Agency in Hierarchies: Middle Managers and Performance Evaluations*

Henrique Castro-Pires†

November 18, 2021

(Click here for the most recent version)

Abstract

This paper 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 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. Also, subjective performance evaluations are valuable if and only if the verifiable performance measure is more informative than the non-verifiable one. Finally, I show that the principal may benefit from reducing transparency in the organization, as the cost of implementing the desired efforts can strictly decrease when the manager has less information about the worker’s effort.

*I thank Jeroen Swinkels, George Georgiadis, Alessandro Pavan, Humberto Moreira, Andrés Espitia, Michael Powell, Cassiano Alves, Miguel Talamas, Gabriel Ziegler, and Bruno Barsanetti for the discussions and feedback. I also thank seminar participants at Northwestern University for helpful comments.

†Managerial Economics and Strategy, Kellogg School of Management, Northwestern University - Evanston, IL, USA - email: henrique.pires@kellogg.northwestern.edu.
1 Introduction

In classic principal-agent models, the principal directly deals with her agents. However, in practice, the top and the bottom of organizations 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 their own direct productive activity and for motivating their subordinates. Managers are closer to the rank-and-file and usually 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. 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 paper studies the optimal joint design of incentives and rating scales in hierarchies.

I analyze this question within a moral-hazard-in-teams model, in which a risk-neutral principal designs a compensation scheme for two risk-averse agents, a manager and a worker. The principal wishes to incentivize both the manager and the worker to exert productive effort. 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 and the worker decide whether to exert effort at a private cost. Efforts stochastically generate team output, and a non-verifiable signal about the worker’s effort privately observed by the manager. Given the output and the private signal realization, the manager decides how to rate the worker’s performance. A crucial feature of the model is the manager’s dual role: exerting productive effort and reporting on worker’s performance.

As an example, consider a retail chain store designing the compensation scheme for its managers and salespeople. Suppose that a given store has one manager and one salesperson. The firm

---

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 paper, Lazear et al. (2015) estimate supervisor effects on firm outcomes and find them to be large. By 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.

2 There exists a broad range of empirical studies on subjective performance evaluations in a variety of 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).
wants to motivate both the manager and the salesperson to exert effort, and it observes sales, which are affected by the manager’s and salesperson’s effort. The manager interacts daily with the salesperson and has more information about the salesperson’s effort than 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 both on sales and on the performance report provided by the manager.

A key challenge in finding the optimal contracts in the model I study is that the information observed by the principal directly depends on the compensation scheme. In most moral hazard problems, the distribution of signals observed by the principal depends only on efforts but not on the contracts directly. Hence, one can solve for the optimal contracts by using the Grossman and Hart (1983) approach. That is, one can fix the effort levels the principal wishes to implement and minimize the expected payments. In my model, the manager’s report about the worker’s effort directly depends on the payment schemes. Hence, one must determine not only the optimal contracts for given information available to the principal but also what information the principal obtains given the contracts in place. I first show that the manager’s reporting strategy must take a simple binary form, and given the information generated under that simple structure, I characterize the optimal contracts.

The main finding is that 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 — would choose the report in which she gets the highest wage. Hence, the manager’s compensation must be independent of her report about the worker’s performance.\(^3\)

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

   \(^3\)The intuition is similar to [MacLeod (2003)](MacLeod2003), in which the evaluator (there the principal, here the manager) must be indifferent between any report. MacLeod (2003) achieves indifference through money burning, while here it requires not conditioning the manager’s payments on her report.
As the principal must motivate the manager to exert effort, she must pay the manager conditional on team output (the verifiable public signal). Both the manager’s and the worker’s effort affect the output distribution. In particular, higher effort by either player shifts probability mass to higher output realizations. As the manager gets larger payments when output is higher, she benefits from higher worker’s effort.

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

   Steps 1 and 2 imply that the manager benefits from inducing higher worker’s effort (step 2) but does not internalize the cost to provide 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 the strongest effort incentives, regardless of how much this costs to the firm. The performance evaluation strategy that generates the strongest incentives for effort is such that 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.

   Although, I state the model in the context of performance reviews, it applies to any situation in which the principal delegates the decision about an agent’s reward to another better-informed agent. For instance, the results are consistent with a firm that has no formal performance review system but lets the manager decide whether or not to promote her subordinate. Even beyond the design of compensation schemes, one might expect the report of non-verifiable information to take a binary structure in instances in which the informed agent cares about the uninformed agent’s action but does not bears compensation costs. For instance, consider an environmental-auditor (an NGO or an environmental-certifier) who wants firms to be environmentally-conscious. Suppose that this auditor is hired (by the regulator, for example) to evaluate firms’ practices. Similar to the manager in my model, the auditor wants to create strong incentives for environment-friendly practices (effort) but she does not design firms’ penalties and rewards. The best the auditor can do is to only use extreme reports - the lowest and the highest - generating the strongest possible incentives for green practices.

   Many incentive schemes used in practice take a binary structure. “Up-or-out” systems and discretionary single-valued bonuses are prominent examples. The result that managers concentrate their evaluations on few points of the rating scale is consistent with empirical findings in the lit-

---

4A standard monotone likelihood ratio assumption on the output distribution generates this feature.
erature. [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 few points in the scale, often in only two. In a more recent paper, [Frederiksen et al. (2020)] use personnel data from a sizable Scandinavian service sector firm and show that two out of the five possible points in the rating scale concentrated over 90% of the reports.

In a binary performance review system, the manager reports whether the worker’s performance was good or bad. The manager’s preferred performance evaluation strategy can be described by a cutoff function. Conditional on output, the manager provides a good evaluation if the signal she privately observes is sufficiently high. The 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. 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 that 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 provides a good report even if the private signal realization is the worst possible. Similarly, the manager provides a bad report regardless of the private signal realization when output is low. In this case, the performance evaluation report does not convey any additional information beyond what is conveyed by output. In those situations, optimal contracts do not use subjective performance evaluations.\footnote{This result is consistent with the view that subjective measures are helpful only when the objective measure is not precise enough.}

When considering cases in which subjective performance evaluations are valuable, one can ask how informed about worker’s effort the principal wishes the manager to be. A less informed man-
ager has a worse assessment of the worker’s action; however, less information might attenuate the principal and manager’s conflicts of interest. The manager observes a private non-verifiable non-fully-informative signal \( z \) with support \( Z := [\underline{z}, \overline{z}] \subset \mathbb{R} \) about the worker’s effort. The principal would like the manager to fully reveal the signal realization. However, in equilibrium, the manager uses at most two ratings: the highest and the lowest. Then, I ask whether reducing the manager’s information about the worker’s effort can improve the principal’s payoff. In particular, I ask whether the principal is better off if instead of observing the realization of \( z \), she observes the realization in a coarser partition of \( Z \). I refer to coarser partitions as less transparent organizations. I show that reducing transparency from the full-transparency benchmark strictly benefits the principal.

Despite having a less informed manager, the principal implements given effort at a cheaper cost. To understand the result, note that reducing transparency decreases the conflict of interests between the manager and the principal 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 high and too rarely when output is low. A coarser partition with two distinct pooling regions - extremely low and extremely high realizations - reduces implementation costs. By censoring the extreme signal realizations, the principal increases the probability of a good report when output is low and decreases the probability of a good report when output is high. This censoring increases the expected worker’s payment for low output realizations and decreases for high output realizations, improving risk-sharing and decreasing compensation costs.

Finally, the paper proceeds as follows. The next section discusses the related literature. Section 2 presents the baseline model. Section 3 solve for the optimal contracts. Section 4 analyzes how informed the principal wants the manager to be about the worker’s effort. Section 5 shows that the optimality of binary performance systems is robust to alternative model specifications: continuous effort levels, multiple workers, substitute efforts and allowing for screening. Finally, Section 6 concludes. Proofs are relegated to the appendix.

---

6I denote random variables in bold.
7For instance, a coarser partition would be \( \{ [\underline{z}, z_1); [z_1, z_2); (z_2, \overline{z}] \} \) for \( z_1, z_2 \in Z \) and \( z_1 < z_2 \). Instead of observing the realization directly, the manager only observes which of the three intervals the realization belongs to.
8The full-transparency benchmark is the partition \( \{ \{ z \} | z \in Z \} \).
1.1 Related Literature

This paper belongs to the literature on optimal provision of incentives for groups of agents, pioneered by Lazear and Rosen (1981) and Holmström (1982). In those papers, 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); Maestri (2012); Maestri (2014); Fuchs (2015); Deb et al. (2016); Letina et al. (2020); Cheng (2021); Ishiguro and Yasuda (2021) — examine optimal contracting in the presence of subjective performance evaluations.

Subjective performance measures are usually modeled as non-verifiable private signals. The non-verifiability creates incentives for non-truthful reporting. Most studies analyzing moral hazard problems with subjective performance evaluations fit in two categories based on who conducts the performance review: the principal (the residual claimant) or an agent. When the principal is the performance reviewer, the temptation to renege on payments is the main friction. After effort has been conducted, the principal has the incentive to provide a low-performance review and save on payments to the agent. Previous literature dealt with that type of limited commitment through repeated interactions - MacLeod and Malcomson (1989); Levin (2003); Fuchs (2007); 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.

The second strand of literature focuses on the case in which a monitor has the sole task of evaluating workers based on non-verifiable information. However, the monitor is not the residual claimant. In those settings, the temptation to renege on payments ceases to be the main friction. The monitor is assumed to have biased preferences in favor of one (Prendergast and Topel (1996)) or all (Letina et al. (2020)) workers. After signals have been realized, the monitor does not have incentives to punish poor performance, which harms incentives ex-ante. When dealing with this issue, the principal must use contracts that force the monitor to punish workers — such as a loser-
gets-nothing tournament. There are two major differences in my setting. First, the manager is not biased towards the worker. Second, the manager not only reports on worker’s performance but also exerts productive effort. The conflict of interest between principal and manager arises because the manager is not the residual claimant but endogenously benefits from higher worker’s effort.

This paper also contributes to the literature on the interaction between verifiable signals and peer monitoring in moral hazard settings. Baker et al. (1994), Che and Yoo (2001), Rayo (2007) and Deb et al. (2016) study such interaction in relational contracts among risk-neutral agents. My contribution differs from theirs in two important aspects. First, I show that subjective performance evaluations might be useful even in a static setting. Second, the static approach allows me to tractably introduce risk-averse preferences. The assumption that agents are risk-neutral eliminates risk-sharing concerns and mechanically generates simpler incentive schemes. I show that even with risk-averse agents, the use of subjective performance ratings must take a simple binary structure.

In a closely related paper, Cheng (2021) analyzes the use of self and peer reviews in incentive contracts and how the degree of subjectivity — measured by the correlation between self and peer perception of performance — shapes optimal contracts. In her paper, subjective evaluations are useful because they can be cross-checked against self-evaluations (and vice-versa). In my setting, it is not feasible to cross-check manager’s subjective performance evaluation reports. However, they are still useful. In particular, the manager has incentives to motivate the worker to exert as much effort as possible.

This paper is also related to the literature on incentives and supervision pioneered by Tirole (1986) and further explored by Laffont (1990), Faure-Grimaud et al. (2003) among others. Those papers focus on how the possibility of collusion between the manager and the worker affects optimal contracts. They assume the manager and the worker can establish fully enforceable side-contracts on top of what has been agreed with the principal. In my context, those side contracts are unfeasible, and there is no room for collusion between manager and worker.

This paper 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

---

11 Brown and Heywood (2005) and Addison and Belfield (2008) provide empirical evidence that performance evaluations are often used for employees with short expected tenure.

12 See Mookherjee (2012) for a survey.
endogenous monitoring. Rahman (2012) shows that the principal can provide the monitor with incentives to exert costly monitoring effort by randomly and secretly allowing the worker to shirk. Strulovici (2021a,b) analyzes the importance of intrinsic ethical motives for successful monitoring by intermediaries. 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 incentives to reveal the 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.

2 Model

A risk-neutral principal who hires a manager (she; denoted with $M$) and a worker (he; denoted with $W$), both risk-averse. First, the principal proposes contracts to the worker and the manager. The contract specifies 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 simultaneously decide whether or not to accept the contracts. If any of the two rejects, both get their respective outside options $\bar{u}_W$ and $\bar{u}_M$. If both accept, the worker privately draws an effort cost $c_i$ from a commonly known distribution $G$ with full-support on $\mathbb{R}^+$. The manager’s effort cost is deterministic and equal to $c_M > 0$. Knowing their respective effort costs, the manager and the worker simultaneously decide whether to exert effort, denoted by $a_i \in \{0, 1\}$ for $i \in \{W, M\}$.

Given the worker’s and the manager’s effort decisions, two signals are realized: verifiable output $y$ and a non-verifiable signal $z$. Output is verifiable, observed by all players, and has a finite support $Y := \{y_0, \ldots, y_n\} \subset \mathbb{R}$. Conditional on efforts, the probability of an output realization $y$ is denoted by $p(y|a_M, a_W) \in (0, 1]$, and with cdf denoted by $P(y|a_M, a_W)$. The non-verifiable signal $z$ is privately observed by the manager, and has support $Z := [z_l, z_u] \subset \mathbb{R}$. I assume it admits a density $q(z|a_W)$ with full-support, and has a cdf denoted by $Q(z|a_W)$.

I assume that $p$ and $q$ are common knowledge and higher signal realizations are associated with

---

13Henceforth, I denote random variables in bold and their realizations in regular font.

14The non-verifiable signal $z$ being a continuous random variable is a matter of convenience. The results also hold if $z$ is a discrete random variable or $y$ is continuous.
higher effort. Formally, \( p \) and \( q \) satisfy the following property:

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

1. \( \frac{p(y|a_W,1)}{p(y|a_W,0)} \) is strictly increasing in \( y \) for any \( a_W \in \{0, 1\} \).
2. \( \frac{p(y|a_M,1)}{p(y|a_M,0)} \) is strictly increasing in \( y \) for any \( a_M \in \{0, 1\} \).
3. \( \frac{q(z|1)}{q(z|0)} \) is strictly increasing in \( z \).

Knowing her own effort choice, and after observing the realizations of \( y \) and \( z \), the manager decides how to evaluate the worker’s performance. That is, the manager chooses a message from the evaluation set \( E \). Formally, the manager’s evaluation decision can be represented by \( \sigma : \{0, 1\} \times Y \times Z \rightarrow \Delta(E) \). I denote by \( \sigma(e|a_M, y, z) \) the probability of message \( e \) given \( a_M, y, z \). Then, payments are realized.

The timing of the model can be summarized as:

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

Contracts are a triple: a finite set of possible performance ratings \( E \), payments for the manager \( \pi_M : Y \times E \rightarrow \mathbb{R}_+ \) and for the worker \( \pi_W : Y \times E \rightarrow \mathbb{R}_+ \) contingent on output and performance evaluation. 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 allowed. In Section 5, I solve for the case in which the manager fully observes the worker’s contract choice. All results remain qualitatively unaltered.

---

\(^{15}\) Restricting attention to finite performance ratings is not necessary but avoids technical complications. What I need is \( \max_{e \in E} \{ \pi_i(y, e) \} \) and \( \min_{e \in E} \{ \pi_i(y, e) \} \) to be well defined for each \( i \in \{W, M\} \) and \( y \in Y \).
The worker and the manager are risk-averse and their ex-post payoffs given efforts, contracts, realized output, performance evaluation, and effort costs are:

\[ U_i(\pi_i, y, e, c_i, a_i) := u_i(\pi_i(y, e)) - c_i a_i \quad \text{for } i \in \{M, W\}. \]

Payments must be above a lower bound normalized to zero. I assume that \( u_W : \mathbb{R}_+ \rightarrow \mathbb{R} \) and \( u_M : \mathbb{R}_+ \rightarrow \mathbb{R} \) are strictly increasing, strictly concave, twice continuously differentiable and \( \lim_{x \rightarrow +\infty} u_i(x) = +\infty \). The manager and the worker have outside options \( \bar{u}_M > u_M(0) \) and \( \bar{u}_W > u_M(0) \), respectively.

The principal’s payoff is output \( y \) minus the payments to the agents (worker and manager). That is,

\[ U_P(y, e) := y - \pi_W(y, e) - \pi_M(y, e). \]

After the manager and worker have accepted the contracts, they play an extensive form game, in which they each choose effort levels, signals are realized, and the manager chooses an evaluation report. I use subgame perfect Nash equilibrium as the solution concept.

I also assume that the manager can freely dispose of output.

**Assumption 2.** [Managerial free disposal] The manager can costlessly decrease output.

Assumption 2 implies that manager’s payments must be increasing in \( y \). If they were not, the manager would decrease output to the highest paying output level. This assumption restricts our focus to monotonic contracts, which simplify the analysis.\(^{16}\) This assumption can be dispensed without changing the results. However, dropping Assumption 2 brings additional complication to the analysis, but no gain in economic insight.\(^{17}\)

As usual in the literature (e.g., Winter (2004), Halac et al. (2021)), I focus on the case in which efforts are complements. I assume that the cdf \( P \) is submodular.

**Assumption 3.** For all \( y \in Y \), \( P(y|\cdot, \cdot) \) is submodular. That is, for all \( y \in Y \)

\[ P(y|1, 1) + P(y|0, 0) \leq P(y|1, 0) + P(y|0, 1). \]

\(^{16}\)It is also a common assumption in the literature, e.g., Innes (1990).

\(^{17}\)Without Assumption 2 I would have to divide the analysis into two cases: 1- conjecture that the manager benefits from higher worker’s effort, 2- conjecture that the manager is worse off with higher worker’s effort. Then, I’d show that the second conjecture cannot hold at the optimal contracts. Assumption 2 assures the manager benefits from higher worker’s effort and allows me to skip this detour. In the proofs, I point out exactly where I use Assumption 2 and how the approach would change if it was absent.
3 Baseline Model Analysis

3.1 Preliminaries: Effort Complementarity

In this subsection, I show that Assumption 3 is equivalent to effort complementarity. I first, formally define effort complementarity.

**Definition 1.** I say that $p$ satisfies **effort complementarity** if for any increasing $b : Y \rightarrow \mathbb{R}$,

$$\sum_y b(y) \left[ p(y|1, 1) - p(y|0, 1) \right] \geq \sum_y b(y) \left[ p(y|1, 0) - p(y|0, 0) \right].$$

Efforts are complements if for any increasing reward function, one agent’s higher effort increases the incentives for the other agent to exert effort. Note that as the inequality must hold for any increasing function, this definition can be interpreted as an ordinal notion of complementarity. Even if we do not know the intensity of how much better higher values of $y$ are compared to a lower levels, one agent’s gain from high effort is larger when the other agent exerts high effort. The next result shows that Assumption 3 is equivalent to effort complementarity.

**Proposition 1.** The distribution $p$ satisfies effort complementarity if and only if Assumption 3 holds.

To the best of my knowledge, this is a novel definition and characterisation of effort complementarity in environments with stochastic and non-binary outcomes. Throughout the paper, I maintain Assumption 3. In Section 5, I discuss what would change with substitute efforts.

---

18I state and prove the result for two agents, which is the relevant case for this paper. However, the result extends to any number of agents and to continuous output. Suppose there are $N \geq 2$ agents. Let $p : Y \times \{0, 1\}^N \rightarrow [0, 1]$ denote the output’s distribution, and $a, \tilde{a} \in \{0, 1\}^N$. Efforts are complements in the sense that for any increasing $b : Y \rightarrow \mathbb{R}$

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

if and only if $P(y|\cdot)$ is submodular for any $y \in Y$. 

12
3.2 Worker’s Effort Choice

I first analyze the worker’s effort choice. That is, fixing contracts \((E, \pi_W, \pi_M)\), the manager’s effort choice \(a_M\) and evaluation strategy \(\sigma\), the worker’s expected utility is

\[ V_W(a_W, a_M, \pi_W, \pi_M, \sigma, c) := \mathbb{E}[u_W(\pi_W(y, e))|a_M, a_W, \sigma] - a_W c. \]

The worker’s best response is given by a cutoff. That is, given the contracts and conjectures about the manager’s actions (effort and evaluation), the worker exerts high effort if \(c\) is smaller than

\[ c_W(\pi_W, a_M, \sigma) = \mathbb{E}[u_W(\pi_W(y, e))|a_M, a_W = 1, \sigma] - \mathbb{E}[u_W(\pi_W(y, e))|a_M, a_W = 0, \sigma]. \]

From now on, when I talk about worker’s effort I refer to a cutoff \(c_W\). That is, given contracts \((E, \pi_W, \pi_M)\) an equilibrium is described by a triple \((a_M, c_W, \sigma)\), where \(a_M\) denotes the manager’s effort, \(c_W\) the worker’s effort cost cutoff and \(\sigma\) represents the distribution of manager’s evaluation reports.

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 paper is that managers do not internalize the compensation cost and wish to motivate as much effort as possible from their subordinates. 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. In Section 5, I analyze a version of the model with deterministic costs but continuous efforts. The main trade-offs and results continue to hold.

3.3 Principal’s Problem

Following the [Grossman and Hart (1983)] approach, I can state the principal’s problem as choosing contracts \((E, \pi_W, \pi_M)\) and recommending actions \((c_W, a_M, \sigma)\) such that those actions are implemented at the minimal possible cost. Now, I show that there is no loss in restricting attention to contracts in which \(\pi_M\) does not depend on the performance report \(e\).

I have not imposed almost any structure on the set of performance ratings \(E\), nor on how they map to payments. Hence, different contracts might generate the same relevant outcomes in equilibrium. For example, if we add performance ratings that are never used in equilibrium, we do not 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 \((E, \pi_W, \pi_M)\) and \((\tilde{E}, \tilde{\pi}_W, \tilde{\pi}_M)\) are \textbf{outcome-equivalent} if for any equilibrium under \((\tilde{E}, \tilde{\pi}_W, \tilde{\pi}_M)\) there exists an equilibrium under \((E, \pi_W, \pi_M)\) with the same effort choices, output and payments distribution and vice-versa.

Lemma 1. For any contract \((E, \pi_W, \pi_M)\) 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, effort costs are sunk. The manager only uses reports that maximize her payments. Hence, if the principal conditions the manager’s payments on her reports, the manager chooses the reports which maximize her payments, not necessarily reflecting the private information about the worker’s effort.\footnote{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.}

Remark 1. Lemma\cite{Lem} implies that we can — without loss — restrict attention to contracts in which \(\pi_M\) does not depend on the performance evaluation. From now on, I restrict attention to such contracts. Then, after signals have been realized, the manager is indifferent between any report. Hence, any reporting strategy is sequentially rational.

Traditionally, on agency models, we break indifferences in favor of the principal. 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 rating, the manager would always provide that rating, regardless of her private signal. Hence, all informational content of the performance evaluation would be lost.

I focus on a different equilibrium selection criterion: the manager’s preferred equilibrium.\footnote{There is also a growing literature on unique implementation. For instance, see \cite{Winter} 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.} The manager is ex-post — after the effort choices — indifferent between any report, but not ex-ante. Different report policies generate different incentives for the worker to exert effort. As the worker’s effort affects output distribution, and the manager’s payment depends on output, the manager may not be indifferent between different report policies ex-ante.\footnote{Manager’s preferred equilibrium selection is similar to the credible threats refinement proposed by \cite{zhu}.} As the manager can decide how to
We can think about the principal’s problem as choosing contracts and recommending actions (a cutoff \(c_W, a_M\) and \(\sigma\)) such that they constitute the manager’s preferred equilibrium. It is more convenient to work on the space of utilities than on the space of payments. Define \(v_M(y) \in [u_M(0), +\infty)\) and \(v_W(y) \in [u_W(0), +\infty)\) as

\[
v_M(y) := u_M(\pi_M(y)),
\]

and

\[
v_W(y,e) := u_W(\pi_W(y,e)),
\]

and \(\varphi_W := u_W^{-1}\) and \(\varphi_M := u_M^{-1}\).

Denote \(\Gamma(E,v_W,v_M)\) as the set of subgame perfect Nash equilibria given contracts \((E,v_W,v_M)\). The principal minimizes the cost of implementing given effort levels \((c_W,a_M = 1)\). I focus on cases where those effort levels can be implemented, and the minimum payment constraint (payments being positive) is slack (thereon assumed). Appendix B proves the existence of optimal contracts and finds conditions for the minimum payment constraint to be slack (the outside options must be sufficiently high).

The principal’s problem can be written as

\[
\min_{E,\sigma,v_W,v_M} \left\{ \mathbb{E} \left[ \varphi^W(v_W(y,e)) + \varphi^M(v_M(y)) | c_W, a_M = 1, \sigma \right] \right\}
\]

subject to

\[
v_M(\cdot) \text{ increasing,} \tag{MON}
\]

\[
\mathbb{E}[v_M(y) | c_W, a_M = 1] - c_M \geq \bar{u}_M, \tag{IR_M}
\]

\[
\mathbb{E}[v_W(y,e) | c_W, a_M = 1, \sigma] - \mathbb{E}[c | c \leq c_W] \geq \bar{u}_W, \tag{IR_W}
\]

\[
\mathbb{E}[v_M(y) | c_W, a_M = 1] - \mathbb{E}[v_M(y) | c_W, a_M = 0] \geq c_M, \tag{IC_M}
\]

\[
\mathbb{E}[v_W(y,e) | a_W = 1, a_M = 1, \sigma] - \mathbb{E}[v_W(y,e) | a_W = 0, a_M = 1, \sigma] = c_W, \tag{IC_W}
\]

and

\[
\mathbb{E}[v_M(y) | c_W, a_M = 1] - c_M \geq \mathbb{E}[v_M(y) | \bar{c}_W, \bar{a}_M] - c_M(\bar{a}_M), \quad \forall (c_W, \bar{a}_M, \bar{\sigma}) \in \Gamma(E,v_W,v_M). \tag{EQ-selection}
\]

\[^{22}\text{If the manager’s payments are exogenously restricted to be invariant to her report about the worker’s performance (to avoid a conflict of interest, or legal liability), then the manager’s preferred equilibrium selection criterion would be equivalent to assuming that the manager can commit to an evaluation strategy.}\]
The principal chooses contracts, and a recommended evaluation strategy such that: manager’s payments are increasing (the \( MON \) constraint); the worker and the manager are willing to accept the contracts (\( IR_W \) and \( IR_M \) constraints) and follow the recommended actions (\( IC_W \) and \( IC_M \) constraints); and we select the manager’s preferred equilibrium (\( EQ - selection \) constraint).

This problem differs from canonical moral hazard problems in two important ways: first, the non-standard equilibrium selection constraint. Second, the manager, besides choosing effort, also chooses the performance evaluation strategy.

### 3.4 Performance Evaluation Strategy

As the manager’s payments do not depend on the performance report, any report is sequentially rational. Given contracts \((E, v_W, v_M)\), different \(\sigma\)’s are associated with different equilibria in \(\Gamma(E, v_W, v_M)\). I now characterize the performance evaluation strategy of the manager’s preferred equilibrium.

Note that given effort choices \((c_W, a_M)\), and a manager’s contract \(v_M\), we can write the manager’s expected payoff as

\[
V_M(c_W, a_M, v_M) = \sum_y v_M(y)[p(y|a_M, 1)G(c_W) + [p(y|a_M, 0)(1 - G(c_W)))] - a_Mc_M
\]

\[
= \sum_y v_M(y)p(y|a_M, 0) + G(c_W) \sum_y v_M(y)[p(y|a_M, 1) - p(y|a_M, 0)] - a_Mc_M.
\]

As the principal wishes to implement high managerial effort, any solution to the principal’s problem must be such that the manager’s payment is not constant. I begin focusing on such contracts.

**Lemma 2.** Take any contract such that \(v_M(\cdot)\) is not constant. Then, the manager benefits from higher worker’s effort. That is, \(V_M(c_W, a_M, v_M)\) is strictly increasing in \(c_W\).

As the manager payments are non-constant and 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, but 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. Then, the manager’s favorite equilibrium evaluation strategy must

\[23\text{In a related context, Benson (2015) documents empirical evidence of sales managers increasing their subordinates incentives beyond what is desirable by the firm by shifting sales across periods or changing sales targets.}\]
maximize the worker’s effort cost cutoff. I now characterize the effort-maximizing evaluation strategy.

For the manager to maximize the worker’s effort incentives, 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 \(v_W\), manager’s effort \(a_M = 1\) and performance evaluation \(\sigma\), the worker’s effort cost cutoff is given by

\[
c_W(v_W, a_M = 1, \sigma) = \mathbb{E}[v_W(y, e)|a_W = 1, a_M = 1, \sigma] - \mathbb{E}[v_W(y, e)|a_W = 0, a_M = 1, \sigma]
\]

\[
= \int_{Z} \sum_{Y} \sum_{E} v_W(y, e) \sigma(e|a_M = 1, y, z) [q(z|1)p(y|1,1) - q(z|0)p(y|1,0)] dz.
\]

The \(\sigma\) that maximizes \(c_W(v_W, 1, \sigma)\) is the one that reports the highest paying message when \(q(z|1)p(y|1,1) > q(z|0)p(y|1,0)\), and the lowest paying message otherwise. Let \(\sigma^*\) be such evaluation strategy. That is, define

\[
z^*(y) := \inf\{z \in Z : q(z|1)p(y|1,1) \geq q(z|0)p(y|1,0)\}
\]

and let \(\bar{e}_y \in \arg\max\{v_W(y, \hat{e})\}\) and \(e_y \in \arg\min\{v_W(y, \hat{e})\}\). Then, \(\sigma^*\) is given by

\[
\sigma^*(a_M, y, z) = \begin{cases} 
\delta_{e_y} & \text{if } z < z^*(y) \\
\delta_{e_y} & \text{otherwise,}
\end{cases}
\]

where \(\delta_e\) denotes the Dirac measure centered on \(e\).

**Proposition 2.** Take any contract \((E, v_W, v_M)\) such that there exists an equilibrium with \(a_M = 1\). For any \((\hat{e}, 1, \hat{\sigma}) \in \Gamma(E, v_W, v_M)\), there exists an alternative equilibrium \((e^*, 1, \sigma^*)\) such that \(V_M(e^*, 1, v_M) \geq V_M(\hat{e}, 1, v_M)\). The inequality is strict if \(\hat{\sigma}\) generates a different distribution of payments to the worker than \(\sigma^*\).

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.

\footnote{Let \(z^*(y) = z\) if the set \(\{z \in Z : q(z|1)p(y|1,1) \geq q(z|0)p(y|1,0)\}\) is empty.
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. Hence, proposition \(^2\) implies that the manager’s preferred evaluation strategy uses only the highest and the lowest paying ratings at each output level.

The previous result is consistent with the empirical evidence that documents that very few points of performance evaluation scales are used in practice. [Frederiksen et al. (2017)] shows that this is a common feature documented by several articles using personnel data from different firms. In a different 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.

As a consequence of Proposition \(^2\), the principal can 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 \(\{g,b\}\).

**Remark 2.** The previous result 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 \(\sigma^*\) directly on payments. The meaningful content of the previous result is that given 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 payment conditional on output take a cutoff form. It takes a high value if the reported \(z\) is high enough and a low value otherwise.

**Remark 3.** The constructed \(\sigma^*\) implies that the performance evaluation is \(g\) if and only if

\[
q(z|1)p(y|1,1) \geq q(z|0)p(y|1,0),
\]

or equivalently if \(z \geq z^*(y)\). Note that \(p\) and \(q\) monotone likelihood ratio properties imply a decreasing \(z^*(y)\). Hence, the set of private signal realizations that generate a good report is larger for higher output realizations than for lower realizations. That is, the manager is more lenient in her report when output is high.
3.5 Cost-minimizing Contracts

Note that if we re-examine the principal’s problem I have already established two of her choices. The optimal set of performance evaluations $E = \{b, g\}$ and the recommended evaluation strategy $\sigma^*$. We can re-write the principal’s problem as

$$C(c_W, a_M = 1) := \min_{(v_W, v_M)} \mathbb{E}\left[ \varphi_M(v_M(y)) + \varphi_W(v_W(y, e)) \right| a_M = 1, c_W, \sigma^*]$$

subject to $(IR_M), (IR_W), (IC_M), (IC_W), (EQ - selection), (MON)$, and $v_W(y, g) \geq v_W(y, b)$ for all $y \in Y$.

The constraint requiring worker’s payments to be higher when the evaluation is good assures that the report $g$ has the desired good performance meaning, and the report $b$ has the poor performance connotation. The constraint assures the report $g$ is associated with the highest payment and $b$ with the lowest for each output realization $y \in Y$.

The analysis proceeds in the following way: first, I solve a relaxed problem without the equilibrium selection, the monotonicity, and the $v_W(y, g) \geq v_W(y, b)$ constraints. Second, I show that the relaxed problem solution is monotonic, and that it satisfies $v_W(y, g) \geq v_W(y, b)$ for all $y \in Y$. Third, I show that it satisfies the equilibrium selection constraint.

Note that $p$ and $q$ fully characterize $\sigma^*$, which has two important implications: first, we can treat the information structure of the signals observed by the principal as exogenous. That is, the principal observes $(y, e)$ which are distributed according to $p$, $q$, and $\sigma^*$, and the latter is characterized by the former two. Second, we can solve the optimal compensation problem for the manager and worker separately. The objective function is additively separable, and the constraints are unrelated after we fix effort choices and $\sigma^*$.

Let’s begin with the manager’s compensation, but we first rename variables to transform the problem in a canonical moral hazard problem. Define,

$$p^{cw}(y|a_M) := p(y|a_M, 1)G(c_W) + (1 - G(c_W))p(y|a_M, 0),$$

as the probability of output $y$ when effort levels are given by $c_W$ and $a_M$. Let,

$$s^M(y) := \frac{p^{cw}(y|1) - p^{cw}(y|0)}{p^{cw}(y|1)}.$$

be the manager’s score when output is $y$ and the worker is exerting effort given by the cutoff $c_W$. As in traditional moral hazard problems, the score is a sufficient statistic for manager’s compensation.
We can then write manager’s compensation minimization problem as

\[
\min_{v_M} \sum_y \Phi_M(v_M(y)) p^{cw}(y|1) \tag{2}
\]

subject to

\[
\sum_y v_M(y) p^{cw}(y|1) - c_M \geq \bar{u}_M \tag{IR_M}
\]

\[
\sum_y v_M(y) s^M(y) p^{cw}(y|1) \geq c_M. \tag{IC_M}
\]

The problem above is a standard moral hazard problem. The objective function is convex, and the constraints are linear. We can solve it by minimizing pointwise.

**Lemma 3.** The solution to problem \([2]\) is given by

\[
\Phi'_M(v^*_M(y)) = \lambda^M + \mu^M s^M(y), \tag{3}
\]

where \(\lambda^M > 0\) and \(\mu^M > 0\) are the respective dual multipliers associated with \((IR_M)\) and \((IC_M)\).

The solution takes the familiar Holmström–Mirrlees contract-form, in which the score is a sufficient statistic for payments. I now solve the worker’s compensation problem.

Similarly to the manager’s case, I rename variables to write the problem in the usual way. First, let \(f^{cw}(y,e)\) be the probability of observing \((y,e)\) given realized cost cutoff \(c_w\), high managerial effort and evaluation strategy \(\sigma^*\). That is,

\[
f^{cw}(y,b) := G(c_w)p(y|1,1)Q(z^*(y)|1) + [1 - G(c_w)]p(y|1,0)Q(z^*(y)|0)
\]

\[
f^{cw}(y,g) := G(c_w)p(y|1,1)[1 - Q(z^*(y)|1)] + [1 - G(c_w)]p(y|1,0)[1 - Q(z^*(y)|0)],
\]

where \(Q\) denotes \(z\)’s cumulative distribution.

Let the worker’s effort score for each signal pair \((y,e)\) with strictly positive \(f^{cw}(y,e)\) be

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

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

Then, we can write the worker’s compensation problem as

\[
\min_{v_W} \sum_y \left[ \Phi_W(v_W(y,g)) f^{cw}(y,g) + \Phi_W(v_W(y,b)) f^{cw}(y,b) \right] \tag{4}
\]
subject to
\[ \sum_Y \left[ v_W(y, g) f_W^c(y, g) + v_W(y, b) f_W^c(y, b) \right] - \mathbb{E}[c | c \leq c_W] \geq \bar{u}_W \] \hspace{1cm} (IR_W)

\[ \sum_Y \left[ v_W(y, g) s_W^c(y, g) + v_W(y, b) s_W^c(y, b) \right] = c_W. \] \hspace{1cm} (IC_W)

Once again, I minimize pointwise.

**Lemma 4.** The solution to problem (4) is given by
\[ \phi'_W(v^*_W(y, i)) = \lambda^W + \mu^W s_W(y, i) \quad \text{for all } y \in Y \text{ and } i \in \{b, g\}, \] \hspace{1cm} (5)

where \( \lambda^W \) and \( \mu^W \) are the respective dual multipliers associated with (IR_W) and (IC_W). Also, \( \lambda^W \) and \( \mu^W \) are both strictly positive.

When I fixed \( \sigma^* \) as the evaluation strategy, I implicitly assumed that the message \( g \) was associated with higher worker’s payment, which I now check. Note that \( \mu^W > 0 \) and \( s_W(y, g) \geq s_W(y, b) \) implies \( v_W(y, g) \geq v_W(y, b) \), which assures that the performance report \( g \) generates higher payments than \( b \).

The next step is assuring that the solution to manager’s compensation relaxed problem satisfies the relaxed monotonicity constraint. That is, \( v^*_M \) is increasing. Note that the solutions to both relaxed problems — \( v^*_M \) and \( v^*_W \) — are increasing in \( y \) if and only if \( s^M \) and \( s^W \) are increasing in \( y \). I, then, assume that \( p \) and \( q \) are such that the scores are increasing. As the scores are fully characterized by the \( p \) and \( q \), assuming that they are increasing is an assumption on primitives.

**Assumption 4.** \( p \) and \( q \) are such that \( s^M(\cdot), s^W(\cdot, b) \) and \( s^W(\cdot, g) \) are increasing.

By fixing \( \sigma^* \), I have imposed that the manager chooses the evaluation strategy that maximizes her payoff among equilibria with \( a_M = 1 \). I still have not shown that it is the best for the manager among all possible equilibria. In principle, there could be a better equilibrium with \( a_M = 0 \). Proposition 3 shows that the effort complementarity assumption assures there is no such better equilibrium. Hence, the (EQ – selection) constraint is satisfied and Lemmas 3 and 4 characterize the cost-minimizing contracts.

**Proposition 3.** The cost-minimizing compensation schemes are given by
\[ \phi'_M(v^*_M(y)) = \lambda^M + \mu^M s^M(y), \] \hspace{1cm} (6)
and
\[ \varphi'_W (v^*_W (y, i)) = \lambda^W + \mu^W s^W (y, i) \quad \text{for all } y \in Y \text{ and } i \in \{b, g\}. \quad (7) \]

Where \( \lambda^W, \mu^W, \lambda^M \) and \( \mu^M \) are characterized by \((IR_W), (IC_W), (IR_M)\) and \((IC_M)\).

Most of the work in proving Proposition 3 has already been done. The last step is to assure that given the contracts characterized by Lemmas 3 and 4, the manager does not prefer an equilibrium in which she exerts low effort. The concern would be the worker exerting high effort more often in an equilibrium with low manager’s effort. The assumption of effort complementarity assures that such an equilibrium does not exist. In the equilibrium \((c^*_W, 1, \sigma^*)\), the manager is indifferent between working and shirking. By effort complementarity, the manager strictly prefers to work whenever \( \hat{c}^*_W > c^*_W \). Hence, the low manager’s effort cannot be a best response to a higher worker’s effort cost cutoff. Therefore, the equilibrium selection constraint is slack. In the last section, I present an extension discussing what would happen with substitute efforts. Due to equilibrium selection concerns, the principal needs to provide stronger incentives for the manager to exert effort. The \((IC_M)\) becomes the slack constraint.

Below, I plot an 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 \( z < z^*(y) < \bar{z} \). Hence, there is informational content on the performance evaluation. The principal delegates to the manager the decision about the worker’s compensation. The manager chooses at each \( y \) whether the worker gets the wage described by the square (blue) curve or the circle (red) curve. Relating to our running sales example, the manager decides whether the manager gets a bonus beyond the sales commission or not.

Note also, that give the optimal contracts, the manager’s favorite equilibrium is the only one consistent with van Damme (1989)’s forward induction definition. Any other equilibrium generates a payoff strictly lower than the manager’s outside option.
Note that in this example, there are only one of the curves for extreme values of output. This feature arises whenever the manager’s report becomes invariant to her private signal realization for sufficiently extreme output levels. For example, suppose that for a given output realization \( \hat{y} \) we have \( q(z|1)p(\hat{y}|1,1) < q(z|0)p(\hat{y}|1,0) \). Then, whenever \( y = y \), the manager reports a bad performance, regardless of the private signal realization. Also, the same holds for any \( y < \hat{y} \). Similarly, if for a given \( \tilde{y} \) we have \( q(z|1)p(\tilde{y}|1,1) > q(z|0)p(\tilde{y}|1,0) \), then the report is good for any \( y \geq \tilde{y} \), regardless of the private signal.

### 3.6 When are subjective performance evaluations valuable?

Subjective performance evaluations are present in a wide range of 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. That is, although \( z \) is informative about \( a_W \), the evaluation \( e \) does not generate any additional information beyond what is already conveyed by \( y \). Output might be a sufficient statistic for \( (y,e) \).

I first define what an objective and a subjective compensation system are and what it means for subjective evaluations to be valuable.

**Definition 3.** I say that a contract \( \pi_W \) is objective at \( y \), if \( \pi_W \) does not depend on the performance
evaluation. That is, \( \pi_W(y,e) = \pi_W(y,\hat{e}) \) for all \( e, \hat{e} \in E \).

Definition 4. I say that a contract \( \pi_W \) is objective if it is objective at every \( y \in Y \).

Definition 5. Fix the effort levels the principal wants to implement \( (c_W,a_M=1) \) with \( c_W > 0 \). I say that subjective performance evaluations are valuable if there exists subjective contracts such that the implementation cost is strictly lower than the implementation cost under any objective contract.

Proposition 4. Subjective performance evaluations are not valuable if and only if for every \( y \in Y \)

\[
q(z|1)p(y|1,1) \leq q(z|0)p(y|1,0) \text{ or } q(z|1)p(y|1,1) \geq q(z|0)p(y|1,0).
\] (8)

By the informativeness principle ([Holmström (1979)]), subjective performance evaluations are valuable if and only if they bring additional information about \( a_W \) beyond what is conveyed by \( y \).

Suppose (8) holds. Hence, for each given realized \( y \) the performance evaluation is always \( b \) or always \( g \), regardless of \( z \)'s realization. Therefore, \( y \) is a sufficient statistic for \( (y,e) \). That is, the subjective performance evaluation does not bring any additional information. Hence, it is not valuable.

If (8) does not hold, then there exists a \( y \) such that \( z^*(y,1) \in (z,\bar{z}) \). That is, the performance evaluation informs whether \( z \geq z^*(y,1) \) or not. Hence, \( y \) is not a sufficient statistic for \( (y,e) \) and subjective performance evaluations are valuable. This proves Proposition 4.

Proposition 4 states that if output generates a sufficiently dispersed likelihood ratio, the manager's performance evaluation is invariant to \( z \)'s realization. If each output realization is very informative about the worker's effort, the manager only considers output when evaluating the worker. Then, the performance evaluation does not convey any information beyond what is conveyed by \( y \).

This result implies that the principal uses subjective performance evaluations when the non-verifiable signal is sufficiently informative about the worker's effort compared to the verifiable measure. If the verifiable measure is sufficiently informative (in the sense that it generates disperse likelihood ratios), the principal does not benefit from introducing subjective performance evaluations. However, there is scope for beneficial subjective performance evaluations if the verifiable measure is imprecise relative to the non-verifiable one.
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 affect the information flow. A manager might, for example, better monitor her subordinate in an open space office plan. This section extends the 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 choose the distribution of the manager’s private signal $z$, 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 4. Suppose the principal can choose any distribution for the manager’s private signal $z$ and wishes to implement $(c_W, a_M = 1)$. The principal prefers the fully-informed manager; manager’s contract is characterized by (3) and worker’s payments are given by

$$v^*_W(y, b) = \bar{u}_W + E[c|c \leq c_W] - G(c_W)c_W \quad \text{and} \quad v^*_W(y, b) = v^*_W(y, g) + c_W.$$ 

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 $a_W$ directly, that is, she reports $g$ if $a_W = 1$ and $b$ if $a_W = 0$. Thus, under full information, the principal’s and the manager’s interests are fully aligned. The principal can act as if she observed effort directly. $(IC_W)$ 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 under a non-fully-informative signal. The principal would like the report to reveal all information possible about $a_W$. While the manager’s reporting strategy pools all signals with a positive score at the highest paying message and all signals with a negative score at the lowest paying message. I now address the question of whether reducing transparency could help aligning 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 instead of fully observing \( z \)?

I now 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 perfectly observes the realized \( z \). Otherwise, the principal can choose any other partition of \( [z, \bar{z}] \). For instance, take arbitrary \( \hat{z}, \check{z} \in (z, \bar{z}) \). The principal could choose a partition \( \{ [z, \hat{z}]; (\hat{z}, \check{z}); [\check{z}, \bar{z}] \} \). That is, the manager observes only in which subset of the partition the realized \( z \) is, but not the realization of \( z \) itself. Graphically, we could represent this partition as

\[
\begin{array}{c}
\hat{z} \\
X_0 \quad X_1 \quad \check{z} \\
\check{z} \\
\end{array}
\]

where the manager would see only the realization of \( x \), but not \( z \).

Take primitive distributions \( p \) and \( q \) satisfying Assumptions 1 and 4, 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 realization of output under full-transparency.

**Assumption 5.** \( p \) and \( q \) jointly satisfy:

1. \( q(\cdot|a_W) \) is continuous for any \( a_W \in \{0, 1\} \).
2. \( z^*(y) \in (\hat{z}, \check{z}) \) for all \( y \in Y \).

Note that both \( z^* \) is fully determined by the primitive distributions \( p \) and \( q \). Hence, Assumption 5 can be stated directly in terms of the distributions as well. The first item is assumed for technical convenience. The second assures that under full-transparency subjective evaluations are valuable for any output realization. If subjective performance evaluations were not valuable even under full-transparency, then they would not be valuable if the manager had less information. Assumption 5 rules out those uninteresting cases in which there is no scope for subjective performance evaluations.
Proposition 5. Full-transparency is not optimal. In particular, the principal strictly benefits from pooling extreme signals. That is, there exists $\tilde{z}$ and $\check{z}$ such that the cost of enforcing $(c_W, a_M = 1)$ is strictly lower if the manager’s information is given by the partition $\left\{ [z, \tilde{z}]; \{z\} \in (\tilde{z}, \check{z}); [\check{z}, z] \right\}$ instead of the finest partition $\left\{ \{z\} \in \{z\} \right\}$.

Proposition 5 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 reason why the principal benefits from reducing transparency is that the manager and 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 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.

When proving this result, the first step is to notice that in comparison to the full transparency case, there is no loss in pooling extreme private signals. The fully-transparent partition and the partition below generate the same performance reports.

![Partition Diagram]

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, while 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 by increasing the first cut-off by a small amount and decreasing the second by also a small amount the principal is strictly better off. 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 distributions 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 effect of $\varepsilon$ 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 $\varepsilon$ has three effects: first, there is a direct effect. For fixed contracts, a change in $\varepsilon$ changes the distribution of signals and expected payments. Second, after changing $\varepsilon$, the principal must adjust payments such that the worker is still willing to participate. Third, the principal must adjust payments such that $c_W$ is still implemented as worker’s effort cost cutoff. I construct $\kappa$ such that the second effect is zero. As the manager chooses the evaluation strategy that to maximizes $c_W$, the third effect is second-order at $\varepsilon = 0$. Hence, only the direct effect remains.

Note that the direct effect of an increase in $\varepsilon$ is to decrease the probability of a good report when output is high ($y_n$), and to increase the probability of a bad report when output is low ($y_0$). Comparing how it impacts 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 Extensions:

This section extends the analysis to different model specifications and discusses the role of a few assumptions. In all extensions, the central trade-off remains as the manager benefiting from higher worker’s effort but not internalizing the compensation costs. Consequently, regardless of the performance rating scale, the manager only uses the extreme performance reports. Hence, binary performance ratings remain optimal in all extensions explored.
5.1 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 performance ratings result is being driven by the binary effort assumption. In this subsection, I show that the relation is not so direct.

Let effort choices be continuous, $a_M, a_W \in [0,1)$. As before denote by $p(\cdot | a_M, a_W)$ and $q(\cdot | a_W)$ the distribution of each of the two signals. I assume that $p$ and $q$ are twice continuously differentiable. Denote by $p_M(\cdot | a_M, a_W), p_W(\cdot | a_M, a_W)$ and $q_W(\cdot | a_W)$ the partial derivatives with respect to $a_M$ and $a_W$. I assume $p$ and $q$ satisfy the following monotone likelihood ratio properties.

Assumption 6. $p$ and $q$ are such that

1. $p_M(y | a_M, a_W) \neq p(y | a_M, a_W)$ is strictly increasing in $y$ for all $(a_M, a_W) \in (0,1)^2$.
2. $p_W(y | a_M, a_W) \neq p(y | a_M, a_W)$ is strictly increasing in $y$ for all $(a_M, a_W) \in (0,1)^2$.
3. $q_W(z | a_W) \neq q(z | a_W)$ is bounded and strictly increasing in $z$ for all $(a_M, a_W) \in (0,1)^2$.

As usual in moral hazard problems with continuous efforts, I need convexity assumptions over the distribution of signals to assure 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 7. $p$ and $q$ are such that

1. $p$ is linear in $a_W$.
2. $Q(z | \cdot) p(y | a_M, \cdot)$ is convex for any $z \in Z$, $y \in Y$ and $a_M \in [0,1]$.

As efforts are now continuous, we already have a notion of effort intensity and we do not need to assume a stochastic worker’s effort cost. Let $c_i : [0,1) \rightarrow \mathbb{R}_+$ denote agent’s $i \in \{M, W\}$ effort cost. I assume that both $c_i$’s are common knowledge, strictly increasing and strictly convex. To avoid concerns about the corner, assume $\lim_{a_i \rightarrow 1} c'_i(a_i) = +\infty$.

Take any given contracts $(v_M, v_W, E)$, where $v_M : Y \rightarrow \mathbb{R}$ and $v_W : Y \times E \rightarrow \mathbb{R}$ are non-constant. For given effort levels and performance evaluation strategy, the manager’s expected utility is given
by
\[ V_M(a_W, a_M, v_M) := \sum_y v_M(y)p(y|a_M, a_W) - c_M(a_M), \]
while the worker’s expected utility is given by
\[ V_W(a_W, a_M, v_W, \sigma) := \int_Z \sum_y \sum_e v_W(y, e) \sigma(e|a_M, y, z)p(y|a_M, a_W)q(z|a_W)dz - c_W(a_W). \]

Note that the manager’s expected payoff is still increasing in the worker’s effort. That is,
\[ \frac{\partial V_M}{\partial a_W}(a_W, a_M, v_M) = \sum_y v_M(y)p_W(y|a_M, a_W) > 0. \]

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

**Proposition 6.** For any given non-constant \( v_W \) and \( \hat{a}_M \), the evaluation strategy that maximizes the worker’s effort uses only the highest and the lowest-paying messages with positive probability.

The proof proceeds as follows: I start with an arbitrary \( \sigma \) that with strictly positive probability uses messages that are not the highest nor the lowest-paying ones. Then, I construct a \( \tilde{\sigma} \) only uses the highest and lowest paying messages and increases the worker’s effort. Suppose the manager chooses a given effort level \( \hat{a}_M \) and uses an evaluation strategy \( \hat{\sigma} \) as described before. Denote the worker’s best response to \( \hat{a}_M \) and \( \hat{\sigma} \) by \( \hat{a}_W \). It must satisfy the following first-order condition
\[ \frac{\partial V_W}{\partial a_W}(\hat{a}_W, \hat{a}_M, v_W, \hat{\sigma}) = 0. \]

Define \( \hat{z} : Y \times [0, 1]^2 \to Z \) such that
\[ \frac{p_W(y|a_M, a_W)}{p(y|a_M, a_W)} = -\frac{q_W(\hat{z}(y, a_M, a_W)|a_W)}{q(\hat{z}(y, a_M, a_W)|a_W)}. \]

Let \( \tilde{\sigma} \) be
\[ \tilde{\sigma}(a_M, y, z) := \begin{cases} \delta_{\hat{e}_y} & \text{if } z < \hat{z}(y, a_M, a_W), \\ \delta_{\hat{e}_y} & \text{otherwise.} \end{cases} \]

Remember that \( \hat{e}_y \) is the lowest-paying and \( \hat{e}_y \) the highest-paying message when output is \( y \). One can check that
\[ \frac{\partial V_W}{\partial a_W}(\hat{a}_W, \hat{a}_M, v_W, \hat{\sigma}) > \frac{\partial V_W}{\partial a_W}(\hat{a}_W, \hat{a}_M, v_W, \tilde{\sigma}) = 0. \]

\footnote{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 6.}
Hence, if the manager uses $\tilde{\sigma}$ instead of $\hat{\sigma}$, the worker has the incentive to locally increase his effort from $\tilde{a}_W$. Then, I show that with $\tilde{\sigma}$ the worker’s payoff is strictly concave in $a_W$, which implies that the best response to $\tilde{\sigma}$ cannot be lower than $\tilde{a}_W$. Assumption 7 and $c''_W(\cdot) > 0$ imply such strict concavity. Finally, I show that existence of an effort maximizing evaluation policy.

### 5.2 Multiple workers

I extend the analysis to multiple workers. I assume there are $I \in \mathbb{N}$ workers and one manager. Each worker is denoted by $i \in \{1, \ldots, I\}$. As before, each worker and the manager decide whether or not to accept the contract, whether or not to exert effort, and the manager decides how to evaluate workers. If any worker or the manager rejects the contract, everyone gets their outside option $\bar{u}_M$ and $\bar{u}_iW$. If they all accept the contracts, then the manager decides whether or not to exert effort $a_M \in \{0, 1\}$ at a private cost $c_M$. Each worker $i$ draws an effort cost $c^i$ from a distribution $G^i$ with full support on $\mathbb{R}_+$. Knowing their own effort cost, each worker decides whether to exert effort or not.

Given workers’ and manager’s effort choices, $2I$ signals are realized: each worker’s output $y^i$ and managerial subjective perception of each worker’s effort $z^i$. As before, each output is distributed according to $p^i(y^i|a_M, a^i_W) \in [0, 1]$. That is, given $a^i_W$ and $a_M$, the probability of a given realization $y^i$ is denoted by $p^i(y^i|a_M, a^i_W) > 0$. Each managerial subjective perception $z^i$ is non-verifiable, privately observed by the manager and distributed according to $q^i(z^i|a^i_W)$. That is, given effort choice $a^i_W$, the distribution of $z^i$ is described by a probability density function $q^i(z^i|a^i_W)$. I also assume $q^i$’s are uniformly bounded away from zero and have bounded likelihood ratio. I assume Assumptions [1] and [3] hold for each $y^i$ and $z^i$.

Contracts are a finite set of performance ratings $E^i$ for each worker $i \in \{1, \ldots, I\}$, payments for the manager $\pi_M: \times_{i=1}^I (Y^i \times E^i) \rightarrow \mathbb{R}_+$, and payments for each worker $\pi^i_W: \times_{i=1}^I (Y^i \times E^i) \rightarrow \mathbb{R}_+$, where $\times_{i=1}^I (Y^i \times E^i)$ denotes the cross product of $(Y^i \times E^i)$ for each $i \in \{1, \ldots, I\}$.

---

27The imposed contract space is not without loss. It implies that worker’s $i$ evaluation only affects worker’s $i$ compensation. The restriction can be motivated to avoid issues with sabotage (it creates incentives for workers to undermine their peers) and favoritism (the manager being biased towards one worker in comparison to others). For instance, [Roch et al. (2007)] documents workers perceiving absolute performance measures (taking pre-defined standards) as fairer than relative (in comparison to peers). Another example is how the use of stacked rankings is pointed as one of the causes for Microsoft poor performance in the 2000s. See Eichenwald, “Microsoft’s Lost Decade”, *Vanity Fair*, July 24, 2012.
Proposition 7. The cost minimizing contracts are given by

Manager’s preferences are unchanged, and each worker’s preferences are described by

\[ U_W^i(\pi_W, y, e, c_W^i, a_W) := u_W^i(\pi_W(y, e)) - c_W^i a_W \] for \( i \in \{1, \ldots, l\} \).

Denote by \( \bar{y}^i \in Y^l := \times_{i=1}^l Y^i \) a vector of output realizations, and by \( \bar{e}^i \in E^l := \times_{i=1}^l E^i \) a vector of performance reports. Principal’s payoff is given by

\[ U_P(\bar{y}^i, \bar{e}^i) := \sum_{i=1}^l \left[ y^i - \pi_W(\bar{y}^i, \bar{e}^i) \right] - \pi_M(\bar{y}^i, \bar{e}^i). \]

The principal wishes to implement high manager’s effort and a vector \((c_W^1, \ldots, c_W^l) \gg 0\) of effort cost cutoffs for the workers at the cheapest possible cost. One can replicate the analysis for each worker \( i \) separately. First, note that the manager’s compensation must not depend on her report about any worker’s performance. Second, the manager benefits from each worker’s higher effort. Third, note that one worker’s compensation and effort choice does not affect other worker’s incentives. For given contracts, manager’s effort \( a^M = 1 \) and performance evaluation strategy \( \sigma \), each worker’s effort cost cutoff is given by

\[ c_W^i(y^i, 1, 1, \sigma) = \mathbb{E}\left[v_W(y^i, e^i) | a_W = 1, a^M = 1, \sigma\right] - \mathbb{E}\left[v_W(y^i, e^i) | a_W = 0, a^M = 1, \sigma\right] \]

\[ = \int_{\mathcal{Y}^l} \int_{\mathcal{E}^l} v_W(y^i, e^i) \sigma(e^i | a_W = 1, \bar{y}^i, \bar{e}^i) \left[q^i(z^i | 1)p^i(y^i | 1, 1) - q^i(z^i | 0)p^i(y^i | 1, 0)\right] dz^1 \ldots dz^l. \]

Let \( \sigma^{i*} \in \text{argmax}_{\sigma \in \Delta(E^l)} \{c_W^i(y^i, 1, 1, \sigma)\} \). Which is given by the manager reporting the highest paying message if \( q^i(z^i | 1)p^i(y^i | 1, 1) \geq q^i(z^i | 0)p^i(y^i | 1, 0) \) and the lowest otherwise.

I can now define the scores in the multiple workers problem. Each worker’s score remains the same as before (one for each \( i \), defined using \( \sigma^{i*} \)). The only change occurs in the manager’s score, which is now a function of the entire output vector \( \bar{y}^i \). With a slight abuse of notation, let

\[ s^M(\bar{y}^i) := 1 - \frac{\Pi_{i=1}^l \left[ p^i(y^i | 0, 1)G^i(c_W^i) + (1 - G^i(c_W^i)) p^i(y^i | 0, 0) \right]}{\Pi_{i=1}^l \left[ p^i(y^i | 1, 1)G^i(c_W^i) + (1 - G^i(c_W^i)) p^i(y^i | 1, 0) \right]}. \]

One can replicate the analysis done in previous sections and characterize the optimal contracts. The following proposition states the result.

**Proposition 7.** The cost minimizing contracts are given by
• **Manager’s compensation:**

\[
\phi'_{ \lambda_M}(v_M(\vec{y})) = \lambda_M + \mu_M s^M(\vec{y}) \quad \text{for all } \vec{y} \in \prod_{i=1}^{I} Y^i,
\]

(9)

where \( \lambda_M, \mu_M \) are the respective multipliers associated with \((IR_M)\) and \((IC_M)\).

• **Workers’ compensation:**

\[
\phi'_{ \lambda_W}(v_W(y^i, e^i)) = \lambda_W + \mu_W s^W(y^i, e^i) \quad \text{for all } y^i \in Y^i, \quad e^i \in \{b, g\}, \quad \text{and } i \in \{1, \ldots, I\},
\]

(10)

where \( \lambda_W, \mu_W \) are the respective multipliers associated with \((IR^i_W)\) and \((IC^i_W)\).

5.3 **Substitute efforts**

The analysis so far assumed complementary efforts. Although being a natural assumption, it raises the question of what would happen in the case of substitute efforts. I now impose the assumption of effort substitutability instead of complementarity.

**Definition 6.** I say that **efforts are strong substitutes** if for any non-constant and increasing \( b : Y \to \mathbb{R} \),

\[
\sum_Y b(y) \left[ p(y|1,1) - p(y|0,1) \right] < \sum_Y b(y) \left[ p(y|1,0) - p(y|0,0) \right].
\]

Efforts are strong substitutes if for any increasing reward function, one agent’s higher effort strictly increases the incentives for the other agent to exert effort. The strengthening requirement eliminates the case in which efforts are both substitutes and complements. That is, it eliminates the case in which \[ p(y|1,1) - p(y|0,1) = p(y|1,0) - p(y|0,0) \]. For such case, all previous results directly apply.

**Assumption 8.** For all \( y \in Y \), \( P(y|\cdot,\cdot) \) is supermodular. That is, for all \( y \in Y \)

\[
P(y|1,1) - P(y|0,1) \geq P(y|1,0) - P(y|0,0),
\]

and the inequality is strict for at least one \( y \in Y \).

Similarly to Proposition 1, one can show that Assumption 8 implies strong effort substitutability. Throughout this section, I maintain Assumptions 1, 2, 4 and 8.
The analysis remains unaltered up to Proposition 2. In particular, the result that binary performance evaluations achieves the best the principal can hope for, remains valid. The assumption of effort complementarity is only used to show that the solution to the relaxed problem satisfies the equilibrium selection constraint. The concern with equilibrium selection was the manager preferring an equilibrium in which she exerted low effort, and in response, the worker would exert an even higher effort. The assumption of effort complementarity ruled that out, while effort substitutability does not.

Under effort substitutability, the approach of relaxing \((EQ - selection)\) and imposing \((IC_M)\) does not work. When \((IC_M)\) binds in the equilibrium with high managerial effort, and efforts are substitutes, the manager might strictly prefer an equilibrium in which she exerts low effort. By exerting lower effort, the manager might increase the incentive for the worker to exert higher effort and benefit from that. Formally, Lemma 8 (in the proof of Proposition 3) does not hold with substitute efforts.

Therefore, under the assumption of substitute efforts, the manager receives steep incentives, such that \((IC_M)\) is slack but \((EQ - selection)\) binds. The equilibrium selection constraint, however, is difficult to deal with. For instance, it is also affected by \(v_W\) directly, which implies that we would not be able to solve manager and worker problems separately.

I do not characterize the optimal contracts, but I provide an upper bound to the cost minimization problem. The issue the principal faces under substitute efforts is that the manager could prefer not to exert effort to generate incentives for even higher effort from the worker. To circumvent that, the principal can offer sufficiently high-powered incentives to the manager, such that she would be willing to exert high effort even when she is sure that the worker exerts high effort. This approach assures that there only exits equilibria with high managerial effort.

The manager’s problem would now be given by

\[
\min_{v_M(y)} \sum_{Y} \phi_M(v_M(y)) p^W(y|1)
\]

subject to

\[
\sum_{Y} v_M(y) p^W(y|1) - c_M \geq \bar{u}_M \quad (IR_M)
\]

\[
\sum_{Y} \frac{v_M(y) p(y|1,1) - p(y|0,1)}{p^W(y|1)} p^W(y|1) \geq c_M, \quad (Strong - IC_M)
\]

---

28 The rationale is similar to moral hazard problems with unique implementation. It is, however, not unique implementation because there is still multiplicity regarding the evaluation strategy.
where \((IR_M)\) denotes manager’s participation constraint, and \((Strong - IC_M)\) denotes the constraint that assures high effort as a dominant strategy for the manager.

The solution to (11) is given by the Holmström-Mirrlees type of contracts

\[
\varphi'_M(\hat{v}_M(y)) = \hat{\lambda}^M + \hat{\mu}^M \frac{p(y|1,1) - p(y|0,1)}{p^w(y|1)},
\]

where \(\hat{\lambda}^M\) and \(\hat{\mu}^M\) are the respective dual multipliers associated with \((IR_M)\) and \((Strong - IC_M)\). The worker’s problem remains unaltered.

As contracts \(\hat{v}_M\) are constructed such that there is no equilibrium with \(a_M < 1\), the equilibrium selection constraint is satisfied. Hence, the contracts characterized by (5) and (12) satisfy all constraints of the original problem. Providing an upper-bound for the cost of enforcing \(a_M = 1\) and \(c^W\). It is also interesting to notice that \((IC_M)\) is slack. That is, the manager faces very steep incentives, to prevent her from slacking off and dumping the workload on the worker.

### 5.4 Screening

Throughout the analysis, I did not allow the principal to screen the worker regarding costs. Communication between worker and principal was ruled out and as the cost of effort is realized after the contract is signed, there would be no way to screen. In this section, I resume to the baseline version where Assumptions 1-2 hold, but now I allow for the possibility of screening. In particular, I show that the binary performance rating system remains optimal.

There are two main changes from the baseline model: first, the principal can offer a menu of contracts to the worker. Second, I invert the timing of signing the contract and the cost realization. That is, now the worker first observes his cost and then decides which contract to accept. The worker’s contract choice is assumed to be observed by the manager. The new timing is
At any point $t$, the worker observes $c$.

At $t=1$, the principal offers a menu of contracts.

At $t=2$, the manager and the worker choose contracts.

At $t=3$, the manager observes the worker’s contract choice and decides whether to leave.

At $t=4$, with probability $\varepsilon > 0$, the worker’s effort cost is re-drawn from $G$.

At $t=5$, the manager chooses $a_M$ and the worker chooses $a_W$.

At $t=6$, the manager observes $y$ and $z$.

At $t=7$, the manager chooses a performance evaluation $e \in E$.

At $t=8$, payments are realized.

A contract is a triple $(E, v_W, v_M)$ where $E$ is a finite set, $v_W : Y \times E \to \mathbb{R}_+$ and $v_M : Y \times E \to \mathbb{R}_+$. A menu of contracts is a collection of contracts. The principal now may screen the worker, that is, she offers a set of choices from which the worker chooses his favorite. For each worker’s contract choice, the manager and the worker still play an extensive form game in which each of them chooses effort, and after the realization of $y$ and $z$, the manager chooses a performance evaluation report.

The role of re-drawing the cost with probability $\varepsilon$ is to guarantee that the manager is never sure of how strong the incentives must be to generate high worker’s effort. Without the re-drawing, the manager could potentially fully learn worker’s cost. If that’s the case, there would be an amount of incentive that would be enough for worker’s high effort, and the manager would be indifferent between providing additional incentives or not. With the re-drawn probability, regardless of how small, worker’s cost can always be higher, and extreme performance evaluations strictly benefit the manager. As $\varepsilon$‘s role is only to break indifferences, it could be arbitrarily small. In Appendix C, I characterize minimal implementation costs when $\varepsilon$ tends to zero.

Note that as before, if the manager’s payments depend on her evaluation report she uses only messages that assure her the highest payment. That is, Lemma 1 still holds. We can restrict attention to menus such that for every contract in the menu, the manager’s compensation is independent of her performance report.

Fix a contract menu $\left\{(E^k, v_W^k, v_M^k)\right\}_{k \in K}$ where contracts are indexed by $k \in K \subseteq \mathbb{R}_+$. Fix a contract choice out of this menu. That is, the worker has chosen contract $\tilde{k}$. Given the contract $\tilde{k}$ and the worker’s equilibrium conjecture of the manager’s actions (effort and performance evaluation), the worker chooses high effort if his effort cost is smaller than the expected gains from effort. The manager observes the worker’s contract choice and updates her belief about the worker’s effort.
cost. Denote by $\tilde{\zeta}_k$ the manager’s belief about worker’s effort given contract choice $\tilde{k}$. That is, after observing contract choice $\tilde{k}$, the manager believes that the worker exerts high effort with probability $\tilde{\zeta}_k$.

Note that $\tilde{\zeta}_k < 1$, for any contract and any equilibrium. As the worker’s effort cost is re-drawn from $G$ with probability $\varepsilon > 0$, and $G$ has full support on $\mathbb{R}_+$, there is always a strictly positive probability that the worker’s effort cost is higher than the expected gains from effort. For each given contract choice, we still have the issue of multiple equilibria. As before, I show that the manager’s preferred equilibrium is the one that has the highest probability of worker’s effort, namely the highest $\tilde{\zeta}_k$. For a given $k$, the manager’s expected payoff is given by

$$\sum_Y v^k_M(y) [p(y|a_M, 1) \tilde{\zeta}_k + p(y|a_M, 0)(1 - \tilde{\zeta}_k)] - a_M c_M$$

$$= \sum_Y v^k_M(y) p(y|a_M, 0) + \tilde{\zeta}_k \sum_Y v^k_M(y) [p(y|a_M, 1) - p(y|a_M, 0)] - a_M c_M.$$ 

Note that for each given $k$, the manager’s payoff is increasing in the probability of high effort. Hence, the manager benefits from maximizing worker’s high effort probability. Note that $\sigma^*$ maximizes such probability. Hence, the performance evaluation strategy $\sigma^*$ generates manager’s preferred equilibrium given contract choice $k$. One could replicate the arguments in Proposition 2 for each given contract and then conclude that the principal can restrict attention to binary performance evaluation contracts.

### 6 Conclusion

In most firms, the residual claimant is far removed from the rank-and-file. Relations between profit-maximizing principals and lower-ranked workers usually rely on intermediaries, such as mid-level managers. Those mid-level managers often have responsibilities that include both daily productive activities and subjective assessment of their subordinates’ performances.

In this paper, I analyze how managers’ incentives to perform relate to their incentives to evaluate their subordinates’ performances. Often, firms base their workers’ compensation on objective team measures such as output or sales, and on managerial subjective evaluations. This study — to the best of my knowledge — is the first to discuss an important agency friction in settings with subjective evaluations: the evaluator — here the manager — is not the residual claimant, but cares about the incentives a given evaluation strategy generates.
Middle managers are often under incentive contracts. For example, retail store managers are frequently paid based on sales. This type of compensation scheme generates the incentive for managers to motivate their subordinates to work hard. However, managers do not pay their subordinates from their own pocket. They want to motivate high effort from workers but do not internalize compensation costs. My model suggests that managers have the incentive to only use extreme performance ratings in order to generate the strongest possible incentives for their subordinates to exert effort.

I characterize the cost-minimizing contracts in a setting with a principal, a manager and a worker. In which, the manager has the dual role of exerting productive effort and evaluating worker’s performance. Optimal contracts take a familiar Holmström-Mirrlees form, in which subjective performance evaluation takes an appealing simple binary structure. Managers get paid only based on a verifiable measure (output), while worker’s compensation depends both on the verifiable outcome and a binary performance report from the manager. I also show that subjective performance evaluations are useful if and only if the non-verifiable information is sufficiently more informative (in a likelihood ratio sense) than the verifiable information. Finally, I show that the principal may benefit from reducing the information the manager has about the worker’s action.

The agency literature has focused mostly on the top (shareholders-CEO) or bottom (manager-worker) of the organization. Even when analyzing manager-worker relations, managers are usually assumed to be fully aligned with the firm’s interests. That is, they are assumed to be profit-maximizing. Most organizations, however, have a longer hierarchical structure, and understanding the incentives of mid-level managers seem to be of great importance and with a lot to be explored. This paper has focused on one aspect of that richer hierarchical structure: the incentives for accurate subjective performance evaluations.

References


## Appendix

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

**Lemma 5** (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 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 6.** 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 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 not constant, then the inequality is strict.

**Proof of Lemma 6**. 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 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. Now for the second part (strict inequality), let \( F \) and \( M \) be non-constant. As \( F \) is non-constant, it cannot be always equal to zero. Then, define two disjoint sets \( T^- := \{t \in T | F(t) < 0\} \) and \( T^+ := \{t \in T | F(t) \geq 0\} \). Define \( \tilde{M} : T \to \mathbb{R} \) as

\[
\tilde{M}(t) := \begin{cases} 
\max \{M(t)\} & \text{if } t \in T^- , \\
\min \{M(t)\} & \text{if } t \in T^+ .
\end{cases}
\]
Note that $\tilde{M}$ is an increasing function. Also, note that for all $t \in T$, we have $F(t)M(t) \geq F(t)\tilde{M}(t)$ and for some $t$ the inequality must be strict. Hence,

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

The first inequality is direct from the construction of $\tilde{M}$. The second comes from the first part of the lemma.

\[\square\]

**Proof of Proposition** Note that for any $b$

$$\sum_{y} b(y) \left[ p(y|1, 1) - p(y|0, 1) \right] \geq \sum_{y} b(y) \left[ p(y|1, 0) - p(y|0, 0) \right]$$

is equivalent to

$$\sum_{y} b(y) \left[ \frac{p(y|1, 1)}{2} + \frac{p(y|0, 0)}{2} \right] \geq \sum_{y} b(y) \left[ \frac{p(y|1, 0)}{2} + \frac{p(y|0, 1)}{2} \right].$$

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

$$\left[ \frac{p(.|1, 1)}{2} + \frac{p(.|0, 0)}{2} \right] \fo\left[ \frac{p(.|1, 0)}{2} + \frac{p(.|0, 1)}{2} \right].$$

Therefore, efforts are complements if and only if for all $y \in Y$

$$P(y|1, 1) + P(y|0, 0) \leq P(y|1, 0) + P(y|0, 1),$$

and the proof is concluded.  \[29\]

\[\square\]

**Proof of Lemma** Fix a contract $(E, \pi_W, \pi_M)$. 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 his payment. That is, if $\sigma(e|a_M, y, z) > 0$, then $\pi_M(y,e) = \max_{\hat{e} \in E} \{ \pi_M(y, \hat{e}) \}$. In fact, any report policy that satisfy this property can be part of an equilibrium.

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 $\tilde{E}_y := \arg\max_{e \in E} \{ \pi_M(y, \hat{e}) \}$ and $\tilde{E} := \cup_{y \in Y} \tilde{E}_y \subseteq E$. I now need to construct payments $\tilde{\pi}_M : Y \times \tilde{E} \to \mathbb{R}_+$ and $\tilde{\pi}_W : Y \times \tilde{E} \to \mathbb{R}_+$ that are outcome-equivalent to our original contract.

\[29\] I thank Humberto Moreira for great suggestions that simplified this proof.
Let $\tilde{\pi}_M(y, e) := \max_{\hat{e} \in \hat{E}} \{\pi_M(y, \hat{e})\}$, which does not depend on $e$. Let, for each $y \in Y$, $e_y$ be an arbitrary given element of $\hat{E}_y$. Define,

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

Note that the set of payments to the worker (among the ones the manager is willing to choose from) has not changed. We now show that both contracts 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 distributions. Let $(a_M, c_W, \sigma)$ be an equilibrium of the original contract. Note that if $\sigma(e|a_M, y, z) > 0$, then $e \in \hat{E}$, $\pi_M(y, e) = \tilde{\pi}_M(y, e)$ and $\pi_W(y, e) = \tilde{\pi}_W(y, e)$. Let $\tilde{\sigma}(e|a_M, y, z) = \sigma(e|a_M, y, z)$ for all $e \in \hat{E}_y$. Note that $\tilde{\sigma}$ generates the same payments distribution as $\sigma$. Hence, $(a_M, c_W, \tilde{\sigma})$ is an equilibrium of the alternative contract with the same output and payments distributions as $(a_M, c_W, \sigma)$.

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

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

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

\[\square\]

**Proof of Lemma 2** Manager’s payoff is given by:

$$V_M(c, a_M, v_M) = \sum_Y v_M(y)p(y|a_M, 0) + G(c) \sum_Y v_M(y)[p(y|a_M, 1) - p(y|a_M, 0)] - e_M,$$

which is strictly increasing in $c$ if and only if $\sum_Y v_M(y)[p(y|a_M, 1) - p(y|a_M, 0)] > 0$.

By Assumption 2, $v_M$ is increasing. By (MLRP), $[p(y|a_M, 1) - p(y|a_M, 0)]$ is never first strictly positive then strictly negative and $\sum_{y \in Y}[p(y|a_M, 1) - p(y|a_M, 0)] = 0$. By Lemma 6,

$$\sum_Y v_M(y)[p(y|a_M, 1) - p(y|a_M, 0)] > 0.$$  (13)
Proof of Proposition

Fix contracts \((E, v_W, v_M)\). We are going to show that any equilibrium is dominated (for the manager) by another one with binary performance evaluation. Fix the manager action \(a_M = 1\) that is part of an equilibrium \((\hat{\sigma}, 1, \hat{\sigma})\).

By Lemma 6, \(V_M(c^*, 1, v_M) \geq V_M(\hat{\sigma}, 1, v_M)\) if and only if \(c^* \geq \hat{\sigma}\). I now find \(c^* \geq \hat{\sigma}\) such that \((c^*, 1, \sigma^*) \in \Gamma(E, v_W, v_M)\).

The only way performance evaluations affect the manager’s payoff is through the worker’s effort. The evaluation strategy \(\sigma^*\) maximizes worker’s effort cutoff. That is, I maximize

\[
V_W(a_W = 1, a_M, v_W, \sigma) - V_W(a_W = 0, a_M, v_W, \sigma) = \int \sum_{Y} \sum_{\hat{Z}} v_W(y, e) \sigma(e|a_M, y, z) \left[q(z|1)p(y|a_M, 1) - q(z|0)p(y|a_M, 0)\right] dz,
\]

by choosing \(\sigma\). Note that \(\sigma^*\) maximizes the expression above.

Denote by \(c^*\), worker’s best response given \(\sigma^*\). Note that \(c^* \geq \hat{\sigma}\), hence \(V_M(c^*, 1, v_M) \geq V_M(\hat{\sigma}, 1, v_M)\).

I still need to show that \((c^*, 1, \sigma^*)\) is an equilibrium. By construction, \(c^*\) is the cost cutoff of worker’s best response to \(\sigma^*\) and to \(a_M = 1\). I now show that \(a_M = 1\) is manager’s best response to \(c^*\). Note that

\[
V_M(c^*, 1, v_M) - V_M(c^*, 0, v_M) = \sum_Y v_M(y) [p(y|1, 0) - p(y|0, 0)] + G(c^*) \sum_Y v_M(y) \left\{[p(y|1, 1) - p(y|0, 1)] - [p(y|1, 0) - p(y|0, 0)]\right\} - c_M \\
\geq \sum_Y v_M(y) [p(y|1, 0) - p(y|0, 0)] + G(\hat{\sigma}) \sum_Y v_M(y) \left\{[p(y|1, 1) - p(y|0, 1)] - [p(y|1, 0) - p(y|0, 0)]\right\} - c_M \\
= V_M(\hat{\sigma}, 1, v_M) - V_M(\hat{\sigma}, 0, v_M) \geq 0.
\]

The first inequality comes from effort complementarity and Lemma 6. Effort complementarity assures the curly bracket term is never strictly positive then strictly negative. An increasing \(v_M\) and Lemma 6 assure the term multiplying \(G(c^*)\) is positive. The second inequality comes from the fact that \((c, 1, \sigma)\) was an equilibrium. Hence, \(a_M = 1\) is a best response to \(c^*\). That is, \((c^*, 1, \sigma^*)\) is an equilibrium.

If \(\hat{\sigma}\) and \(\sigma^*\) generate different worker’s payment distribution, then \(c^* > \hat{\sigma}\) which implies

\[
V_M(c^*, 1, v_M) - V_M(c^*, 0, v_M) > V_M(\hat{\sigma}, 1, v_M) - V_M(\hat{\sigma}, 0, v_M) \geq 0.
\]
Proof of Lemma 3. We can write problem (2)’s Lagrangian as
\[ \mathcal{L}^M = \sum_Y \left[ \phi_M(v_M(y)) + \lambda^M(v_M(y) - c_M - \bar{a}_M) + \mu^M(v_M(y)s^M(y) - c^M) \right] p^{cw}(y|1). \] (14)

By minimizing pointwise, I get
\[ \phi'_M(v^*_M(y)) = \lambda^M + \mu^M s^M(y), \] (15)
where \( \lambda^M \) and \( \mu^M \) are the respective dual multipliers associated with \( (IR_M) \) and \( (IC_M) \).

Now I show that \( \lambda^M \) and \( \mu^M \) are strictly positive and \( v^*_M(y) \) is increasing. Suppose \( \mu^M = 0 \). Define \( \rho^M = \phi^{-1}_M \). Then, \( v_M(y) = \rho^M(\lambda^M) \). Which implies
\[ \sum_Y v_M(y)s^M(y)p^{cw}(y) = \rho^M(\lambda^M) \sum_Y s^M(y)p^{cw}(y) = 0 < c_M. \]
A violation of \( (IR_M) \). Hence, \( \mu^M > 0 \).

Suppose we perturb the optimal contracts by subtracting a small \( \epsilon > 0 \) to every \( v^*_M(y) \). The multipliers must be such that this change increases the value of the Lagrangian. That is, taking the first-order condition with respect to \( \epsilon \) and evaluating at \( \epsilon = 0 \)
\[ -\sum_Y \phi'_M(v_M(y))p^{cw}(y) + \lambda^M \geq 0. \]
As \( \phi'_M > 0, \lambda^M > 0. \)

Proof of Lemma 4. Analogous to the manager’s case.

Proof of Proposition 3. I prove Proposition 3 in two steps, divided into two lemmas.

Lemma 7. Let \( \hat{c}_W \) be the cutoff implemented in an equilibrium in which \( a_M \neq 1 \). If \( \hat{c}_W \leq c_W \), then the manager prefers the equilibrium with \( a_M = 1 \).

Proof. The difference between an equilibrium in which \( a_M = 1 \) and \( \hat{a}_M < 1 \) is
\[ V_M(c_W, 1, v_M) - V_M(\hat{c}_W, \hat{a}_M, v_M) \]
\[ \frac{\gamma}{\gamma} \text{As and So,} \]

Hence, whenever the term multiplying the manager strictly mixes, that is, equality from the fact that \((\text{MLRP})\), Lemma 6 and \(\hat{\gamma} > \gamma\),

\begin{align*}
\sum_{y} v_M(y) \left[ (p(y|1,1) - p(y|0,1)) - (p(y|1,0) - p(y|0,0)) \right] & - c_M \\
- \hat{a}^M \left\{ \sum_{y} v_M(y) \left[ (p(y|1,1) - p(y|0,1)) - (p(y|1,0) - p(y|0,0)) \right] - c_M \right\} \\
+ [G(c_w) - G(\hat{c}_w)] \sum_{y} v_M(y) [p(y|0,1) - p(y|0,0)] \\
\geq -\hat{a}^M \left\{ \sum_{y} v_M(y) \left[ (p(y|1,1) - p(y|0,1)) - (p(y|1,0) - p(y|0,0)) \right] - c_M \right\} \\
+ [G(c_w) - G(\hat{c}_w)] \sum_{y} v_M(y) [p(y|0,1) - p(y|0,0)] \\
= [G(c_w) - G(\hat{c}_w)] \sum_{y} v_M(y) [p(y|0,1) - p(y|0,0)] \geq 0.
\end{align*}

The first equality is an algebraic manipulation. The first inequality comes from \((JC_M)\). The second equality from the fact that \(\hat{a}^M\) is an equilibrium choice. There are two possibilities: \(\hat{a}^M = 0\); or the manager strictly mixes, that is, \(\hat{a}^W \in (0,1)\). If \(\hat{a}^M = 0\), the second equality trivially holds. If \(\hat{a}^M \in (0,1)\), the manager must be indifferent between exerting high and low effort. Which implies the term multiplying \(\hat{a}^M\) must be equal to zero. The last inequality results from \(v_M\) being increasing, \(p\) satisfying \((\text{MLRP})\), Lemma 6[3] and \(\hat{c}_w \leq c_w\).

**Lemma 8.** There is no equilibrium with \(\hat{c}_w > c_w\) and \(a_M < 1\).

**Proof.** Take any \(\hat{c}_w > c_w\). Suppose \(\sum_{y} v_M(y) \left[ (p(y|1,1) - p(y|0,1)) - (p(y|1,0) - p(y|0,0)) \right] > 0\), then

\[ G(\hat{c}_w) \sum_{y} v_M(y) \left[ (p(y|1,1) - p(y|0,1)) - (p(y|1,0) - p(y|0,0)) \right] + \sum_{y} v_M(y) [p(y|1,0) - p(y|0,0)] > G(c_w) \sum_{y} v_M(y) \left[ (p(y|1,1) - p(y|0,1)) - (p(y|1,0) - p(y|0,0)) \right] + \sum_{y} v_M(y) [p(y|1,0) - p(y|0,0)] = c_M.
\]

Hence, whenever \(\hat{c}_w > c_w\) and \(\sum_{y} v_M(y) \gamma(y) > 0\), the manager strictly prefers to exert high effort.\footnote{Recall that \(\gamma(y) = [p(y|1,1) - p(y|0,1)] - [p(y|1,0) - p(y|0,0)]\)}

So, \(a_M < 1\) cannot be part of an equilibrium. It remains to show that \(\sum_{y} v_M(y) \gamma(y) > 0\). Let

\[ \bar{y} := \max \{ y \in Y : \gamma(y) \leq 0 \}, \]

and

\[ \bar{v}_M(y) := \begin{cases} v_M(y) & \text{if } y \leq \bar{y} \\ v_M(\bar{y}) & \text{otherwise.} \end{cases} \]

As \(\gamma\) strictly single crosses zero from below and \(v_M\) is strictly increasing

\[ \sum_{y} v_M(y) \gamma(y) > \sum_{y} \bar{v}_M(y) \gamma(y). \]

\footnote{Recall that \(\gamma(y) = [p(y|1,1) - p(y|0,1)] - [p(y|1,0) - p(y|0,0)]\).}
As \( v_M \) is increasing, so it is \( \tilde{v}_M \). By Assumption 3 and Lemma 6, \( \sum_Y v_M(y) \gamma(y) \geq 0 \). Hence,

\[
\sum_Y v_M(y) \gamma(y) > \sum_Y \tilde{v}_M(y) \gamma(y) \geq 0.
\]

Finally, we can conclude that the solution to the relaxed problem satisfies the equilibrium selection constraints and is the optimal compensation scheme.

Proof of Remark 4. Note that the manager’s contract as constructed in the previous section does not depend on the distribution of \( z \). Hence, it is as characterized by (3).

Now regarding worker’s contract. One way to represent a fully informative signal is to let \( z \sim U[\frac{z - z^*}{2}, z] \) if \( a_W = 0 \) and \( z \sim U[\frac{z - z^*}{2}, z] \) if \( a_W = 1 \). Any realization below \( \frac{z - z^*}{2} \) fully reveals low effort and above fully reveals high effort. Note that under full information \( \sigma^* \) is given by reporting \( g \) with probability 1 when \( a_W = 1 \) and \( b \) with probability 1 otherwise. That is, under full-information, the manager’s report is fully informative about 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 assures participation is given by

\[
v^*_W(y, b) = \bar{u}_W + \mathbb{E}[c|c \leq c_W] - G(c_W) c_W \quad \text{and} \quad v_W(y, g) = v^*_W(y, b) + c_W.\]

Proof of Proposition 5. Given 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, we can focus solely on the worker’s compensation.

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

\[
p(y|1, 1)Q(z^*(y)|1) < p(y|1, 0)Q(z^*(y)|0),
\]

\[
p(y|1, 0)[1 - Q(z^*(y)|0)] < p(y|1, 1)[1 - Q(z^*(y)|1)] \quad \text{for all } y \in Y.
\]
As the inequalities are strict, the principal can slightly increase the upper bound of the lowest set of the partition and decrease the lower bound of the highest set of the partition and keep the same evaluations. That is, take \( \kappa > 0 \) and a small \( \varepsilon > 0 \). Let the manager’s information structure be given by

\[
\left\{ [z, z^*(y_n) + \varepsilon]; \{ z \} \subset (\varepsilon(y_n), z^*(y_0)]; [z^*(y_0) - \kappa \varepsilon, z] \right\}.
\]

For \( \varepsilon \) sufficiently small, the manager remains to provide a bad evaluation whenever she gets the signal associated with the lowest set of the partition and the good evaluation if she sees the signal associated with the highest set.

For a given \( \kappa \) and \( \varepsilon \), we have an information partition. The principal can choose the optimal contracts that enforce efforts \( (c_W, a_M = 1) \) for that given information structure. Denote the minimal cost as \( C(\varepsilon, \kappa) \). If \( \varepsilon = 0 \) we are at the full transparency case. I am going to show that there exists a \( \kappa > 0 \) such that \( \frac{dC}{d\varepsilon}(0, \kappa) < 0 \). By increasing \( \varepsilon \) slightly from zero, the principal strictly reduces the cost.

Let \( v_W \) denote the optimal contracts when \( \varepsilon = 0 \). When we increase \( \varepsilon \) we increase the probability of good evaluation when output is \( y_0 \) and decrease the probability at \( y_n \). I construct \( \kappa \) such that the marginal effect of \( \varepsilon \) 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.

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

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

\[
C(\varepsilon, \kappa) = \min_{v_W} \left\{ \sum_y \sum_{e \in \{g, b\}} \left[ \phi_W(v_W(y, e)) - \lambda^W v_W(y, e) - \mu^W s^W(y, e) + \mu^W c_W + \lambda^W (\bar{a}_W + \mathbb{E}[c|c \leq c_W]) \right] f^{cw}(y, e) \right\}.
\]

Note that \( s, f^{cw}, \lambda^W \) and \( \mu^W \) all depend on the information structure. Hence, on \( \varepsilon \). However, when differentiating at \( \varepsilon = 0 \), we can apply the Milgrom and Segal (2002)’s envelope theorem.\(^{31}\) Therefore, we can keep \( v_W \) fixed and we do not need to worry about how \( \lambda^W \) and \( \mu^W \) change with

\(^{31}\)In Appendix B I show that one can bound \( v_W(y, e) \) from above, which assure the conditions for Milgrom and Segal (2002)’s Theorem 1.
\[ \frac{dC}{d\varepsilon}(0, \kappa) = \frac{dC}{d\varepsilon}(0, \kappa) = \sum_{\epsilon \in \{g, b\}} \sum_{y \in \{y_n, y_n\}} \left\{ \left[ \phi_W(v_W(y, e)) - \lambda^W v_W(y, e) \right] \frac{d\phi_W}{d\varepsilon}(y, e) - \mu^W \frac{d\phi_W}{d\varepsilon}(y, e) \right\} \]

\[ = \sum_{\epsilon \in \{g, b\}} \sum_{y \in \{y_n, y_n\}} \left\{ \left[ \phi_W(v_W(y, e)) - \lambda^W v_W(y, e) \right] \frac{d\phi_W}{d\varepsilon}(y, e) - \mu^W \frac{d\phi_W}{d\varepsilon}(y, e) \right\} \]

\[ = \frac{p(y_n|1, 1)q(z^+(y_n)|1)G(c_w) + (1 - G(c_w))p(y_n|1, 0)q(z^+(y_n)|0)}{v_W(y_0, g) - v_W(y_0, b)} \times \left[ \frac{\phi_W(v_W(y_0, g)) - \phi_W(v_W(y_0, b))}{v_W(y_0, g) - v_W(y_0, b)} - \frac{\phi_W(v_W(y_n, g)) - \phi_W(v_W(y_n, b))}{v_W(y_n, g) - v_W(y_n, b)} \right] < 0. \]

Where the first two terms are strictly positive and the last is strictly negative from convexity of \( \phi_W \) and the fact that \( v_W(y_n, e) > v_W(y_0, e) \) for \( e \in \{g, b\} \), which is a direct implication from \( s^W(y_n, e) > s^W(y_0, e) \).

**B Appendix — Existence and Slackness of the Minimum Payment Constraint**

Existence of a solution to (1) is assured by existence in (2) and (4). As shown before, problem (1) can be relaxed and solved by approaching (2) and (4) separately. The additional constraints were then verified in the main text. The remaining steps are to assure existence of solutions to (2) and (4) and that the minimum payment constraint is slack.

**Proposition 8.** For any \( \tilde{u}_M \in (u_M(0), +\infty) \), \( \tilde{u}_W \in (u_W(0), +\infty) \), \( c_M \in \mathbb{R}_{++} \) and \( c_W \in \mathbb{R}_{++} \) there exists solutions to problems (2) and (4).

**Proof.** Note that

\[ v_M(y) = \begin{cases} \tilde{u}_M & \text{if } y < y_n \\ \tilde{u}_M + 1 + \frac{c_M}{p(y_n|1, 1) - p(y_n|1, 0)} & \text{otherwise,} \end{cases} \]

and

\[ v_W(y, e) = \begin{cases} \tilde{u}_W & \text{if } y \neq y_n \\ \tilde{u}_W + 1 + \frac{c_W}{p(y_n|1, 1) - p(y_n|1, 0)} & \text{otherwise} \end{cases} \]

51
satisfy \((IR_M), (IC_M), (IR_W)\) and \((IC_W)\) with slackness at a finite cost. Hence, we can bound payments from above. We have compact constraint sets and continuous objective functions. The existence of a minimum follows from Weierstass’ theorem.

**Proposition 9.** For any \(c_M \in \mathbb{R}_{++}\) and \(c_W \in \mathbb{R}_{++}\), there exists \(\hat{u}_M \in (u_M(0), +\infty)\) and \(\hat{u}_W \in (u_W(0), +\infty)\) such that for any \(\bar{u}_M > \hat{u}_M\) and \(\bar{u}_W > \hat{u}_W\) the solution to (2) and (4) are such that

\[
v_M(y) > u_M(0) \quad \text{for all } y \in Y,
\]

and

\[
v_W(y, e) > u_W(0) \quad \text{for all } y \in Y, e \in \{b, g\}.
\]

**Proof.** We show that if the minimum payment remains binding when \(\bar{u}_i (i \in \{M, W\})\) increases, then the incentive compatibility constraint becomes slack. A contradiction.

Let’s start with manager’s problem. Take an increasing sequence \(\bar{u}_{M_k} \in (u_M(0), +\infty)\) such that

\[
\lim_{k \to +\infty} \bar{u}_{M_k} = +\infty.
\]

Let \(v_{M_k}\) be associated optimal contract.

For \(\bar{u}_{M_k}\) high enough \((IR_M)\) must bind. That is,

\[
\sum_{Y} v_{M_k}(y)p_{c_W}(y|1) = c_M + \bar{u}_{M_k}, \quad (16)
\]

The optimal payment considering the minimum payment constraint is given by

\[
\phi'_M(v_{M_k}(y)) = \lambda_k^M + \mu_k^M s^M(y) + \phi_k^M(y),
\]

where \(\phi_k^M(y)\) is the minimum payment multiplier and it is strictly bigger than zero only if \(v_{M_k}(y) = u_M(0)\). Note that as \(s^M(y)\) is increasing in \(y\), the smallest payment happens at \(y_0\). That is, \(v_{M_k}(y_0) \leq v_{M_k}(y)\) for all \(y \in Y\). If the minimum payment constraint binds, it must be the case that \(v_{M_k}(y_0) = u_M(0)\).

Suppose for the sake of obtaining a contradiction that there exists a subsequence \(\bar{u}_{M_{k_j}}\) and \(J \in \mathbb{N}\) such that \(v_{M_{k_j}}(y_0) = u_M(0)\) for all \(k_j > J\).

By (16)

\[
\sum_{y > y_0} v_{M_{k_j}}(y)p_{c_W}(y|1) = c_M + u_{M_{k_j}} - u_M(0)p_{c_W}(y_0|1). \quad (17)
\]
As the manager is risk averse, the cheapest payments that would still satisfy the equation above and have $v_{MK_j}(y_0) = u_M(0)$ would be a flat payment for all other output realizations, that is,

$$
\bar{v}_{MK_j}(y) = \begin{cases} 
  u_M(0) & \text{if } y = y_0, \\
  \frac{u_{MK_j} + c_{MK_j} - u_M(0)p^w(y_0|1)}{1 - p^w(y_0|1)} & \text{otherwise.}
\end{cases}
$$

As $\bar{v}_{MK_j}$ is cheaper than any other contract that satisfies (IR$_M$) and has the lowest payment binding, it must be optimal if it satisfies (IC$_M$). I now show that for large enough $u_{MK_j}$, $\bar{v}_{MK_j}$ not only satisfies (IC$_M$) but it does with slack. Note that

$$
\sum_y \bar{v}_{MK_j}(y) \left[ p^w(y|1) - p^w(y|0) \right] = \sum_y \bar{v}_{MK_j}(y) p^w(y|1) - \sum_y \bar{v}_{MK_j}(y) p^w(y|0) = \bar{u}_{MK_j} + c_M + \frac{u_{MK_j} + c_{MK_j} - u_M(0)p^w(y_0|1)}{1 - p^w(y_0|1)} [1 - p^w(y_0|0)]
$$

$$
= c_M + \left[ c_{MK_j} - u_M(0)p^w(y_0|1) \right] \frac{1 - p^w(y_0|0)}{1 - p^w(y_0|1)} + \bar{u}_{MK_j} \left[ 1 - \frac{1 - p^w(y_0|0)}{1 - p^w(y_0|1)} \right]
$$

As $\frac{1 - p^w(y_0|0)}{1 - p^w(y_0|1)} < 1$, for $\bar{u}_{MK_j}$ large enough, the equation above is strictly higher than $c_M$. The principal could increase $\bar{v}_{MK_j}(y_0)$ and decrease $\bar{v}_{MK_j}(y)$ for $y > y_0$ such that participation and incentive compatibility are still satisfied. Additionally, the new contract would improve risk sharing and decrease expected payments. A contradiction. Hence, it does not exist a subsequence $\bar{u}_{MK_j}$ and $J \in \mathbb{N}$ such that $v_{MK_j}(y_0) = u_M(0)$ for all $k_j > J$. Therefore, for $\bar{u}_M$ large enough the minimum payment constraint is slack.

For the worker I take a similar approach. Take an increasing sequence $\bar{u}_{Ak} \in (u_W(0), +\infty)$ such that $\lim_{k \to +\infty} \bar{u}_{Ak} = +\infty$. Let $v_{Ak}$ be associated optimal contract.

For $\bar{u}_{Ak}$ high enough (IR$_W$) must bind. That is,

$$
\sum_{Y} \sum_{E} v_{Ak}(y, e) f^w(y, e) = \bar{u}_{Ak} + \mathbb{E}[c | c \leq c_W].
$$

The optimal payment considering the minimum payment constraint is given by

$$
\phi'_W(v_{M_k}(y, e)) = \lambda_k^W + \mu_k^W s^W(y, e) + \phi_k^W(y, e),
$$

where $\phi_k^W(y, e)$ is the minimum payment multiplier and it is strictly bigger than zero only if $v_{Ak}(y, e) = u_W(0)$.

Let $\bar{y} = \text{argmin}_{Y} \{ s^W(y, b) \}$. The realization $(\bar{y}, b)$ must be the lowest paying one. Hence, if the minimum payment constraint binds it must be the case that $v_{Ak}(\bar{y}, b) = u_W(0)$.
Suppose there exists a subsequence \( \bar{a}_{Ak_j} \) and \( J \in \mathbb{N} \) such that \( v_{Ak_j}(\tilde{y}, b) = u_W(0) \) for all \( k_j > J \).

The cheapest payments that would still satisfy \((IR_A)\) and have \( v_{Ak_j}(\tilde{y}, b) = u_W(0) \) would be a flat payment for all other output and performance report realizations, that is,

\[
\bar{v}_{Ak_j}(y) = \begin{cases} 
  u_W(0) & \text{if } (y, e) = (\tilde{y}, b), \\
  \frac{\bar{a}_{Ak} + \mathbb{E}[c|c \leq c_W] - u_W(0)f_W(\tilde{y},b)}{1 - f_W(\tilde{y},b)} & \text{otherwise.}
\end{cases}
\]

As \( \bar{v}_{Ak_j} \) 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 \( u_{Ak_j}, \bar{v}_{Ak_j} \) not only satisfies \((IC_W)\) but it does with slack. A contradiction because \((IC_W)\) must bind at the optimum.

Note that

\[
\sum_Y \sum_E \bar{v}_{Ak_j}(y, e)[f(y, e|1) - f(y, e|0)] = u_W(0) \left[ f(\tilde{y}, b|1) - f(\tilde{y}, b|0) \right] \frac{1}{1 - f_W(\tilde{y}, b)} + \mathbb{E}[c|c \leq c_W] + \bar{a}_{Ak_j} \left[ f(\tilde{y}, b|0) - f(\tilde{y}, b|1) \right] \frac{1}{1 - f_W(\tilde{y}, b)}.
\]

As \([f(\tilde{y}, b|0) - f(\tilde{y}, b|1)] > 0\), for \( \bar{a}_{Ak_j} \) large enough, the equation above is strictly higher than \( c_W \). A contradiction. Hence, it does not exist a subsequence \( u_{Ak_j} \) and \( J \in \mathbb{N} \) such that \( v_{Ak_j}(\tilde{y}, b) = u_W(0) \) for all \( k_j > J \). Therefore, for \( \bar{a}_W \) large enough the minimum payment constraint is slack. \( \square \)

C Appendix C — Continuous Efforts

Proof of Proposition[\textcolor{red}{6}] Take given contracts \((E, v_M, v_W)\) and a given manager’s effort level \( \hat{a}_M \). Denote the worker’s best response to \((a_M, \sigma)\) by \( \hat{a}_W(a_M, \sigma) \). I split the proof in two cases.

Case 1: \( \hat{a}_W(\hat{a}_M, \sigma) = 0 \) for all \( \sigma \). Then, the worker’s effort is zero regardless of the evaluation policy and the result trivially holds.

Case 2: there exists \( \sigma \) such that \( \hat{a}_W(\hat{a}_M, \sigma) > 0 \).

Take an arbitrary \( \hat{\sigma} \) such that \( \hat{a}_W(\hat{a}_M, \hat{\sigma}) > 0 \) and \( \hat{\sigma} \) sends with strictly positive probability messages that are not the highest and the lowest-paying\textsuperscript{32} Denote by \( \hat{a}_W := \hat{a}_W(\hat{a}_M, \hat{\sigma}) \). As \( \lim_{x \to 1} c'_W(x) = +\infty \), the effort level \( \hat{a}_W \) must satisfy the following first-order condition

\[
\int_Z \sum_Y \sum_E v_W(y, e) \hat{\sigma}(e|\hat{a}_M, y, z) \left[ \frac{p_W(y|\hat{a}_M, \hat{a}_W)}{p(y|\hat{a}_M, \hat{a}_W)} + \frac{q_W(z|\hat{a}_W)}{q(z|\hat{a}_W)} \right] p(y|\hat{a}_M, \hat{a}_W) q(z|\hat{a}_W) dz = c'_W(\hat{a}_W).
\]

\textsuperscript{32}If such \( \hat{\sigma} \) does not exists, the result again trivially holds.
Define \( \hat{\varepsilon} : Y \times [0, 1]^2 \to Z \) such that

\[
\frac{p_W(y|a_M, a_W)}{p(y|a_M, a_W)} = -\frac{q_W(\hat{\varepsilon}(y, a_M, a_W)|a_W)}{q(\hat{\varepsilon}(y, a_M, a_W)|a_W)}
\]

Let \( \hat{\sigma} \) be

\[
\hat{\sigma}(a_M, y, \hat{e}) := \begin{cases} 
\delta_{\hat{e}_y} & \text{if } z < \hat{\varepsilon}(y, a_M, a_W), \\
\delta_{\hat{e}_y} & \text{otherwise.}
\end{cases}
\]

Remember that \( \hat{e}_y \) is the lowest-paying and \( \hat{e}_y \) the highest-paying message when output is \( y \). Note that

\[
\frac{\partial V_W}{\partial a_W}(a_M, a_W, y, z) = \int \sum_{y \in \mathcal{Y}} v_W(y, \hat{e}) \hat{\sigma}(e|a_M, y, z) \left[ \frac{p_W(y|a_M, a_W)}{p(y|a_M, a_W)} + \frac{q_W(z|a_W)}{q(z|a_W)} \right] p(y|a_M, a_W) q(z|a_W) dz - c_W(a_W)
\]

\[
> \int \sum_{y \in \mathcal{Y}} v_W(y, e) \hat{\sigma}(e|a_M, y, z) \left[ \frac{p_W(y|a_M, a_W)}{p(y|a_M, a_W)} + \frac{q_W(z|a_W)}{q(z|a_W)} \right] p(y|a_M, a_W) q(z|a_W) dz - c_W(a_W)
\]

\[
= \frac{\partial V_W}{\partial a_W}(a_M, a_W, y, z) = 0.
\]

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

I now show that \( V_W(., \hat{a}_M, v_M, \hat{\sigma}) \) is strictly concave. Given the evaluation strategy \( \hat{\sigma} \), the worker’s payoff is given by

\[
V_W(a_M, \hat{a}_M, v_M, \hat{\sigma}) = \sum_y \left\{ v_W(y, \hat{e}_y) \left[ 1 - Q(\hat{\varepsilon}(y, \hat{a}_M, \hat{a}_W)|a_W) \right] + v_W(y, \hat{e}_y) Q(\hat{\varepsilon}(y, \hat{a}_M, \hat{a}_W)|a_W) \right\} p(y|\hat{a}_M, a_W) - c_W(a_W).
\]

Hence,

\[
\frac{\partial^2 V_W}{\partial a_W^2}(a_M, \hat{a}_M, v_M, \hat{\sigma}) = -\frac{\partial c_W}{\partial a_W}(a_W)
\]

\[-\sum_y \left[ v_W(y, \hat{e}_y) - v_W(y, \hat{e}_y) \right] \left[ Q_W(\hat{\varepsilon}(y, \hat{a}_M, \hat{a}_W)|a_W) p(y|\hat{a}_M, a_W) + 2 Q_W(\hat{\varepsilon}(y, \hat{a}_M, \hat{a}_W)|a_W) p_W(y|a_W) \right] < 0.
\]

The first square bracket is positive by construction, while the second is positive by Assumption 7.

Therefore, for any arbitrary \( \hat{\sigma} \), I have constructed a better evaluation strategy \( \hat{\sigma} \). It remains to show that there exists a \( \sigma^* \) that maximizes \( \hat{a}_w(\hat{a}_M, \sigma) \).

33 I denote by \( Q_W(z|a_W) := \frac{\partial q_W}{\partial a_W}(z|a_W) \) and \( Q_W(z|a_W) := \frac{\partial^2 q_W}{\partial a_W^2}(z|a_W) \).

55
I say that a performance evaluation strategy $\sigma$ has a a cutoff form if there exists a cutoff function $\bar{z} : Y \times [0, 1] \rightarrow Z$ such that

$$\sigma(a_M, y, z) := \begin{cases} \delta_{E_y} & \text{if } z < \bar{z}(y, a_M), \\ \delta_{E_y} & \text{otherwise.} \end{cases}$$

Note that for each $\sigma$, there exists a $\sigma'$ with a cutoff form such that $\bar{a}_W(a_M, \sigma) \leq \bar{a}_W(a_M, \sigma')$. Hence, one can restrict attention to performance evaluations with a cutoff form. Also, for a given manager’s effort choice $a_M$, a performance evaluation with a cutoff form can be described by a vector of cutoffs for each $y$. As each $\bar{z}(y, a_M)$ 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.

\section{Appendix D — Optimal Menu}

I now focus on the limit case in which $\varepsilon$ goes to zero. That is, worker’s cost is re-drawn with negligible probability, and $\sigma^*$ is the performance evaluation strategy. Incentive compatibility for the worker refers to truthful reporting and obedience. Let $k(c)$ be the contract assigned to a type that reports $c$. Let $\alpha_M(k(c))$ be the effort recommended to the manager after seeing the worker chose contract $k(c)$. Let $\alpha_W(c)$ be the effort recommended to the worker who reported $c$ if his type does not change.\footnote{Note that $\alpha_W$ is measurable with respect to $c$, while $\alpha_M$ is measurable only with respect to $k(c)$. The reason is that the manager does not necessarily see the reported $c$, but she sees the contract choice $k(c)$, while the worker observes his own cost.}

As argued in the main text, the evaluation strategy is given by $\sigma^*$, and the performance ratings are binary. Hence, each contract $k$ specifies payments to the worker and the manager for each of the $2(n+1)$ possible realizations. There are $(n+1)$ possible output realizations and 2 potential performance reviews. Therefore, I can denote $v^k_W \in \mathbb{R}^2_{(n+1)}$ and $v^k_M \in \mathbb{R}^{(n+1)}$ as the payments associated with contract $k$.

For each given contract $k$, define as $S(k)$ the expected worker’s utility payment if he does not exert effort and by $D(k)$ the expected utility payment if he does. That is,

$$S(k) := \int_Z \sum_Y \sum_E v^k_W(y, e) \sigma^*(e|y, z) p(y|\alpha_M(k), 0) q(z|0) dz,$$

$$D(k) := \int_Z \sum_Y \sum_E v^k_M(y, e) \sigma^*(e|y, z) p(y|\alpha_M(k), 0) q(z|0) dz.$$
and

\[ D(k) := \int \sum_{y} \sum_{e} v_W(y, e) \sigma^*(e|y, z)p(y|\alpha_M(k), 1)q(z)dz. \]

Now suppose that a worker with contract \( k \) gets his cost re-drawn. He chooses to exert effort if and only if the re-drawn cost \( \bar{c} \leq [D(k) - S(k)] \). Hence, given contract \( k \), the worker’s expected utility conditional on the cost being re-drawn is given by

\[ R(k) := D(k) G(D(k) - S(k)) + S(k) \left[ 1 - G(D(k) - S(k)) \right] - \mathbb{E}[c|c \leq (D(k) - S(k))]. \]

Each type \( c \) chooses a contract in the menu. Denote the contract chosen by type \( c \) by \( k(c) \). worker’s incentive compatibility requires that the truthful type reporting and obedience. That is,

\[ (1 - \varepsilon) \left[ \alpha_W(k(c)) D(c) + (1 - \alpha_W(c)) S(c) - \alpha_W(c) c \right] + \varepsilon R(k(c)) \geq (1 - \varepsilon) \left[ a_W D(k) + (1 - a_W) S(k) - a_W c \right] + \varepsilon R(k) \quad \forall c, k \text{ and } a_W. \]

\[ (IC'_W) \]

Participation also must hold for each contract offered, that is

\[ (1 - \varepsilon) \left[ \alpha_W(c) D(k(c)) + (1 - \alpha_W(c)) S(k(c)) - \alpha_W(c) c \right] + \varepsilon R(k(c)) \geq \bar{a}_W. \]

\[ (IR'_W) \]

**Lemma 9.** worker’s recommended effort \( \alpha_W(c) \) must be a cutoff function. That is, \( \alpha_W(c) = 1 \) if \( c \leq c^* \) and zero otherwise.

**Proof of Lemma 9.** Take any menu satisfying \( (IR'_W) \) and \( (IC'_W) \). Define the utility of type \( c \) that chooses contract \( \hat{k} \) and exerts effort \( a_W \) in case his cost is not re-drawn as

\[ V_W(\hat{c}, c, a_W) := (1 - \varepsilon) \left[ (1 - a_W) S(\hat{k}(\hat{c})) + a_W [D(\hat{k}(\hat{c})) - c] \right] + \varepsilon R(\hat{k}(\hat{c})). \]

Suppose \( \alpha_W(c) > 0 \). Note that \( \alpha_W(c) > 0 \) implies that \( V(c, c, 1) \geq V(c, c, 0) \) and that \( V(c, c, 1) \) is the best utility a worker with type \( c \) can achieve. Take any \( c' < c \)

\[ V_W(c, c', 1) = (1 - \varepsilon) \left[ (1 - a_W) S(k(c)) + [D(k(c)) - c'] \right] + \varepsilon R(k(c)) \]

\[ > (1 - \varepsilon) \left[ (1 - a_W) S(k(c)) + [D(k(c)) - c] \right] + \varepsilon R(k(c)) \]

\[ = \ V(c, c, 1) \geq V(\hat{c}, c, 0) \quad \forall \hat{c} \in \mathbb{R}_+, \]

including \( \hat{c} = c' \). Note also, that \( V_W(c', c', 0) = V_W(c', c, 0) \). Hence,

\[ V_W(c, c', 1) > V_W(c, c, 1) \geq V_W(c', c, 0) = V_W(c', c', 0). \]

Hence, it must be the case that \( \alpha_W(c') = 1. \)
I now minimize the cost of implementing high manager’s effort and a worker’s effort cutoff \( c_W \). That is,

\[
\alpha_W(c) = \begin{cases} 
1 & \text{if } c \leq c_W \\
0 & \text{otherwise.} 
\end{cases}
\]

**Lemma 10.** There is no loss in restricting attention to menus with at most two contracts.

**Proof.** Take a contract menu \( \mathbf{K} \) that implements effort cost cutoff \( c_W \) and \( a_M = 1 \). By \( \mathcal{IC}_W \), for \( c \leq c_W \)

\[
V_W(c, c, 1) = (1 - \varepsilon) \left[ D(k(c)) - c \right] + \varepsilon R(k(c)) \geq (1 - \varepsilon) \left[ D(\hat{k}) - c \right] + \varepsilon R(\hat{k}) \quad \forall \hat{k} \in \mathbf{K}.
\]

Hence, for all types \( c \leq c_W \) it must be the case that \( k(c) \in K_1 := \argmax_{k \in \mathbf{K}} \{(1 - \varepsilon)D(\hat{k}) + \varepsilon R(\hat{k})\} \).

By \( \mathcal{IC}_W \), for \( c > c_W \)

\[
V_W(c, c, 0) = (1 - \varepsilon)S(k(c)) + \varepsilon R(k(c)) \geq (1 - \varepsilon)S(\hat{k}(\varepsilon)) + \varepsilon R(\hat{k}(\varepsilon)).
\]

Hence, for all types \( c \leq c_W \) it must be the case that \( k(c) \in \argmax_{\hat{k} \in \mathbf{K}_2 := \mathbf{K}} \{(1 - \varepsilon)S(\hat{k}) + \varepsilon R(\hat{k})\} \).

The manager observes worker’s contract choice \( k \) and decides whether to participate and to exert effort. Note that we can compute manager’s belief about worker’s probability of effort for each given contract. That is,

\[
\zeta_{k(c)} = \begin{cases} 
(1 - \varepsilon) + \varepsilon G(D(k(c)) - S(k(c))) & \text{if } c \leq c_W \\
\varepsilon G(D(k(c)) - S(k(c))) & \text{otherwise.} 
\end{cases}
\]

The manager’s incentive and participation constraints for each contract chosen by the worker are

\[
\sum_y v^k_M \left[ \zeta_{k(c)} p(y|1, 1) + (1 - \zeta_{k(c)}) p(y|1, 0) \right] - c_M \geq \bar{u}_M, \quad (IR^k_M) \\
\sum_y v^k_M \left[ \zeta_{k(c)} [p(y|1, 1) - p(y|0, 1)] + (1 - \zeta_{k(c)}) [p(y|1, 0) - p(y|0, 0)] \right] \geq \bar{c}_M, \quad (IC^k_M)
\]

For a given worker’s contract choice \( k \), define the principal’s expected implementation costs as

\[
\mathcal{C}(k, \varepsilon) := \int \sum_y \sum_z \left[ \varphi_W(v^k_W(y, \varepsilon)) + \varphi_M(v^k_M(y)) \right] \left[ \zeta_{k(c)} \sigma^*(e|y, z)p(y|1, 1) + (1 - \zeta_{k(c)}) \sigma^*(e|y, z)p(y|1, 0) \right] dz.
\]

Now instead of offering the full menu \( \mathbf{K} \), let the principal offer a menu with two contracts \{\( k_1, k_2 \)\}, where

\[
k_1 \in \argmin_{\hat{k} \in \mathbf{K}_1} \{\mathcal{C}(k, \varepsilon)\} \quad \text{and} \quad k_2 \in \argmin_{\hat{k} \in \mathbf{K}_2} \{\mathcal{C}(k, \varepsilon)\}.
\]

58
Note that \( (IR_k(c), IC_k(c), IR'_W) \) and \( (IC'_W) \) are all satisfied. Also, the expected payments have decreased, since \( k_1 \) and \( k_2 \) are the cheapest among the ones chosen on path from the menu \( K \).

Note that increasing the number of contracts chosen on path by some type \( c \) increases the manager’s information about the worker’s effort. Additional information to the manager imposes additional \( (IR_k(c), IC_k(c)) \) constraints.

The principal has two options: offering two contracts, one chosen by types that will not exert effort and the other by types that will, or offering a single contract — which provides no information to the manager about worker’s effort cost. On the one hand, the first option improves risk-sharing regarding worker’s compensation. On the other hand, it requires additional \( (IR_k(c), IC_k(c)) \) constraints, which need to be satisfied for each contract \( k \), increasing compensation costs to the manager.

I characterize the best contract in each of those classes — a single or two contracts — in the limit case in which \( \varepsilon = 0 \). The optimum is then given by the cheapest among the two classes. Let’s first characterize the single contract optimum, which is similar to Section 2. The only difference is that now worker’s participation must be satisfied after observing cost \( c \). The program in this case is

\[
\bar{\mathcal{C}}(c_W, 1) := \min_{v_W, v_M} \left\{ \mathbb{E} \left[ \varphi^W(v_W(y, e)) + \varphi^M(v_M(y)) | c_W, a_M = 1, \sigma^* \right] \right\} \tag{19}
\]

subject to

\[
\sum_Y v_M(y)p^{c_W}(y|1) - c_M \geq \bar{u}_M, \quad (IR_M)
\]

\[
\int_Z \left[ v_W(y, g)(1 - Q(z^*(y)|0)) + v_W(y, b)Q(z^*(y)|0) \right] p(y|1, 0) \geq \bar{u}_W, \quad (IR_W)
\]

\[
\sum_Y v_W(y, g)s^W(y, g)f^{c_W}(y, g) + v_W(y, b)s^W(y, b)f^{c_W}(y, b) = c_W. \quad (IC_W)
\]

\[
\sum_Y v_M(y)s^M(y)p^{c_W}(y|1) \geq c_M. \quad (IC_M)
\]

I solve the problem above by minimizing pointwise. The solution is described in the Lemma below.

**Lemma 11.** The solution to (19) is given by

\[
\varphi'_M(\tilde{v}_M(y)) = \lambda^M + \bar{\mu}^M s^M(y), \quad (20)
\]
and

\[ \varphi_W^t(\hat{v}_W^*(y, i)) = \hat{\lambda}^W + \hat{\mu}^W s^W(y, i) \quad \text{for all } y \in Y \text{ and } i \in \{b, g\}. \quad (21) \]

Where \( \hat{\lambda}^W \), \( \hat{\mu}^W \), \( \hat{\lambda}^M \) and \( \hat{\mu}^M \) are the respective dual multipliers associated with \((\hat{I}R_W), (\hat{I}C_W), (\hat{I}R_M)\) and \((\hat{I}C_M)\).

The alternative for the principal is to offer two contracts: \( k_1 \), chosen by worker types \( c \leq c_W \), and \( k_2 \), chosen by types \( c > c_W \).

The principal’s problem when offering two contracts can be written as

\[
\mathcal{G}(c_W, 1, \varepsilon) := \min_{v^k_W, v^k_M, \varepsilon} \left\{ \mathcal{G}(k_1, \varepsilon)G(c_W) + \left[ 1 - G(c_W) \right] \mathcal{G}(k_2, \varepsilon) \right\}
\]

subject to

\[
\sum_y v^k_M(y) \left[ p(y|1, 1) \zeta_k + p(y|1, 0)(1 - \zeta_k) \right] - c_M \geq \bar{u}_M, \quad \forall k \in \{k_1, k_2\}, \quad (\hat{I}R^k_M)
\]

\[
\sum_y v^k_M(y) s^M(y, k, \varepsilon) \left[ p(y|1, 1) \zeta_k + p(y|1, 0)(1 - \zeta_k) \right] \geq c_M \quad \forall k \in \{k_1, k_2\}, \quad (\hat{I}C^k_M)
\]

\[
(1 - \varepsilon) [D(k_1) - c_W] + \varepsilon R(k_1) \geq \bar{u}_W, \quad (\hat{I}R^k_W)
\]

\[
(1 - \varepsilon) S(k_2) + \varepsilon R(k_2) \geq \bar{u}_W, \quad (\hat{I}C^k_W)
\]

\[
D(k_1) - S(k_1) = c_W, \quad (\hat{I}C^k_W)
\]

\[
D(k_2) - S(k_2) \leq c_W. \quad (\hat{I}C^k_W)
\]

One can show by the Maximum Theorem that \( \mathcal{G}(c_W, 1, \varepsilon) \) is continuous in \( \varepsilon \). The first step is to show that one can artificially bound payments. The argument is standard and omitted here. Then I show that the objective function and the correspondence that describe the feasible contracts are continuous. Denote by \( \mathcal{Y}(\varepsilon) \subset \mathbb{R}^{6(n + 1)} \) the correspondence such that if \( (v^k_W, v^k_M, v^k_M) \in \mathcal{Y}(\varepsilon) \), then it satisfies all the constraints in problem (22). Note that as \( |D(k_2) - S(k_2)| \leq c_W \), then \( \zeta_k \) is continuous in \( \varepsilon \). Hence, both the objective function and all the constraints are continuous in \( \varepsilon \).

As \( \mathcal{G}(c_W, 1, \varepsilon) \) is continuous, I solve for \( \varepsilon = 0 \), which is equal to \( \lim_{\varepsilon \rightarrow 0^+} \mathcal{G}(c_W, 1, \varepsilon) \). Before presenting the solution, I now define the appropriate score functions for this case.

\[
s^M(y, k_1) = \frac{[p(y|1, 1) - p(y|0, 1)]}{p(y|1, 1)}, \quad s^M(y, k_2) = \frac{[p(y|1, 0) - p(y|0, 0)]}{p(y|1, 0)},
\]

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

60
\[ \hat{s}^W(y, b) := \frac{p(y|1,1)Q(z^*(y)|1) - p(y|1,0)Q(z^*(y)|0)}{p(y|1,1)Q(z^*(y)|1)}. \]

The solution is then described in the Lemma below.

**Lemma 12.** For \( \varepsilon = 0 \), the solution to (22) is given by

\[ \phi'_M(\hat{v}^k_M(y)) = \hat{\lambda}^M_k + \hat{\mu}^M_s^M(y, k) \quad \text{for all } k \in \{k_1, k_2\}, \quad (23) \]

and

\[ \phi'_W(\hat{v}^k_W(y, i)) = \hat{\lambda}^W_{k_1} + \hat{\mu}^W_{k_1} \hat{s}^W(y, i) \quad \text{for all } y \in Y \text{ and } i \in \{b, g\}, \quad (24) \]

and

\[ v^k_W(y, e) = \bar{u}^W \quad \text{for all } y \in Y \text{ and } i \in \{b, g\}, \quad (25) \]

Where \( \hat{\lambda}^W_{k_1}, \hat{\mu}^W_{k_1}, \hat{\lambda}^M_{k_1}, \hat{\mu}^M_{k_1}, \hat{\lambda}^M_{k_2}, \hat{\mu}^M_{k_2} \) are the respective dual multipliers associated with \( (\hat{IR}^k_W), (\hat{IC}^k_W), (\hat{IR}^k_M), (\hat{IC}^k_M) \) and \( (\hat{IC}^k_M) \).

The minimum of the two options gives the cost minimizer menu.

**Proposition 10.** Suppose Assumptions 7-3 hold, that the principal wants to implement worker’s effort cutoff \( c_W > 0 \) and high effort for the manager. The minimum cost when \( \varepsilon \) tends to zero is given by

\[ \lim_{\varepsilon \to 0} \left\{ \mathcal{C}^*(c_W, 1, \varepsilon) \right\} = \min \left\{ \hat{\mathcal{C}}(c_W, 1, 0), \bar{\mathcal{C}}(c_W, 1) \right\}. \]