Agency in Hierarchies: Middle Managers and Performance Evaluations*

Henrique Castro-Pires†

August 16, 2023

Abstract

This article studies the optimal joint design of incentives and performance rating scales in a principal-manager-worker hierarchy. The principal wants to motivate the manager and the worker to exert unobservable effort. Given the effort choices, two signals are realized: public and verifiable team output and a non-verifiable signal about the worker's effort, privately observed by the manager. The principal may try to elicit the manager's private information by requiring her to evaluate the worker's performance. Payments depend on team output and the manager's evaluation. I show that the principal can achieve no more than what is feasible with a binary rating scale. I also characterize when subjective evaluations are valuable and find conditions for when the principal benefits from reducing organizational transparency and for when a forced ranking outperforms individual performance evaluation systems. JEL: D23, D82, D86, J41.

Keywords: Moral hazard, subjective performance evaluations, middle manager, organizational transparency.

1 Introduction

In classic principal-agent models, the principal directly deals with her agents. However, in practice, organizations' top and bottom are far apart. For example, a retail store's headquarters (or even shareholders) have little (if any) direct contact with the salespeople. Often, the relationship is intermediated by managers who are responsible for motivating their subordinates.[1] Managers are closer to the rank-and-file and usually

*I am grateful to an Editor and three anonymous referees for helpful suggestions. I thank Jeroen Swinkels, George Georgiadis, Alessandro Pavan, Humberto Moreira, Andrés Espitia, Michael Powell, Daniel Barron, Cassiano Alves, Miguel Talamas, Gabriel Ziegler, and Bruno Barsanetti for helpful comments and feedback. I also thank seminar participants at various places. Any remaining errors are mine only.

†University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom - email: h.castro-pires@surrey.ac.uk.

[1]For a discussion on middle managers' role in the organization see Huy, "In Praise of Middle Managers", *Harvard Business Review*, September 01, 2001; Jaser, "The Real Value of Middle Managers", *Harvard Business Review*, June 07, 2021; and Blow, "Why companies need middle managers", *The Economist*, October 02, 2021. In a related article, Lazear et al. (2015) estimate large supervisor effects on firm outcomes. Using data from a large services company, they find that the average boss adds about 1.75 times as much output as the average worker.

have better information about workers' behavior than higher ranks in the organization. Many firms try to elicit this additional information by requiring managers to evaluate their subordinates' performance.[2] Furthermore, when this is the case, firms typically use coarse performance rating scales such as 1-5 stars or broad categories such as "Unsatisfactory — Satisfactory — Outstanding" performance. This article studies the optimal joint design of incentives and rating scales in hierarchies.

I analyze this topic within a moral-hazard model, in which a risk-neutral principal designs a compensation scheme for two agents, a weakly risk-averse manager, and a strictly risk-averse worker. The principal wishes to incentivize the worker to exert productive effort and the manager to truthfully report on the worker's performance. Payments can be conditioned on two performance measures: publicly observable and verifiable team output and subjective managerial evaluation. The principal offers contracts consisting of a set of performance ratings and mappings from output and ratings to payments. The manager and the worker simultaneously decide whether to accept or reject contracts. If any of them rejects, all players get their outside options. If they both accept, the manager announces how she will evaluate the worker as a function of her information, and the worker decides whether to exert effort at a private cost. Effort stochastically generates output and a non-verifiable signal privately observed by the manager. Given the output and the private signal realization, the manager decides how to rate the worker's performance.

For example, consider a retail chain store designing a compensation scheme for its managers and salespeople. Suppose that a given store has one manager and one salesperson. The firm wants to motivate the salesperson to exert effort, and it observes sales. The manager interacts daily with the salesperson and has more information about the salesperson's effort than the higher ranks of the chain store. The firm may require the manager to evaluate the salesperson's performance to elicit this extra information. Payments for both parties can depend on sales and the performance report the manager provided.

This paper focuses on the agency conflict between the evaluator (the manager) and the residual claimant (the principal). While the residual claimant cares about compensation costs, namely the size of the wage bill, the manager does not pay the worker from her own pocket and is only concerned with providing strong incentives for effort. The first main finding is that, due to this agency conflict, the principal can achieve no more than what is feasible with a binary performance rating scale. Even if she sets up a richer rating scale, the manager has the incentive to use only the highest and the lowest-paying messages. The intuition behind the result is better understood when decomposed into three steps:

1. **The principal cannot condition the manager's payments on her reports about the worker's performance:**

   The goal of using subjective performance evaluation is to get information about the worker's effort. If the principal pays the manager distinct amounts for different reports, the manager — instead of reporting accurate information about the worker's effort — will choose the report in which she gets the highest wage. Hence, the manager's compensation must be independent of her report about the

[2] There exists a broad range of empirical studies on subjective performance evaluations in various fields such as accounting, management science, personnel psychology, and economics. See Kampkoetter and Sliwka (2016) for a broad survey. For a survey focused on the empirical literature in economics, see Frederiksen et al. (2017).

worker's performance.[3]

2. **The manager benefits from higher worker's effort:**

   I assume the manager directly benefits from higher output. Such benefits can come from career concern — having better observable outcomes generating better outside offers — or even as an unmodeled incentive contract to the manager, whom the principal might also wish to motivate to exert unobservable effort. As a higher worker's effort generates higher output levels and the manager's payment cannot depend on what she reports (step 1), the manager strictly benefits from a higher worker's higher effort.

3. **The manager provides the strongest possible incentives to the worker:**

   Steps 1 and 2 imply that the manager benefits from inducing a higher worker's effort (step 2) but does not internalize the cost of providing stronger incentives to the worker, as payments to the worker are paid by the principal, not the manager (step 1). As a result, the manager wishes to provide the worker with the strongest effort incentives, regardless of how much this costs the firm. The performance evaluation strategy that generates the strongest incentives for effort is the one in which the manager reports the message delivering the lowest payment to the worker for low enough signals and the highest-paying one otherwise. As the manager uses only the highest and the lowest-paying messages, the principal can do no better than designing a binary performance rating scale.

Many incentive schemes used in practice take such a binary structure. "Up-or-out" systems and discretionary single-valued bonuses are prominent examples. The result that managers concentrate their evaluations on a few rating scale points is also consistent with empirical findings in the literature. Frederiksen et al. (2017) review different studies using personnel data from several firms and document that they all share the feature of concentrated ratings in a few points in the scale, often in only two. In a more recent article, Frederiksen et al. (2020) use personnel data from a sizeable Scandinavian service sector firm and show that over 90%

In a binary performance review system, the manager reports whether the worker's performance is good or bad. A cutoff function can describe the manager's preferred performance evaluation strategy. Conditional on output, the manager reports a good performance if the signal she privately observes is sufficiently high. I show that such a cutoff is decreasing in output. A decreasing cutoff implies that the manager is more lenient in her subjective evaluations when output is high. That is, she requires a lower minimal private signal to report good performance.

After characterizing the manager's preferred evaluation strategy, I find the cost-minimizing contracts to implement the desired effort levels as in a traditional moral hazard problem. Applying Holmström (1979)'s informativeness principle, I characterize when performance evaluations are valuable. The informativeness principle states that optimal contracts link any signal that provides information about actions to payments.

[3] The intuition is similar to MacLeod (2003), in which the evaluator (there the principal, here the manager) must be indifferent between any report. MacLeod (2003) achieves indifference through money burning, whereas here, it requires not conditioning the manager's payments on her report.

One could then expect subjective performance evaluations to be always valuable. However, they are valuable only if the performance report provides additional information beyond what the verifiable performance measure already provides. If the verifiable performance measure is very informative about the worker's action, the manager's report might never reveal any information about her private signal. For instance, suppose there are two possible output realizations: high and low. Suppose that the probability of high output when the worker exerts high effort is close to one, and in the case of low effort, it is close to zero. By observing a high output realization, the manager is sufficiently convinced that the worker has exerted high effort and reports a good performance even if the private signal realization is the worst possible. Similarly, when output is low, the manager provides a bad report regardless of the private signal realization. In this case, the performance evaluation report does not convey any additional information beyond what the output conveys. In those situations, optimal contracts do not use subjective performance evaluations.

When considering cases in which subjective performance evaluations are valuable, one can ask how informed the principal wishes the manager to be about the worker's effort. A less informed manager has a worse assessment of the worker's action; however, less information might attenuate the principal and manager's conflicts of interest. In my second main result, I show that reducing transparency in the organization — meaning reducing how informed the manager is about the worker's effort — might be strictly beneficial for the principal.

The central intuition is that reducing transparency decreases the manager's and principal's conflict of interest regarding performance evaluations. The manager wishes to maximize the worker's effort regardless of compensation costs. As a result, the manager provides a good report too often (in the principal's perspective) when output is sufficiently high and too rarely when output is sufficiently low. By censoring part of the information the manager observes, the principal can increase the probability manager reports a good performance when output is low and decrease the probability of a good report when output is high. This censoring increases the expected worker's payment for low-output realizations and reduces it for high-output realizations, improving risk-sharing and decreasing expected compensation costs.

The paper's final part extends the analysis to the case of multiple workers and allows the principal to use a forced ranking. I show that forced rankings outperform individual performance evaluation mechanisms when the manager's private information is noisy, while individual evaluations are better when the manager is very informed. While forced rankings allow the principal to mitigate agency conflicts with the manager, they provide only a relative performance measure while wasting information about how well each worker did.

Finally, the article proceeds as follows. The next section discusses the related literature. Section 2 presents the baseline model. Section 3 solves for the optimal contracts. Section 4 analyzes how informed the principal wants the manager to be about the worker's effort. Section 5 studies the trade-off between individual performance evaluations and forced ranking when there are multiple workers. Finally, Section 6 discusses variations on the assumptions and avenues for future research. Proofs are relegated to the appendix.

1.1 Related Literature

This article belongs to the literature on optimal provision of incentives, pioneered by Ross (1973), Mirrlees (1999), and Holmström (1979). In those articles, the principal chooses a compensation scheme based on verifiable signals about workers' actions (effort). Further contributions — e.g., Baker et al. (1994); MacLeod (2003); Levin (2003); Fuchs (2007, 2015); Maestri (2012, 2014); Deb et al. (2016); Cheng (2021); Ishiguro and Yasuda (2021) — examine optimal contracting in the presence of subjective performance evaluations.[4]

Subjective performance measures are usually modeled as non-verifiable private signals. When the information is non-verifiable, the performance reviewer might have the incentive to misreport their information. In particular, when the principal is the performance reviewer, she is tempted to renege on payments. After effort has been exerted, the principal has the incentive to provide a low-performance review and save on payments to the agent.[5] Previous literature dealt with this type of limited commitment through repeated interactions — MacLeod and Malcomson (1989); Levin (2003); Fuchs (2007); Fong and Li (2017); Zhu (2018) — through money burning — MacLeod (2003); Kambe (2006) — through a feedback effect increasing incentives for future effort — Zábojník (2014) — or assuming costly justification of performance ratings — Lang (2019). In my model, the residual claimant is not the performance reviewer. Hence, the incentive to renege on high payments is not present.

This article also discusses how informative the principal wants the signal about the worker's effort to be. Firms often have control over what managers observe about their subordinate's actions. I find conditions under which the principal strictly benefits from reducing the manager's information. The exercise of endogeneizing the informativeness of the signal relates to the literature on endogenous monitoring. Rahman (2012) shows that the principal can incentivize the monitor to exert costly monitoring effort by randomly and secretly allowing the worker to shirk. Ostrizek (2021) shows that the principal might prefer a non-fully informative structure to prevent the agent from learning about his own ability. Gershkov and Winter (2015) analyze how complementarity in the production function affects how costly formal monitoring relates to informal peer monitoring. Georgiadis and Szentes (2020) and Li and Yang (2020) allow the principal to choose the information structure at a given cost directly. My contribution is not about costly monitoring. The focus is not on information acquisition but on the incentives of an intermediary to reveal non-verifiable information after it has been acquired. I allow the principal to change the information structure costlessly and show that decreasing the manager's information reduces compensation costs.

Finally, Section 5 contrasts the advantages of forced rankings and individual performance evaluations in a setting with two workers, which connects the current paper to the literature on moral hazard in teams, pioneered by Lazear and Rosen (1981) and Holmström (1982). Two closely related papers are Prendergast and Topel (1996) and Letina et al. (2020), which study cases where the performance reviewer is biased in

[4]See Lazear and Oyer (2012) and Holmström (2017) for surveys.

[5]This stands in contrast to Alchian and Demsetz (1972) classic argument for making the monitor the residual claimant. In their analysis, they assume that the measure of the agent's performance is verifiable. The primary friction there is to generate incentives for the supervisor to monitor the agent. Here, I assume that the measure is non-verifiable. Hence, the main friction is the accurate revelation of information, not acquisition. In this case, the separation between monitoring and ownership is beneficial. See Rahman (2012) for a related discussion.

favor of one or all workers. In my model, the agency conflict between the principal and the manager does not arise from a bias in favor or against any agent but from the fact that the manager is not paying the wage bill from her own pocket.

2 Model

A risk-neutral principal designs contracts for a risk-averse worker (he, denoted by W) and a manager (she, denoted by M). First, the principal proposes contracts to the manager and the worker. Contracts specify a finite set of performance ratings (E) and payments that depend on realized output and performance evaluations, as described later. Second, the manager and the worker decide whether to accept or reject the contract. If any of them rejects, the game ends, and both players get their respective outside options valued at $\bar{u}_M, \bar{u}_W \in \mathbb{R}$. If both accept the contract, the worker then privately draws an effort cost $\mathbf{c}$ from a commonly known distribution G with full support on $\mathbb{R}_+$. Knowing his effort cost, the worker then decides whether or not to exert effort, denoted by $a \in \{0,1\}$.

Given the worker's effort decision, two signals are realized: output $\mathbf{y}$ and a non-verifiable signal $\mathbf{z}$. Output is verifiable, observed by all players, and has finite support $Y := \{y_0, \ldots, y_n\} \subset \mathbb{R}$. Conditional on effort, the probability of an output realization y is denoted by $p(y|a) \in (0,1)$, with c.d.f. denoted by $P(y|a)$. The non-verifiable signal $\mathbf{z}$ denotes any soft information the manager observes, but the principal does note. That is, $\mathbf{z}$ is privately observed by the manager, and has support $Z := [\underline{z}, \bar{z}] \subset \mathbb{R}$. I assume it admits a density $q(.|a)$ and denote its c.d.f. by $Q(.|a)$.[6]

I assume that p and q are common knowledge, and higher signal realizations are associated with higher effort. Formally, the likelihood ratio between high and low effort is strictly increasing in the signal realization.

Assumption 1. *[Monotone Likelihood Ratio Property (MLRP)]*

1. $\frac{p(y|1)}{p(y|0)}$ *is strictly increasing in y.*
2. $\frac{q(z|1)}{q(z|0)}$ *is strictly increasing in z.*

Before the worker's effort choice, the manager announces how she will evaluate his performance from a finite set of possible performance ratings E. Formally, the announced evaluation strategy $\sigma : Y \times Z \to \Delta(E)$ denotes the distribution of performance reports conditional on each signal realization. That is, the manager announces that if the realized signals are (y,z), she will report a rating e with probability $\sigma(e|y,z)$. There is then an essential distinction between what the manager announces and how she actually behaves. After observing the realizations of $\mathbf{y}$ and $\mathbf{z}$, the manager decides what performance rating to report. I denote

[6] The non-verifiable signal $\mathbf{z}$ being a continuous random variable and the verifiable being discrete is a matter of convenience. The results also hold if $\mathbf{z}$ is a discrete random variable or if $\mathbf{y}$ is continuous.

by $\sigma_r : Y \times Z \to \Delta(E)$ the distribution of reports the manager actually uses. The later defined equilibrium concept imposes that the announced performance evaluation strategy must match the realized one.

The timing of the model can be summarized as follows:

- t=0, the principal offers contracts
- t=1, the manager and the worker accept/reject contracts
- t=2, the manager announces her evaluation strategy σ to the worker
- t=3, the worker observes c and chooses effort
- t=4, the manager observes y and z
- t=5, the manager chooses a performance report $e \in E$
- t=6, payments are realized.

Contracts are a triple: a finite set of possible performance ratings E, and payments for the manager and the worker $\pi_i : Y \times E \to \mathbb{R}_+$ for $i \in \{M, W\}$ contingent on output and performance evaluation, where $\pi_i(\cdot, e)$ must be increasing for all $i \in \{M, W\}$ and $e \in E$.[7] Note that there are two constraints imposed on the set of contracts: first, I restrict attention to payments that are increasing in output. One justification for the restriction to monotonic contracts is assuming the agents can freely dispose of output or sabotage production.[8] Second, note that the worker's effort cost is realized after the contract is signed, and I do not allow the worker to report his cost to the principal. That is, screening is not permitted. However, the main qualitative results are maintained as long as the principal cannot communicate directly with the worker. In Section 6, I discuss the implications of such a restriction.

The worker is strictly risk-averse, and his ex-post payoffs given effort, contract, realized output, performance evaluation, and effort cost is $U_W(\pi_W, y, e, c, a) := u_W\big(\pi(y, e)\big) - c \cdot a$. Where $u_W : \mathbb{R}_+ \to \mathbb{R}$ is strictly increasing, strictly concave, twice continuously differentiable and $\underset{x \to +\infty}{lim} u_W(x) = +\infty$. I also assume that the worker's outside option is non-trivial, in the sense that $\bar{u}_W > u_W(0)$.

The manager's payoff is given by $U_M(\pi_M, y, e) := u_M\big(\pi_M(y, e) + b(y)\big)$, where $u_M : Y \to \mathbb{R}_+$ is strictly increasing, weakly concave, and $b : Y \to \mathbb{R}_+$ is strictly increasing. That is, the manager's utility is higher the higher the payments she receives and the higher the output. For simplicity, the function b is assumed to be exogenous and denotes direct benefits the manager receives from the output. Such direct benefits from higher output can arise from career concerns (managers with larger output realizations have better future career opportunities) or from direct incentives the manager might face.[9] The principal's payoff is given by output minus the payments to the agents (manager and worker). That is, $U_P(y, e) := b(y) - \pi_M(y, e) - \pi_W(y, e)$. Finally, the manager's outside option is given by $\bar{u}_M > u_M(0)$.

[7] Restricting attention to finite performance ratings is not necessary but simplifies the analysis. The issue is assuring there exists a continuation equilibrium at each information set. A weaker restriction would be to assume that payments have a well-defined maximum and minimum. That is, $\underset{e \in E}{max}\{\pi_i(y, e)\}$ must exist for each $i \in \{M, W\}$ and $y \in Y$.

[8] It is also a common assumption in the literature, e.g., Innes (1990).

[9] In an earlier working paper version, I assumed the manager also exerted productive effort and would need to be incentivized by the principal. As a result, the manager's equilibrium payoff endogenously becomes increasing in output.

Suppose the manager and the worker accept the contract. In that case, they play an extensive form game where the manager announces an evaluation strategy, the worker chooses effort, signals are realized, and then the manager decides what rating to report. The solution concept I use is the Weak Perfect Bayesian Equilibrium (weak-PBE), in which the manager's evaluation strategy announced to the worker matches her actual reporting strategy. Given contracts (E, π_M, π_W), denote by σ the announced evaluation strategy, by $\sigma_r : Y \times Z \to \Delta(E)$ the realized distribution on evaluations, and by $\alpha : \mathbb{R}_+ \times \Delta(E) \to \{0,1\}$ the worker's effort choice as a function of the effort cost c and manager's announced evaluation distribution.

Definition 1. *I say that $(\alpha, \sigma, \sigma_r)$ is an **equilibrium** if*

1. $\alpha(c,\sigma) \in \underset{a\in\{0,1\}}{argmax}\, \mathbb{E}\big[U_W(\pi_w, \mathbf{y}, \mathbf{e}, c, a)|\sigma, a\big]$;
2. $\sigma \in \underset{\tilde{\sigma}\in\Delta(E)}{argmax}\, \mathbb{E}\big[U_M(\pi_M, \mathbf{y}, \mathbf{e})|\tilde{\sigma}, \alpha\big]$;
3. $\sigma_r(e|y,z) > 0$ *implies* $U_M(\pi_M, y, e) \geq U_M(\pi_M, y, \hat{e})$ *for all* $\hat{e}, e \in E$*,* $y \in Y$*, and* $z \in Z$*;*
4. $\sigma = \sigma_r$.

Item 1 assures that the worker chooses his effort optimally given the contract, the announced evaluation strategy, and his effort cost. Item 2 implies that the manager announces her preferred evaluation strategy, considering that the evaluation strategy affects the worker's incentives for effort. Item 3 imposes that any evaluation strategy actually used must be sequentially rational. Item 4 imposes that, in equilibrium, the announced evaluation matches the actual one. In summary, the equilibrium concept imposes sequential rationality on both players and assumes the manager selects the evaluation strategy — among the sequentially rational ones — that generates the highest manager's continuation equilibrium payoff.

Finally, I add a regularity condition that ensures higher signal realizations are associated with higher output. This feature is a direct consequence of the monotone likelihood ratio property when the signal about effort is unidimensional. Here, however, the principal observes output and the manager's report. The regularity condition below assures that higher output is more strongly associated with higher effort for any given observed report.

Assumption 2. *[Regularity conditions]* $\frac{q(z|0)}{q(z|1)} \cdot \frac{Q(z|1)}{Q(z|0)}$ *and* $\frac{q(z|0)}{q(z|1)} \cdot \frac{(1-Q(z|1))}{(1-Q(z|0))}$ *are decreasing on* $(\underline{z}, \overline{z})$.

Assumption 2 imposes two monotone hazard ratio properties on Q. It assumes the ratios of hazard rates $\Big[q(\cdot|0)/(1-Q(\cdot|0))\Big] \Big/ \Big[q(\cdot|1)/(1-Q(\cdot|1))\Big]$ and reversed hazard rates $\Big[q(\cdot|0)/Q(\cdot|0)\Big] \Big/ \Big[q(\cdot|1)/Q(\cdot|1)\Big]$ are both decreasing. Later, I discuss in detail the role of such an assumption and how it can be relaxed. Moreover, to ensure that Assumptions 1-2 can be simultaneously satisfied, I present two classes of examples. For both examples, let $Z = [0,1]$, $Q(z|0) = z$, and P be such that (MLRP) is satisfied.[10]

Example 1. *Let* $Q(z|1) = exp(\beta(z-1))z$*, for* $\beta \in [1, +\infty)$*. Then, Assumptions 1-2 are satisfied.*

Example 2. *Let* $Q(z|1) = z(1+z)/2$*. Then, Assumptions 1-2 are satisfied.*

[10] As the signal $\mathbf{z}$ is not directly payoff relevant to any player (principal, manager, nor worker) it is without loss of generality assuming that $Q(z|0) = z$. It is simply a renormalization so that z represents the quantile of the distribution when there is no worker's effort.

3 Model Analysis

3.1 Worker's Effort Choice

I first analyze the worker's effort choice. Fixing contracts (E, π_W, π_M), and the evaluation strategy the worker believes the manager is going to use $\hat{\sigma}$, the worker's expected utility is $V_W(a, \pi_W, \hat{\sigma}, c) := \mathbb{E}\Big[u_W(\pi_W(\mathbf{y}, \mathbf{e}))|a, \hat{\sigma}\Big] - a \cdot c$.

Note that the worker's best response is given by a cutoff. Given the contracts and the conjecture about the manager's evaluation strategy, the worker exerts high effort if c is smaller than

$$c_W(\pi_W, \hat{\sigma}) = \mathbb{E}\Big[u_W(\pi_W(\mathbf{y}, \mathbf{e}))|a = 1, \hat{\sigma}\Big] - \mathbb{E}\Big[u_W(\pi_W(\mathbf{y}, \mathbf{e}))|a = 0, \hat{\sigma}\Big].$$

From now on, I refer to the worker's effort as a cost cutoff c_W. That is, given contracts (E, π_W, π_M) an equilibrium is described by (c_W, σ), where c_W denotes the worker's effort cost cutoff and σ represents both the manager's announced and actually used evaluation strategy.

Moral hazard problems with binary efforts are convenient due to the simplicity of their incentive compatibility constraints. However, the notion of higher effort is restrained to exerting effort or not. The key insight of this article is that managers do not internalize the compensation cost and wish to motivate their subordinates to exert as much effort as possible. Hence, a notion of effort intensity is needed. By introducing a stochastic effort cost, I allow for a granular notion of effort (the cost cutoff) while keeping the model tractable.[11]

3.2 Principal's Problem

Following the Grossman and Hart (1983) approach, I focus on the problem of implementing a given effort cost cutoff c_W at the minimal possible cost to the principal. That is, the principal minimizes expected payments to the manager and the worker by choosing contracts (E, π_W, π_M) and recommending actions (c_W, σ) subject to both agents being willing to accept the contracts and follow the recommendations. In particular, I assume the principal wants to implement a non-trivial cost cutoff for the worker ($c_W > 0$).

3.2.1 Verifiable Signals Benchmark

Before tackling the problem described above, it is helpful first to analyze a benchmark case in which all signals are publicly observed and verifiable ($\mathbf{z}$ is public and verifiable). When all signals are verifiable, the principal does not need to request a report from the manager since she has no private information. In such

[11] In Section 6 (with a detailed analysis in Appendix C), I discuss a version of the model with deterministic costs but continuous efforts. The main trade-offs and results continue to hold.

a setting, the principal's problem becomes a canonical moral hazard problem, where the principal chooses mappings from realized verifiable signals to payments while assuring participation and obedience.

In this case, the manager's payments must simply assure participation. That is, the principal must pay the manager the smallest fixed amount $\hat{\pi} \in \mathbb{R}_+$ such that $\mathbb{E}[u_M(b(y)+\hat{\pi}_M)|c_W] \geq \bar{u}_M$.[12] Moreover, as the principal directly observes all information, the evaluation strategy and its announcement are vacuous.

The optimal worker's incentive scheme also takes a standard form. By the canonical arguments in the moral hazard literature, one can show that the worker's contract takes the form

$$\frac{1}{u_W'(\hat{\pi}_W(y,z))} = \hat{\lambda}_W + \hat{\mu}_W \hat{s}(y,z),$$

where $(\hat{\lambda}_W, \hat{\mu}_W) \in \mathbb{R}^2_+$ are the respective dual multipliers associated with participation and incentive compatibility, and $s(y,z)$ is the score associated with realization (y,z). The score is an increasing function of the likelihood ratio of a given realization pair (y,z).[13] Note that payments to the worker vary with the realizations of both signals and pairs with a higher likelihood ratio are associated with higher payments[14].

Note that even if **z** was not directly observed by the principal but it was by both the manager and the worker, the principal could request the manager and worker to simultaneously report **z** and punish them if their report disagrees. One could, then, implement the same outcome as when **z** is public. This implementation requires, however, that the worker can observe **z** and directly communicate with the principal. In the remainder of the paper, I analyze the case in which **z** is the manager's private information, and the principal can communicate with the manager but not with the worker directly. Here, the hierarchy imposes a constraint on the communication inside the organization. The manager is an intermediary, as she is the only agent capable of communicating both with the organization's top (principal) and bottom (worker).

3.2.2 Privately Informed Manager

I resume our main case of interest in which **z** is an unverifiable signal privately observed by the manager and that direct communication between principal and worker is unfeasible. Such a restriction on communication was directly embedded in the definition of contracts, as payments depended only on output and manager's reports.

The first step is to show that there is no loss in restricting attention to contracts in which π_M does not depend on the performance report e. As I have not imposed almost any structure on the set of performance ratings E, nor on how they map to payments, different contracts might generate the same relevant outcomes in equilibrium. For example, if one adds performance ratings that are never used in equilibrium, one does not

[12]This assumes the manager is necessary for the firm to operate. If the manager's role is only monitoring, the principal might prefer not even to satisfy her participation constraint.

[13]Given the cutoff c_M the principal wishes to implement; the score is given by $\hat{s}(y,z) = \frac{p(y|1)f(z|1)-p(y|0)f(z|0)}{G(c_W)p(y|1)f(z|1)+(1-G(c_w))p(y|0)f(z|0)}$.

[14]Similarly, if there existed only one signal (**y** or **z**) which was publicly observable and verifiable, then payments would be conditioned only on such signal via the score.

affect effort choices, payments, and output distributions. To deal with this innocuous form of multiplicity, I define a concept of equivalence between contracts.

Definition 2 (Outcome equivalence)**.** *Two contracts* (E, π_W, π_M) *and* $(\tilde{E}, \tilde{\pi}_W, \tilde{\pi}_M)$ *are* ***outcome-equivalent*** *if for any equilibrium under* $(\tilde{E}, \tilde{\pi}_W, \tilde{\pi}_M)$ *there exists an equilibrium under* (E, π_W, π_M) *with the same effort choices, output and payments' joint distributions, and vice-versa.*

Lemma 1. *For any contract* (E, π_W, π_M) *there exists an outcome-equivalent one* $(\tilde{E}, \tilde{\pi}_W, \tilde{\pi}_M)$ *in which* $\tilde{\pi}_M$ *does not depend on the performance evaluation.*

After y and z are realized, the effort has already been executed, and the manager chooses the reports that maximize her payment. If the principal conditions the manager's payments on her reports, the manager chooses the reports that maximize her payments, not necessarily reflecting the private information about the worker's effort.[15]

The intuition for this result is reminiscent of MacLeod (2003). In his article, the principal has the dual role of reviewer and residual claimant. That is, she directly observes the unverifiable information and decides how to evaluate and pay the worker. The information is unverifiable, so the principal must be indifferent between reports; otherwise, only the lowest-paying report would be used on the equilibrium path. In MacLeod (2003), the reviewer's indifference among multiple reports is achieved through money burning. Here, as the reviewer is the manager rather than the principal, one can generate such indifference without money burning, which would not be beneficial in the present setting.

The main message of Lemma 1 is that sequential rationality implies that the principal can do no better than choosing contracts in which π_M does not depend on the reported performance evaluation. From now on, I restrict attention to such contracts. Note then that after signals have been realized, the manager is indifferent between reports, implying that any evaluation strategy is sequentially rational. Therefore, when the manager's payments do not depend on her report, the manager is always willing to follow her announced evaluation strategy. Hence, one can directly set $\sigma_r = \sigma$ when looking for an equilibrium.

Remark 1. *Most agency models do not distinguish announcing strategies from following them. Moreover, whenever there are multiple sequentially rational strategies, they break indifferences in favor of the principal*[16]*. However, the argument favoring the principal's preferred equilibrium usually relies on the fact that the principal could break the indifference by adjusting the payments by an arbitrarily small amount. In this case, that is not true. If the principal increases the manager's payment for a given reported rating, the manager would always report such rating, regardless of her private signal. Hence, all informational content of the performance evaluation would be lost.*

By augmenting the game with the evaluation strategy announcing stage and using the equilibrium concept

[15] note that the principal cannot make the manager the residual claimant and still use her private information. When the manager is the residual claimant, her best response after effort has been exerted is to report the lowest paying worker's performance — keeping the highest residual for herself.

[16] In this model, if one removes the evaluation announcing stage and selects the principal's preferred weak-PBE, the manager would report her non-verifiable information truthfully at no extra cost to the principal, as ex-post she is indifferent between any report.

defined above, my model generates a different weak-PBE selection: the manager's preferred weak-PBE.[17] *The manager is ex-post — after the effort choices — indifferent between any reporting strategy, but not ex-ante. Different reporting strategies generate different incentives for the worker to exert effort. As the worker's effort affects the output distribution, and the manager's payoff depends on output, the manager may not be indifferent between different reporting strategies ex-ante.*[18] *The manager then announces and follows her preferred evaluation strategy.*[19]

One can think about the principal's problem as choosing contracts and recommending actions (a cutoff c_W and σ) such that they constitute an equilibrium. The principal then minimizes the cost of implementing a given cutoff $c_W > 0$ subject to obedience and minimum payment constraints. I focus on cases where the minimum payment constraints (payments being positive) are slack (thereon assumed)[20].

Denote by $\Gamma(E, \pi_W, \pi_M)$ the set of continuation weak-PBE for a given contract (E, π_W, π_M) after both agents have accepted to participate. The principal's problem can be written as

$$\min_{E,\sigma,\pi_W,\pi_M} \left\{ \mathbb{E}\Big[\pi_W(\mathbf{y},\mathbf{e}) + \pi_M(\mathbf{y}) | c_W, \sigma\Big] \right\}$$

subject to

$$\mathbb{E}\big[u_M\big(\pi_M(\mathbf{y}) + b(\mathbf{y})\big)\big|c_W\big] \geq \bar{u}_M, \tag{IR_M}$$

$$\mathbb{E}\big[u_W\big(\pi_W(\mathbf{y},\mathbf{e})\big)\big|c_W, a=1, \sigma\big] - \mathbb{E}\big[\mathbf{c}\big|c \leq c_W\big] \geq \bar{u}_W, \tag{IR_W}$$

$$\mathbb{E}\big[u_W\big(\pi_W(\mathbf{y},\mathbf{e})\big)\big|a=1, \sigma\big] - \mathbb{E}\big[u_W\big(\pi_W(\mathbf{y},\mathbf{e})\big)\big|a=0, \sigma\big] = c_W, \tag{IC_W}$$

$$\mathbb{E}\big[u_M\big(\pi_M(\mathbf{y}) + b(y)\big)\big|c_W, \sigma\big] \geq \mathbb{E}\big[u_M\big(\pi_M(\mathbf{y})\big)\big|\tilde{c}_W, \tilde{\sigma}\big] \quad \forall (\tilde{c}_W, \tilde{\sigma}) \in \Gamma(E, \pi_W, \pi_M). \tag{IC_M}$$

The principal chooses contracts and a recommended evaluation strategy such that: the worker and the manager are willing to accept the contracts ((IR_W) and (IR_M) constraints), the worker is willing to follow the recommended effort level ((IC_W) constraint), and the manager announces and follows the recommended evaluation strategy ((IC_M) constraint).

The main distinction from the canonical moral hazard problem is the non-standard (IC_M) constraint. Given the contracts in place, the principal must recommend an evaluation strategy that is part of the manager's favorite continuation equilibrium. I characterize the optimal incentive contracts in three steps. First, I show the manager strictly benefits from higher worker's effort. Second, I characterize the manager's preferred evaluation strategy for all feasible contracts. Third, I find the optimal contracts.

[17] There is also a growing literature on unique implementation. For instance, see Winter (2004), Halac et al. (2021) and related references therein. Subjective performance evaluations cannot be used for unique implementation. As the manager must be indifferent between any report after signals have been realized, there are always multiple equilibria.

[18] Manager's preferred weak-PBE is similar to the *credible threats refinement* proposed by Zhu (2018).

[19] If the manager's payments are exogenously restricted to be invariant to her report about the worker's performance (to avoid a conflict of interest or legal liability), then the manager's preferred equilibrium selection criterion would be equivalent to assuming that the manager can commit to an evaluation strategy.

[20] Appendix B proves the existence of optimal contracts and finds conditions for the minimum payment constraints to be slack (the outside options must be sufficiently high).

3.3 Manager's Preferred Evaluation Strategy

The first observation is that for any manager's contract, the manager strictly benefits from higher worker's effort. Note that given c_W, and a manager's contract π_M, one can write the manager's expected payoff as

$$\begin{aligned} V_M(c_W,\pi_M) &= \sum_Y u_M\big(\pi_M(y)+b(y)\big)\left[p(y|1)G(c_W)+p(y|0)(1-G(c_W))\right] \\ &= \sum_Y u_M\big(\pi_M(y)+b(y)\big)p(y|0)+G(c_W)\sum_Y u_M\big(\pi_M(y)+b(y)\big)\left[p(y|1)-p(y|0)\right]. \end{aligned}$$

Lemma 2. *Take any contract such that $\pi_M(\cdot)$ is non-decreasing. Then, the manager strictly benefits from higher worker's effort. That is, $V_M(c_W,\pi_M)$ is strictly increasing in c_W.*

As the manager's payoff is strictly increasing in output, and higher worker's effort increases the likelihood of higher output realizations, the manager strictly benefits from higher worker's effort. The manager wants to motivate high effort from the worker. However, as the manager does not pay the worker from her own pocket, she does not take into account how much it costs to compensate the worker.[21] Then, the manager's preferred evaluation strategy must maximize the worker's effort cost cutoff.

Proposition 1. *Take any contract (E,π_W,π_M) such that $\pi_M(\cdot)$ is non-decreasing. Let $\underline{e}_y \in \underset{\tilde{e}\in E}{argmin}\{\pi_W(y,\tilde{e})\}$ and $\overline{e}_y \in \underset{\tilde{e}\in E}{argmax}\{\pi_W(y,\tilde{e})\}$ be, respectively, the lowest and highest paying messages given an output realization $y\in Y$. Define the cutoff*

$$z^*(y) := inf\{z\in Z : q(z|1)p(y|1) \geq q(z|0)p(y|0)\}.^{22}$$

Then, the manager's preferred evaluation strategy is

$$\sigma^*(a_M,y,z) = \begin{cases} \delta_{\underline{e}_y} & if\ z < z^*(y) \\ \delta_{\overline{e}_y} & otherwise, \end{cases}$$

where δ_e denotes the Dirac measure centered on e. Moreover, the distribution of payments is unique. Any manager's preferred evaluation strategy generates the same distribution of payments as σ^*.

Note that for the manager to maximize the worker's incentives for effort, she must generate the highest expected utility difference between high and low effort. In doing so, the manager provides the highest compensation to any pair of signals (y,z), which are more likely to realize under high effort than under low effort. Also, the manager reports the lowest paying performance whenever the realized signal pair (y,z) is associated with low effort. Formally, for a given contract π_W and performance evaluation σ, the worker's

[21] In a related context, Benson (2015) documents empirical evidence of sales managers increasing their subordinates incentives beyond what is desirable by the firm by shifting sales across periods or changing sales targets.

[22] Let $z^*(y) = \underline{z}$ if the set $\{z\in Z : q(z|1)p(y|1) \geq q(z|0)p(y|0)\}$ is empty.

effort cost cutoff is given by

$$\begin{aligned} c_W(\pi_W,\sigma) =& \mathbb{E}\Big[u_W\big(\pi_W(\mathbf{y},\mathbf{e})\big)\big|a=1,\sigma\Big] - \mathbb{E}\Big[u_W\big(\pi_W(\mathbf{y},\mathbf{e})\big)\big|a=0,\sigma\Big] \\ =& \int_Z \sum_Y \sum_E [u_W\big(\pi_W(y,e)\big)\sigma(e|y,z)\Big[q(z|1)p(y|1) - q(z|0)p(y|0)\Big]dz. \end{aligned}$$

The σ that maximizes $c_W(\pi_W,\sigma)$ is the one that reports the highest paying message when $q(z|1)p(y|1) > q(z|0)p(y|0)$, and the lowest paying message otherwise.

Note that the manager sends a performance report after observing (y,z). If a pair (y,z) is more likely to appear when the worker has exerted high effort, that pair can be seen as a "good signal". By increasing the payments for such realization pairs, the manager increases the difference between the worker's expected utility when exerting effort and not, which raises the worker's cost cutoff for high effort. Similarly, if a pair (y,z) is more likely to realize when the worker has exerted low effort, decreasing payments increases incentives for effort. Proposition 1 highlights the agency conflict between principal and manager: while the first is concerned with the wage bill, the second only wants to motivate as much effort as possible. As a result, the manager uses only the highest and the lowest paying ratings at each output level.

Proposition 1 is consistent with the empirical evidence that documents that very few points of performance evaluation scales are used in practice. Frederiksen et al. (2017) highlights that this feature is documented by several articles using personnel data from different firms. In a separate article, Frederiksen et al. (2020) study performance evaluations in a large Scandinavian service sector firm. They show that two out of five possible points on the scale concentrate over 90% of all performance ratings. However, it is essential to highlight that Proposition 1 states that only the lowest and the highest *paying* messages will be used, but it is silent about which ratings are used if multiple ratings provide the same payments. For instance, suppose a given firm has a five-point evaluation scale — one to five stars. If payments are strictly increasing in the evaluation, the result implies that only ratings one and five would be used. However, if, for instance, the payments are such that low performers (ratings one and two) are punished while high performers (ratings four and five) receive a fixed bonus, then the prediction would only be that rating three would never be used.

Perhaps, however, the central message of Proposition 1 is that the principal can — without loss — restrict attention to binary performance evaluation scales. Even if the principal tries to implement a more granular scheme, the manager only uses the two extremes. Therefore, there is no loss in restricting attention to contracts with only two messages. From now on, I refer to them as good (g) and bad (b) performance ratings.

Corollary 1. *The principal can do no better than what is feasible with binary messages, good and bad* $\{g,b\}$.

The restriction to binary rating scales generates straightforward mechanisms in which the manager's discretion over the worker's payment boils down to simply deciding whether the worker gets a bonus of a pre-determined size or not. Such a binary structure of the compensation is consistent with the commonly observed fixed-bonus contracts or up-or-out systems (Murphy (1999) and Holmström (2016)), which are particularly prevalent in consulting and banking; see PageExecutive (2020) and Charles Aris Inc. (2022) for surveys.

Remark 2. *Corollary 1 states no loss in restricting attention to binary performance evaluation systems, but it does not imply that the performance review scale must be binary. For example, the principal could reproduce any binary performance review system with a direct revelation contract by replicating* σ^* *directly on payments. The meaningful content of the previous result is that conditional on an output realization* y*, the realized payments must take at most two different levels. For instance, when using a direct truthful revelation contract, the worker's compensation conditional on output take a cutoff form. It takes a high value if the reported* z *is high enough and a low value otherwise.*

Another direct implication of Proposition 1 relates to how lenient the manager is. Given the binary structure of the incentive scheme, the manager reports a good performance if the unverifiable private signal realization is high enough. Note, however, that "high enough" depends on the output realization. In particular, the manager is less demanding the higher the output.

Corollary 2. *The manager is more lenient the larger the output. That is, the cutoff* $z^*(y)$ *for reporting* g *is decreasing in* y*.*

Note that the constructed σ^* implies that the performance evaluation is g if and only if

$$\frac{q(z|1)}{q(z|0)} \geq \frac{p(y|0)}{p(y|1)}, \tag{1}$$

or equivalently if $z \geq z^*(y)$. Note that the monotone likelihood ratio assumption (Assumption 1) implies the left-hand-side of (1) is increasing in z while the right-hand is decreasing in y. Hence, the lower the output realization, the higher the manager's cutoff for providing a good performance evaluation. That is, the manager is more lenient in her reporting when output is high.

3.4 Cost-minimizing Contracts

I have already established two of the principal's choices: the optimal set of performance evaluations $E = \{b, g\}$, and the recommended evaluation strategy σ^*. By replacing such choices in the principal's problem, one can re-write it as

$$\mathscr{C}(c_W) := \min_{(\pi_W, \pi_M)} \quad \mathbb{E}\Big[\pi_M(\mathbf{y}) + \pi_W(\mathbf{y}, \mathbf{e}) \Big| c_W, \sigma^*\Big] \tag{2}$$

subject to (IR_M), (IR_W), (IC_W), and $\pi_W(y, g) \geq \pi_W(y, b)$ for all $y \in Y$.

The constraint requiring worker's payments to be higher when the evaluation is good ensures that the report g has the desired good performance meaning, and the report b has the poor performance connotation. It ensures the report g is associated with the highest payment and b with the lowest for each output realization $y \in Y$, which implies that (IC_M) is satisfied.

Note that p and q fully characterize σ^*, which has an important implication: one can treat the information structure of the signals observed by the principal as exogenous. That is, the principal observes $(\mathbf{y}, \mathbf{e})$ which are distributed according to p, q, and σ^*, and the latter is characterized by the former two. As usual, in moral hazard problems, it is useful to define the score associated with a pair $(\mathbf{y}, \mathbf{e})$. First, define the probability of observing (y, e) if both agents follow the recommendations (meaning effort cutoff c_W and evaluation strategy σ^*) as $f : Y \times E \to [0, 1]$. That is, let

$$f(y,b) := G(c_W)p(y|1)Q(z^*(y)|1) + [1 - G(c_W)]p(y|0)Q(z^*(y)|0)$$

$$f(y,g) := G(c_W)p(y|1)[1 - Q(z^*(y)|1)] + [1 - G(c_W)]p(y|0)[1 - Q(z^*(y)|0)].$$

Then, for each pair (y, e) such that $f(y, e) > 0$, define the score as

$$s(y,b) := \frac{p(y|1)Q(z^*(y)|1) - p(y|0)Q(z^*(y)|0)}{f(y,b)},$$

$$s(y,g) := \frac{p(y|1)[1 - Q(z^*(y)|1)] - p(y|0)[1 - Q(z^*(y)|0)]}{f(y,g)}.$$

The score is an increasing function of the likelihood ratio of signals, and the higher the score, the more the signal realization is associated with high effort. Under Assumptions 1 and 2, $s(\cdot, i)$ is increasing for each $i \in \{b, g\}$, which implies that higher output realizations are more strongly associated with high effort. Using this notation, I can now describe the cost-minimizing compensation schemes.

Proposition 2. *Suppose Assumptions 1-2 hold. The cost-minimizing manager's compensation scheme is given by*

$$\pi_M^*(y) = \bar{\pi}_M = inf\left\{x \in \mathbb{R}_+ : \mathbb{E}\left[u_M\left(x + b(y)\right)\middle|c_W\right] \geq \bar{u}_M\right\} \text{ for all } y \in Y. \tag{3}$$

Moreover, there exists $(\lambda, \mu) \in \mathbb{R}_{++}$ such that the cost-minimizing worker's compensation scheme is given by

$$\frac{1}{u_W'\left(\pi_W^*(y,i)\right)} = \lambda + \mu \cdot s(y,i) \quad \text{for all } y \in Y \text{ and } e \in \{b, g\}, \tag{4}$$

where (λ, μ) are such that (IR_W) and (IC_W) hold with equality.

Proposition 2 implies that the principal sets the manager's utility at the lowest possible level that satisfies the participation constraint. Even though the principal controls transfers to the manager, she cannot solve their agency conflict. The non-verifiable nature of the manager's information prevents the principal from being able to condition the manager's payments on her reports. Hence, the best the principal can do is to offer the manager a flat wage.[23]

However, the worker must be incentivized, and his payment depends both on the performance evaluation and the output realization. Note that the optimal compensation scheme takes the canonical Holström-Mirlees form, in which the score of the signal observed by the principal is a sufficient statistic. Below, I plot an

[23] When the manager is strictly risk-averse, the flat wage is uniquely optimal. However, all results remain valid when the manager is risk neutral; the only change is that there would be multiple optimal π_M's.

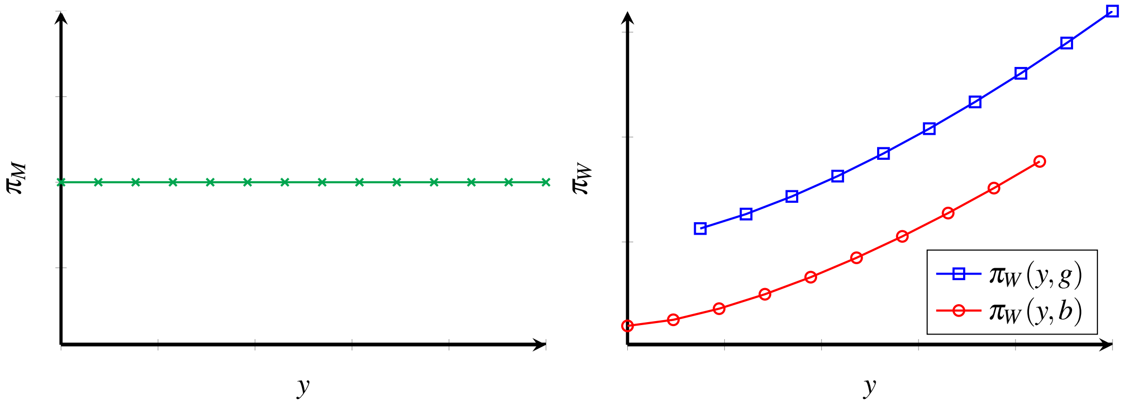


Figure 1: Illustration of optimal compensation schemes

example of an optimal compensation scheme as a function of output. The manager's payment depends only on output and is increasing in y. The worker's compensation is characterized by two curves, one when the performance evaluation is bad and the other when it is good. For intermediate values of y, the performance report informs the principal whether $\underline{z} < z^*(y) < \bar{z}$. Hence, for each output realization, the principal delegates to the manager the choice of whether the worker gets the wage described by the square (blue) curve or the circle (red) curve. Relating it to the earlier retail store example, the manager simply decides whether the worker gets a bonus beyond the sales commission.

Note that in this example, only one of the curves for extreme output values exists. This feature arises whenever the manager's report becomes invariant to her private signal realization for sufficiently extreme output levels. For example, suppose that there exists $\hat{y} \in Y$ such that $q(\bar{z}|1)p(\hat{y}|1) < q(\bar{z}|0)p(\hat{y}|0)$. Then, whenever the realized output is below $\hat{y}$, the manager reports a bad performance, regardless of her private signal realization. Similarly, if there exists $\tilde{y}$ such that $q(\underline{z}|1)p(\tilde{y}|1) > q(\underline{z}|0)p(\tilde{y}|0)$, then the report is good for any $y \geq \tilde{y}$, regardless of the private signal realization.

3.5 When are subjective performance evaluations valuable?

Subjective performance evaluations are present in many situations, but not always. Holmström (1979) shows that the principal would like to condition payments on any non-redundant information about efforts. One could then expect subjective evaluations to be always valuable. I show here that the performance evaluation might not be informative, despite the primitive private non-verifiable signal being. Although $\mathbf{z}$ is informative about a_W, the evaluation $\mathbf{e}$ does not generate any additional information beyond what is already conveyed by $\mathbf{y}$. That is, the output might be a sufficient statistic for $(\mathbf{y}, \mathbf{e})$.

Note that in the framework presented here, output $\mathbf{y}$ can be seen as hard information — meaning public and verifiable — while $\mathbf{z}$ can be seen as soft information — meaning private and non-verifiable. Depending on the industry or sector, each type of signal might be more informative about effort than the other. I will refer to a setting as *hard-information-intensive* (or HII) when output is more informative than the managers' private non-verifiable information. For instance, assembly lines in manufacturing or sales departments can be seen as HII settings since the number of units produced or sold conveys most of the relevant information. I then refer to settings as *soft-information-intensive* (or SII) when the soft information is sufficiently informative to reverse the inference about the effort that would be made based solely on output. SII settings are the ones in which output is very noisy or the context in which the actions are taken matters greatly. Management consulting or research and development activities can be thought of as examples of SII settings. Formally, HII and SII are defined based on the relationship between the likelihood ratios of the verifiable and the non-verifiable signals.

Definition 3. *A setting is* ***hard-information-intensive (HII)*** *if for every $y \in Y$*

$$q(\overline{z}|1)p(y|1) \leq q(\overline{z}|0)p(y|0) \text{ or } q(\underline{z}|1)p(y|1) \geq q(\underline{z}|0)p(y|0). \tag{5}$$

A setting is ***soft-information-intensive (SII)*** *if it is not HII.*

I then define objective and subjective compensation systems; and what it means for subjective evaluations to be valuable.

Definition 4. *I say that a contract π_W is* ***objective*** *if π_w does not depend on the performance report. That is, for all $y \in Y$, $\pi_W(y,e) = \pi_W(y,\hat{e})$ for all $e, \hat{e} \in E$.*

Definition 5. *I say that a contract π_W is* ***subjective*** *if it is not objective.*

Definition 6. *Fix the effort level the principal wants to implement $c_W > 0$. I say that* ***subjective performance evaluations are valuable*** *if there exists a subjective contract such that the implementation cost is strictly lower than the implementation cost under any objective contract.*

Proposition 3. *Subjective performance evaluations are valuable if and only if the setting is soft-information-intensive.*

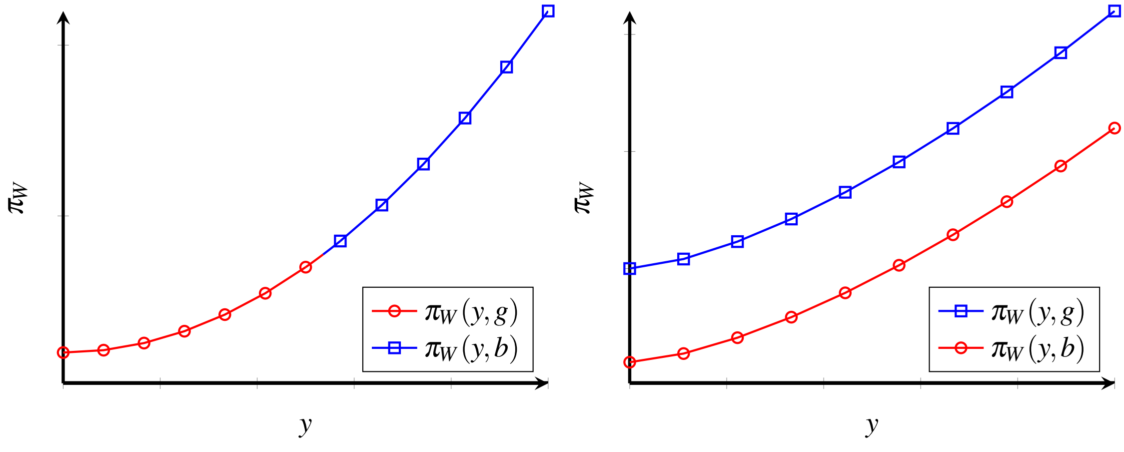


Figure 2: HII vs. SII setting

Proposition 3 states that the principal only benefits from using subjective reports in SII settings. Hence, one should observe subjective evaluations being more widely used in such cases, while in HII settings, one should observe contracts that only condition on output. While seemingly very intuitive, the result seems to be at odds with the information principle, which states that any additional information is valuable, regardless of how noisy. However, Proposition 3 actually stems from the information principle. The central idea is that any information *the principal obtains* is valuable. However, whether the manager's reports convey additional information on top of what output conveys depends on whether it is a HII or SII setting.

On the one hand, in an HII setting, the output generates such dispersed likelihood ratios that the manager's performance evaluation is invariant to the soft information. The manager provides a good report if the output is sufficiently high and a bad otherwise, regardless of her private information. Hence, the performance evaluation does not convey any information beyond what $\mathbf{y}$ conveys. On the other hand, in an SII setting, the manager's report provides information beyond output and, hence, is valuable.

Figure 2 contrasts the two possibilities: the plot on the left side displays a case in which the manager always reports a bad performance for low output realizations and a good performance for high realizations (HII setting). Hence, the evaluation does not convey additional information beyond what the output reveals. The plot on the right side represents the other extreme, in which the manager's evaluation conveys additional information for any output realization (SII setting).

4 Organizational Transparency

In the analysis so far, I have taken the information structure describing what the manager observes about the worker's effort as exogenous. When setting up the organization structure and production processes, a firm might affect how much information one employee observes about the other. For instance, the physical architecture of the workplace might influence the information flow. A manager might, for example, better monitor her subordinate in an open-space office plan. In another example, the firm might be able to decide whether or not to allow the employees to work remotely, and if they do, choose how much information the manager can observe about their subordinates, such as e-mail activity and log-in trackers[24].

This section extends the previous analysis to understand how informed a principal would like the manager to be about her subordinate's actions. Given that manager's and principal's interests are not fully aligned, I ask whether the principal could mitigate the conflict of interest by reducing the information available to the manager.

I start with a benchmark case in which the principal could choose an arbitrarily informed manager. That is, I let the principal select the distribution of the manager's private signal $\mathbf{z}$, including the possibility of a fully-informed manager who perfectly observes effort. Then, I show that a fully-informed manager is optimal from the principal's perspective.

Remark 3. *Suppose the principal can choose any distribution for the manager's private signal* $\mathbf{z}$ *and wishes*

[24]For a discussion on monitoring remote workers see Kurkowski (2021), "Monitoring Remote Workers: The Good, The Bad And The Ugly", *Forbes*, December 08, 2021

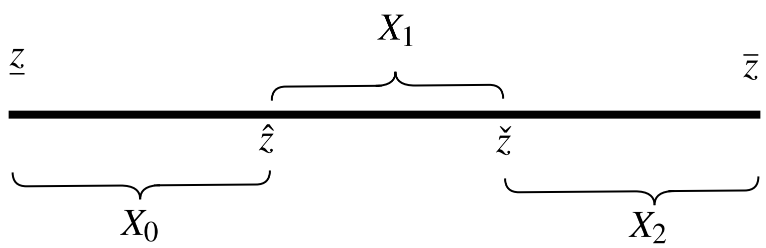
to implement $c_W > 0$. The principal prefers the fully-informed manager, the manager's contract is characterized by (3), and the worker's payments are given by

$$u_W\left(\pi_W^*(y,b)\right) = \bar{u}_W - G(c_W)\Big[c_W - \mathbb{E}[\mathbf{c}|c \le c_W]\Big] \quad \textit{and} \quad u_W\left(\pi_W^*(y,g)\right) = u_W\left(\pi_W^*(y,b)\right) + c_W.$$

In earlier sections, I have shown that the manager would use only two messages: the highest-paying when it is more likely that the worker has exerted effort, and the lowest-paying when it is more likely that he has not. Under full information, the manager observes a directly, that is, she reports g if $a = 1$ and b if $a = 0$. Thus, the principal's and the manager's interests are fully aligned under full information. The principal can act as if she observed effort directly. Then, binding (IC_W) and (IR_W) characterize the optimal payments.

Suppose now that the fully-informative signal is not feasible. Note that the manager's and principal's interests are misaligned. The principal would like the report to reveal all information possible about a_W. In contrast, the manager's reporting strategy pools all signals with a positive score at the highest-paying message and all with a negative score at the lowest. I now address the question of whether reducing transparency could help to align incentives. Would the principal benefit from reducing the informational content on the manager's private signal? In particular, can the principal benefit from the manager observing the realization in a coarser partition of the signal space instead of fully observing $\mathbf{z}$'s realization?

I allow the principal to choose a coarser partition of Z as the manager's information. For example, if the principal chooses full transparency, the manager observes the realized z perfectly. Otherwise, the principal can choose any other partition of $[\underline{z}, \overline{z}]$. For instance, take arbitrary $\hat{z}, \check{z} \in (\underline{z}, \overline{z})$. The principal could choose a partition $\left\{[\underline{z}, \hat{z}]; (\hat{z}, \check{z}); [\check{z}, \overline{z}]\right\}$. That is, the manager only observes in which subset of the partition the realized z is, but not the realization of z itself. Graphically, one can represent this partition as



where the manager would see only the realization of $\mathbf{X}$, but not $\mathbf{z}$.

Take the primitive distributions p and q satisfying Assumptions 1 and 2, and a desired effort level $c_W > 0$ to be enforced. Given her information, the manager provides a good evaluation if the score of the observed signals (including output) is positive and a bad performance evaluation otherwise. I further assume that subjective performance evaluations are informative for any output realization under full transparency.

Assumption 3. *p and q jointly satisfy:*

1. *$q(.|a)$ is continuous for any $a \in \{0,1\}$.*

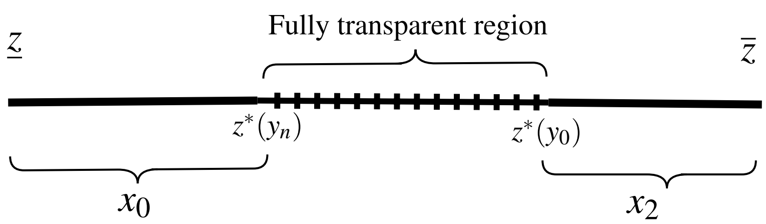
2. $z^*(y) \in (\underline{z}, \overline{z})$ *for all* $y \in Y$.

Note that $z^*(\cdot)$ is fully determined by the primitive distributions p and q. Hence, Assumption 3 can also be stated directly in terms of the distributions. The first item is assumed for technical convenience. The second ensures that the setting is SII and that under full-transparency subjective evaluations are valuable for any output realization.[25] If subjective performance evaluations were not valuable even under full transparency, then they would not be valuable if the manager had even less information. Assumption 3 rules out those uninteresting cases in which there is no scope for subjective performance evaluations.

Proposition 4. *Suppose Assumptions 1-3 hold. Then, full transparency is not optimal. In particular, the principal strictly benefits from pooling extreme signals. That is, there exists* $\hat{z}$ *and* $\check{z}$ *such that the cost of enforcing a given* $c_W > 0$ *is strictly lower if the manager's information is given by the partition* $\left\{[\underline{z}, \hat{z}]; \{z\}_{z \in (\hat{z}, \check{z})}; [\check{z}, \overline{z}]\right\}$ *instead of the finest partition* $\left\{\{z\}_{z \in [\underline{z}, \overline{z}]}\right\}$.

Proposition 4 states that the principal strictly benefits from reducing the manager's information. One of the manager's roles is to monitor the worker's effort. The principal benefits from reducing transparency because the manager and the principal's interests are not fully aligned. The manager wants to maximize incentives for the worker's effort regardless of risk-sharing. When deciding how to evaluate the worker, the manager relies not only on her private signal but also on the output realization. As a result, the manager reports a good performance more often (too often from the principal's perspective) when output is high and more rarely (too rarely from the principal's perspective) when output is low. By reducing transparency and censoring extreme private signals, the principal reduces the likelihood of a bonus when output (and worker's compensation) is high and increases the likelihood of a bonus when output (and worker's compensation) is low, which improves risk-sharing and reduces compensation costs. Note that the fact that the principal wishes to insure the worker while the manager does not stems directly from the key agency conflict between manager and principal in the basic framework: the evaluator (manager) does not foot the wage bill.[26]

When proving this result, the first step is to notice that there is no loss in pooling all signals below $z^*(y_n)$ and (in a separate region) pooling all signals above $z^*(y_0)$. The fully-transparent partition and the partition below generate the same performance reports.



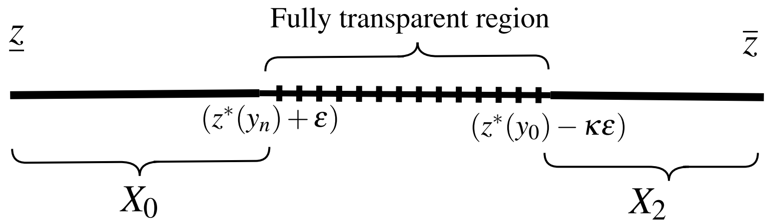
In this partition, the manager observes whether $z \leq z^*(y_n)$; whether $z \geq z^*(y_0)$ or z perfectly in between. The first set X_0 corresponds to z's such that the manager would provide a bad evaluation regardless of output,

[25] SII only requires that subjective evaluations are valuable for some output realization, not necessarily all.

[26] note also that such an implication is not related to the manager's risk attitude. Regardless of whether the manager is risk-neutral or risk-averse, as she does not pay the worker, she does not benefit from smoothing out the worker's compensation.

whereas the set X_2 includes only the z's for which the evaluation would be good regardless of output. This coarser partition generates the same performance reports as the fully transparent one. Hence, the same compensation costs to the principal.

Next, I show that the principal is strictly better off by increasing the first cutoff and decreasing the second by a small amount. That is, there exists $\varepsilon > 0$ and $\kappa > 0$ such that the compensation costs are strictly smaller under the following partition:



The difficulty in comparing different information structures is that the worker's contract depends on the distribution of scores. By changing the cutoff by a small amount, I can apply the Envelope Theorem and look just at the direct effect. I look at the marginal impact of ε on compensation costs evaluated at $\varepsilon = 0$ and show that it is strictly negative. By pooling extreme signals a little further, the principal strictly reduces costs.

Increasing ε has three effects: first, there is a direct effect. For fixed contracts, a change in ε changes the distribution of signals and expected payments. Second, after changing ε, the principal must adjust payments so the worker is still willing to participate. Third, the principal must adjust payments such that c_W is still implemented as the worker's effort cost cutoff. I construct κ such that the second effect is zero. As the manager chooses the evaluation strategy that maximizes c_W, the third effect is second-order at $\varepsilon = 0$. Hence, only the direct effect remains. The direct effect of an increase in ε is to decrease the probability of a good report when output is high (y_n) and to increase the likelihood when output is low (y_0). Comparing how it impacts the principal's costs boils down to a trade-off between a higher chance of the worker getting a bonus under low versus high output. As the worker is strictly risk-averse, it is cheaper to provide bonuses when output is low (y_0) than when output is high (y_n).

5 Multiple Workers and Forced Rankings

The analysis in the single-worker case points in the direction that the principal would benefit from restricting the evaluation strategies the manager can use. When the manager oversees multiple workers, a simple — and widely used — form of restricting the manager's evaluations is to require the manager to rank workers. Forced rankings can help tie the manager's hands but might also be detrimental because someone must be ranked at the bottom and receive a low payment even when all workers had a good performance (or someone must be at the top despite everyone performing poorly). This section attempts to understand such a trade-off and find conditions under which forced rankings outperform individual evaluations and vice-versa.

Suppose that there are k workers $i \in K := \{1,\dots,k\}$ instead of one. Output depends on both workers' efforts, where $p(y|\sum_{i\in K} a_i)$ denotes the probability of output y conditional on efforts $a := (a_1,\dots,a_k) \in \{0,1\}^k$. I assume that p satisfies MLRP for each effort conditional on the other workers' efforts and that efforts are complements.[27] Besides output, the manager also observes one non-verifiable, identically distributed, and conditionally independent private signal about each worker's performance denoted by $\mathbf{z_i} \sim Q(\cdot|a_i)$.

Assumption 4. *P, Q and G satisfy:*

1. *p satisfies MLRP:* $p(y|m+1)/p(y|m)$ *is strictly increasing in y for all* $m \in \{0,\dots,k-1\}$.
2. *Efforts are complements: the p.d.f.* $P(y|\sum_{i\in K} a_i)$ *is submodular in* (a_1,a_2) *for all* $y \in Y$.
3. *q satisfies MLRP:* $q(\cdot|1)/q(\cdot|0)$ *is increasing.*
4. *G is strictly concave.*

I then compare two classes of mechanisms: the first is the *individual performance mechanisms (IP-mechanisms)*, in which one worker's compensation does not depend on the reports about other workers. That is, the compensation to each agent is given by $\pi_i(y,e_i)$, where e_i denotes the evaluation of i's performance. The second class is the *forced ranking mechanisms (FR-mechanism)* in which the manager is required to simply rank the workers.[28] I then present sufficient conditions over the information structure under which each of those classes outperforms the other.

The first result states that if the non-verifiable information is sufficiently informative, the principal is better off using an individual performance evaluation mechanism. Before formally stating the result, I define what I mean by "sufficiently informative". I say that a distribution $Q^{FI} : [\underline{z},\bar{z}] \times \{0,1\} \to [0,1]$ is fully-informative if it admits a density q^{FI} and there exists $z_0 \in (\underline{z},\bar{z})$ such that the density $q^{FI}(z|1) > 0$ if and only if $z > z^0$, and $q^{FI}(z|0) > 0$ if and only if $z < z^0$. That is, by observing z, one can infer with probability one whether the effort was high or low. I say that a distribution Q is sufficiently informative if its density is sufficiently close to q^{FI}.

Proposition 5. *Suppose the principal wants to implement symmetric effort cost cutoffs* $(c,\dots,c) \in \mathbb{R}^k_{++}$. *Consider a sequence of problems, each with the non-verifiable information distribution denoted by* Q_n *and satisfying Assumption 4. Suppose that the sequence of densities* $\{q_n\}_{n\in\mathbb{N}}$ *converge almost everywhere to a fully informative density* q^{FI}. *Then, there exists* $N \in \mathbb{N}$ *such that for any* $n > N$, *the principal prefers an individual performance evaluation mechanism over any forced ranking mechanism.*

Proposition 5 states that IP-mechanisms outperform forced rankings if the manager's private non-verifiable information is sufficiently informative about effort. On the one hand, when using a FR-mechanism, the most

[27] It is usual in the literature (e.g., Winter (2004), Halac et al. (2021)) to focus on the case in which efforts are complements. Moreover, in the Appendix, I provide an ordinal notion of effort complementarity equivalent to the submodularity condition assumed below.

[28] Consistently with the single worker case, I select the manager's preferred equilibrium whenever there are multiple continuation equilibria.

information the principal can hope to get from the manager's report is the order of the non-verifiable signals z_i's. However, such a ranking does not reveal whether those workers had all very high, low, or disparate signal realizations. On the other hand, when using an IP-mechanism, the manager's reports reveal whether each worker's performance was sufficiently high. Intuitively, FR-mechanisms provide information about relative performance, while IP-mechanisms mainly provide information about absolute performance. As the non-verifiable information gets precise, the value of such absolute performance information becomes larger than that of relative performance. I then show that this ordering is reversed when the manager's private information becomes noisier.

Proposition 6. *Suppose the principal wants to implement effort cost cutoffs $(c,\dots,c) \in \mathbb{R}^k_{++}$, that Assumption 4 holds, and that for each $y \in Y$ one of the following two conditions holds*

$$\min_{m\in\{0,\dots,k-1\}} \left\{ \frac{p(y|m+1)}{p(y|m)} \right\} > \frac{q(\underline{z}|0)}{q(\underline{z}|1)} \quad or \quad \max_{m\in\{0,\dots,k-1\}} \left\{ \frac{p(y|m+1)}{p(y|m)} \right\} < \frac{q(\overline{z}|0)}{q(\overline{z}|1)}. \tag{6}$$

Then, the principal is strictly better off by using a forced ranking mechanism over any individual performance mechanism.

The likelihood ratio of a given signal realization conveys how much more likely one is to observe a given signal realization under high versus low effort. If the manager's private information is very noisy, then the likelihood of observing a given realization does not depend much on whether the worker has exerted high effort or not. Hence, the likelihood ratio of any signal $q(z|1)/q(z|0)$ will be close to one. Condition (6) is akin to the HII definition but in a setting with multiple workers. Hence, one can interpret Proposition 6 as forced rankings dominating individual performance evaluations in HII settings. As argued in Proposition 3, when the manager's non-verifiable signal is noisy, and she evaluates each worker individually, she will not condition her report on her private information. Therefore, individual performance reports will not reveal any additional information beyond what is already conveyed by output. Note, however, that if the principal uses a forced ranking mechanism, she can at least get information on how workers' performances compare to each other. A forced ranking allows the principal to elicit information that would not be feasible under an IP-mechanism.

There is an extensive debate on the pros and cons of using forced rankings or, more generally, relative performance evaluations. Among the pros, the most prominent one is that relative performance evaluations are a way to filter out external factors that affect all workers simultaneously. Among the cons, it is often mentioned fairness concerns or favoritism. Propositions 5 and 6 abstract from such considerations and point towards costs and benefits related to the conflict of interests between the evaluator (manager) and the residual claimant (principal). Forced rankings limit the control of managers over the performance reports, which, on the one hand, reduces the manager's ability to use evaluations in an undesirable way but, on the other hand, limits the amount of information that can be provided.

6 Discussion

This section describes ways to enrich my model, points to how the analysis changes, and identifies a few questions for future research.

Continuous efforts.— One might wonder whether the binary performance evaluation result directly stems from the binary effort assumption. It does not. Even if the worker's effort is a continuous variable, I show that the manager still uses only the two extreme messages as long as the first-order approach is valid.

Denote by p_a and q_a the derivative of the probability mass and density of $\mathbf{y}$ and $\mathbf{z}$ with respect to an effort level $a \in [0,1]$. The worker's effort choice first-order condition is

$$\frac{\partial V_W}{\partial a}(\hat{a}, \pi_W, \hat{\sigma}) = \int_Z \sum_Y \sum_E u_W\left(\pi_W(y,e)\right)\hat{\sigma}(e|y,z)\left[\frac{p_a(y|\hat{a})}{p(y|\hat{a})} + \frac{q_a(z|\hat{a})}{q(z|\hat{a})}\right] p(y|\hat{a})q(z|\hat{a})dz - c'(\hat{a}) = 0.$$

Then, similarly to the binary effort case, the manager maximizes incentives for effort by reporting the highest paying message when $\left[p_a/p + q_a/q\right]$ is positive, and the lowest paying when it is negative. In Appendix C, I provide the formal description of the model with continuous efforts and conditions that assure the validity of the first-order approach in this setting.

Manager Exerting Productive Effort.— Throughout the paper, I have assumed the manager exogenously benefits from higher output realizations (the function b is exogenous). Such an assumption can be justified as a reduced-form model capturing direct incentives to the manager regarding output. One could augment the baseline model to make the manager's payments fully endogenously determined. For instance, suppose that, besides evaluating the worker, the manager must also be motivated to exert productive effort. That is, suppose that the output distribution depends not only on the worker's effort but also on the manager's. If the principal wants to motivate the manager to exert effort, she must condition the manager's payments on output. In particular, if higher output levels are associated with higher effort (in a likelihood ratio sense), then the manager's payments must be increasing in output, which was replicated by the exogenous increasing b in the baseline model.[29]

Screening.— In the main analysis, I have restricted contracts to depend only on the realized output and the reported performance. In particular, I did not allow the principal to offer a menu of contracts to the worker, who would then choose her preferred option conditional on her cost realization. This modeling choice is consistent with the lack of menu offerings in labor relations but raises the question of whether the results would differ when considering a larger contracting space.

The answer crucially depends on whether the manager observes which contract the worker has chosen. For instance, suppose that the principal designs a menu of contracts the worker can choose from, but the manager knows the worker's choice. For example, that would be the case if the manager were actively hiring

[29]The full characterization of optimal contracts also depends whether higher effort from one agent (manager or worker) increases or decreases the incentives to the other. That is, the characterization of the optimal contracts depends on whether efforts are complements or substitutes. However, the main message of using binary performance reports remains valid. In an earlier version of the paper, I analyze such a case. See Chapter 1 of Brasiliense de Castro Pires (2022) for details.

the worker. In such a case, one could replicate the baseline analysis in this paper contract by contract. The principal would still not be able to condition the manager's payments in her reports, and, for each contract, the worker could choose, the manager would use only the highest and lowest paying messages.[30] However, if the manager did not observe the worker's contract choice, the principal could use this information to cross-check the manager's performance report. Workers who choose steeper contracts are more likely to exert high effort and, hence, have a higher chance of generating higher signals. The principal could then pay more to the manager when she reports a high (low) performance and the worker has chosen a steep (flat) contract.

This discussion around screening sheds light on the role of the hierarchy as a communication structure inside the organization. When the top and bottom of the organization do not communicate directly, which often happens in practice, the manager has a crucial role as an intermediary. Such a role, combined with agency conflicts between the principal and manager, affects the information flow in the organization, the distribution of performance reports, and the shape of the optimal incentive schemes.

Large Number of Workers.— In Section 5, I show that forced rankings might outperform individual performance evaluations. A natural question is whether the principal could use a forced ranking mechanism to approximate the verifiable signals benchmark (observable $\mathbf{z}$) when the number of workers is sufficiently large. The central intuition — akin to Jackson and Sonnenschein (2007) — would be that when the number of workers is large, a worker's ranking position would approximately reveal his non-verifiable signal. The caveat is that such a claim is valid only if the manager prefers to rank workers truthfully. However, establishing that the manager prefers to rank workers truthfully is far from trivial.[31] It is possible, for instance, that the manager would be better off by providing stronger incentives to one worker by only ranking such a worker in the highest or lowest positions than by reporting truthfully. Characterizing conditions under which the manager ranks workers truthfully is left as an exciting avenue for future research. Addressing this question might speak to applications beyond the one discussed here, for instance, the incentives of self-interested third-party tournament committees to rank participants truthfully.

A Appendix

A.1 Preliminary results

The following Lemma (Beesack (1957)) is central to my analysis.

Lemma 3 (Beesack's inequality)**.** *Let $r : X \to \mathbb{R}$ be an integrable function with domain an interval $X \subseteq \mathbb{R}$. Assume that r is never first strictly positive and then strictly negative and that $\int_X r(x)dx = 0$. Then, for any*

[30] For a detailed analysis, see Chapter 1 of Brasiliense de Castro Pires (2022).

[31] Formally, the manager is choosing a distribution over rankings conditional on a vector of signal realizations to maximize her expected payoff.

increasing function $h : X \to \mathbb{R}$ such that rh is integrable,

$$\int_X r(x)h(x)dx \geq 0.$$

Beesack's inequality implies the following Lemma, which I use extensively.

Lemma 4. *Let $T = \{t_1, \ldots, t_n\}$ be a finite subset of $\mathbb{R}$ and $F : T \to \mathbb{R}$. Assume F is never first strictly positive and then strictly negative and that $\sum_{t \in T} F(t) = 0$. Then, for any increasing function $M : T \to \mathbb{R}$*

$$\sum_{t \in T} F(t)M(t) \geq 0.$$

Furthermore, if F is strictly increasing and M is non-constant, then the inequality is strict.

Proof of Lemma 4. Let $X = [0,1)$ and $r : X \to \mathbb{R}$, $h : X \to \mathbb{R}$ be such that

$$r(x) = F(t_i) \text{ if } x \in \left[\frac{(i-1)}{n}, \frac{i}{n}\right) \text{ for } i \in \{1, \ldots, n\},$$

$$h(x) = M(t_i) \text{ if } x \in \left[\frac{(i-1)}{n}, \frac{i}{n}\right) \text{ for } i \in \{1, \ldots, n\}.$$

Note that h is increasing. Also, r is never first strictly positive and then strictly negative and

$$\int_X r(x)dx = \frac{1}{n} \sum_{t \in T} F(t) = 0.$$

By Beesack's inequality

$$\sum_{t \in T} F(t)M(t) = n \int_X r(x)h(x)dx \geq 0.$$

We are done with the first part of the Lemma. For the second part (strict inequality), let F be strictly increasing and M non-constant. Define two sets $T^- := \{t \in T | F(t) < 0\}$ and $T^+ := \{t \in T | F(t) \geq 0\}$. Note that as F is strictly increasing, both sets are non-empty and disjoint. Define

$$\tilde{M} := \frac{1}{2}\left[\underset{\tilde{t} \in T^-}{max}\{M(\tilde{t})\} + \underset{\tilde{t} \in T^+}{min}\{M(\tilde{t})\}\right].$$

Note that $F(t)M(t) \geq F(t)\tilde{M}$ for all $t \in T$. Also, note that as $M(t)$ is non-constant, the previous inequality must be strict for at least some $t \in T$. Hence,

$$\sum_{t \in T} F(t)M(t) > \sum_{t \in T} F(t)\tilde{M} = 0.$$

The inequality comes from the construction of $\tilde{M}$ and the fact that $M(\cdot)$ is not constant, whereas the equality comes from $\sum_{t \in T} F(t) = 0$. □

A.2 Main Results

Proof of Lemma 1. Fix a contract (E, π_W, π_M). After y and z are realized, the manager chooses a performance evaluation report. In any equilibrium, the manager must only assign strictly positive probability to reports that maximize her payment. That is, if $\sigma(e|a_M, y, z) > 0$, then $\pi_M(y,e) = \underset{\hat{e}\in E}{max}\{\pi_M(y,\hat{e})\}$. In fact, any report policy that satisfies this property is sequentially rational.

I construct an outcome-equivalent alternative contract $(\tilde{E}, \tilde{\pi}_W, \tilde{\pi}_M)$ in which $\tilde{\pi}_M$ is independent of the performance evaluation.

Let $\tilde{E}_y := \underset{\hat{e}\in E}{argmax}\{\pi_M(y,\hat{e})\}$, $\tilde{E} := \cup_{y\in Y}\tilde{E}_y \subseteq E$. I now construct the payment functions $\tilde{\pi}_M : Y \times \tilde{E} \to \mathbb{R}_+$ and $\tilde{\pi}_W : Y \times \tilde{E} \to \mathbb{R}_+$ of the outcome-equivalent contract.

Let $\tilde{\pi}_M(y,e) := \underset{\hat{e}\in E}{max}\{\pi_M(y,\hat{e})\}$, which does not depend on e. Let, for each $y \in Y$, e_y be an arbitrary given element of $\tilde{E}_y$. Define,

$$\tilde{\pi}_W(y,e) := \begin{cases} \pi_W(y,e) \text{ if } e \in \tilde{E}_y \\ \pi_W(y,e_y) \text{ otherwise.} \end{cases}$$

Note that the set of payments to the worker (among the ones the manager is willing to choose from) has stayed the same. I now show that both contracts (E, π_W, π_M) and $(\tilde{E}, \tilde{\pi}_W, \tilde{\pi}_M)$ are outcome-equivalent.

First, I show that for any equilibrium in the original contract, there exists an equilibrium in the alternative contract with the same output and payments' joint distribution. Let (c_W, σ) be an equilibrium under the original contract. Note that if $\sigma(e|y,z) > 0$, then $e \in \tilde{E}$, $\pi_M(y,e) = \tilde{\pi}_M(y,e)$ and $\pi_W(y,e) = \tilde{\pi}_W(y,e)$. Let $\tilde{\sigma}(e|y,z) = \sigma(e|y,z)$ for all $e \in \tilde{E}$. Note that $\tilde{\sigma}$ generates the same payment distribution as σ. Hence, $(c_W, \tilde{\sigma})$ is an equilibrium of the alternative contract with the same output and payments' joint distribution as in (c_W, σ).

Finally, I show that for any equilibrium in the alternative contract, an outcome-equivalent one exists in the original contract. Let $(\tilde{c}_W, \tilde{\sigma})$ be an equilibrium under the alternative contract. Let

$$\sigma(e|y,z) = \begin{cases} 0 & \text{if } e \in E \setminus \tilde{E}_y \\ \tilde{\sigma}(e|y,z) & \text{if } e \in \tilde{E}_y \setminus \{e_y\} \\ \tilde{\sigma}(e_y|y,z) + \sum_{\hat{e}\in\tilde{E}\setminus E_y} \tilde{\sigma}(\hat{e}|y,z) & \text{if } e = e_y \end{cases}$$

For every realization of (y,z), the constructed σ generates the same distribution of payments to the worker as $\tilde{\sigma}$. Then, $(\tilde{c}_W, \sigma)$ is an equilibrium of the original contract with the same output and payments' joint distribution as under $(\tilde{c}_W, \tilde{\sigma})$. □

Proof of Lemma 2. The manager's payoff is given by:

$$V_M(c,\pi_M) = \sum_Y u_M\big(\pi_M(y)+b(y)\big)p(y|0) + G(c)\sum_Y u_M\big(\pi_M(y)+b(y)\big)[p(y|1)-p(y|0)],$$

which is strictly increasing in c if and only if $\sum_Y u_M\big(\pi_M(y)+b(y)\big)[p(y|1)-p(y|0)] > 0$.

By free-disposal, $\pi_M(\cdot)$ must be weakly increasing, while $b(\cdot)$ is strictly increasing. By (MLRP), $[p(y|1)-p(y|0)]$ is strictly increasing and $\sum_{y\in Y}[p(y|1)-p(y|0)] = 0$. By Lemma 4,

$$\sum_Y u_M\big(\pi_M(y)+b(y)\big)[p(y|1)-p(y|0)] > 0. \tag{7}$$

□

Proof of Proposition 1. Fix contracts (E,π_W,π_M). By Lemma 2, $V_M(c^*,\pi_M) \geq V_M(\hat{c},v_M)$ if and only if $c^* \geq \hat{c}$. I now find the evaluation strategy that generates the highest effort cost cutoff. That is, I maximize

$$\begin{aligned} &V_W(a=1,\pi_W,\sigma) - V_W(a=0,\pi_W,\sigma) \\ =& \int_Z \sum_Y \sum_E u_W\big(\pi_W(y,e)\big)\sigma(e|y,z)\Big[q(z|1)p(y|1) - q(z|0)p(y|0)\Big]dz, \end{aligned}$$

by choosing σ. Note that σ^* maximizes the expression above. Moreover, the distribution of payments is unique. The evaluation strategy σ maximizes $u_W\big(\pi_W(y,e)\big)\big[q(z|1)p(y|1)-q(z|0)p(y|0)\big]$ pointwise. Hence, any other distribution generates a different distribution of $u_W\big(\pi_W(y,e)\big)$ creates strictly weaker incentives for the worker's effort, and it is worse for the manager. □

Proof of Proposition 2. I prove Proposition 2 in three steps. First, I solve for the manager's compensation scheme. Second, I solve a relaxed problem in which I drop the constraints $\pi_M(y,b) \leq \pi_M(y,g)$ for all $y \in Y$, and, third, I check the dropped constraints are satisfied in the solution to the relaxed problem.

Note that problem (2) can be written as

$$\min_{\pi_M(\cdot),\pi_W(\cdot)} \sum_{y\in Y}\sum_{e\in\{b,g\}} \Big[\pi_M(y)+\pi_W(y,e)\Big]f(y,e) \tag{8}$$

subject to

$$\sum_Y u_M(\pi_M(y)+b(y))\big(f(y,b)+f(y,g)\big) \geq \bar{u}_M \tag{IR_M}$$

$$\sum_Y \sum_{e\in\{b,g\}} u_W(\pi_W(y,e))f(y,e) - \int_0^{c_W} c\,dG(c) \geq \bar{u}_W, \tag{IR_W}$$

$$\sum_Y \sum_{e\in\{b,g\}} u_W(\pi_W(y,e))s(y,e)f(y,e) = c_W, \tag{IC_M}$$

and $\pi_W(y,g) \geq \pi_W(y,b)$ for all $y \in Y$. By standard arguments, the solution to this relaxed problem is

Step 1: optimal manager's compensation scheme.

Take an arbitrary increasing manager's payment $\pi_M : Y \to \mathbb{R}_+$ that satisfies (IR_M). Let the expected payment be denoted by $\tilde{\pi}_M := \sum_Y \pi_M(y)\big(f(y,b)+f(y,g)\big)$. Note that if the principal pays the manager $\tilde{\pi}_M$ for any output realization instead of π_M, the expected payment does not change. Moreover, as π_M is increasing, the random variable $[b(y)+\tilde{\pi}_M]$ is a mean-preserving contraction of $[b(y)+\pi_M(y)]$. As u_M is concave, this implies that

$$\sum_Y u_M\big(b(y)+\tilde{\pi}_M\big)\big(f(y,b)+f(y,g)\big) \geq \sum_Y u_M\big(b(y)+\pi_M(y)\big)\big(f(y,b)+f(y,g)\big) \geq \bar{u}_M,$$

which implies that participation would still be satisfied under the flat contract. Moreover, if the inequality is strict (which occurs if π_M is non-constant and the manager risk averse), the principal could strictly decrease $\tilde{\pi}$ and strictly reduce her costs. Therefore, the optimal compensation to the manager is the minimum constant payment that guarantees the manager's participation. That is, $\pi_M^*(y) = \bar{\pi}_M$ for all $y \in Y$, where

$$\bar{\pi}_M = inf\big\{x \in \mathbb{R}_+ : \mathbb{E}\big[u_M\big(x+b(y)\big)\big|c_W\big] \geq \bar{u}_M\big\}.$$

Step 2: relaxed problem and the optimal worker's compensation scheme.

The second step consists of dropping the constraints $\pi_M(y,b) \leq \pi_M(y,g)$ for all $y \in Y$ and solving the best worker's compensation scheme. Note that for a given cost cutoff c_W and a given performance evaluation σ^*, characterizing the worker's optimal compensation scheme becomes a standard moral hazard problem. That is,

$$\min_{\pi_W(\cdot)} \sum_Y \sum_{e\in\{b,g\}} \pi_W(y,e)f(y,e) \tag{9}$$

subject to

$$\sum_Y u_W(\pi_W(y,e))f(y,e) - \int_0^{c_W} cdG(c) \geq \bar{u}_W, \tag{IR_W}$$

$$\sum_Y \sum_{e\in\{b,g\}} u_W(\pi_W(y,e))s(y,e)f(y,e) = c_W. \tag{IC_W}$$

The existence of a solution to this relaxed problem is shown in Appendix B, Proposition 8. When proving Proposition 8, I show that complementary slackness holds and, hence, the problem has an equivalent Lagrangian formulation. One can write problem (9)'s Lagrangian as

$$\mathscr{L} = \sum_{y\in Y}\sum_{e\in\{b,g\}} \left[\pi_W(y,e) - \lambda\left[u_W\left(\pi_W(y,e)\right) - \int_0^{c_W} cdG(c) - \bar{u}_W\right] - \mu^W\left[u_W\left(\pi_W(y,e)\right)s(y,e) - c_W\right]\right] f(y,e). \tag{10}$$

By minimizing pointwise, I get

$$\frac{1}{u_W'\left(\pi_W^*(y,e)\right)} = \lambda + \mu \cdot s(y,e), \tag{11}$$

where λ and μ are the respective dual multipliers associated with (IR_W) and (IC_W).

Step 3: showing that $\lambda, \mu > 0$ and that $\pi_W^*(y,g) \geq \pi_W^*(y,b)$ for all $y \in Y$.

Suppose $\mu = 0$. Let $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ be such that $\rho\left(1/u_W'(x)\right) = x$. Then, $\pi_W^*(y) = \rho(\lambda)$. Which implies

$$\sum_Y \sum_{e \in \{b,g\}} u_W\left(\pi_W(y,e)\right) s(y,e) f(y,e) = \rho(\lambda) \sum_Y \sum_{e \in \{b,g\}} s(y,e) f(y,e) = 0 < c_W.$$

A violation of (IC_W). Hence, $\mu > 0$.

Suppose one perturbs the optimal contracts by subtracting a small $\varepsilon > 0$ to every $\pi_W(y,e)$. The multipliers must be such that this change increases the value of the Lagrangian. That is, taking the first-order condition with respect to ε and evaluating at $\varepsilon = 0$

$$-1 + \lambda \sum_Y \sum_{e \in \{b,g\}} u_W'(\pi_W(y,e)) f(y,e) \geq 0.$$

As $u_W'(x) > 0$ for all $x \in \mathbb{R}_+$, $\lambda > 0$.

Finally, as $\rho(\cdot)$ is strictly increasing, $\pi_W^*(y,g) \geq \pi_W^*(y,b)$ if and only if $s(y,g) \geq s(y,b)$. However, Assumption 1 implies that $s(y,b) \leq 0 \leq s(y,g)$. □

Proof of Proposition 3. When we fix the manager's performance evaluation strategy, the design of incentives to the worker becomes a standard moral hazard problem. By the informativeness principle (Holmström (1979) and Shavell (1979)), any signal is valuable if and only if it brings non-redundant information about the effort. In my model's context, subjective performance evaluations are valuable if and only if they bring additional information about a beyond what is conveyed by $\mathbf{y}$.

Suppose it is a HII setting. Note that (5) implies that $z^*(y) \in \{\underline{z}, \bar{z}\}$ for all $y \in Y$. Hence, for each given realized y, the performance evaluation is always b or always g, regardless of $\mathbf{z}$'s realization. Therefore, $\mathbf{y}$ is a sufficient statistic for $(\mathbf{y}, \mathbf{e})$. That is, the subjective performance evaluation does not bring any additional information. Hence, it is not valuable.

Now, suppose it is an SII setting. Then, there exists $y \in Y$ such that (5) does not hold. That is, there exists $y \in Y$ such that $z^*(y) \in (\underline{z}, \bar{z})$. The performance evaluation informs whether $z \geq z^*(y)$ or not. Hence, $\mathbf{y}$ is not a sufficient statistic for $(\mathbf{y}, \mathbf{e})$, and subjective performance evaluations are valuable. □

A.3 Organizational Transparency

Proof of Remark 3. Note that the manager's contract, as constructed in the previous section, does not depend on the distribution of $\mathbf{z}$. Hence, it is characterized by (3).

Now regarding the worker's contract: one way to represent a fully informative signal is to let $\mathbf{z} \sim U[\underline{z}, \frac{\bar{z}-\underline{z}}{2}]$ if $a = 0$ and $\mathbf{z} \sim U[\frac{\bar{z}-\underline{z}}{2}, \bar{z}]$ if $a = 1$. Any realization below $\frac{\bar{z}-\underline{z}}{2}$ fully reveals low effort, and above fully reveals high effort. Note that under full information σ^* is given by reporting g with probability 1 when $a = 1$ and b

with probability 1 otherwise. That is, under full information, the manager's report is fully informative about the worker's effort and achieves the best the principal can attain. It is then easy to see that the cheapest contract that implements c_W and ensures participation is given by

$$u_W\left(\pi_W^*(y,b)\right) = \bar{u}_W - G(c_W)\Big[c_W - \mathbb{E}[\mathbf{c}|c \le c_W]\Big] \quad \text{and} \quad u_W\left(\pi_W^*(y,g)\right) = u_W\left(\pi_W^*(y,b)\right) + c_W.$$

□

Proof of Proposition 4. Conditional on full transparency, the optimal compensation scheme is as described in previous sections. Where a good evaluation is provided if $z > z^*(y)$, and a bad one otherwise. Note that the manager's compensation does not depend on the information he observes about z, hence when analyzing the effect of transparency, one can focus solely on the worker's compensation.

Note that $z^*(y)$ characterizes the distribution of scores observed by the principal. If instead of full transparency (the manager being endowed with the finest partition), the manager observes a signal from the partition $\Big\{[\underline{z}, z^*(y_n)]; \{z\}_{z\in(z^*(y_n),z^*(y_0))}; [z^*(y_0), \bar{z}]\Big\}$, the performance evaluations would be exactly the same. As $z^*(y)$ is characterized by $p(y|1)q(z^*(y)|1) = p(y|0)q(z^*(y)|0)$, then

$$\begin{aligned} p(y|1)Q(z^*(y)|1) &< p(y|0)Q(z^*(y)|0), \\ p(y|0)[1-Q(z^*(y)|0)] &< p(y|1)[1-Q(z^*(y)|1)] \quad \text{for all } y \in Y. \end{aligned}$$

As the inequalities are strict, the principal can slightly increase the upper bound of the lowest set of the partition, decrease the lower bound of the highest set of the partition, and keep the same evaluations. That is, take $\kappa > 0$ and $\varepsilon > 0$. Let the manager's information structure be given by

$$\Big\{[\underline{z}, z^*(y_n)+\varepsilon]; \{z\}_{z\in(z^*(y_n),z^*(y_0))}; [z^*(y_0)-\kappa\varepsilon, \bar{z}]\Big\}.$$

For ε sufficiently small, the manager provides a bad evaluation whenever she gets the signal associated with the lowest set of the partition and a good evaluation if she sees the signal associated with the highest set.

For a given κ and ε, we have an information partition. The principal can choose the optimal contracts that enforce effort c_W for that given information structure. Denote the minimal cost as $C(\varepsilon,\kappa)$. If $\varepsilon = 0$, we are at the full transparency case. I show that there exists a $\kappa > 0$ such that $\frac{dC}{d\varepsilon}(0,\kappa) < 0$. By increasing ε slightly from zero, the principal strictly reduces the cost.

Let π_W^* denote the optimal contracts when $\varepsilon = 0$. When we increase ε, we increase the probability of good evaluation when output is y_0 and decrease the probability at y_n. I construct κ such that the marginal effect of ε on the worker's expected utility is zero at zero. That is, the additional probability of a good evaluation when output is low exactly offsets the utility loss from the lower probability of a good evaluation when output is high.

$$\kappa := \frac{\left[u_W\left(\pi_W^*(y_n,g)\right) - u_W\left(\pi_W^*(y_n,b)\right)\right]}{\left[u_W\left(\pi_W^*(y_0,g)\right) - u_W\left(\pi_W^*(y_0,b)\right)\right]} \cdot \frac{p(y_n|1)q(z^*(y_n)|1)G(c_W) + (1-G(c_W))p(y_n|0)q(z^*(y_n)|0)}{p(y|1)q(z^*(y_0)|1)G(c_W) + (1-G(c_W))p(y_0|0)q(z^*(y_0)|0)}.$$

$C(\varepsilon,\kappa)$ is given by

$$C(\varepsilon,\kappa)=\min_{\pi_W}\left\{\sum_{Y}\sum_{e\in\{g,b\}}\left[\pi_W(y,e)-\lambda u_W\left(\pi_W(y,e)\right)-\mu s(y,e)+\mu c_W+\lambda\left(\bar{u}_W+\mathbb{E}[\mathbf{c}|c\leq c_W]\right)\right]f(y,e)\right\}.$$

Note that s, f, λ, and μ all depend on the information structure. Hence, on ε. However, when differentiating at $\varepsilon=0$, we can apply the Milgrom and Segal (2002)'s envelope theorem.[32] Therefore, we can keep π_W fixed and we do not need to worry about how λ and μ change with ε (f and s we need though). That is,

$$\frac{dC}{d\varepsilon}(0,\kappa)=\frac{\partial C}{\partial\varepsilon}(0,\kappa)=\sum_{y\in\{y_0,y_n\}}\sum_{e\in\{g,b\}}\left\{\left[\pi_W^*(y,e)-\lambda u_W\left(\pi_W^*(y,e)\right)\right]\frac{df(y,e)}{d\varepsilon}-\mu\frac{df(y,e)s(y,e)}{d\varepsilon}\right\}.$$

Note that as $q(\cdot|a)$ is continuous (Assumption 3), $z^*(\cdot)$ is such that $p(y|1)q(z^*(y)|1)=p(y|0)q(z^*(y)|0)$. Hence,

$$\frac{df(y,e)s^W(y,e)}{d\varepsilon}\Big|_{\varepsilon=0}=p(y|1)q(z^*(y)|1)-p(y|0)q(z^*(y)|0)=0.$$

As the manager chooses the evaluation strategy to maximize the worker's effort, any small perturbation in the information structure has, at most, a second-order effect on incentives.

The second observation is that κ was constructed such that the marginal effect of ε on the worker's expected utility was zero. That is,

$$\lambda\sum_{y\in\{y_0,y_n\}}\sum_{e\in\{g,b\}}u_W\left(\pi_W^*(y,e)\right)\frac{df(y,e)}{d\varepsilon}\Big|_{\varepsilon=0}=0.$$

Therefore only the direct effect remains. That is,

$$\begin{aligned}\frac{dC}{d\varepsilon}(0,\kappa)=&\left[p(y_n|1)q(z^*(y_n)|1)G(c_W)+(1-G(c_W))p(y_n|0)q(z^*(y_n)|0)\right]\\&\times\frac{\left[\pi_W^*(y_n,g)-\pi_W^*(y_n,b)\right]\left[\pi_W^*(y_0,g)-\pi_W^*(y_0,b)\right]}{\left[u_W\left(\pi_W^*(y_0,g)\right)-u_W\left(\pi_W^*(y_0,b)\right)\right]}\\&\times\left[\frac{u_W\left(\pi_W^*(y_n,g)\right)-u_W\left(\pi_W^*(y_n,b)\right)}{\pi_W^*(y_n,g)-\pi_W^*(y_n,b)}-\frac{u_W\left(\pi_W^*(y_0,g)\right)-u_W\left(\pi_W^*(y_0,b)\right)}{\pi_W^*(y_0,g)-\pi_W^*(y_0,b)}\right]<0.\end{aligned}$$

Where the first two terms are strictly positive, and the last is strictly negative from the strict concavity of u_W and the fact that $\pi_W^*(y_n,e)>\pi_W^*(y_0,e)$ for $e\in\{b,g\}$, which is a direct implication from $s(y_n,e)>s(y_0,e)$. □

A.4 Multiple Workers

Assumption 4 states as *effort complementarity* as a submodularity condition on the distribution P. In this section, I first present an ordinal definition of effort complementarity and then show it is equivalent to the

[32]In Apppendix B I show that one can bound $\pi_W(y,e)$ from above, which ensure the conditions for Milgrom and Segal (2002)'s Theorem 1.

condition assumed on Assumption 4.

Definition 7. *I say that p satisfies* ***ordinal effort complementarity*** *if for any increasing* $h : Y \to \mathbb{R}$, $i \in K$ *and* $\hat{a} \geq a$

$$\sum_Y h(y)\left[p\left(y\middle|1+\sum_{j\neq i}\hat{a}_j\right)-p\left(y\middle|\sum_{j\neq i}\hat{a}_j\right)\right] \geq \sum_Y h(y)\left[p\left(y\middle|1+\sum_{j\neq i}a_j\right)-p\left(y\middle|\sum_{j\neq i}a_j\right)\right], \tag{12}$$

and the inequality is strict for any strictly increasing b.

Efforts are ordinal complements if, for any increasing reward function, one worker's incentive to exert effort increases when we increase other workers' efforts. Note that as the inequality must hold for any increasing function, this definition can be interpreted as an ordinal notion of complementarity. Even if one does not know the intensity of how much better higher values of y are compared to lower levels, one agent's gain from high effort is larger when other agents exert high effort.

Proposition 7. *The distribution p satisfies ordinal effort complementarity if and only if* $P\left(y\middle|\sum_{i\in K}a_i\right)$ *is submodular in a for all* $y \in Y$.

Proof. Note that for any h, (12) is equivalent to

$$\sum_Y h(y)\frac{\left[p\left(y\middle|1+\sum_{j\neq i}\hat{a}_j\right)+p\left(y\middle|\sum_{j\neq i}a_j\right)\right]}{2} \geq \sum_Y h(y)\frac{\left[p\left(y\middle|\sum_{j\neq i}\hat{a}_j\right)+p\left(y\middle|\sum_{j\neq i}a_j\right)\right]}{2}.$$

Hence, the inequality holds for any increasing b if and only if

$$\left[\frac{p\left(\cdot\middle|1+\sum_{j\neq i}\hat{a}_j\right)}{2}+\frac{p\left(\cdot\middle|\sum_{j\neq i}a_j\right)}{2}\right] \succeq_{FOSD} \left[\frac{p\left(\cdot\middle|\sum_{j\neq i}\hat{a}_j\right)}{2}+\frac{\left(\cdot\middle|\sum_{j\neq i}a_j\right)}{2}\right],$$

which is equivalent to

$$P\left(y\middle|1+\sum_{j\neq i}\hat{a}_j\right)+P\left(y\middle|\sum_{j\neq i}a_j\right) \leq P\left(y\middle|\sum_{j\neq i}\hat{a}_j\right)+P\left(\cdot\middle|\sum_{j\neq i}a_j\right) \quad \forall y \in Y. \tag{13}$$

Finally, (13) holding for any $\hat{a} \geq a$ and any $i \in K$ is equivalent to P being submodular in a.[33]

The second part is about the strictness for any strictly increasing h. The same argument applies, but it requires that for some $y \in Y$, the inequality (13) is strict to ensure strict first-order stochastic dominance. □

To the best of my knowledge, this is a novel definition and characterization of effort complementarity in environments with stochastic and non-binary outcomes.[34] I then present the proofs for the other results in

[33] I thank Humberto Moreira for a great suggestion that simplified this proof.

[34] The assumptions that efforts enter p additively or that Y is discrete are not necessary for this result. More generally, one could consider $p : Y \times \{0,1\}^N \to [0,1]$ and equivalently define efforts as ordinal complements when for any increasing $b : Y \to \mathbb{R}$

$$\sum_Y b(y)[p(y|a\wedge\tilde{a})+p(y|a\vee\tilde{a})] \geq \sum_Y b(y)[p(y|a)+p(y|\tilde{a})].$$

the Multiple Workers section.

Proof of Proposition 5. I show that as q_n approximates q^{FI}, the principal can approximate her payoff when she directly observes effort by using an IP-mechanism. In contrast, when using a FR-mechanism, she has a strictly higher implementation cost.

If the principal could directly observe the workers' efforts, the best she could do is pay a fixed amount that assures participation and offer a bonus in case of high effort, which would assure workers with effort cost below c prefer to exert effort. In contrast, workers with costs above the threshold would not exert. That is, the cost-minimizing contract would map the effort of each worker to their own payment and would satisfy $u_W\big(\pi^{FI}(e_i=0)\big)=\bar{u}-G(c)\big(c-\mathbb{E}[\tilde{c}|\tilde{c}\leq c]\big)$ and $u_W\big(\pi^{FI}(e_i=1)\big)=u_W\big(\pi^{FI}(e_i=0)\big)+c$.[35]

Now that I have established the full-information benchmark, I consider the sequence of problems associated with the sequence of distributions q_n converging to q^{FI}.[36] It is useful first to establish the upper-hemicontinuity of the correspondence of the manager's optimal evaluation strategies with respect to the distribution of non-verifiable information and workers' contracts. I first proceed to establish such a property.

In the setting with multiple workers, a mechanism consists of a set of performance evaluations and payments to the manager $\pi_M : Y\times E\to\mathbb{R}_+$ and each worker $\pi_i : Y\times E\to\mathbb{R}_+$ for $i\in K$. I denote by $\mathscr{M}:=(\pi_M,\pi_1,\ldots,\pi_k)$ a given mechanism. Given a mechanism in place, the manager chooses an evaluation strategy $\sigma : Y\times Z^k\to\Delta(E)$. For each mechanism and each evaluation strategy, the manager's payoff is given by

$$V_M(\sigma,\mathscr{M}):=\max_{(c_1,\ldots,c_k)\in\Gamma(\sigma,\pi_1,\ldots,\pi_k)}\mathbb{E}\big[U_M(\pi_M,\mathbf{y},\mathbf{e})|\sigma,c_1,\ldots,c_k\big],$$

where $(c_1,\ldots,c_k)$ denotes the effort cost cutoffs chosen by each worker, $\Gamma(\sigma,\pi_1,\ldots,\pi_k)$ denotes the set of continuation equilibria when the manager has announced evaluation strategy σ, and the maximum operator implies that, as in the single worker case, the manager's preferred continuation equilibrium is selected.

Denote by $\Theta\big(Q,\mathscr{M}\big)$ the set or manager's optimal evaluation strategies given that the distribution of the non-verifiable information is Q and the mechanism in place is $\mathscr{M}$. That is,

$$\Theta\big(Q,\mathscr{M}\big):=\underset{\sigma:Y\times Z^2\to\Delta(E)}{argmax}\;V_M(\sigma,\mathscr{M}).$$

Lemma 5. *Fix a finite set of performance ratings E. Then, Θ is upper hemicontinuous.*

Proof. Note that for any $(\sigma,\mathscr{M})$, the set $\Gamma(\sigma,\pi_1,\ldots,\pi_k)$ is compact valued and upper hemicontinuous in the weak topology. Moreover, for a fixed set of performance ratings E, the function $\mathbb{E}\big[U_M(\pi_M,\mathbf{y},\mathbf{e})|\sigma,c_1,\ldots,c_k\big]$

[35] As before, the manager's payment must be a fixed amount to guarantee participation. I keep π_M fixed at such a level for the rest of this proof.

[36] Note that even in the limit, there is a distinction between the full-information benchmark (when the principal directly observes efforts) and the case with a fully informative signal (the manager perfectly observes effort).

is continuous with respect to $(\sigma, \pi_1, \dots, \pi_k)$ in the weak topology. Hence, V_M is upper semicontinuous in $(\sigma, \pi_1, \dots, \pi_k)$. As V_M is upper semicontinuous and the set of possible performance evaluations is compact valued, then Θ is upper hemicontinuous. □

Suppose that for each $n \in \mathbb{N}$, the principal uses the following individual performance evaluation mechanism: the manager reports a good (g) or a bad (b) performance for each worker. The payments to each worker are independent of output and are given by

$$u_W\big(\pi_{w,n}(b)\big) = \bar{u} \quad \text{and} \quad u_W\big(\pi_{w,n}(g)\big) = \bar{u} + c + \varepsilon_n,$$

where

$$\varepsilon_n := c \cdot \frac{1 - \big[Q_n(z_0|1) - Q_n(z_0|0)\big]}{\big[Q_n(z_0|1) - Q_n(z_0|0)\big]}.$$

Note that as $n \to \infty$, ε_n converges to zero and $\pi_{w,n}$ converges to paying $\pi^{FI}(0)$ when $z_i < z_0$ and $\pi^{FI}(1)$ otherwise. Hence, if the evaluation strategy used by the manager converges to truthtelling (reporting g if $z_i > z_0$ and b otherwise), then the principal's payoff converges to the full-information case.

Note that under a Q^{FI} and workers' payments π^{FI}, the manager's unique best response is to report a good performance for each worker who has generated a signal $z_i > z_0$ and a bad performance otherwise.[37]

As $\Theta\big(Q^{FI}, (\pi_M, \pi^{FI}, \dots, \pi^{FI})\big)$ has a unique maximizer, then all sequences $\tilde{\sigma}_n \in \Theta\big(Q_n, (\pi_M, \pi_{w,n}, \dots, \pi_{w,n})\big)$ converge to truthtelling. Therefore, as n grows, the individual performance evaluation mechanism with workers' contracts given by $\pi_{w,n}$ implement effort cutoffs that approximate $(c, \dots, c)$ and at a cost that approximates the full-information benchmark.

The last step is observing that under a forced ranking mechanism, the principal's payoff is bounded away from the full-information case even when $Q = Q^{FI}$. When the manager is fully informed, her private information has cardinality four: each worker has worked or shirked. However, the forced ranking message space E has cardinality two: the manager only reports who is ahead. Therefore, it is impossible to provide full information to the principal. As information is strictly valuable, the principal is strictly worse under a forced ranking mechanism than under full information. As a sequence of individual performance evaluation mechanisms can approximate full information as Q_n converge to Q^{FI}, then there exists N such that for all $n > N$, the principal strictly prefers individual performance evaluations over forced ranking mechanisms.

□

Proof of Proposition 6. The proof is divided into two parts: first, I show that when (6) holds, no information is transmitted to the principal in an individual performance evaluation mechanism. Hence, when restricted to IP-mechanisms, the principal cannot do better than conditioning payments only on output. Second, I show that when using a forced ranking, the principal can partially elicit the non-verifiable information and,

[37] As before, the manager benefits from higher workers' efforts. The strategy that maximizes effort incentives is to report a good performance when the effort is high and a low performance when the effort is low. The argument is the same as in Remark 3.

hence, does strictly better than with IP-mechanisms.

Lemma 6. *Suppose (6) holds and the principal offers an individual performance evaluation mechanism* $(\{b,g\}^k, \pi_1, \dots, \pi_k)$. *Then, the manager's evaluation report does not depend on the non-verifiable information.*

Proof. As in the single-worker case, the manager strictly benefits from increasing the effort of each worker. Denote by c_{-i}, z_{-i} and π_{-i} the vector of cost cutoffs, non-verifiable signal realizations, and contracts for all workers other than worker i. Moreover, denote by $h(z_{-i}, m|c_{-i})$ the probability density of z_{-i} being realized and exactly m workers other than i exerting effort given the respective cost cutoffs c_{-i}. For instance, if there are only two workers $h(z_{-i}, 0|c_{-i}) = (1-G(c_{-i}))q(z_i|0)$ and $h(z_{-i}, 1|c_{-i}) = G(c_{-i})q(z_i|1)$.

Note that if workers other than i have a cutoff $c_{-i} \in \mathbb{R}_+^{k-1}$, then worker's i cost cutoff is

$$c_i(\sigma) = \sum_Y \int_{Z^{k-1}} \int_Z \Bigg\{ \Big[u_W\big(\pi_i(y,g)\big)\sigma_i(g|y,z_i,z_{-i}) + u_W\big(\pi_i(y,b)\big)\big(1-\sigma_i(g|y,z_i,z_{-i})\big)\Big] \cdot \left[\sum_{m=0}^{k-1} h(z_{-i},m|c_{-i})\big[p(y|m+1)q(z_i|1) - p(y|m)q(z_i|0)\big] \right] \Bigg\} dz_i dz_{-i}.$$

Note that when (6) holds, then for each $y \in Y$, the expression inside the large square brackets has all terms being strictly positive or strictly negative, regardless of c_i. That is, irrespective of the cost cutoff of the other workers and the vector of non-verifiable signals, y determines whether the score is positive or negative. Hence, only y affects the manager's evaluation. As the manager uses an evaluation strategy that does not condition on $(\mathbf{z_i}, \mathbf{z_{-i}})$'s realizations, no information beyond $\mathbf{y}$ is conveyed by the manager's report. □

Lemma 6 and the information principle imply that the principal cannot improve over the optimal objective contract when constrained to IP-mechanisms. Now, I need to show that the principal can do strictly better by using a forced ranking mechanism than by only conditioning on output (and, hence, strictly better than using an individual performance evaluation mechanism). To prove this claim, I first show that the principal can improve over the optimal objective contract by requesting a forced ranking when there are two workers ($k = 2$). Then, I extend the result beyond two workers by fixing the contracts of $(k-2)$ workers and conditioning the payment of the remaining two on how they are ranked relative to each other.

Lemma 7. *Suppose that there are two workers, Assumption 4 holds, and the principal wants to implement* $(c,c) \in \mathbb{R}^2_{++}$. *Then, the principal is strictly better off by using a forced ranking over any objective contract.*

Proof. Denote by $\tilde{\pi} : Y \to \mathbb{R}_+$ the symmetric cost-minimizing workers' contract that implements $(c,c) \in \mathbb{R}^2_{++}$ when the principal only observes output and does not request any performance evaluation. Suppose now that the principal could observe an additional binary signal disclosing which worker had the highest non-verifiable signal realization. That is, suppose the principal could observe whether $z_i > z_j$. By the informativeness principle, we know that the principal could strictly reduce the implementation cost by increasing the payment of the worker with the highest z and decreasing the payment of the worker with the lowest z by a small amount.[38] Therefore, if the manager is willing to truthfully rank workers' performances when the bonus from being ranked first is sufficiently small, and payments are sufficiently close to $\tilde{\pi}$, then there exists a forced ranking mechanism that implements (c,c) at a strictly lower cost than any individual performance evaluation system.

Consider a perturbation of $\tilde{\pi}$ in which the worker ranked first (a good report) gets $\hat{\pi}(y,g) = \tilde{\pi}(y) + \varepsilon/2$ and the worker ranked last (a bad report) $\hat{\pi}(y,b) = \tilde{\pi}(y) - \varepsilon/2$. In such a perturbed mechanism, the manager is restricted to giving one good and one bad evaluation. That is, it is a forced ranking mechanism that is ε close to $\tilde{\pi}$. In particular, by the argument above, it strictly outperforms $\tilde{\pi}$ if the manager ranks workers truthfully. I now show that the manager reports truthfully under such a mechanism.

The first step is to show that the manager prefers to enforce symmetric effort cost cutoffs. Consider a pair of effort cutoffs (c_1,c_2) where, without loss of generality, $c_1 > c_2$. I show that the manager is better off if the workers have the symmetric cutoff pair $\big((c_1+c_2)/2,(c_1+c_2)/2\big)$ instead. The manager's payoff as a function of effort cost cutoffs is

$$\begin{aligned}
V_M(c_1,c_2) =& \sum_Y u_M\big(\pi_M + b(y)\big)\Big[G(c_1)G(c_2)p(y|2) + \big(G(c_1)+G(c_2)\big)p(y|1) + \big(1-G(c_1)\big)\big(1-G(c_2)\big)p(y|0)\Big] \\
=& \sum_Y u_M\big(\pi_M + b(y)\big)\Big[G(c_1)G(c_2)[p(y|2)+p(y|0)-2p(y|1)] + \big(G(c_1)+G(c_2)\big)[p(y|1)-p(y|0)]\Big] \\
\leq& \sum_Y u_M\big(\pi_M + b(y)\big)\Big[\Big(G\Big(\frac{c_1+c_2}{2}\Big)\Big)^2[p(y|2)+p(y|0)-2p(y|1)] + 2G\big(\frac{c_1+c_2}{2}\big)[p(y|1)-p(y|0)]\Big] \\
=& V_M\Big(\frac{c_1+c_2}{2},\frac{c_1+c_2}{2}\Big).
\end{aligned}$$

By MLRP (item 1 of Assumption 4) and the fact that b is increasing, the term multiplying $(G(c_1)+G(c_2))$ is positive.[39] Also, by effort complementarity and the fact that b is increasing, one can apply the Beesack's inequality and assure that the term multiplying $G(c_1)G(c_2)$ is also positive. Then, one can establish the inequality above by the fact that G is concave.

I then show that starting from an evaluation strategy that generates a pair of asymmetric effort cost cutoffs, one can construct a symmetric evaluation strategy that enforces a symmetric cutoff that is strictly above the average of the initial cutoffs. Hence, the manager would be better off using the constructed symmetric strategy.

Suppose the manager uses an evaluation strategy σ that generates a pair of cost cutoffs $c_1 \neq c_2$. Each c_i is determined by the difference between expected utility from payments when i exerts effort versus when he

[38]The construction of such perturbation is analogous to the one described by Holmström (1979) in the proof of the informativeness principle (Proposition 3).

[39]In the proof of Lemma 7, MLRP is only used here to assure that such a term is positive. Note, however, that first-order stochastic dominance would be enough. Hence, the requirement for this Lemma can be weakened to imposing that $P(\cdot|1+x) \succeq_{FOSD} P(\cdot|x)$ for all $x \in (0,1)$.

does not for a given cutoff of the other worker. That is,

$$c_i(\sigma,c_j)=\sum_Y\int_Z\int_Z\Big[u_W\big(\hat{\pi}(y,b)\big)+\varepsilon\sigma_i(g|y,z_i,z_j)\big)\Big]\Big\{q(z_j|0)\big[p(y|1)q(z_i|0)-p(y|0)q(z_i|0)\big]$$

$$+G(c_j)\Bigg[q(z_j|1)\big[p(y|2)q(z_i|1)-p(y|1)q(z_i|0)\big]-q(z_j|0)\big[p(y|1)q(z_i|0)-p(y|0)q(z_i|0)\big]\Bigg]\Bigg\}dz_idz_j.$$

Note that q is symmetric across workers. Hence, one can add $c_1(\sigma,c_2)+c_2(\sigma,c_1)$ and use the symmetry by switching z_1 and z_2 in the expression of $c_2(\sigma,c_1)$. Then, one can write

$$\begin{aligned}
&\frac{c_1+c_2}{2}=\\
&=\sum_Y\int_Z\int_Z\Big[u_W\big(\hat{\pi}(y,b)\big)+\varepsilon\Big(\frac{\sigma_1(g|y,z_1,z_2)+\sigma_2(g|y,z_2,z_1)}{2}\Big)\Big]q(z_2|0)\big[p(y|1)q(z_1|0)-p(y|0)q(z_1|0)\big]dz_1dz_2\\
&\quad+\frac{G(c_1)+G(c_2)}{2}\cdot\sum_Y\int_Z\int_Z\Big\{u_W\big(\hat{\pi}(y,b)\big)\Big[q(z_2|1)\big[p(y|2)q(z_1|1)-p(y|1)q(z_1|0)\big]\\
&\qquad\qquad-q(z_2|0)\big[p(y|1)q(z_1|0)-p(y|0)q(z_1|0)\big]\Big]dz_1dz_2\Big\}\\
&\quad+\varepsilon\cdot\sum_Y\int_Z\int_Z\Big\{\Big[\frac{\sigma_1(g|y,z_1,z_2)G(c_1)+\sigma_2(g|y,z_2,z_1)G(c_2)}{2}\Big]\\
&\qquad\cdot\Bigg[q(z_2|1)\big[p(y|2)q(z_1|1)-p(y|1)q(z_1|0)\big]-q(z_2|0)\big[p(y|1)q(z_1|0)-p(y|0)q(z_1|0)\big]\Bigg]\Bigg\}dz_1dz_2.
\end{aligned}$$

Note that the first term (first line after the equality) of the right-hand side does not depend on (c_1,c_2). Moreover, as $\hat{\pi}(\cdot,b)$ is strictly increasing and efforts are complements, the term multiplying $\big(G(c_1)+G(c_2)\big)/2$ is strictly positive. As G is strictly convex, by replacing (c_1,c_2) by their average, the second term of the right-hand side strictly increases. For a small enough ε, the effect on the last term does not matter. That is, the right-hand side strictly increases when one replaces c_1 and c_2 by their average.

Finally, replace the asymmetric evaluation strategy σ with a symmetric one $\hat{\sigma}$ where the manager randomizes with uniform probability which worker is indexed by 1 or 2 and replicates σ. That is,

$$\hat{\sigma}_1(g|y,z_1,z_2)=\hat{\sigma}_2(g|y,z_2,z_1)=\frac{\sigma_1(g|y,z_1,z_2)+\sigma_2(g|y,z_2,z_1)}{2}.$$

By replacing $\hat{\sigma}$ into the equation above when ε is sufficiently small, one gets

$$\frac{c_1+c_2}{2}<c_i\big(\hat{\sigma},c_j(\hat{\sigma})\big).$$

Hence, when the manager uses a evaluation strategy $\hat{\sigma}$, the continuation equilibrium effort cutoffs is larger than the average $(c_1+c_2)/2$. Then, the manager is strictly better off using the symmetric evaluation strategy $\hat{\sigma}$. Among symmetric evaluation strategies, the one that generates the strongest incentives for effort is reporting as the highest-ranked, the worker with the highest z. Therefore, the manager is willing to report truthfully, and the forced ranking mechanism constructed is strictly better than any individual performance evaluation mechanism. □

The final step is to extend the result beyond the two workers' case. Suppose that $k > 2$, the principal wishes to implement $(c,\dots,c) \in \mathbb{R}^k_{++}$, and that Assumption 4 holds. Fix the effort cutoffs of $k-2$ workers at c (without loss, I fix the cutoff of all workers with index $i > 2$). Suppose all workers with an index strictly above 2 exert effort if their realized cost is above c. The vector of efforts is then a random vector, in which each entry i is equal to one if the realized $c_i \leq 0$ and is zero otherwise. Define $\tilde{p}(y|a_1+a_2)$ as the probability mass of output realization y, conditional on (a_1,a_2). That is, $\tilde{p}(y|a_1+a_2) := \mathbb{E}_a\big[p(y|\mathbf{a})\big|a_1,a_2\big]$, where the expectation is taken with respect to $\mathbf{a}$.

Note that if I show that when p satisfies Assumption 4, then $\tilde{p}$ satisfies the conditions required for Lemma 7, then the principal can construct an FR-mechanism that is strictly better than any objective contract. In other words, one could replicate the construction in the proof of Lemma 7 and generate a strict improvement to the principal. The principal could request a forced ranking report, keep the payments to all workers with $i > 2$ invariant to the reported ranking, and condition the payments of workers 1 and 2 on who among them is ranked above. By Lemma 7, such a construction would be strictly better than any mechanism that only conditions in output. It remains to show that $\tilde{p}$ satisfies the requirements for Lemma 7.

The condition to verify in Lemma 7 is the validity Assumption 4 under $\tilde{p}$. Note that items 3 and 4 do not depend on output distribution. Therefore, one needs only to verify 1 and 2. As pointed out in the proof of Lemma 7, item 1 is stronger than necessary. That is, one needs only first-order stochastic dominance and not necessarily MLRP. Define $\tilde{P}(y|a_1+a_2) := \sum_{x\leq y}\tilde{p}(x|a_1+a_2) = \mathbb{E}_a\big[P(y|\mathbf{a})\big|a_1,a_2\big]$. Then, I must show that $\tilde{P}(\cdot|x+1) \succeq_{FOSD} \tilde{P}(\cdot|x)$ for all $x \in \{0,1\}$ and that $\tilde{P}(y|a_1+a_2)$ is submodular in (a_1,a_2) for all $y \in Y$.

Claim 1. $\tilde{P}(\cdot|x+1) \succeq_{FOSD} \tilde{P}(\cdot|x)$ *for all* $x \in \{0,1\}$.

Proof. As P satisfies MLRP, it satisfies FOSD. As the expectation operator preserves FOSD, then $\tilde{P}(\cdot|x+1) \succeq_{FOSD} \tilde{P}(\cdot|x)$ for all $x \in \{0,1\}$. □

Claim 2. $\tilde{P}(y|a_1+a_2)$ *is submodular in* (a_1,a_2) *for all* $y \in Y$.

Proof. As for every y the function $P(y|\sum_{i\in K}a_i)$ is submodular in a, it is submodular in (a_1,a_2) for each $(a_3,\dots,a_k)$. As the expectation operator preserves submodularity when supermodularity holds pointwise, then $\tilde{P}(y|a_1+a_2)$ is submodular in (a_1,a_2). □

Therefore, one can apply Lemma 7, and the principal strictly benefits from using a forced ranking. □

B Appendix — Existence and Slackness of the Minimum Payment Constraint

Existence of a solution to (2) is ensured by the existence in (8). As shown before, problem (2) can be relaxed and solved by approaching (9). The additional constraints were then verified in the main text. The remaining step is to ensure the existence of a solution to (9) and that the minimum payment constraint is slack.

Proposition 8. *For any $\bar{u}_W \in (u_W(0), +\infty)$ and $c_W \in \mathbb{R}_{++}$ there exists a solution to problem (9).*

Proof. First, further relax the problem (9) by imposing (IC_W) as an inequality instead of equality. That is, impose

$$\sum_{y\in Y}\sum e\in\{b,g\}u_W\big(\pi_W(y,e)\big)s(y,e)f(y,e)\geq c_W \tag{IC'_W}$$

instead of (IC_W). Note that if there exists a solution to this further relaxed problem in which (IC'_W) holds with equality, then such a solution also solves the problem (9).

We first show that there exists a solution to the further relaxed problem. Let $\tilde{\pi}_W$ be such that

$$u_W\big(\tilde{\pi}_W(y,e)\big)=\begin{cases}\bar{u}_W & \text{if } y\neq y_n\\ \bar{u}_W+1+\frac{c_W}{p(y_n|1)-p(y_n|0)} & \text{otherwise}\end{cases}$$

satisfy (IR_W) and (IC'_W) with slackness at a finite cost. Hence, one can bound payments from above. The choice set of the further relaxed problem is compact, and the objective functions are continuous. The existence of a minimum follows from Weierstass' theorem.

Finally, note that in any solution, (IC'_W) must hold with equality. Suppose for the sake of obtaining a contradiction that there is a solution to the relaxed problem $\tilde{\pi}_W$ such that

$$\sum_{y\in Y}\sum_{e\in\{b,g\}}u_W\big(\tilde{\pi}_W(y,e)\big)s(y,e)f(y,e)>c_W.$$

Let $\pi_W(y,e)=\gamma\tilde{\pi}_W(y,e)+\theta$, where $(\gamma,\theta)\in(0,1)\times\mathbb{R}$ are such that

$$\sum_{y\in Y}\sum_{e\in\{b,g\}}u_W\big(\gamma\pi_W(y,e)+\theta\big)s(y,e)f(y,e)=c_W,$$

and

$$\sum_{y\in Y}\sum_{e\in\{b,g\}}u_W\big(\gamma\pi_W(y,e)+\theta\big)f(y,e)=\sum_{y\in Y}\sum_{e\in\{b,g\}}u_W\big(\pi_W(y,e)\big)f(y,e).$$

By construction, π_W satisfies all constraints of the further relaxed problem. Moreover, note that the random variable $u_W\big(\pi_W(y,e)\big)$ is a mean-preserving contraction of $u_W\big(\tilde{\pi}_W(y,e)\big)$. Hence,

$$\sum_{y\in Y}\sum_{e\in\{b,g\}}\pi_W(y,e)f(y,e)<\sum_{y\in Y}\sum_{e\in\{b,g\}}\tilde{\pi}_W(y,e)f(y,e).$$

A contradiction. □

Proposition 9. *For any $c_W \in \mathbb{R}_{++}$, there exists $\hat{u}_W \in (u_W(0), +\infty)$ such that for any $\bar{u}_W > \hat{u}_W$ the solution to (9) is such that*

$$u_W\big(\pi^*_W(y,b)\big)>u_M(0)\quad \textit{for all } y\in Y.$$

Proof. I show that if the minimum payment remains binding when $\bar{u}_W$ increases to $+\infty$, then the incentive compatibility constraint becomes slack — a contradiction.

Take an increasing sequence $\bar{u}_{W_k} \in (u_W(0), +\infty)$ such that $\lim_{k\to+\infty} \bar{u}_{W_k} = +\infty$. Let π_{W_k} be the associated optimal contract.

For $\bar{u}_{W_k}$ high enough (IR_W) must bind. That is,

$$\sum_Y \sum_E u_W\left(\pi_{W_k}(y,e)\right) f(y,e) = \bar{u}_{W_k} + \mathbb{E}[\mathbf{c}|c \leq c_W]. \tag{14}$$

The optimal payment considering the minimum payment constraint is given by

$$\frac{1}{u'_W\left(\pi_{W_k}(y,e)\right)} = \lambda_k + \mu_k s(y,e) + \phi_k(y,e),$$

where $\phi_k(y,e)$ is the minimum payment multiplier, and it is strictly bigger than zero only if $u_W\left(\pi_{W_k}(y,e)\right) = u_W(0)$.

Let $\tilde{y} = \underset{y\in Y}{argmin}\{s(y,b)\}$. The realization $(\tilde{y},b)$ must be the lowest paying one. Hence, if the minimum payment constraint binds, it must be the case that $u_W\left(\pi_{W_k}(\tilde{y},b)\right) = u_W(0)$.

Suppose there exists a subsequence $\bar{u}_{W_{k_j}}$ and $J \in \mathbb{N}$ such that $u_W\left(\pi_{W_{k_j}}(\tilde{y},b)\right) = u_W(0)$ for all $k_j > J$.

The cheapest payments that would still satisfy (IR_W) and have $u_W\left(\pi_{W_{k_j}}(\tilde{y},b)\right) = u_W(0)$ would be a flat payment for all other output and performance report realizations, that is,

$$u_W\left(\tilde{\pi}_{W_{k_j}}(y,e)\right) = \begin{cases} u_W(0) \text{ if } (y,e) = (\tilde{y},b), \\ \frac{\bar{u}_{W_{k_j}} + \mathbb{E}[\mathbf{c}|c\leq c_W] - u_W(0) f(\tilde{y},b)}{1-f(\tilde{y},b)} \text{ otherwise.} \end{cases}$$

As $\tilde{\pi}_{W_{k_j}}$ is cheaper than any other contract that satisfies (IR_W) and has the lowest payment binding, it must be optimal if it satisfies (IC_W). I now show that for large enough $u_{W_{k_j}}$, $\tilde{\pi}_{W_{k_j}}$ not only satisfies (IC_W) but it does with slack. This a contradiction because (IC_W) must bind at the optimum.

Note that

$$\sum_Y \sum_E u_W\left(\tilde{\pi}_{W_{k_j}}(y,e)\right) s(y,e) f(y,e) = -s(\tilde{y},b) f(\tilde{y},b) \left[\frac{\bar{u}_{W_{k_j}} + \mathbb{E}[\mathbf{c}|c \leq c_W] - u_W(0) f(\tilde{y},b)}{1 - f(\tilde{y},b)} - u_M(0)\right].$$

As $f(\tilde{y},b) > 0$ and $s(\tilde{y},b) < 0$, for $\bar{u}_{W_{k_j}}$ large enough, the equation above is strictly higher than c_W. A contradiction. Hence, it does not exist a subsequence $\bar{u}_{W_{k_j}}$ and $J \in \mathbb{N}$ such that $u_W\left(\pi_{W_{k_j}}(\tilde{y},b)\right) = u_W(0)$ for all $k_j > J$. Therefore, the minimum payment constraint is slack for $\bar{u}_W$ large enough. □

C Continuous Efforts

The convenience of working with binary efforts is the simple characterization of incentive compatibility for each agent, which is given by a single inequality constraint. One could ask whether the binary effort assumption drives the binary performance ratings result. In this subsection, I show that it is not the case.

Let the worker effort choice be continuous, $a \in [0,1)$. As before, denote by $p(.|a)$ and $q(.|a)$ the distribution of each signal. I assume that p and q are twice continuously differentiable. Denote by $p_a(.|a)$, and $q_a(.|a)$ the partial derivatives with respect to a. I assume p and q satisfy the following monotone likelihood ratio properties.

Assumption 5. *p and q are such that*

1. $\frac{p_a(y|a)}{p(y|a)}$ *is strictly increasing in y for all* $a \in (0,1)$.
2. $\frac{q_a(z|a)}{q(z|a)}$ *is bounded and strictly increasing in z for all* $a \in (0,1)$.

As usual, in moral hazard problems with continuous efforts, I need convexity assumptions over the distribution of signals to ensure that the worker's payoff is concave in his effort choice. Such assumptions allow me to characterize the worker's effort choice by the first-order condition.

Assumption 6. *p and q are such that*

1. *p is linear in a.*
2. $Q(z|\cdot)p(y|\cdot)$ *is convex for any* $(y,z) \in Y \times Z$.

Let $c : [0,1) \to \mathbb{R}_+$ denote the worker's effort cost function. I assume that c is common knowledge, strictly increasing and strictly convex. To avoid concerns about corner choices, I assume that $\lim_{a \to 1} c'(a) = +\infty$.

Take any given contracts (E, π_M, π_W). For given effort levels and performance evaluation strategy, the manager's expected utility is given by

$$V_M(a, \pi_M) := \sum_Y u_M\big(\pi_M(y) + b(y)\big)\, p(y|a),$$

whereas the worker's expected utility is given by

$$V_W(a, \pi_W, \sigma) := \int_Z \sum_Y \sum_E u_W\big(\pi_W(y,e)\big)\,\sigma(e|y,z)\, p(y|a)q(z|a)dz - c(a).$$

Note that the manager's expected payoff is still increasing in the worker's effort. That is,

$$\frac{\partial V_M}{\partial a}(a, \pi_M) = \sum_Y u_M\big(\pi_W(y) + b(y)\big)\,\frac{p_a(y|a)}{p(y|a)}\, p(y|a) > 0.$$

Hence, the manager wants to create the strongest possible incentives for the worker's effort.[40] The performance evaluation strategy that maximizes the worker's effort uses only the highest and the lowest-paying reports.

Proposition 10. *Suppose Assumptions 5-6 hold. Then, the evaluation strategy that maximizes the worker's effort uses only the highest and the lowest-paying messages with strictly positive probability.*

Proof. At the end of Appendix C. □

The proof proceeds as follows: I start with an arbitrary $\hat{\sigma}$ which uses — with strictly positive probability — messages that are not the highest nor the lowest-paying ones. Then, I construct a $\tilde{\sigma}$ that only uses the highest and the lowest paying messages and increases the worker's effort.

Suppose the manager chooses an evaluation strategy $\hat{\sigma}$ as described before. Denote the worker's best response to $\hat{\sigma}$ by $\hat{a}$. It must satisfy the following first-order condition

$$\frac{\partial V_W}{\partial a}(\hat{a}, \pi_W, \hat{\sigma}) = \int_Z \sum_Y \sum_E u_W\left(\pi_W(y,e)\right)\hat{\sigma}(e|y,z)\left[\frac{p_a(y|\hat{a})}{p(y|\hat{a})} + \frac{q_a(z|\hat{a})}{q(z|\hat{a})}\right] p(y|\hat{a})q(z|\hat{a})dz - c'(\hat{a}).$$

As in the baseline version of the model, I construct an alternative evaluation strategy $\tilde{\sigma}$ that sends the highest paying message when the signal pair (y,z) has a positive score and the lowest paying message otherwise. In the continuous effort version, the score of a given pair (y,z) is given by

$$\left[\frac{p_a(y|a)}{p(y|a)} + \frac{q_a(z|a)}{q(z|a)}\right].$$

Hence, one can define $\hat{z} : Y \times [0,1] \to Z$ such that

$$\frac{p_a(y|a)}{p(y|a)} = -\frac{q_a(\hat{z}(y,a)|a)}{q(\hat{z}(y,a)|a)}$$

and let $\tilde{\sigma}$ be

$$\tilde{\sigma}(y,z) := \begin{cases} \delta_{\underline{e}_y} \text{ if } z < \hat{z}(y,\hat{a}), \\ \delta_{\overline{e}_y} \text{ otherwise,} \end{cases}$$

where $\underline{e}_y$ is the lowest-paying and $\overline{e}_y$ the highest-paying message when output is y. One can check that

$$\frac{\partial V_W}{\partial a}(\hat{a}, \pi_W, \tilde{\sigma}) > \frac{\partial V_W}{\partial a}(\hat{a}, \pi_W, \hat{\sigma}) = 0.$$

Hence, if the manager uses $\tilde{\sigma}$ instead of $\hat{\sigma}$, the worker has the incentive to locally increase his effort from $\hat{a}$. The final step is to show that with $\tilde{\sigma}$ the worker's payoff is strictly concave in a, which implies that the best response to $\tilde{\sigma}$ cannot be lower than $\hat{a}$. Assumption 6 and $c''(\cdot) > 0$ imply such strict concavity.

[40] Note that v_M is increasing and non-constant, and $p_w(.|a_M, a_W)/p(.|a_M, a_W)$ is strictly increasing. Hence, the inequality is a direct implication of Lemma 4.

C.1 Proofs for the Continuous Effort Case

Proof of Proposition 10. Take given contracts (E, π_M, π_W). Denote the worker's best response to σ by $\breve{a}(\sigma)$. I split the proof into two cases.

Case 1: $\breve{a}(\sigma) = 0$ for all σ. Then, the worker's effort is zero regardless of the evaluation policy, and the result trivially holds.

Case 2: there exists σ such that $\breve{a}(\sigma) > 0$.

Take an arbitrary $\hat{\sigma}$ such that $\breve{a}(\hat{\sigma}) > 0$ and $\hat{\sigma}$ sends with strictly positive probability messages that are not the highest and the lowest-paying. Denote by $\hat{a} := \breve{a}(\hat{\sigma})$. As $\lim_{x \to 1} c'(x) = +\infty$, the effort level $\hat{a}$ must satisfy the following first-order condition

$$\int_Z \sum_Y \sum_E u_W\left(\pi_W(y,e)\right) \hat{\sigma}(e|y,z) \left[\frac{p_a(y|\hat{a})}{p(y|\hat{a})} + \frac{q_W(z|\hat{a})}{q(z|\hat{a})}\right] p(y|\hat{a}) q(z|\hat{a}) dz = c'(\hat{a}).$$

Define $\hat{z} : Y \times [0,1] \to Z$ such that

$$\frac{p_a(y|a)}{p(y|a)} = -\frac{q_W(\hat{z}(y,a)|a)}{q(\hat{z}(y,a)|a)}$$

Let $\tilde{\sigma}$ be

$$\tilde{\sigma}(y,z) := \begin{cases} \delta_{\underline{e}_y} & \text{if } z < \hat{z}(y,\hat{a}), \\ \delta_{\overline{e}_y} & \text{otherwise.} \end{cases}$$

Remember that $\underline{e}_y$ is the lowest-paying and $\overline{e}_y$ is the highest-paying message when output is y. Note that

$$\begin{aligned} \frac{\partial V_W}{\partial a}(\hat{a}, \pi_W, \tilde{\sigma}) &= \int_Z \sum_Y \sum_E u_W\left(\pi_W(y,e)\right) \tilde{\sigma}(e|y,z) \left[\frac{p_a(y|\hat{a})}{p(y|\hat{a})} + \frac{q_a(z|\hat{a})}{q(z|\hat{a})}\right] p(y|\hat{a}) q(z|\hat{a}) dz - c'(\hat{a}) \\ &> \int_Z \sum_Y \sum_E u_W\left(\pi_W(y,e)\right) \hat{\sigma}(e|y,z) \left[\frac{p_a(y|\hat{a})}{p(y|\hat{a})} + \frac{q_a(z|\hat{a})}{q(z|\hat{a})}\right] p(y|\hat{a}) q(z|\hat{a}) dz - c'(\hat{a}) \\ &= \frac{\partial V_W}{\partial a}(\hat{a}, \pi_W, \hat{\sigma}) = 0. \end{aligned}$$

Hence, if the manager uses $\tilde{\sigma}$ instead of $\hat{\sigma}$, the worker has the incentive to locally increase his effort from $\hat{a}$. If $V_W(., \pi_W, \tilde{\sigma})$ is strictly concave, then $\breve{a}(\tilde{\sigma}) > \hat{a}$.

I now show that $V_W(., \pi_W, \tilde{\sigma})$ is strictly concave. Given the evaluation strategy $\tilde{\sigma}$, the worker's payoff is given by[41]

$$V_W(a, \pi_W, \tilde{\sigma}) = \sum_Y \left\{ u_W\left(\pi_W(y, \overline{e}_y)\right) \left[1 - Q(\hat{z}(y,\hat{a})|a)\right] + u_W\left(\pi_W(y, \underline{e}_y)\right) Q(\hat{z}(y,\hat{a})|a) \right\} p(y|a) - c(a).$$

Hence,

$$\frac{\partial^2 V_W}{\partial a^2}(a, \pi_W, \tilde{\sigma}) = -c''(a)$$

[41] I denote by $Q_a(z|a) := \frac{\partial Q}{\partial a}(z|a_W)$ and $Q_{aa}(z|a) := \frac{\partial^2 Q}{\partial a^2}(z|a)$.

$$-\sum_{Y}\left[u_W\left(\pi_W(y,\bar{e}_y)\right)-u_W\left(\pi_W(y,\underline{e}_y)\right)\right]\left[Q_{aa}\left(\hat{z}(y,\hat{a})|a\right)p(y|a)+2Q_a\left(\hat{z}(y,\hat{a})|a\right)p_a(y|a)\right]<0.$$

The first square bracket is positive by construction, whereas the second is positive by Assumption 6. Therefore, for any arbitrary $\hat{\sigma}$, I have constructed a better evaluation strategy $\tilde{\sigma}$. It remains to show that there exists a σ^* that maximizes $\check{a}(\sigma)$.

I say that a performance evaluation strategy σ has a cutoff form if there exists a cutoff function $\check{z}: Y \to Z$ such that

$$\sigma(y,z) := \begin{cases} \delta_{\underline{e}_y} \text{ if } z < \check{z}(y), \\ \delta_{\bar{e}_y} \text{ otherwise.} \end{cases}$$

Note that for each σ, there exists a σ' with a cutoff form such that $\check{a}(\sigma) \leq \check{a}(\sigma')$. Hence, one can restrict attention to performance evaluations with a cutoff form. Also, a performance evaluation with a cutoff form can be described by a vector of cutoffs, one for each y. As each $\check{z}(y)$ belongs to a compact set Z, the set of performance evaluations with a cutoff form is compact. Finally, as the worker's effort choice is characterized by the first-order condition, it is continuous in the cutoff vector. Therefore, there exists an effort maximizing performance evaluation strategy with a cutoff form. □

References

Alchian, A. A. and H. Demsetz (1972). Production, information costs, and economic organization. *The American Economic Review 62*(5), 777–795.

Baker, G., R. Gibbons, and K. J. Murphy (1994). Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics 109*(4), 1125–1156.

Benson, A. (2015). Do agents game their agents' behavior? evidence from sales managers. *Journal of labor economics 33*(4), 863–890.

Brasiliense de Castro Pires, H. (2022). *Essays on Contract Theory*. Ph. D. thesis. Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2023-03-08.

Charles Aris Inc. (2022). Strategy Consulting Compensation Study. *Report*.

Cheng, C. (2021). Moral hazard in teams with subjective evaluations. *The RAND Journal of Economics 52*(1), 22–48.

Deb, J., J. Li, and A. Mukherjee (2016). Relational contracts with subjective peer evaluations. *The RAND Journal of Economics 47*(1), 3–28.

Fong, Y.-F. and J. Li (2017). Information revelation in relational contracts. *The Review of economic studies 84*(1 (298)), 277–299.

Frederiksen, A., L. B. Kahn, and F. Lange (2020). Supervisors and performance management systems. *Journal of Political Economy 128*(6), 2123–2187.

Frederiksen, A., F. Lange, and B. Kriechel (2017). Subjective performance evaluations and employee careers. *Journal of Economic Behavior & Organization 134*, 408–429.

Fuchs, W. (2007, September). Contracting with repeated moral hazard and private evaluations. *American Economic Review 97*(4), 1432–1448.

Fuchs, W. (2015, February). Subjective evaluations: Discretionary bonuses and feedback credibility. *American Economic Journal: Microeconomics 7*(1), 99–108.

Georgiadis, G. and B. Szentes (2020). Optimal monitoring design. *Econometrica 88*(5), 2075–2107.

Gershkov, A. and E. Winter (2015, April). Formal versus informal monitoring in teams. *American Economic Journal: Microeconomics 7*(2), 27–44.

Grossman, S. J. and O. D. Hart (1983). An analysis of the principal-agent problem. *Econometrica 51*(1), 7–45.

Halac, M., E. Lipnowski, and D. Rappoport (2021, March). Rank uncertainty in organizations. *American Economic Review 111*(3), 757–86.

Holmström, B. (1982). Moral hazard in teams. *The Bell Journal of Economics 13*(2), 324–340.

Holmström, B. (2016). Pay for Performance and Beyond. *Nobel Prize Lecture*.

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics 10*(1), 74–91.

Holmström, B. (2017, July). Pay for performance and beyond. *American Economic Review 107*(7), 1753–77.

Huy, Q. (2001, September). In praise of middle managers. *Harvard business review 79*(8), 72—9, 160.

Innes, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory 52*(1), 45–67.

Ishiguro, S. and Y. Yasuda (2021, May). Moral Hazard and Subjective Evaluation. *Working paper*, 1–45.

Jackson, M. O. and H. F. Sonnenschein (2007). Overcoming incentive constraints by linking decisions. *Econometrica 75*(1), 241–257.

Jaser, Z. (2021, June). The real value of middle managers. *Harvard business review*.

Kambe, S. (2006, March). Subjective Evaluation In Agency Contracts. *The Japanese Economic Review 57*(1), 121–140.

Kampkoetter, P. and D. Sliwka (2016). The complementary use of experiments and field data to evaluate management practices: The case of subjective performance evaluations. *Journal of Institutional and Theoretical Economics 172*(2), 364–389.

Kurkowski, H. (2021, December). Monitoring remote workers: The good, the bad, and the ugly. *Forbes*.

Lang, M. (2019). Communicating subjective evaluations. *Journal of Economic Theory 179*, 163–199.

Lazear, E. P. and P. Oyer (2012). Personnel economics. In R. Gibbons and J. Roberts (Eds.), *The Handbook of Organizational Economics*. Princeton University Press.

Lazear, E. P. and S. Rosen (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy 89*(5), 841–864.

Lazear, E. P., K. L. Shaw, and C. T. Stanton (2015). The value of bosses. *Journal of Labor Economics 33*(4), 823–861.

Letina, I., S. Liu, and N. Netzer (2020). Delegating performance evaluation. *Theoretical Economics 15*(2), 477–509.

Levin, J. (2003, June). Relational incentive contracts. *American Economic Review 93*(3), 835–857.

Li, A. and M. Yang (2020). Optimal incentive contract with endogenous monitoring technology. *Theoretical Economics 15*(3), 1135–1173.

MacLeod, W. B. (2003, March). Optimal contracting with subjective evaluation. *American Economic Review 93*(1), 216–240.

MacLeod, W. B. and J. M. Malcomson (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica 57*(2), 447–480.

Maestri, L. (2012, April). Bonus payments versus efficiency wages in the repeated principal-agent model with subjective evaluations. *American Economic Journal: Microeconomics 4*(3), 34–56.

Maestri, L. (2014). The efficiency of bonus-or-terminate incentive schemes under subjective evaluations. *Games and Economic Behavior 87*, 412–418.

Milgrom, P. and I. Segal (2002). Envelope theorems for arbitrary choice sets. *Econometrica 70*(2), 583–601.

Mirrlees, J. (1999). The theory of moral hazard and unobservable behaviour: Part i. *Review of Economic Studies 66*(1), 3–21.

Murphy, K. (1999). Executive compensation. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (1 ed.), Volume 3, Part B, Chapter 38, pp. 2485–2563. Elsevier.

Ostrizek, F. (2021, December). Vague by Design: Performance Evaluation and Learning from Wages. *Working paper*, 1–43.

PageExecutive (2020). Banking and Asset Management Salary Survey. *Report*.

Prendergast, C. and R. H. Topel (1996). Favoritism in organizations. *Journal of Political Economy 104*(5), 958–978.

Rahman, D. (2012, May). But who will monitor the monitor? *American Economic Review 102*(6), 2767–97.

Ross, S. (1973). The economic theory of agency: The principal's problem. *American Economic Review 63*(2), 134–39.

Shavell, S. (1979). Risk sharing and incentives in the principal and agent relationship. *The Bell Journal of Economics 10*(1), 55–73.

Winter, E. (2004). Incentives and discrimination. *The American Economic Review 94*(3), 764–773.

Zhu, J. Y. (2018, November). A foundation for efficiency wage contracts. *American Economic Journal: Microeconomics 10*(4), 248–88.

Zábojník, J. (2014). Subjective evaluations with performance feedback. *The RAND Journal of Economics 45*(2), 341–369.