KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE OF ECONOMIC RESEARCH

Discussion Paper No. 1099

“Monitoring and Collusion in Subjective Evaluations”

Masanori Hatada

November 2023

KYOTO UNIVERSITY

KYOTO, JAPAN

Monitoring and Collusion in Subjective Evaluations *

Masanori Hatada

November 26, 2023

Abstract

This study investigates the effect of hiring a monitor to observe an agent's behavior in situations where a principal can use subjective measures of his performance. We assume there is a possibility of collusion between the agent and the monitor, i.e., the agent can promise the monitor a monetary payment and get the monitor to make a false report. We derive the optimal contracts in the principal-agent model and the optimal collusion-proof contracts in the principal-monitor-agent model and compare them. The analysis shows that the optimal collusion-proof contracts with the monitor can reduce the agent's rent and burnt money that occur in the optimal contract without the monitor to zero. Furthermore, we also find that, under the optimal collusion-proof contracts, the amount paid to the monitor is never greater than the above payment reductions. Thus, hiring a monitor benefits the principal despite the possibility of collusion, which implies that monitors play vital roles in contracts with subjective evaluations.

*I am grateful to Tadashi Sekiguchi for his very helpful discussions and comments.

1 Introduction

When hiring workers to do a task, employers often cannot provide objective measures of the outcome at the time of writing contracts or paying compensation. For example, let us consider the case in which a person hires an expert to produce a product, such as software, which requires a high level of knowledge. The value of such a product will become explicit only after the client uses it for a long time, and it is generally difficult for the client without the expertise to prepare objective measures of the product quality in advance and write them into the contract. Alternatively, the extent to which employees contribute to the firm value is often complex and ambiguous, and it is difficult to measure this with objective measures (Thiele, 2011). In such situations, employers often use their subjective evaluations to determine how much to pay workers (Prendergast, 1999).

On the other hand, there is another method in which, instead of relying solely on subjective evaluations, they hire monitors who can more accurately observe workers' performances and also use their reports to determine compensation. It is common in society to hire another expert to report on the quality of technical products or to have supervisors report on the subordinaries' performances. In most cases, reporting the quality of the product or the extent of the worker's effort is not as costly a task as producing the product or putting forth the effort. Hence, the monitor seems beneficial to the employer since she can obtain more accurate information almost for free.

However, it is unclear whether hiring a monitor is the better choice for the employer because the monitor may collude with the worker. For example, the monitor might receive money from the worker and report to the employer that the quality of the product is good, even though the actual quality is poor. In this case, the employer must pay additional compensation to prevent collusion, which costs the employer. In particular, if the employer can only use subjective evaluations, preventing collusion is expected to be a more troublesome problem than if objective ones are available. Thus, it is important to ask what kind of contract the employer should draft to prevent collusion and whether the monitor will benefit the employer even in light of the possibility of collusion.

In order to answer these questions, this paper analyzes two models: one in which a principal and an agent play, and one in which a monitor, in addition to them, also plays. Through these analyses, we obtain two main results. First, hiring the monitor allows the employer to write contracts where she does not have to pay the agent's rent and the costs inherent in subjective evaluation. Second, as a result, even allowing for the possibility of collusion, hiring the monitor is always optimal for the employer. We describe the characteristics of the model and the result below.

In the first analysis, we consider the game, played by the risk-neutral principal and agent, with minimum wage constraints. The agent chooses either a high effort or a low effort, and the principal obtains noisy and subjective signals about the action. Subjective signals here mean that they are private information of the principal, i.e., they are not observed by the agent or a third party. In common with the second model, we set the specific model to derive the optimal contract explicitly.

The characteristics of the optimal contracts in this model are that the principal gives the agent rent and burns money on the equilibrium path. The latter feature is specific to contracts with subjective evaluation and has already been mentioned in many other studies (MacLeod, 2003; Kambe, 2006). When signals about the action are private information for the principal, she can make false statements about the signals she receives. Thus, even if the principal receives a signal that suggests the agent's effort, she can claim to have received a signal indicating that he did not make an effort to keep the payment low. The rational agent would anticipate it and thus would not make efforts in the first place. To avoid this, the principal must convince the agent that she will report the truth. It is made possible by money burning. When the principal wants to vary the amount the agent receives depending on the signal to incentivize the agent, the principal pays a fixed amount regardless of the signal and commits to burning the differences. Since this burnt money becomes a literal cost that goes to no one, contracts based on subjective evaluations generally lead to inefficient consequences.

In the second model, we consider the game in which, in addition to the two players above, the risk-neutral monitor also plays: the monitor observes the agent's action or performance perfectly at no cost and sends a monitoring report to the principal. Hence, in this game, the principal can write the contracts that rely on monitoring reports from the monitor and the subjective signals she receives. However, we assume the monitor can manipulate the information she observes without cost and may collude with the agent. Therefore, the principal must draft the contracts that prevent collusion between the agent and the monitor, which we call the collusion-proof contracts.

We find that one of the optimal collusion-proof contracts in the model has the property that, on the equilibrium path, whether to reward the agent and monitor depends solely on the monitoring report. Specifically, the contract is characterized as follows: (1) The agent is rewarded if and only if the monitor reports a high effort, regardless of the principal's signal. (2) The monitor is rewarded when she reports a high effort, regardless of the signal; when she reports a low effort (this is off the equilibrium path), she is rewarded only if the principal receives the worst signal.

Finally, we compare the two games above and show that hiring the monitor reduces the principal's payment because the optimal collusion-proof contracts allow the principal and monitor to reduce the agent's rent and the money the principal would burn to zero. As noted above, in the optimal collusion-proof contract, the monitoring report alone determines whether to reward each player on the path, so the principal's noisy subjective signal does not affect this. Thus, there is no need for the principal to pay rent to the agent, and to burn money since it eliminates the problems inherent in the subjective evaluation discussed above. The latter is remarkable. The literature on contracts based on objective evaluations stated that the delegation of the monitoring to the monitor is profitable, mainly because the principal can recover the agent's rent (Tirole, 1993; Strausz, 1997). On the other hand, our study shows that, under subjective evaluation, hiring the monitor can also revive burnt money and expand the economic pie, which suggests that monitors play vital roles in contracts with subjective evaluations.

This study is related to the literature on monitoring and collusion in organizations (Tirole, 1986, 1993; Strausz, 1997; Che et al., 2021). Some of this literature is associated with subjective evaluations or unverifiable signals, of which Deb et al. (2016) and Thiele (2011) are the closest to our study.

Deb et al. (2016) study the optimal dynamic contract with multiple agents where compensation is based on public performance signals and privately reported peer evaluations. Deb et al. (2016) analyze a static model by the monitor (supervisor) and agent as a baseline model and shows that the principal achieves the first-best outcome. There are two main differences between our study and this one. First, they rule out the possibility of collusion throughout the paper, while our study considers the possibility of collusion. Deb et al. (2016) assume that the monitor receives private signals, including noise, about the agent's action, and it is generally challenging to consider the collusion problem in such a model. On the other hand, our study assumes that the monitor can perfectly observe the agent's action, and thus the agent knows what the monitor sees. Therefore, our model allows to study the collusion problem between the agent and monitor. Second, their study does not take into account the minimum wage constraints. As this study clarifies, the minimum wage constraints are an essential assumption in discussing the use of monitors. Due to the constraints, our analysis shows that hiring the monitor is beneficial for the principal, but does not achieve first-best.

On the other hand, Thiele (2011) studies the optimal relational contract in games played by potentially colluding monitor and agent. The difference between ours and this one is the quality of the signal. Thiele (2011) assumes that signals are unverifiable but shared among players and hence addresses this as a long-term relationship problem. In our study, we assume that the information is completely private for the principal and monitor.

Also, Section 2 of our study addresses the model of subjective evaluation without the monitor as a baseline model, and related to this are MacLeod (2003) and Kambe (2006). MacLeod (2003) studies the contract with subjective evaluations without minimum wage constraints and shows that signal compression occurs in the optimal contract. In particular, MacLeod (2003) shows that the optimal contracts have the properties that the agent is punished when the worst signal is realized and he is rewarded when all other signals are realized. On the other hand, Kambe (2006) works on the contract with subjective evaluations with minimum wage constraints and shows that signal compressions do occur in the case of the existence of minimum wage constraints, but there is not always just one signal for punishment.

Since we are working on the model with minimum wage constraints, we should clarify the similarities and differences between our results without a monitor and ones in Kambe (2006). Kambe (2006) analyzes a general setting and derives general characteristics of optimal contracts in subjective evaluations. Our model without a monitor is a special case of the general model of Kambe (2006), and in fact our results in Subsection 2.2 are special case of the Kambe's result. However, Kambe (2006) is not explicit about the specific features of the optimal contract in our model. Thus, our paper contributes to the literature on optimal contracts in subjective evaluations by explicitly deriving optimal contracts with discrete signals.

This paper demonstrates the positive effect of introducing an additional player in a subjective evaluation environment. A study showcasing a similar effect is Ishiguro and Yasuda (2023). Ishiguro and Yasuda (2023) show that in situations involving multiple agents in subjective evaluations, the efficiency loss associated with subjective assessments, which would exist in a single-agency optimal contract, disappears. In contrast, this paper establishes that the elimination of efficiency loss in the optimal contract occurs even when adding a monitor to a model with a single agent.

The rest of the paper is organized as follows. Section 2 studies the contract without a monitor as a benchmark. Section 3 studies the contract with a monitor. Section 4 examines the role of the monitor under subjective evaluation by comparing the results of the two games. Section 5 concludes and discusses extensions.

2 The Game without the monitor

2.1 The Model

First, as a benchmark, we study a game played by only two risk-neutral players: a principal and an agent[1]. As mentioned in the introduction, we consider a situation where the principal can only evaluate the agent's action on subjective measures. Since subjective measures are unverifiable and unobservable, this model has a different setup from models of objective evaluations. Including this explanation, the following will discuss the environment, contract, and payoff.

The principal offers a contract to an agent to perform a task. If he accepts it, he chooses an action $a \in \{a_l, a_h\}$, where a_l means a low effort or performance and a_h means a high effort or performance. If the agent chooses the action a_i, the cost of effort is c_i, and we assume $c_h > c_l > 0$. When the agent chooses the action a, he can produce the output $Y(a)$. The output $Y(a)$ takes time to be determined and is unknown until the end of the game[2]. The principal cannot observe the agent's action or the output but receives a signal $k \in \{1, \dots, K\}$, whose probability depends on his action. When the agent chooses a_i, the principal receives the signal k with probability $P(k|a_i)$. We assume that the probability distributions $P(\cdot|a_h)$ and $P(\cdot|a_l)$ have the full supports, $P(k|a) > 0$ for any a and any k, and

[1] As noted in the introduction, the model in the section is a special case of Kambe (2006). Especially, Lemmas 1 through 3 in our paper are subsumed by Lemma 2 in Kambe (2006). However, whereas Kambe (2006) characterizes optimal contracts from these lemmas by approximating the signals as continuous, our study explicitly characterizes optimal contracts with discrete signals and minimum wage constraints in Propositions 1 and 2.

[2] We can also assume that the output is Y_k, which depends on the subjective signal the principal receives. In this case, the output Y_k is observable for the principal but not the others.

satisfy the monotonicity of the likelihood ratio,

$$\frac{P(k+1|a_h)}{P(k+1|a_l)} > \frac{P(k|a_h)}{P(k|a_l)}$$

for any $k \in \{1, \ldots, K-1\}$.

While the probability distributions are common knowledge, we assume that the signal that the principal receives is her subjective evaluation, which implies that the others cannot observe or verify it. Thus, the signal itself is not contractable information. We also assume that the agent and a third party do not get signals that correlate with the principal's evaluation.

Now, we define the contract. Since the signal is the principal's private information, she cannot draft the contract which depends on it. Then, we assume that she makes a report r after she receives the signal. The principal determines her payment and the agent's receipt separately. Let $\{v_r, w_r\}$ be the contract, where v_r is the amount of money the principal pays and w_r is the amount of money the agent receives when she reports r, and we assume that the contract satisfies the minimum wage constraints, which implies $v_r \geq w_r \geq 0$ for all r. We also assume that the principal can commit to burn the difference $v_r - w_r \geq 0$.

We define the principal's report strategy as the function $r(\cdot)$ whose domain and range are both $\{1, 2, \ldots, K\}$. The image $r(k)$ is the report of the principal when she receives the signal k. When the agent chooses the action a and the principal chooses the report stragtegy $r(\cdot)$ and receives the signal k, the agent's payoff is $u_A = w_{r(k)} - c(a)$ and the principal's payoff is $u_P = Y(a) - v_{r(k)}$. Similarily, when the agent chooses the action a and the principal chooses the report strategy $r(\cdot)$, the expected payoff functions are given by,

$$Eu_A = \sum_{k=1}^{K} P(k|a) w_{r(k)} - c(a),$$

$$Eu_P = Y(a) - \sum_{k=1}^{K} P(k|a) v_{r(k)},$$

respectively. If the agent rejects the contract, he obtains the reservation utility normalized to 0.

The timing of the game is as follows:

- The principal offers the contract $\{v_r, w_r\}$ to the agent.
- The agent decides whether to accept it. If the agent rejects it, he gets the reservation utility 0, ending the game.
- The agent chooses his action a.
- The principal receives the signal k and makes a report $r = r(k)$. Then, the principal pays v_r and the agent receives w_r.

2.2 The Analysis of the Game

In this subsection, we analyze the game defined above. We restrict our analysis to contracts that elicit good performance a_h from the agent. Since the model is specific, we can derive a closed-form solution, as we will see below. We proceed in the following order: the formulation of the cost minimization problem, the derivation of the lemmas indicating the properties of the optimal contract, and the derivation of the optimal contract.

First, we formulate the cost minimization problem. By the revelation principle, we can focus on equilibria that report $r(k) = k$ for any k without loss of generality. Furthermore, for the principal to report truthfully and minimize her payment, for any k, $v_k = \max_{k'} w_{k'}$ must hold. If $\min_{k'} v_{k'} < \max_{k'} w_{k'}$, then the principal would report $k \in \arg\min_{k'} v_{k'}$ regardless of the actual signal. If $\min_{k'} v_{k'} > \max_{k'} w_{k'}$, then the principal would minimize her payment by changing each v_k to $\max_{k'} w_{k'}$.

Thus, we can concentrate on $\{w_k\}$ instead of $\{v_k, w_k\}$, with $v_k = \max_{k'} w_{k'}$ for each k.

The problem can be formulated as follows:

$$\min \max_{k'} w_{k'},$$

$$\sum_{k=1}^{K} P(k|a_h)w_k - c_h \geq 0, \quad (1)$$

$$\sum_{k=1}^{K} \Delta P_k w_k - \Delta c \geq 0, \quad (2)$$

$$w_k \geq 0, \forall k \in \{1, ..., K\}, \quad (3)$$

where $\Delta P_k = P(k|a_h) - P(k|a_l)$ for any k and $\Delta c = c(a_h) - c(a_l)$. The constraint (1) is the individual rationality condition (IR condition), the constraint (2) is the incentive compatibility condition (IC condition), and the constraints (3) are the minimum wage conditions.

Before the derivation of the optimal contract, we state the necessary conditions for $\{w_k\}$ to be the optimal contract. In the following, the smallest k that satisfies $P(k|a_h) \geq P(k|a_l)$ is denoted by k^*.

Lemma 1 *If $\{w_k\}$ is an optimal contract, the IC condition is binding.*

Proof. Suppose that a contract $\{w_k\}$ would satisfy the IC condition strictly. Then, we can always reduce the principal's payment while keeping all constraints hold by setting the new contract $\{w_k^*\}$ to, for an arbitrarily small $\varepsilon > 0$,

$$w_k^* = \begin{cases} w_k - \frac{\varepsilon}{\sum_{k \in \overline{k}} P(k|a_h)} & (k \in \overline{k}) \\ w_k + \frac{\varepsilon}{\sum_{k \notin \overline{k}} P(k|a_h)} & (k \notin \overline{k}) \end{cases},$$

where $\overline{k} = \arg\max_{k'} w_{k'}$. Note that, since the IC condition is satisfied strictly, w_k for any $k \in \overline{k}$ is strictly positive and there always exists a constant ε such that $w_k^* \geq 0$ for any k. ■

Lemma 2 *If $\{w_k\}$ is an optimal contract, then $w_k = \max_{k'} w_{k'}$ for any k such that $P(k|a_h) > P(k|a_l)$.*

Proof. Suppose that in an optimal contract, there would exist some k satisfying both $P(k|a_h) > P(k|a_l)$ and $w_k < \max_{k'} w_{k'}$. In this case, if we increase the corresponding w_k by $\max_{k'} w_{k'} - w_k$, we can satisfy the IC condition strictly, keeping the principal's payment unchanged and the IR condition hold. Therefore, from Lemma 1, this contradicts the assumption that $\{w_k\}$ is an optimal contract. ■

Lemma 3 *If $\{w_k\}$ is an optimal contract, then there is at most one κ satisfying*

$$0 < w_\kappa < \max_{k'} w_{k'}.$$

Furthermore, $w_k = \max_{k'} w_{k'}$ for all $k > \kappa$ and $w_k = 0$ for all $k < \kappa$.

Proof. Suppose otherwise in an optimal contract $\{w_k\}$. Then, there would exist s and $t < s$ such that $w_s < \max_{k'} w_{k'}$ and $w_t > 0$. Note that $k^* \geq s > t$ from Lemma 2. Hence, $\Delta P_s \leq 0$ and $\Delta P_t < 0$. Let ε_s and ε_t be positive constants satisfying

$$\frac{\Delta P_s}{\Delta P_t} < \frac{\varepsilon_t}{\varepsilon_s} \leq \frac{P(s|a_h)}{P(t|a_h)},$$

$$\varepsilon_s \leq \max_{k'} w_{k'} - w_s,$$

$$\varepsilon_t \leq w_t.$$

Increasing w_s by ε_s and decreasing w_t by ε_t can make the IC condition satisfy strictly while keeping the principal's payment unchanged and the IR condition hold. Therefore, from Lemma 1, this contradicts the assumption that $\{w_k\}$ is an optimal contract. ■

From Lemmas 2 and 3, the optimal contract has two or three values. As with MacLeod (2003) and Kambe (2006), these lemmas show that optimal contracts with subjective evaluations compress variations in payment. Dividing the evaluation too finely increases the expected amount of burnt money and the principal's payment. Therefore, the optimal contract distinguishes about only two types of signals: the signal that the agent works hard and the signal that the agent shrinks.

Given these properties, we solve the problem by dividing it into cases based on whether the IR condition is binding or not.

Proposition 1 *Suppose that*

$$\sum_{k \geq k^*} P(k|a_l)c_h \geq \sum_{k \geq k^*} P(k|a_h)c_l \tag{4}$$

holds. If $P(k^*|a_h) \neq P(k^*|a_l)$*, the optimal contract is*

$$(w_1^*, ...w_{k^*-1}^*, w_{k^*}^*, ..., w_K^*) = \left(0, ..., 0, \frac{\Delta c}{\sum_{k \geq k^*} \Delta P_k}, ..., \frac{\Delta c}{\sum_{k \geq k^*} \Delta P_k}\right).$$

If $P(k^*|a_h) = P(k^*|a_l)$*, then the optimal contract is one in which* $w_{k^*}^*$ *satisfies*

$$\max\left\{0, \frac{\sum_{k>k^*}\{P(k|a_h)c_l - P(k|a_l)c_h\}}{P(k^*|a_h)\sum_{k>k^*} \Delta P_k}\right\} \leq w_{k^*}^* \leq \frac{\Delta c}{\sum_{k>k^*} \Delta P_k},$$

and the others are the same as above.

Proof. We show that the contract in Proposition 1 is the solution to the optimization problem without the IR condition, and yet it satisfies the IR condition when the inequality (4) holds.

First, suppose $P(k^*|a_h) \neq P(k^*|a_l)$. We solve the optimization problem without the IR condition. From Lemma 2, we must set w_k for any $k \geq k^*$ to $\max_{k'} w_{k'}$. In addition, w_k for any $k < k^*$ should not be strictly positive, because we can relax the IC condition by changing it to $w_k - \varepsilon$ for sufficiently small ε, which is contrary to Lemma 1. Thus, in order to minimize the payment while satisfying the IC condition, it is desirable to set w_k for any $k < k^*$ to 0 and w_k for any $k \geq k^*$ to $\max_{k'} w_{k'}$. Since the IC condition is established by equal in the optimal contract, solving this equality based on the above yields

$$(w_1, ..., w_{k^*-1}, w_{k^*}, ..., w_K) = \left(0, ..., 0, \frac{\Delta c}{\sum_{k \geq k^*} \Delta P_k}, ..., \frac{\Delta c}{\sum_{k \geq k^*} \Delta P_k}\right).$$

This is the optimal contract, excluding the IR condition, and is the same as the contract in Proposition 1.

Furthermore,

$$\begin{aligned}\sum_{k=1}^{K} P(k|a_h)w_k - c_h &= \frac{\sum_{k\geq k^*} P(k|a_h)\Delta c}{\sum_{k\geq k^*} \Delta P_k} - c_h \\ &= \frac{\sum_{k\geq k^*} \{P(k|a_l)c_h - P(k|a_h)c_l\}}{\sum_{k\geq k^*} \Delta P_k} \\ &\geq 0.\end{aligned}$$

Thus, the contract in Proposition 1 satisifies the IR condition when the inequality (4) holds.

Next, suppose $P(k^*|a_h) = P(k^*|a_l)$. Hence $\sum_{k>k^*} \Delta P_k = \sum_{k\geq k^*} \Delta P_k$. In this case, the value of w_{k^*} does not affect the IC condition. Thus, the contract obtained above, in which only w_{k^*} is changed to satisfy $0 \leq w_{k^*} \leq \Delta c / \sum_{k>k^*} \Delta P_k$, is still the solution to the optimization problem wihtout the IR condition. Therefore, the contract in Proposition 1 is the optimal contract without the IR condition. Furthermore, when w_{k^*} satisfies the range described in Proposition 1,

$$\begin{aligned}\sum_{k=1}^{K} P(k|a_h)w_k - c_h &= \frac{\sum_{k>k^*} P(k|a_h)\Delta c}{\sum_{k>k^*} \Delta P_k} + P(k^*|a_h)w_{k^*} - c_h \\ &= \frac{P(k^*|a_h)w_{k^*} \sum_{k>k^*} \Delta P_k - \sum_{k>k^*}\{P(k|a_h)c_l - P(k|a_l)c_h\}}{\sum_{k>k^*} \Delta P_k} \\ &\geq 0.\end{aligned}$$

Thus, the contract in Proposition 1 satisifies the IR condition. ■

Since the signals k^* or more signals indecate a high effort, it is plausible to make only w_k satisfying $k \geq k^*$ positive in the optimal contract. As explained in the lemmas, a distinctive feature of contracts with subjective evaluation is that in order to minimize the amount of burnt money, all w_k satisfying $P(k|a_h) > P(k|a_l)$ are made constant, eliminating wage variations.

Note the effect of the minimum wage conditions. As in the case of objective evaluations, the minimum wage constraints require the principal to pay the agent rent even if he is risk-neutral. In addition, there are circumstances unique to subjective evaluations. As shown in MacLeod (2003), in the absence of the minimum wage constraints, it is optimal to reward when $k \geq 2$ and to give a considerable punishment only when $k = 1$ to minimize the amount of burnt money. However, the minimum wage constraints require $w_1 \geq 0$, so the principal cannot do it. Therefore, in a case where the inequality (4) holds and it is difficult to make the agent work hard, the difference between reward and punishment must be significant enough, then the principal should treat signals that satisfy $1 \leq k < k^*$ as punishment signals. Hence, the amount of burnt money is more significant when there are minimum wage constraints than when there are no minimum income constraints.

Proposition 2 *Suppose that*

$$\sum_{k \geq k^*} P(k|a_l)c_h < \sum_{k \geq k^*} P(k|a_h)c_l \tag{5}$$

holds. Denote by κ *the largest of the* k' *satisfying* $\sum_{k \geq k'} P(k|a_l)c_h \geq \sum_{k \geq k'} P(k|a_h)c_l$. *The optimal contract is*

$$\begin{aligned} \underline{w} &= 0, \\ w_\kappa &= \frac{\sum_{k \geq \kappa+1}\{P(k|a_h)c_l - P(k|a_l)c_h\}}{P(\kappa|a_l)\sum_{k \geq \kappa+1} P(k|a_h) - P(\kappa|a_h)\sum_{k \geq \kappa+1} P(k|a_l)}, \\ \overline{w} &= \frac{c_h P(\kappa|a_l) - c_l P(\kappa|a_h)}{P(\kappa|a_l)\sum_{k \geq \kappa+1} P(k|a_h) - P(\kappa|a_h)\sum_{k \geq \kappa+1} P(k|a_l)}, \end{aligned}$$

where $\underline{w} \equiv w_k$ *for any* $k < \kappa$, *and* $\overline{w} \equiv w_k$ *for any* $k > \kappa$.

Proof. If the inequality (5) holds, since the solution to the optimization problem without the IR condtion does not satisfy the IR condition, both the IR and IC conditions are binding in the optimal contracts. Thus, from Lemmas 2 and 3, there exist $\overline{w}$, k' and $w_{k'}$ such that

$$\overline{w} \sum_{k \geq k'+1} \Delta P_k + \Delta P_{k'} w_{k'} - \Delta c = 0, \tag{6}$$

$$\overline{w} \sum_{k \geq k'+1} P(k|a_h) + P(k'|a_h) w_{k'} - c_h = 0, \tag{7}$$

$$\overline{w} \geq w_{k'} > 0. \tag{8}$$

Solving the equations (6) and (7), we get

$$w_{k'} = \frac{\sum_{k \geq k'+1} \{P(k|a_h)c_l - P(k|a_l)c_h\}}{P(k'|a_l)\sum_{k \geq k'+1} P(k|a_h) - P(k'|a_h)\sum_{k \geq k'+1} P(k|a_l)}, \tag{9}$$

$$\overline{w} = \frac{c_h P(k'|a_l) - c_l P(k'|a_h)}{P(k'|a_l)\sum_{k \geq k'+1} P(k|a_h) - P(k'|a_h)\sum_{k \geq k'+1} P(k|a_l)}. \tag{10}$$

Finally, we only need to identify k'. From (8),(9) and (10), k' must satisfy

$$P(k'|a_l)c_h - P(k'|a_h)c_l \geq \sum_{k \geq k'+1} \{P(k|a_h)c_l - P(k|a_l)c_h\},$$

$$\sum_{k \geq k'+1} \{P(k|a_h)c_l - P(k|a_l)c_h\} > 0.$$

These inequalities imply that $\sum_{k \geq k'} P(k|a_l)c_h \geq \sum_{k \geq k'} P(k|a_h)c_l$ and $\sum_{k \geq k'+1} P(k|a_h)c_l > \sum_{k \geq k'+1} P(k|a_l)c_h$. Therefore, k' must be κ defined in the proposition. Note that κ always exists because the first inequality holds when $k' = 1$, and the second inequality holds when $k' = k^* - 1$. ∎

The need for the principal to burn money and the compression of evaluations are common to the contracts in Propositions 1 and 2. However, in the situation where it is easier to elicit the agent's effort than that of Proposition 1, it is optimal to set the threshold for rewarding and punishment signals to κ instead of k^*. Lowering the threshold of signals has the disadvantage of taking away the agent's incentive to work hard, but also has the advantage of lowering the expected amount of burnt money. In particular, when $\sum_{k \geq 2} P(k|a_l)c_h < \sum_{k \geq 2} P(k|a_h)c_l$

holds, i.e., when it is easiest to make the agent exert effort, the punishment signal can be only 1, as if there are no minimum wage constraints.

Furthermore, the optimal contract is 3-valued when it holds that the inequality (5), $\kappa \neq 1$ and $\sum_{k \geq \kappa} P(k|a_l)c_h > \sum_{k \geq \kappa} P(k|a_h)c_l$. This feature is not present in models with no minimum wage constraint (MacLeod, 2003) or continuous signals (Kambe, 2006). As already mentioned, when no minimum wage constraint is imposed, in order to make burnt money as small as possible, the optimal contract gives a considerable punishment only when $k = 1$ and rewards the other signals. Also, if the signals received by the principal are continuous, the optimal contract is binary, because the probability of realization of κ is infinitesimal and insignificant.

3 The Game with the Monitor

3.1 The Model

Next, we set up the game in which a risk-neutral monitor participates. The monitor observes the agent's action perfectly at no cost and reports it to the principal, though she potentially colludes with the agent. Since introducing a monitor needs another different setup, this subsection mainly describes the changes from the previous game.

The principal hires a risk-neutral monitor in addition to the agent. After the agent chooses the action, the monitor perfectly observes it without cost and sends a monitoring report $m(a) \in \{l, h\}$ to the principal. We assume that the information the monitor gets is her private information, and she can make false monitoring reports costlessly. Note that the agent knows what the monitor observes because he knows that the monitor can perfectly observe his action. Although the information the monitor gets is private, the monitoring reports are verifiable, and the principal cannot rewrite or hide them.

In addition, we assume that the monitor and the agent can collude. After the monitor observes the action, the agent proposes a side contract $\{z_{rm}\}$ to the monitor. The monitor can choose to accept the side contract or not, and if she accepts, according to the monitoring message m, she gets z_{rm} from the agent at the end of the game.

In this game, the principal receives two types of information: monitoring reports and signals. Thus, the principal can draft the contract conditioning the monitoring report, in addition to the report of the previous game. We denote the amount paid by the principal as v_{rm}, the amount received by the agent as w_{rm}, and the amount received by the monitor as x_{rm} when the principal's report is r and the monitoring report is m.

The monitor's payoff function is $u_M = x_{rm} + z_{rm}$. Since she is risk-netutral, the expected payoff function is $Eu_M = \sum_{k=1}^{K} P(k|a)\{x_{rm} + z_{rm}\}$. Her reservation payoff is also normalized to 0.

The timing of the game is as follows:

- The principal offers the contract $\{v_{rm}, w_{rm}, x_{rm}\}$ to the agent and monitor.
- The agent and monitor decide whether to accept it. If either of them rejects it, they get the reservation payoff, ending the game.
- The agent chooses his action a. At this time, the monitor observes his action.
- The agent proposes the side contract $\{z_{rm}\}$ to the monitor.
- The monitor decides whether or not to accept the side contract, and she sends the monitoring report $m = m(a)$ to the principal. If she accepted the side contract at the previous stage, the agent pays her according to $\{z_{rm}\}$
- The principal receives the signal k and makes the report $r = r(k, m)$. Then she pays v_{rm}, and the agent and monitor receive w_{rm} and v_{rm}, respectively.

3.2 The Analysis of Game

In this subsection, we derive the optimal contract under the condition that each player makes an honest report, preventing collusion between the monitor and the agent. As in the previous game, we proceed in the following order: formulation of the cost minimization problem, derivation of lemmas indicating the properties of the optimal collusion-proof contract, and derivation of the optimal collusion-proof contract.

First, we formulate the cost minimization problem. We restrict our analysis to contracts that all parties tell the truth and the agent chooses a_h. As in the case without the monitor, $v_{r(k)m} = \max_{k'}(w_{r(k')m} + x_{r(k')m})$ must hold for any k and m so that $r(k,m) = k$. In addition, since we consider the equilibrium that the agent chooses a_h and the monitor tells the truth, the principal's payment is $\max_{k'}(w_{k'h} + x_{k'h})$. Therefore, the problem to be solved can be formulated as follows.

$$\min \max_{k'}(w_{k'h} + x_{k'h}),$$

$$\sum_{k=1}^{K} P(k|a_h)w_{kh} - c_h \geq 0, \tag{11}$$

$$\sum_{k=1}^{K} P(k|a_h)w_{kh} - c_h \geq \sum_{k=1}^{K} P(k|a_l)w_{kl} - c_l, \tag{12}$$

$$\sum_{k=1}^{K} P(k|a_h)x_{kh} \geq \sum_{k=1}^{K} P(k|a_h)x_{kl}, \tag{13}$$

$$\sum_{k=1}^{K} P(k|a_l)x_{kl} \geq \sum_{k=1}^{K} P(k|a_l)x_{kh}, \tag{14}$$

$$\sum_{k=1}^{K} P(k|a_h)\{w_{kh} + x_{kh}\} \geq \sum_{k=1}^{K} P(k|a_h)\{w_{kl} + x_{kl}\}, \tag{15}$$

$$\sum_{k=1}^{K} P(k|a_l)\{w_{kl} + x_{kl}\} \geq \sum_{k=1}^{K} P(k|a_l)\{w_{kh} + x_{kh}\}, \tag{16}$$

$$w_{km}, x_{km} \geq 0, \forall k \in \{1, ..., K\}, \forall m \in \{l, h\}. \tag{17}$$

The constraints (11) and (12) are the agent's IR and IC conditions, respectively, based on the premise that the monitor sends the truth monitoring report. The constraints (13) and (14) are the conditions for the monitor not to lie without collusion. The constraints (15) and (16) represent that the sum of the payoffs of the agent and monitor if the monitoring report is accurate is greater than or equal to the sum of the payoffs if the monitoring report is false, which is to prevent collusion between them. The constraints (17) are the minimum wage conditions.

Before going into the main issue, we should mention the optimal contract in two special cases as references. First, we derive the optimal contract for the situation where collusion is exogenously prohibited. When there is no possibility of collusion, the constraints (15) and (16) can be ignored, leaving only equations (13), (14), and (17) for the monitor's wage. These constraints can be satisfied by setting $x_{km} = 0$ for all k and m. Thus, the principal can have the monitor report honestly without paying any compensation. Furthermore, the constraints (11) and (12) for the agent's wage can be satisfied by setting $w_{kh} = c_h$ and $w_{kl} = c_l$ for any k. In other words, if there is no possibility of collusion between the monitor and the agent, the principal pays only c_h for the agent to choose a_h and can achieve the first-best outcome by hiring the monitor.

Second, we consider the optimization problem above with the additional constraints $x_{kh} = 0$ for all k, which we call NPM (no payment for the monitor) problem. Since $x_{kl} = 0$ for any k from the constraint (13) holds obviously, NPM problem is one whose variables are only $\{w_{km}\}$, and we denote the solution of this problem by $\{w_{km}^{NPM}\}$. The following lemma holds for $\{w_{km}^{NPM}\}$.

Lemma 4 $\{w_{kh}^{NPM}\}$ *is the equivalent of* $\{w_k\}$ *defined in Propositions 1 and 2, therefore the principal's payment in any solutions of NPM problem is equal to the payment in the game without the monitor.*

Proof. Since we can derive the constraint (2) from the constraints (12) and (16), $\{w_{kh}^{NPM}\}$ must satisfy the constraints (1), (2) and (3). Thus, if all the other constraints can be satisfied by adjusting $\{w_{kl}\}$, $\{w_{kh}^{NPM}\}$ must be the equivalent of the contract defined in Propositions 1 and 2. In fact, for any w_{kh}, there always exists w_{kl} that satisfies $\sum_{k=1} P(k|a_h)w_{kh} \geq \sum_{k=1} P(k|a_h)w_{kl}$ and $\sum_{k=1}^{K} P(k|a_l)w_{kl} = \sum_{k=1}^{K} P(k|a_l)w_{kh}$ simultaneously (for example, $w_{kl}^{NPM} = w_{kh}^{NPM}$ for any k), so this lemma is valid. ■

This lemma implies that even if there is a possibility of collusion, the employment of the monitor is not a loss for the principal. Since contracts feasible in NPM problem are also feasible in the original problem, the principal can achieve the same gain as she does not hire the monitor by paying only to the agent.

We now derive the solutions to the original optimization problem and examine whether an even better contract exists than the solution to NPM problem. To derive the optimal collusion-proof contracts, we relax the problem as follows. First, let the objective function be $\min(\max_{k'} w_{k'h} + \max_{k'} x_{k'h})$ instead of $\min \max_{k'}(w_{k'h} + x_{k'h})$. Note that the former is equal to or larger than the latter, and further that if w_{kh} and x_{kh} are both monotonically increasing for k, then these two values coincide. That is, if w_{kh} and x_{kh} are both monotonically increasing for k in the solution to the problem where the objective function is relaxed in this sense, then it is a solution to the original problem. Also, as in ordinary contract theory, we once ignore the constraints (14) and (15) in deriving the optimal collusion-proof contract. As shown in the proof of Proposition 3 below, the contract derived in this way may be too broad to satisfy these constraints, so we adjust it to satisfy them in the end.

We list some properties of the solutions to the relaxed optimizeation problem.

Lemma 5 *If $\{w_{km}, x_{km}\}$ is a solution to the relaxed optimization problem, then the constraints (12), (13) and (16) are binding.*

Proof. First, we prove the constraint (13) must be binding in the optimal contract. If the constraint (13) were not binding, by slightly decreasing x_{kh} for any $k \in \arg\max_{k'} x_{k'h}$, we could reduce the value of the relaxed objective function without violating the other constraints. Therefore, if $\{w_{km}, x_{km}\}$ is a solution to the relaxed problem, then the constraint (13) must be binding.

Next, we prove the constraint (16) must be binding in the optimal contract. If the constraint (16) were not binding and if $x_{kl} > 0$ for some k, we could make the constraint (13) not binding by slightly decreasing it. Then, as described above, we can reduce $\max_{k'} x_{k'h}$

without violating the other constraints, which implies we can reduce the value of the relaxed objective function. If $x_{kl} = 0$ for any k, $x_{kh} = 0$ for any k because (13) is binding. From Lemma 4, any solution with $x_{kh} = 0$ for any k has $\{w_{kh}\}$ satisfying (2) with equality. However, substituting (16) to (12), we would have (2) not binding. Therefore, if $\{w_{km}, x_{km}\}$ is a solution to the relaxed problem, then the constraint (16) must be binding.

Finally, we prove the constraint (12) must be binding in the optimal contract. Note that $\sum_{k=1}^{K} P(k|a_l)w_{kl}$ should be greater than or equal to c_l. This is because if $\sum_{k=1}^{K} P(k|a_l)w_{kl} < c_l$, then by increasing w_{kl} slightly, we could make the constraint (16) satisfy stirctly without violating the other constraints, which implies that we can reduce the value of the relaxed objective function in the way shown above. Thus, $\sum_{k=1}^{K} P(k|a_l)w_{kl} \geq c_l$ and the constraint (11) can be ignored. Therefore, if the constraint (12) were not binding, we could decrease $\max_{k'} w_{k'h}$ while keeping the constraints (12) and (16) valid. ■

Lemma 6 *If* $\{w_{km}, x_{km}\}$ *is a solution to the relaxed optimization problem, then for any* $k \geq 2$, $x_{kl} = 0$ *and* $x_{kh} = \max_{k'} x_{k'h}$.

Proof. Suppose that there would exist an optimal contract satisfying $x_{kl} > 0$ for some $k \neq 1$. We arbitrarily take one corresponding k, and let ε_k and ε_1 be positive constants satisfying

$$\frac{P(k|a_l)}{P(1|a_l)} \leq \frac{\varepsilon_1}{\varepsilon_k} < \frac{P(k|a_h)}{P(1|a_h)},$$

$$\varepsilon_k \leq x_{kl}.$$

Increasing x_{1l} by ε_1 and decreasing x_{kl} by ε_k can make the constraint (13) satisfy strictly while keeping the value of the relaxed objective function unchanged and all conditions hold. Therefore, from Lemma 5, it contradicts the assumption that the given contract is optimal.

Next, suppose that there would exist an optimal contract satisfying $\max_{k'} x_{k'h} > x_{kh}$ for some $k \neq 1$. We arbitrarily take one corresponding k, and let ε_{kh} and ε_{1l} be positive constants satisfying

$$\frac{P(k|a_l)}{P(1|a_l)} \leq \frac{\varepsilon_{1l}}{\varepsilon_{kh}} < \frac{P(k|a_h)}{P(1|a_h)},$$

$$\varepsilon_{kh} \leq \max_{k'} x_{k'h} - x_{kh}.$$

Increasing x_{kh} by ε_{kh} and x_{1l} by ε_{1l} can make the constraints (13) satisfy strictly while keeping the value of the relaxed objective function unchanged and all conditions hold. Therefore, from Lemma 5, it contradicts the assumption that the given contract is optimal. ■

Lemma 7 *If $\{w_{km}, x_{km}\}$ is a solution to the relaxed optimizaiton problem, then any contracts $\{w'_{km}, x'_{km}\}$ that modify only x_{1h} and x_{1l} to satisfy $0 \leq x'_{1h} \leq \max_{k'} x_{k'h}$ and $x'_{1l} = x_{1l} + (x'_{1h} - x_{1h})$ are also solutions.*

Proof. The value of $x_{1h} - x_{1l}$, not x_{1h} or x_{1l}, determines whether or not the constraints (13) and (16) are satisfied. Besides, x'_{1l} defined above is always greater than or equal to 0 because $x_{1l} \geq x_{1h}$ from the constraint (13) and Lemmas 5 and 6. Thus, if $\{w_{km}, x_{km}\}$ satisfies all constraints, then $\{w'_{km}, x'_{km}\}$ also satisfies all constraints. Furthermore, x_{1h} and x_{1l} do not affect the value of the relaxed objective function as long as $0 \leq x'_{1h} \leq \max_{k'} x_{k'h}$ is satisfied. Therefore, the lemma holds. ■

The variation in wages received by the monitor is binary from Lemma 6. In other words, as in the previous game, there is signal compression in this game. Lemma 7 also shows that there are multiple optimal collusion-proof contracts.

We obtain the following proposition by deriving the optimal collusion-proof contracts based on the above lemmas.

Proposition 3 *A contract $\{w^*_{km}, x^*_{km}\}$ is the optimal collusion-proof contract if and only if* *it satisfies these conditions:*

$$\overline{w}_h^* \sum_{k\geq 2} P(k|a_h) + P(1|a_h) w_{1h}^* = c_h, \tag{18}$$

$$\sum_{k=1}^{K} P(k|a_l) w_{kl}^* = c_l, \tag{19}$$

$$\sum_{k=1}^{K} P(k|a_h) w_{kl}^* \leq c_h, \tag{20}$$

$$\max\left\{0, \frac{\sum_{k\geq 2} P(k|a_h) c_l - \sum_{k\geq 2} P(k|a_l) c_h}{P(1|a_l) - P(1|a_h)}\right\} \leq w_{1h}^* \leq \overline{w}_h^*, \tag{21}$$

$$w_{kl}^* \geq 0,\ \forall k \in \{1, \ldots, K\}, \tag{22}$$

$$\overline{x}_h^* = \frac{P(1|a_h)\left\{\sum_{k\geq 2}^{K} P(k|a_l)\overline{w}_h^* + P(1|a_l) w_{1h}^* - c_l\right\}}{P(1|a_l) - P(1|a_h)}, \tag{23}$$

$$x_{1h}^* \in \left[0, \frac{P(1|a_h)\left\{\sum_{k\geq 2} P(k|a_l)\overline{w}_h^* + P(1|a_l) w_{1h}^* - c_l\right\}}{P(1|a_l) - P(1|a_h)}\right], \tag{24}$$

$$\underline{x}_l^* = 0, \tag{25}$$

$$x_{1l}^* = \frac{\{1 - P(1|a_h)\}\left\{\sum_{k\geq 2} P(k|a_l)\overline{w}_h^* + P(1|a_l) w_{1h}^* - c_l\right\}}{P(1|a_l) - P(1|a_h)} + x_{1h}^*, \tag{26}$$

where $\overline{w}_h$, $\overline{x}_h$ and $\underline{x}_l$ are w_{kh}, x_{kh} and x_{kl} for any $k \geq 2$, respectively.

Proof. We derive the optimal collusion-proof contract by the following procedure. First, we derive the optimal contract under the relaxed problem. Since it is clear from Lemma 7 that there is a plurality of solutions, we derive the contracts that satisfy $x_{kh} = \max_{k'} x_{k'h}$ for any k and then apply Lemma 7 to obtain all contracts. Finally, we find the contracts that satisfy the constraints (14) and (15) of those contracts, and check those are satisfying the monotoniciy on k.

First, we find the contracts that satisfy the conditions $x_{kh} = \max_{k'} x_{k'h}$ for any k. In the following, we denote by $\overline{x}_h = x_{kh}$ for any k. From Lemmas 5, 6 and 7 and the constraints

(12), (13) and (16), there exist $\overline{x}_h^*$, x_{1l}^* and $\{w_{kh}^*\}$ such that

$$\overline{x}_h^* = \frac{P(1|a_h)(\Delta c - \sum_{k=1}^K \Delta P_k w_{kh}^*)}{P(1|a_l) - P(1|a_h)}, \tag{27}$$

$$x_{1l}^* = \frac{\Delta c - \sum_{k=1}^K \Delta P_k w_{kh}^*}{P(1|a_l) - P(1|a_h)}. \tag{28}$$

We now consider $\{w_{km}^*\}$. We prove that the optimal contracts satisfy $w_{kh}^* = \max_{k'} w_{k'h}^*$ for any $k \geq 2$. Denote by $V(w_{kh})$ the value of the (relaxed) objective function when the contract to the agent is $\{w_{kh}\}$ and the equation (27) is satisfied. In other words,

$$V(w_{kh}) = \max w_{kh} + \max x_{kh} = \overline{w}_h + \frac{P(1|a_h)\left(\Delta c - \sum_{k \in \overline{k}} \Delta P_k \overline{w}_h - \sum_{k \notin \overline{k}} \Delta P_k w_{kh}\right)}{P(1|a_l) - P(1|a_h)},$$

where $\overline{k} = \arg\max_{k'} w_{k'}$and $\overline{w}_h = \max_{k'} w_{k'h}$.

Next, for a contract $\{w'_{kh}\}$, we define $\hat{w}$ as $\hat{w} = \sum_{k \in \overline{k}} P(k|a_h)\overline{w}'_h + \sum_{k \notin \overline{k}} P(k|a_h) w'_{kh}$, and consider the contract $\{\hat{w}_{kh}\}$ such that $\hat{w}_{kh} = \hat{w}$ for all k. Note that, as long as $\overline{x}_h^*$ is determined by the equation (27), if the contract $\{w'_{kh}\}$ satisfies all the constraints, then $\{\hat{w}_{kh}\}$ also satisfies all the constraints. Furthermore, for $\{w'_{kh}\}$ and $\{\hat{w}_{kh}\}$,

$$\begin{aligned} V(w'_{kh}) - V(\hat{w}_{kh}) &= \frac{\overline{w}'_h\{-\Delta P_1 - P(1|a_h)\sum_{k \in \overline{k}} \Delta P_k\} - P(1|a_h)\sum_{k \notin \overline{k}} \Delta P_k w'_{kh} + \Delta P_1 \hat{w}}{-\Delta P_1} \\ &= \frac{\sum_{k \notin \overline{k}} (\overline{w}'_h - w'_{kh})\{P(1|a_l)P(k|a_h) - P(1|a_h)P(k|a_l)\}}{-\Delta P_1} \\ &\geq 0 \end{aligned}$$

is satisfied. As for the last inequality, the equality sign is established if and only if $\overline{k} = \{1, 2, \ldots, K\}$ or $\overline{k} = \{2, 3, \ldots, K\}$. Therefore, we find that the optimal contracts satisfy $w_{kh}^* = \max_{k'} w_{k'h}^*$ for any $k \geq 2$.

Based on the above, we derive the conditions that $\{w_{km}^*\}$ must satisfy. From Lemma 5 and the constraint (12), for $\overline{w}_h^*$ and w_1^*,

$$\overline{w}_h^* = \frac{B - P(1|a_h)w_1^*}{\sum_{k\geq 2} P(k|a_h)},$$

where $B \equiv \sum_{k=1}^{K} P(k|a_l)w_{kl}^* + \Delta c$. Then the following holds:

$$\begin{aligned}
V(w_{kh}^*) &= \overline{w}_h^* + \frac{P(1|a_h)(\Delta c - \sum_{k\geq 2}\Delta P_k \overline{w}_h^* - \Delta P_1 w_1^*)}{P(1|a_l) - P(1|a_h)} \\
&= \frac{B - P(1|a_h)w_1^*}{\sum_{k\geq 2} P(k|a_h)} + \frac{P(1|a_h)\left\{\Delta c - \sum_{k\geq 2}\Delta P_k \frac{B - P(1|a_h)w_1^*}{\sum_{k\geq 2} P(k|a_h)} - \Delta P_1 w_1^*\right\}}{P(1|a_l) - P(1|a_h)} \\
&= \frac{P(1|a_h)\Delta c + B\sum_{k\geq 2}\Delta P_k}{-\Delta P_1}
\end{aligned}$$

From this result, we can find that $V(w_{kh}^*)$ depends only on the value of B as long as (12) is satisfied. Since $V(w_{kh}^*)$ is increasing with respect to B and $\sum_{k=1}^{K} P(k|a_l)w_{kl} \geq c_l$ must hold as shown in the proof of Lemma 5, it is desirable to set $\sum_{k=1}^{K} P(k|a_l)w_{kl}^* = c_l$.

From the above, we get the equations (18),(19) and (23) in the proposition. If we fix w_{1h}^*, $\overline{w}_h^*$ is determined by (18). This in turn determines $\overline{x}_h^*$ in (27). For $\overline{x}_h^*$ to be non-negative,

$$\frac{\sum_{k\geq 2} P(k|a_h)c_l - \sum_{k\geq 2} P(k|a_l)c_h}{P(1|a_l) - P(1|a_h)} \leq w_{1h}^*$$

must hold, and we obtain (21). Substituting the equation (18) into the equation (28) also yields

$$x_{1l}^* = \frac{\{1 - P(1|a_h)\}\left\{\sum_{k\geq 2} P(k|a_l)\overline{w}_h^* + P(1|a_l)w_{1h}^* - c_l\right\}}{P(1|a_l) - P(1|a_h)} + x_{1h}^*.$$

Applying Lemma 7 to this, we obtain the equations (24) and (26) of the proposition. Furthermore, we can easily check w_{kh}^* and x_{kh}^* are both monotonically increasing for k.

Finally, these are adjusted to satisfy the ignored constraints (14) and (15). The contract derived above clearly satisfies the constraint (14). On the other hand, substituting the above contract into the constraint (15), we obtain $\sum_{k=1}^{K} P(k|a_h)w_{kl}^* \leq c_h$. Therefore, adding this to the above conditions is the necessary and sufficient conditions for $\{w_{km}^*, x_{km}^*\}$ to be optimal.

■

Although there are multiple optimal collusion-proof contracts, the agent's expected payoff as well as the principal's expected payoff remains the same regardless of which contract is chosen that satisfies the conditions. On the other hand, the monitor's expected payoff can change depending on which contract is chosen. The contract that is optimal collusion-proof and maximizes the monitor's expected payoff (i.e., Pareto optimal) is uniquely determined for all but w_{kl}.

Corollary 1 *Of the optimal collusion-proof contracts, a Pareto optimal contract is*

$$w_{kh} = c_h, \tag{29}$$

$$w_{kl} = c_l, \tag{30}$$

$$x_{kh} = \frac{P(1|a_h)\Delta c}{P(1|a_l) - P(1|a_h)}, \tag{31}$$

for any k and

$$x_{1l} = \frac{\Delta c}{P(1|a_l) - P(1|a_h)}, \tag{32}$$

$$x_{kl} = 0 \tag{33}$$

for any $k \geq 2$.

Proof. As already pointed out, for any optimal collusion-proof contract, the expected payoffs of the principal and agent are invariant. Therefore, the contract that maximizes the expected payoff of the monitor is the Pareto-optimal optimal collusion-proof contract. Since the expected amount of compensation the monitor receives on the path is $\sum_{k=1}^{K} P(k|a_h)x_{kh}^*$, we can determine $\overline{x}_h^*$ and x_{1h}^* to maximize it. First, x_{1h}^* can be maximized by setting $x_{1h}^* = \overline{x}_h^*$

from inequality (24). Also, note that we can rewrite the equation (23) for

$$\overline{x}_h^* = \frac{P(1|a_h)(\Delta c - \sum_{k\geq 2} \Delta P_k \overline{w}_h^* - \Delta P_1 w_{1h}^*)}{P(1|a_l) - P(1|a_h)},$$

$\overline{x}_h^*$ is maximized when $\overline{w}_h^*$ is minimized and w_{1h}^* is maximized. From the equations (18) and (21), these conditions are satisfied when $\overline{w}_h^* = w_{1h}^* = c_h$ hold, and we get the contract defined in the corollary by substituting this to the equation (23). ■

The feature of the contract is that the payment to the agent and the monitor on the path does not depend on any signals, and the amount of the burnt money is zero. The contract with the agent is the same as when the principal can completely observe the agent's behavior, i.e. the first-best contract. This is because the principal can draft the contract conditional on the actual action of the agent through the monitor's report. Note that the signal received by the principal is less informative than the monitor's report, and the optimal collusion-proof contract above does not depend on the signals. This result is consistent with the Informativeness Principle (Holmström, 1979).

In the contract with the monitor, x_{kh} does not depend on the signal, which implies that he is always rewarded just by sending the monitoring report h. On the other hand, x_{kl} depends on whether the signal is the worst or not, and then the monitor is punished in most cases for sending the monitoring report l. This alone clearly makes it more attractive to the monitor to send the report h, so the wage when the principal receives the worst signal must be high. Such a contract can reduce the money burned on the equilibrium path to zero, including one to the agent. We have already mentioned that the evaluation is compressed to make the amount of burnt money as small as possible. In the optimal contract that includes the monitor, the variation is fully compressed on the equilibrium path, setting the compensation to a single value.

Contracts other than those defined in Corollary 1 can be characterized in comparison to them as follows. First, w_{kh} can differ in the reward amount between $k = 1$ and the other. In this case, the principal apparently loses money because she must burn money, but by making the difference here, it becomes easier to clear the constraint (16). Hence, $\overline{x}_h$, the amount paid to the monitor on the path, can be reduced. In other words, in this case, the increase in the amount of burnt money can be offset by the decrease in payments to the monitor.

Also, for x_{kh}, it is possible to differentiate between $k = 1$ and other signals. In this case, there will naturally be the burnt money again, but this can be adjusted by increasing x_{1l} that will not be paid on the path, as proved in Lemma 7. Thus, this is still the optimal contract for the principal.

4 The Effect of Hiring a Monitor

In this section, we examine the effects of hiring a monitor by comparing the results of the two games above. First, we show that the principal's payment in the collusion-proof contract obtained in Section 3 is not higher than the payment in the contract without the monitor obtained in Section 2, and we briefly discuss the causes of this difference.

In the following, the principal's payments when hiring the monitor and when not hiring the monitor are denoted by V_m and V_n, respectively. First, from Propositions 1, 2, and 3, the following corollary is immediately obtained.

Corollary 2 *V_m is less than or equal to V_n. The equal sign is established if and only if $\sum_{k\geq 2} P(k|a_l)c_h < \sum_{k\geq 2} P(k|a_h)c_l$.*

Proof. Suppose $\sum_{k \geq k^*} P(k|a_l)c_h > \sum_{k \geq k^*} P(k|a_h)c_l$. Then,

$$\begin{aligned}
V_n &= \frac{\Delta c}{\sum_{k \geq k^*} \Delta P_k} \\
&> \frac{\Delta c - \left(\sum_{k \geq k^*} P(k|a_l)c_h - \sum_{k \geq k^*} P(k|a_h)c_l\right)}{\sum_{k \geq k^*} \Delta P_k} \\
&= \frac{\left(1 - \sum_{k \geq k^*} P(k|a_h)\right) \Delta c}{\sum_{k \geq k^*} \Delta P_k} + c_h \\
&= \frac{\sum_{k < k^*} P(k|a_h)\Delta c}{-\sum_{k < k^*} \Delta P_k} + c_h \\
&\geq \frac{P(1|a_h)\Delta c}{P(1|a_l) - P(1|a_h)} + c_h \\
&= V_m.
\end{aligned}$$

Next, suppose $\sum_{k \geq k^*} P(k|a_l)c_h \leq \sum_{k \geq k^*} P(k|a_h)c_l$. Then,

$$\begin{aligned}
V_n &= \frac{c_h P(\kappa|a_l) - c_l P(\kappa|a_h)}{P(\kappa|a_l) \sum_{k \geq \kappa+1} P(k|a_h) - P(\kappa|a_h) \sum_{k \geq \kappa+1} P(k|a_l)} \\
&\geq \frac{c_h P(1|a_l) - c_l P(1|a_h)}{P(1|a_l) \sum_{k \geq 2} P(k|a_h) - P(1|a_h) \sum_{k \geq 2} P(k|a_l)} \\
&= \frac{P(1|a_h)\Delta c}{P(1|a_l) - P(1|a_h)} + c_h \\
&= V_m.
\end{aligned}$$

The inequality holds with the equal sign if and only if $\kappa = 1$, which implies that $\sum_{k \geq 2} P(k|a_l)c_h < \sum_{k \geq 2} P(k|a_h)c_l$. ■

The corollary shows that hiring the monitor is beneficial to the principal. The principal can draft the contract by hiring the monitor as if there were no minimum wage conditions. In the game without the monitor, the minimum wage constraints are hindrances in two ways. The first is that there are cases where the agent gets rent, i.e., solutions for which the IR condition is not binding. The second is specific to subjective evaluation, in that the principal wants to minimize the punishing signal as much as possible, but the minimum wage constraint prevents her from doing so. On the other hand, hiring the monitor allows the principal to solve these problems.

We indicate more clearly that hiring the monitor eliminates the problems associated with the minimum wage conditions. Consider the contract defined in Corollery 1 and let ΔEu_P, ΔEu_M, ΔEu_A, and ΔEBM denote the principal's payoff change, monitor's payoff change, agent's payoff change, and expected the amount of burnt money change between the two games, respectively. Since the expected payment amount of the principal should be equal to the sum of the expected receipts of the other players and the expected amount of burnt money, $\Delta Eu_P + \Delta Eu_M + \Delta Eu_A + \Delta EBM = 0$ always holds. In addition, the sign of each of them can be seen from a straightforward calculation to be $\Delta Eu_P \geq 0$, $\Delta Eu_A \leq 0$, $\Delta Eu_M > 0$, $\Delta EBM < 0$. This result indicates that the principal and monitor recover the agent's rent and the burnt money, which are caused mainly by the minimum wage conditions.

It is simply illustrated by the fact that the payment amount remains the same when $\sum_{k\geq 2} P(k|a_l)c_h < \sum_{k\geq 2} P(k|a_h)c_l$. When $\sum_{k\geq 2} P(k|a_l)c_h < \sum_{k\geq 2} P(k|a_h)c_l$, it is easy to make the agent work hard, and in this case, there is no need to give the agent rent and a single signal is enough to punish him. Hence, from the beginning, there is no effect of the minimum wage constraints. Thus, the payment remains the same whether or not the monitor is hired.

The fact that the principal and monitor recover the agent's rent is often pointed out in the contract theory of objective evaluation and is not unusual. As long as the agent has the rent, hiring the monitor will yield good results for the principal. However, as Corollary 2 shows, under subjective evaluation, even when the agent has no information rents, i.e., if the inequality (5) is valid, the monitor is beneficial for the principal. This is because the amount of burnt money can be recovered by hiring the monitor. It is especially important to note that since ΔEBM is strictly negative, the delegation of monitoring tasks always expands the economic pie. If we follow MacLeod (2003) and interpret burnt money as the cost of conflict between the principal and agent, this result shows that the presence of the monitor can completely resolve the conflict, even if there is a possibility of collusion.

These results show that we may have underestimated the usefulness of delegating the monitoring tasks. For example, in previous studies on contracts with objective evaluation, when the information received by the monitor was soft, there were cases where hiring a monitor is not profitable for the principal (Tirole, 1986, 1993; Strausz, 1997). Indeed, assuming that the principal receives verifiable signals in our model, we can see that when the IR and IC conditions both bind, the principal loses her payoff by having the monitor report truthfully. This result is inconsistent with our casual observations. In most organizations, principals do not evaluate agents based on relatively noisy signals alone, but rather ask someone who has more accurate information about the agent, and his information is often soft. We show that this gap between reality and theory can be bridged if we take into account that principals can only use subjective measures in many cases. In other words, under the assumption that principals can only make subjective evaluations, hiring monitors never hurts efficiency, and like many organizations, we should always hire them.

5 Conclusion

In this paper, we prove the usefulness of monitors in contracts with subjective evaluations by comparing two games. It turns out that hiring the monitor benefits the principal because the principal and monitor can recover the agent's rent and the money that would be thrown away.

The remainder of this section discusses aspects that can be further extended based on the results of this study.

First, we can consider the case where the agent receives signals that correlate with the principal's evaluation. This model assumes that the agent has no information about the principal's subjective evaluation, which is a rather strong assumption. Usually, the agent's perception of performance and the principal's subjective evaluation are more or less correlated. MacLeod (2003) shows that when the agent has some prediction about evaluation, the principal's payment decreases. Therefore, the monitor may not be useful in this situation.

Second, we can consider the case where the monitor cannot perfectly observe. In many situations, it is difficult for the monitor to perfectly observe the agent's performance or behavior. While usually somewhat more accurate than the principal, the observation of the monitor may be imperfect. This naturally reduces the value of the monitor for the principal, thus it is an important task to determine what quality of information the monitor receives makes it worthwhile to hire the monitor.

References

Che, Xiaogang, Yangguang Huang, and Le Zhang. 2021. "Supervisory efficiency and collusion in a multiple-agent hierarchy." *Games and Economic Behavior* 130 425–442. 10.1016/J.GEB.2021.09.003.

Deb, Joyee, Jin Li, and Arijit Mukherjee. 2016. "Relational contracts with subjective peer evaluations." *RAND Journal of Economics* 47 3–28. 10.1111/1756-2171.12116.

Holmström, Bengt. 1979. "Moral Hazard and Observability." *The Bell Journal of Economics* 10 (1): 74–91, `http://www.jstor.org/stable/3003320`.

Ishiguro, Shingo, and Yosuke Yasuda. 2023. "Moral hazard and subjective evaluation." *Journal of Economic Theory* 209 105619. https://doi.org/10.1016/j.jet.2023.105619.

Kambe, Shinsuke. 2006. "Subjective Evaluation in Agency Contracts." *The Japanese Economic Review* 57 121–140.

MacLeod, W. Bentley. 2003. "Optimal Contracting with Subjective Evaluation." *American Economic Review* 93 (1): 216–240. 10.1257/000282803321455232.

Prendergast, Canice. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature* 37 (1): 7–63. 10.1257/jel.37.1.7.

Strausz, Roland. 1997. "Delegation of Monitoring in a Principal-Agent Relationship." *The Review of Economic Studies* 64 (3): 337–357. 10.2307/2971717.

Thiele, Veikko. 2011. "Subjective Performance Evaluations, Collusion, and Organizational Design." *The Journal of Law, Economics, and Organization* 29 (1): 35–59. 10.1093/jleo/ewr021.

Tirole, Jean. 1986. "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations." *Journal of Law, Economics & Organization* 2 181.

Tirole, Jean. 1993. "Collusion and the theory of organizations." In *Advances in Economic Theory: Sixth World Congress*, edited by Laffont, Jean-Jacques Volume 2. 151–213, Cambridge University Press, . 10.1017/CCOL0521430194.003.