac{\partial^2 f}{\partial e^2}(q|e, s), 0\right\} dq \approx -0.121$ *for all* e, $u(\bar{W}) = -\frac{1}{200}$, $C''(e) = \beta$, *and the condition in Lemma 1 is simply* $\beta > \frac{0.121}{200}$.

We also show how equation (8) can be simplified in special cases. Since $K_e^- < 0$, it is easier to satisfy condition (8) when the manager's outside wealth $\bar{W}$ is higher. When $\lim_{c \nearrow +\infty} u(c) = 0$, as with constant absolute risk aversion and CRRA with $\gamma > 1$, the sufficient condition (8) becomes

$$K_e^- u(\bar{W}) < C''(e)$$

for all $e \in (0, \bar{e})$.

Note that Lemma 1 can also be applied to settings in which there is no additional signal. In that case, we have:

$$K_e^+ = \int_{\underline{q}}^{\infty} \max\left\{\frac{\partial^2 f}{\partial e^2}(q|e),\ 0\right\} dq \quad \text{and} \quad K_e^- = \int_{\underline{q}}^{\infty} \min\left\{\frac{\partial^2 f}{\partial e^2}(q|e),\ 0\right\} dq.$$

C Determinants of the Threshold

Without loss of generality, let an increase in c parametrize an increase in the marginal cost of effort that leaves the equilibrium cost of effort $C(\hat{e})$ unchanged. Suppose that $\overline{u} \leq u\left(\bar{W}\right) - C(\hat{e})$, i.e., the IR is nonbinding and $\lambda = 0$. Assuming that the FOA holds, plugging the optimal contract (9) into the IC (4) yields:

$$\sum_s \left[\int_{\underline{q}}^{q_s^*} u\left(\bar{W}\right) \frac{\partial f}{\partial e}(q, s|\hat{e}) dq + \int_{q_s^*}^{\infty} u\left(u'^{-1} \left(1 \middle/ \mu \frac{\frac{\partial f}{\partial e}(q, s|\hat{e})}{f(q, s|\hat{e})} \right) \right) \frac{\partial f}{\partial e}(q, s|\hat{e}) dq \right] = C'(\hat{e}). \tag{70}$$

From equation (28) that describes the thresholds, with $\lambda = 0$:

$$LR_s(q_s^*|\hat{e}) \;=\; \frac{\frac{\partial f}{\partial e}(q_s^*, s|\hat{e})}{f(q_s^*, s|\hat{e})} = \frac{1}{\mu} \frac{1}{u'(\bar{W})} > 0. \tag{71}$$

This implies that $\frac{\partial f}{\partial e}(q, s|\hat{e}) > 0 \; \forall q > q_s^*$ due to MLRP, so that the first derivative of the LHS of the IC in equation (70) with respect to μ is positive (recall that, by construction, $u'^{-1}\left(1/\mu LR_s(q_s^*|\hat{e})\right) = \bar{W}$). Therefore, to maintain incentive compatibility, following an increase in c that raises the RHS of equation (70), the Lagrange multiplier μ must rise for the LHS of equation (70) to increase correspondingly. In turn, using MLRP, $u' > 0$, and equation (71), a higher μ implies a lower q_s^* for any s. In sum, when IR does not bind, a higher c results in a lower q_s^* for any s. Since the likelihood ratio is positive at the threshold (see equation (71)), an increase in c brings the threshold q_s^* closer to the point where the likelihood ratio $LR_s(q|\hat{e})$ is zero. Intuitively, a higher marginal cost of effort requires the principal to offer higher incentives, which she ensures by lowering the threshold.

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html