ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 347

Complete Contracts under Incomplete Information

Gregorio Curello Qianjun Lyu Yimeng Zhang

November 2024 www.econtribute.de

UNIVERSITÄT BONN UNIVERSITÄT ZU KÖLN

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged.

Complete Contracts under Incomplete Information*

Gregorio Curello
University of Mannheim

Qianjun Lyu
University of Bonn

Yimeng Zhang
University of Bonn

November 16, 2024

Latest version is available here

Abstract: We study a moral hazard model in which the output is stochastically determined by both the agent's hidden effort and an uncertain state of the world. We investigate how the contractibility of the ex-post realization of the state affects the principal's incentive to provide information. While detailed information allows the principal to better tailor the effort levels to the revealed states, coarser information enables the principal to base payments on the ex-post realization of states, thereby designing incentive schemes more effectively. Our main result establishes that when the state is contractible, full information is never optimal; however, when the state is not contractible, under certain conditions, full information is optimal.

Keywords: Moral hazard, contractibility, information design, complete contracts, incomplete contracts

*We would like to thank Sarah Auster, Francesc Dilmé, Andreas Kleiner, Stephan Lauermann, Mathijs Janssen, Wing Suen, and seminar participants at HU-Berlin for helpful discussions and suggestions. Q. Lyu acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2126/1–390838866).

1. Introduction

In contract theory, there are two notions of incompleteness. *Informational incompleteness* means the contracting parties cannot tailor their actions precisely to the state of the world. *Contractual incompleteness* means that the contract itself (i.e., the payment scheme) cannot depend on the state. The literature on incomplete contracts mostly does not distinguish between these two types of incompleteness. Both are frequently attributed to the impossibility of fully describing the state, for example, because of unforeseeable contingencies or the costs of preparing contracts. On the other hand, both types of incompleteness can also arise when the state is fully describable but not contractible, particularly within information elicitation mechanisms in the presence of information asymmetry.

In this paper, we explore incompleteness from a perspective that allows for complete contracts, in which the state is perfectly describable and contractible ex ante and ex post. Using this perspective, we attempt to disentangle the effects of informational and contractual incompleteness, and we examine whether the optimal contract endogenously exhibits either form of incompleteness. Our model is built on a moral hazard framework à la Holmström (1979), where a risk-neutral principal incentivizes a risk-averse agent to work on a project. The probability of the project's success depends on both the agent's private effort choice and an uncertain state (which may capture real-world factors such as task potentials, market or macroeconomic conditions, or customer relationships). The effort and state are complementary, meaning that the first-best action is monotone in the state. Because of the complementarity, the realization of the state also provides additional information about the agent's effort choice, conditional on the outcome of the project.

The principal has two instruments through which to influence the agent's effort: information provision and monetary compensation. First, before offering a contract, she can conduct a Blackwell experiment to learn a signal about the state; while disclosing this signal to the agent, she changes the agent's belief about the state (the interim information environment). Second, given the interim information environment, she can commit to the payment scheme, which specifies the wage to be paid to the agent based on the project outcome and the ex-post realization of the state. The contract is *complete* as all elements are contractible (the signal, the project outcome and the ex-post realization of the state), except for the agent's effort choice.

To illustrate, consider a stylized model where a platform (the principal) hires an influencer (the agent) to promote a new product during a livestream. The platform, which has access to a rich dataset, first determines how much data to use in predicting the audience size for the livestream (which is the state). Having made its prediction, the firm offers the influencer a pay-for-performance contract. The influencer decides whether to take the job and, if so, how much effort to expend in preparing for the livestream (i.e., his action). In this scenario, both the predicted and the realized audience size are contractible,[1] so the payments specified in the contract can depend on both of these quantities. Project success is defined as a certain level of sales; we assume the contract offers the influencer a base payment plus a bonus if this level is achieved. The choices faced by the platform are (1) how precisely to predict the audience size and (2) how to design the wage scheme in the contract.

One natural strategy for the platform is to obtain (and disclose) the best possible estimate of the audience size, then attempt to induce a level of effort that is appropriate for the estimated size. Intuitively, the larger the anticipated audience size (i.e., the higher the state), the more effort the platform would like the influencer to exert (this reflects the complementarity between the effort and state assumed in our model). Therefore, in higher states, the platform must offer higher bonuses. This strategy of matching the induced action to the state benefits the platform by enhancing efficiency. (In our analysis, we refer to the corresponding information structure as the *complete-information benchmark*.)

Alternatively, the platform can choose not to utilize its data at all, requiring the influencer to choose his effort level without any knowledge of the audience size. In that case, the influencer will act according to his prior information; the platform has no way to shape his effort choices to fit the state. Nonetheless, this strategy offers the platform the advantage of an *incentive-smoothing* effect: because the ex-post realization of the state is indicative of the influencer's effort choice, following the established *informativeness principle* (Holmström, 1979; Shavell, 1979), the platform can more effectively incentivize a particular level of effort by providing different bonuses for different states. Specifically, it is optimal for the platform to provide strong incentives for effort (i.e., higher bonuses) in high states, while focusing on insuring the influencer (providing relatively high base payments even if the project fails) in lower states. The optimal contract aggregates incentives in this way that reduces costs for the platform. (We refer to the corresponding information structure as the *no-information benchmark*.)

In general, if the platform obtains and discloses too little information about the state, the influencer's effort choice may be heavily distorted, and the resulting efficiency loss may outweigh the cost savings derived from incentive-smoothing. On the other hand, if the platform makes a high-quality prediction about the state, it will be able to tune the influencer's effort choice more accurately to the state (increasing efficiency), but will forgo the benefits of incentive-smoothing. The tension between efficiency and incentive-smoothing is the central subject of this paper. Our main result suggests that it is strictly suboptimal for the platform to maximize efficiency (i.e., to provide complete information).

[1]In a legal dispute, the court can examine the dataset used by the platform for its predictions, and the size of the livestream audience is publicly verifiable.

We show that the optimal contract generally takes the form of a "complete contract under incomplete information". That is, with a complete contract, the principal never finds it optimal to fully reveal the state to guide the agent's effort; rather, she utilizes the ex-post state information to smooth the agent's incentives and minimize his expected wages. The intuition is as follows: under the complete-information benchmark, the optimal wage scheme is designed state by state. That is, *in each state*, the wage contract balances the trade-off between insurance and incentive; however, *across* states, the agent is generally either under- or over-insured. To be more concrete, if we compare the payment schemes for two neighboring states, the ratio between insurance provision (measured by the difference between the agent's marginal utilities from success and failure) and relative informativeness (measured by the difference in likelihood ratio between success and failure) will be different in each state. This means the principal will strictly benefit if she pools the two states and adjusts the associated payment schemes to smooth out the insurance provision and relative informativeness across them.

Our result on the suboptimality of complete information has two important implications. First, it shows how informational incompleteness can arise endogenously. The literature often attributes informational incompleteness to the transaction costs of specifying all contingencies; we provide an alternative rationale by showing that even when the principal *can* fully describe the state to the agent, it may be optimal for her to choose not to. Second, it highlights the role played by contractual completeness: even if the principal does not use information about the state ex ante to guide the agent's actions, she benefits from using it ex post to minimize the expected costs of incentivizing efforts. As explicitly observed in Holmström (2017), the optimal contract must make use of all relevant ex-post information that provides additional statistical data about the agent's action.[2]

In the last section of the paper, we incorporate *contractual incompleteness* as in Aghion and Bolton (1992); that is, we suppose the ex-post realization of the state is not contractible. We show that in this case the conclusion of our main result is reversed: when the principal's payoff from the optimal contract is each state is convex in the state, it is indeed optimal for the principal to disclose the state to the agent to guide his effort. In particular, the lack of contractibility weakens the incentive-smoothing effect, so that the efficiency gains from information provision are dominant.

Literature. Our paper contributes to the literature on principal–agent models for moral hazard problems, pioneered by Mirrlees (1975, 1999) and Holmström (1979). In the canonical models,[3] the principal designs a wage scheme to incentivize the agent to exert a private effort,

[2]As in most moral hazard problems, in equilibrium there is no uncertainty regarding the agent's action. It is simply that the contract is designed as if there were a statistical inference problem.

[3]For a comprehensive survey, see Hart and Holmström (1987). Also, Holmström (2017) gives a more recent overview of pay-for-performance models.

at his own cost. When the action is not contractible, the output-contingent wage scheme is structured as if the principal were conducting a statistical inference regarding the action based on the realized output, which gives rise to the famous incentive–insurance trade-off. In general, one may suppose that the realized output is determined by the agent's action as well as some residual uncertainty in the economic environment.[4] Our model identifies the *contractible* part of the uncertainty as the state of nature and analyzes the interplay between contractibility and the principal's incentive to provide information over this uncertainty.

In addition, our paper derives a novel form of endogenous incompleteness within a complete contracting framework and thus relates to the literature on incomplete contracts; see, for example, Tirole (1999) and Aghion and Bolton (1992). More recently, in their study of mechanism design problems, Bergemann, Heumann and Morris (2022) show that the optimal contract is indeed coarse so as to mitigate distortion from adverse selection. Corrao, Flynn and Sastry (2023) introduce a cost for specifying the extent to which states are contractible; in their model, they show that the optimal contract is incomplete because of the difficulties faced by the principal in describing payoff-relevant outcomes. As discussed earlier, contractual frictions like adverse selection and transaction cost cannot disentangle informational incompleteness and contractual incompleteness.

Jung, Kim and Lee (2022) ask a question similar to ours, except that the state in their model represents the agent's risk attitude and does not directly enter the production function. Since information about the state does not create efficiency (as it does not affect the first-best action), they conclude that less information is always better, which is not necessarily true in our model. Similarly, Kwak (2022) studies the principal's optimal information provision under incomplete contracts (state is not contractible in his model). He focuses on a restrictive scenario where the principal induces the agent to exert the highest effort at all states. In contrast, our paper mainly discusses the tension between providing information to increase efficiency (tailoring efforts to different states) and withholding information to leverage the ex-post contractibility of the state. This perspective is absent in the other papers cited.

In principle, the state of nature in our framework plays two roles. On the one hand, its ex-post realization provides additional information on the agent's private effort choice. On the other hand, it is directly embedded into both players' payoffs through its stochastic effect on the marginal return from the agent's effort. Hence, the central question in our paper is how the principal can balance these two roles when jointly designing the wage scheme and the information environment.

[4] Indeed, this was the approach in earlier works, such as Wilson (1969), Spence and Zeckhauser (1971), Ross (1973), and Harris and Raviv (1979). Mirrlees (1974, 1976) advocated for the parametrized distribution approach, which gained great popularity later on because of its tractability.

The key insight in agency models with moral hazard is the value of information from performance measures other than output. The *informativeness principle*, proposed by Holmström (1979) and Shavell (1979), says that an additional contractible performance measure is valuable if and only if it provides new information about the agent's action conditional on the output. Gjesdal (1982) and Grossman and Hart (1983) employ the Blackwell order to derive a similar principle under weaker conditions. Kim (1995) further simplifies the condition and shows that one information system will outperform another so long as its likelihood ratio distribution is a mean-preserving spread of that of the latter. Chaigneau, Edmans and Gottlieb (2018, 2022) consider the validity of the informativeness principle in models with limited liability, and they discuss how the optimal contract depends on the available informative signals.

Another stream of the literature, on the role of informational control in moral hazard models, has generated a wide range of insights that relate to our paper (Sobel, 1993; Lewis and Sappington, 1997; Raith, 2008; Rantakari, 2008). Jehiel (2015), studying a rather general framework, concludes that when the state space has higher dimension than the action space, the principal generically has no incentive to disclose all the information she has. The common feature of the models studied in these works is the lack of contractibility of the state (and sometimes even of the signal realization). By contrast, our paper mainly focuses on complete contracts.

Lastly, our paper is related to the literature on moral hazard problems with informed principal (Beaudry, 1994; Inderst, 2001; Wagner et al., 2015; Bedard, 2017; Mekonnen, 2021; Clark, 2023). In this literature, the principal has private information about the state, which may be reflected in the contract she designs. In contrast, our principal commits to reveal her private information. This structure simplifies our analysis.

2. An illustrative example

This section gives an example that explains the trade-off faced by the principal, between improving efficiency and smoothing incentives. In this example, we assume the agent is risk-neutral. The principal assigns the agent a project, which may either succeed (outcome H) or fail (outcome L). The output of the project (its value to the principal) is v if it succeeds and 0 if it fails. The probability of success depends on the effort $a \in [0,1]$ that the agent chooses to exert, at a cost $c(a)$, as well as on a state θ, which is drawn from a uniform distribution on $[0,1]$. We denote the probability of success by $h(a,\theta)$. It is assumed increasing and concave in both a and θ, and it is log-supermodular in a and θ; that is, the effort and the state exhibit complementarity. The principal's payoff equals the output of the project net the wage w that she pays to the agent (according to a wage scheme to which she has committed in advance). The agent's payoff is $w - c(a)$. Since the agent is risk-neutral, we impose bounded payments as in Jewitt et al. (2008), i.e., we assume $w \in [0, \bar{w}]$. The agent's reservation utility is 0.

The timeline is as follows. At time 0, the principal commits to an information structure and generates a verifiable signal about the state according to this information structure. At time 1, the principal discloses the signal to the agent and simultaneously commits to a wage scheme (i.e., wage amounts w_H (w_L) to be paid in the event of success (failure) at each state, which can depend on the signal and the state. At time 2, the agent chooses his level of effort. At time 3, the project output is realized, the state is revealed, and the principal and the agent receive their payoffs.

As benchmarks, we consider the two most extreme possible information structures: (1) complete information (in which the signal reveals the state perfectly), and (2) no information (in which the signal is entirely uninformative about the state).

Complete-information benchmark. Under complete information, at the interim stage preceding time 1, the principal observes θ and designs the optimal contract as in the standard moral hazard problem. Specifically, for each state θ, she chooses the cost-minimizing wage scheme, offering wages $w_H(a,\theta)$ for success and $w_L(a,\theta)$ for failure, to incentivize the agent to take effort a. Since the agent's interim IR constraint is slack because of limited liability,[5] the cost-minimizing wage is determined by his interim IC constraint. Therefore $w_H(a,\theta) = \frac{c'(a)}{h_a(a,\theta)}$ and $w_L(a,\theta) = 0$, where $h_a(a,\theta)$ is the marginal productivity of effort a.[6] The principal then maximizes her state-wise expected payoff by choosing the effort $a^*(\theta)$ that solves $\max_a (v - w_H(a,\theta))h(a,\theta)$. The quantity $a^*(\theta)$ represents the second-best effort level in state θ. In general, it varies across states.

From an ex-ante point of view, the complete-information benchmark captures the informational gain from matching the effort with the state in the contracting process. Under this strategy, the principal's ex-ante expected payoff equals $\int_0^1 (v - w_H(a^*(\theta),\theta))h(a^*(\theta),\theta)d\theta$.

No-information benchmark. When there is no information about the state, the principal's problem is to design an optimal output- and state-contingent wage scheme given that the agent will optimize his effort choice based only on the interim information available about the state. Specifically, the principal chooses wages $w^i_H(a,\theta)$ for success and $w^i_L(a,\theta)$ for failure such that the agent is incentivized to exert effort a without knowing the state. Since the agent has different interim information other than the complete-information benchmark, here we add a superscript i to distinguish the two wage schemes. Given that the agent is risk-neutral, the cost-minimizing wage scheme has the following structure: there exists a cutoff state $t(a)$ such

[5]Zero effort satisfies the agent's interim IR constraint.

[6]This is obtained from the first-order approach, since the project output is binary and h is concave in a.

(a) Optimal action (b) Optimal wage scheme

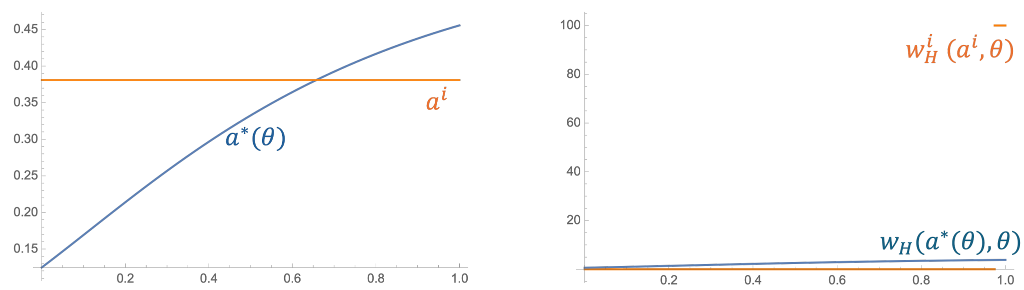


Figure 1: The optimal contract in each benchmark case.

that

$$w_H^i(a,\theta) = \begin{cases} \bar{w} & \text{if } \theta \in [t(a), 1], \\ 0 & \text{if } \theta \in [0, t(a)), \end{cases} \quad \text{and} \quad w_L^i(a,\theta) = 0.$$

Here, $t(a)$ solves $\int_{t(a)}^1 \bar{w} h_a(a,\theta) d\theta = c'(a)$. This "bang-bang" solution means that the principal rewards success with the highest possible wage, but only in the most favorable states. This is the most effective way for her to provide aggregated incentives and so save on expected costs.

The principal maximizes her expected payoff, conditional on the cost-minimizing wage scheme, by choosing the effort a^i that solves $\max_a \int_0^1 \big(v - w_H^i(a,\theta)\big) h(a,\theta) d\theta$. The quantity a^i represents the second-best effort level under the aggregated states. The principal's ex-ante expected payoff is $\int_0^1 \big(v - w_H^i(a^i,\theta)\big) h(a^i,\theta) d\theta$.

Comparison. Now consider the following parametrized example: $v = 100$, $c(a) = a^3$, $h(a,\theta) = 1 - \left(\frac{1}{2}\right)\exp\left[-\left(\frac{1}{2}\right)\left(a\theta + \frac{1}{2}a + 1\right)\right]$, and $\bar{w} = 100$.[7] In Figure 1, we plot the optimal effort and wage for each benchmark case in this example.

In the no-information case, the principal saves on her costs by pooling all the states. The benefit to her of doing so is given by

$$CS := \int_0^1 \big(100 - w_H^i(a^i,\theta)\big) h(a^i,\theta)\, d\theta - \int_0^1 \big(100 - w_H(a^i,\theta)\big) h(a^i,\theta)\, d\theta = 0.8232.$$

Here, the first term is the principal's optimal profit from full pooling, and the second term is her profit from inducing the constant effort a^i by providing state-wise incentives. Since the aggregated IC is easier to satisfy than enforcing IC at every state, $CS > 0$.

In the complete-information case, the principal obtains efficiency gains by matching actions

[7]The comparison of this example remains the same as long as the maximum wage the principal can offer is higher than 4, roughly the success reward for optimal action under state 1.

with states. The benefit of doing so is given by

$$IG := \int_0^1 (100 - w_H(a^*(\theta), \theta))h(a^*(\theta), \theta)\, d\theta - \int_0^1 \big(100 - w_H(a^i, \theta)\big)h(a^i, \theta)\, d\theta = 0.4309.$$

Since $CS > IG$, the no-information option dominates the complete-information option in this example.

However, it is easy to come up with choices of c and h for which $CS < IG$, i.e., providing complete information generates higher expected payoff for the principal than providing no information. In such cases, the principal can improve upon her no-information payoff by introducing a binary signal, e.g., by partitioning the state space into two intervals of equal width and generating a signal that indicates which interval the state lies in. This allows her to induce either of two actions on path, leading to efficiency gains. In appendix B, we provide a model with risk-neutral agent and show that there exists a finite number $N \geq 1$ such that a partition of the state space into N equal-width intervals dominates complete information. In other words, providing complete information is strictly suboptimal, which implies an endogenous informational incompleteness even with a continuous state space.

With risk-neutrality of the agent, the cost-saving effect from the optimal contract is overly powerful, since the principal can set the wage for success at the upper bound and wage for failure at the lower bound to satisfy the agent's incentive constraint. Clearly, such a contract would expose the agent to too much risk if the agent is risk-averse. Nonetheless, the result of sub-optimality of full information is robust to a risk-averse agent. We discuss it in the main model.

3. Model

We consider a moral hazard model in which a (female) principal designs the information environment and the incentive scheme so as to motivate a (male) agent to exert effort on a project.

Production function. There is an ex-post verifiable yet ex-ante uncertain state θ, which is a random variable with cumulative distribution function F on $\Theta = [0, 1]$. We assume F admits a continuously differentiable density f with full support. The project has binary outcome $y \in \{L, H\}$, which we interpret as either success (H) or failure (L). Without loss of generality, we let y represent the (numerical) value of the outcome to the principal. The probability of project success, denoted by $h(a, \theta)$, is a function of the agent's effort level $a \in \mathbb{R}_+$ and the state $\theta \in \Theta$; it is thrice continuously differentiable in each variable. We impose the following assumptions on $h(\cdot, \cdot)$: (1) *monotonicity*, which means an increase in either a or θ implies an increase in h (in other words, $h_a > 0$ and $h_\theta > 0$ for all θ and a); (2) *strong complementarity*, which means

h is log-supermodular in a and θ (i.e., $h_{a\theta}h - h_a h_\theta > 0$); (3) *diminishing marginal returns*, meaning that $h_{aa} < 0$ and $h_{\theta\theta} < 0$.

Let $l_H := \frac{h_a}{h}$ and $l_L := \frac{-h_a}{1-h}$ denote the likelihood ratios for success and failure. The assumptions of monotonicity and strong complementarity imply that l_H is positive and increasing in θ, while l_L is negative and decreasing in θ. That is, conditional on success, the higher the state is, the more likely it is that the agent has exerted greater effort. Note that monotonicity and strong complementarity together imply that h is supermodular.

Payoffs. The agent is risk-averse, and his utility function u is strictly increasing, strictly concave, and thrice continuously differentiable with respect to the wage w that he receives from the principal.[8] Independent of the wage, the agent incurs an effort cost $c(a)$, which is strictly increasing, strictly convex, and twice continuously differentiable with $c(0) = 0$. Hence his net utility is $u(w) - c(a)$. The agent's outside option is zero. The principal is risk-neutral, and her object is to maximize her expected net profit $y - w$. We make the normalization $H = v$ and $L = 0$; that is, the principal values success at v and failure at 0.

Information environment and contract. The principal is uninformed about the state ex ante and performs a Blackwell experiment to learn about it. The experiment specifies a mapping $\mathscr{E} : [0,1] \to \Delta\Sigma$, where Σ denotes the signal realization space.[9] This induces a joint distribution over states and signals, $\Delta(\Theta \times \Sigma)$, such that the marginal distribution over Θ is the prior distribution F. We denote the marginal distribution over Σ by $\mathscr{M}$. When there is no ambiguity, we also use $\sigma(\theta)$ to denote the posterior distribution over Θ after the realization of the signal σ is observed.

After both parties observing σ, the principal formulates a contract. A contract consists of a wage scheme $\{w_y^\sigma\}_{y\in\{L,H\}}$, where $w_y^\sigma : \text{supp}\ (\sigma) \to \mathbb{R}$ maps each realizable state to a wage to be paid by the principal to the agent in the event of outcome y. Since both the signal σ and the state θ are contractible, we refer to such a contract as a *complete contract*.

Timeline. At time 0, the principal commits to an experiment and generates a signal about the state. At time 1, both parties observe the signal, and the principal offers a contract to the agent. If the agent rejects the contract, the game ends and both parties receive a payoff 0 from the outside option. If the agent accepts the contract, the game proceeds to the next period. At time 2, the agent chooses a (private) level of effort. At time 3, the project output is realized, and the state is revealed. Payments are made as specified in the contract.

To analyze this game, following the tradition of Grossman and Hart (1983), we work di-

[8]The case of a risk-neutral agent is discussed separately in Appendix B.

[9]We assume the signal space Σ is endowed with some σ-algebra, such that $\mathscr{E}$ is measurable with respect to the Borel σ-algebra on Θ.

rectly with the agent's utility. Given a realized signal σ and state θ, let $u_y^\sigma(\theta) \equiv u\Big(w_y^\sigma(\theta)\Big)$ be the agent's utility from his wage for output $y \in \{L, H\}$. Denote the inverse of a utility function u by $\omega \equiv u^{-1}$; this function is strictly increasing and convex.

The principal chooses the information structure and wage scheme to maximize her ex ante expected payoff, conditional on the agent's interim IC and interim IR. That is, she solves

$$\max_{\mathcal{E}, \{u_y^\sigma(\cdot)\}, a_\sigma} \int_\Sigma \int_\Theta \Big[\big(v - \omega(u_H^\sigma(\theta))\big)h(a_\sigma, \theta) - \omega(u_L^\sigma(\theta))(1 - h(a_\sigma, \theta))\Big] \mathrm{d}\sigma(\theta) \mathrm{d}\mathscr{M}(\sigma) \tag{P}$$

$$\text{s.t. } a_\sigma \in \arg\max_{a' \in A} \int_\Theta \Big[u_H^\sigma(\theta)h(a', \theta) + u_L^\sigma(\theta)(1 - h(a', \theta))\Big] d\sigma(\theta) - c(a'), \quad \forall \sigma \in \Sigma, \tag{IC$_\sigma$}$$

$$\int_\Theta \Big[u_H^\sigma(\theta)h(a_\sigma, \theta) + u_L^\sigma(\theta)(1 - h(a_\sigma, \theta))\Big] d\sigma(\theta) \geq c(a_\sigma), \quad \forall \sigma \in \Sigma. \tag{IR$_\sigma$}$$

Notice that both (IC$_\sigma$) and (IR$_\sigma$) are imposed at the interim stage. Since the outcome space is binary, it is easy to verify the validity of the first-order approach under the monotonicity and concavity assumptions. Therefore, in what follows, we primarily work with the first-order condition instead of with (IC$_\sigma$):

$$\int \big(u_H^\sigma(\theta) - u_L^\sigma(\theta)\big) h_a(a_\sigma, \theta) d\sigma(\theta) = c'(a_\sigma), \quad \forall \sigma \in \Sigma. \tag{FOC$_\sigma$}$$

Remark. (Another interpretation of the model.) The choice of information environment and contract in our model can also be interpreted as an ex-ante mechanism. Specifically, before observing the state, the principal commits to a mechanism specifying (1) a mapping from states to distributions of effort recommendations, and (2) the wage scheme associated with each effort recommendation. The wage offered depends on the effort recommendation and on the state and outcome realized ex post. The agent is obedient to the action recommendation given the associated wage scheme (this is interim IC). Furthermore, the agent may withdraw from the contract if he has not exerted effort yet (this is interim IR).

4. Complete-Information Benchmark: Holmström (1979) Revisited

In this section we consider the complete-information benchmark, in which the experiment is perfectly revealing. Therefore both parties observe the true state at the interim stage at time 1. In this case, we can solve for the optimal contract for each state separately as in Holmström (1979). Following Grossman and Hart (1983), we decompose the problem into a two-stage program as follows:

$$\max_a \quad \left\{ \max_{u_H, u_L} \quad (v - \omega(u_H))h(a,\theta) - \omega(u_L)(1 - h(a,\theta)) \right\} \tag{P$_\theta$}$$

$$\text{s.t.} \qquad \big(u_H - u_L\big)h_a(a,\theta) = c'(a), \tag{IC$_\theta$}$$

$$u_H h(a,\theta) + u_L\big(1 - h(a,\theta)\big) \geq c(a). \tag{IR$_\theta$}$$

The inner maximization problem is essentially the cost-minimization problem conditional on an effort recommendation. The outer maximization problem yields the optimal effort.

The lemma below specifies the cost-minimizing wage scheme under complete information.

Lemma 1. *In state θ, the optimal wage scheme that induces effort a is $\{\omega(u_H(a,\theta)), \omega(u_L(a,\theta))\}$, where $u_H(a,\theta)$ and $u_L(a,\theta)$ are the solution for (IC$_\theta$) and a binding (IR$_\theta$),*

$$u_H(a,\theta) := c(a) - \frac{c'(a)}{l_L(a,\theta)}, \qquad u_L(a,\theta) := c(a) - \frac{c'(a)}{l_H(a,\theta)}.$$

Moreover, we have $\frac{\partial u_H(a,\theta)}{\partial a} > 0$ and $\frac{\partial u_L(a,\theta)}{\partial a} < 0$, while $\frac{\partial u_H(a,\theta)}{\partial \theta} < 0$ and $\frac{\partial u_L(a,\theta)}{\partial \theta} > 0$.

Though this lemma is obtained through fairly standard arguments, it is worth explaining the underlying economics. Figure 2(a) depicts the agent's utilities under the optimal wage scheme for a fixed effort level. Naturally, the agent gains a higher wage when the project outcome is high. Furthermore, the difference $u_H(a,\cdot) - u_L(a,\cdot)$ is decreasing in θ, which means the principal can provide smaller incentive for the agent to exert effort when the state is higher. This is because the production function is supermodular in the effort and state. Thus, with a higher state, the marginal productivity of effort (h_a) is also higher. This implies that, with a binding (IC$_\theta$), a smaller wage gap (i.e., a smaller gap between the wages for success and for failure) is sufficient to compensate for the marginal cost.

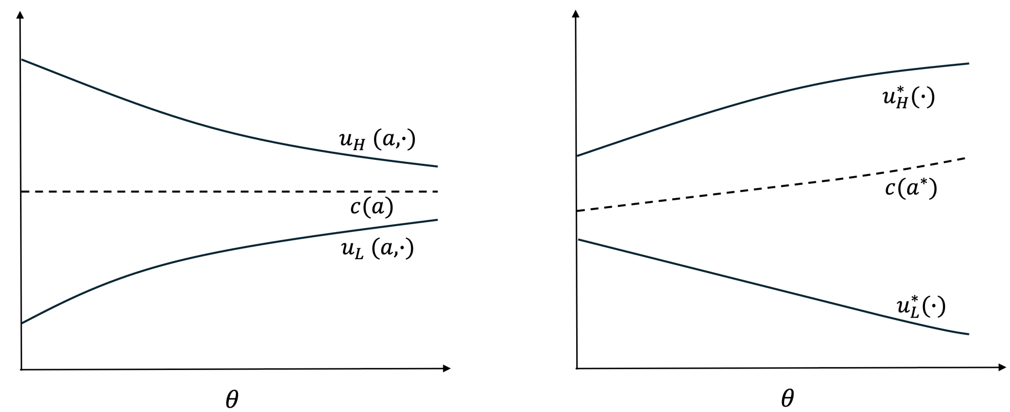
Now we turn to the optimal level of effort. Let $\Pi(a,\theta)$ be the principal's state-wise payoff from inducing effort a using the cost-minimizing wage scheme at state θ:

$$\Pi(a,\theta) = h(a,\theta)\big[v - \omega(u_H(a,\theta))\big] - (1 - h(a,\theta))\big[\omega(u_L(a,\theta))\big].$$

Let $a^*(\theta) = \arg\max_a \Pi(a,\theta)$ denote the optimal effort for state θ; we call a^* the *efficient effort*. The following result provides a sufficient condition for $a^*(\theta)$ to be monotone.

Proposition 1. *Suppose the agent's expected payoff from success, $h \cdot u_H$, is submodular in (a,θ). Then the efficient effort $a^*(\theta)$ is increasing in θ.*

If $h \cdot u_H$ is submodular, then the expected payoff that the principal must provide to com-



(a) Agent's utility under cost-minimizing wage

(b) Agent's utility under state-wise optimal wage

Figure 2: Complete information: constant effort vs. efficient effort.

pensate the agent for exerting greater effort is smaller when the state is higher. In addition, the supermodularity of h implies that the principal intrinsically prefers greater effort at higher states. Therefore the efficient effort level is increasing in the state.

Now, since the efficient effort is increasing in θ, in higher states the principal must enlarge the wage gap to offset the increased marginal cost of the agent's effort. Moreover, since $u_H(a,\theta)$ is increasing in a but decreasing in θ, if $a^*(\theta)$ is sufficiently elastic to the state, then the wage for success could be increasing in θ. A similar argument applies to a decreasing wage for failure. To simplify the notation, we use $\Pi^*(\theta) := \Pi(a^*(\theta),\theta)$ to represent the state-wise efficient profit, and we use $u_H^*(\theta) := u_H(a^*(\theta),\theta)$ and $u_L^*(\theta) := u_L(a^*(\theta),\theta)$ to represent the utilities induced by the state-wise optimal wages.

As shown in Figure 2, if the principal induces the efficient effort state-wise, then the monotonicity of the wage gap may be reversed: in higher states, although marginal productivity is higher, the agent's marginal cost for exerting the efficient effort is also higher.

5. Incomplete Information

Now let us turn to the role of incomplete information in our model. We study the optimal contract under an arbitrary non-degenerate posterior distribution σ.

Formally, for a fixed effort a, we write down the Lagrangian for the cost-minimization problem under the distribution σ:

$$
\begin{aligned}
L := &\int_{\Theta} \Big[\big(v - \omega(u_H^{\sigma}(\theta))\big)h(a,\theta) - \omega\big(u_L^{\sigma}(\theta)\big)(1-h(a,\theta))\Big]\,\mathrm{d}\sigma(\theta) \\
&+ \lambda_{(a,\sigma)}\left(\int_{\Theta}\big(u_H^{\sigma}(\theta) - u_L^{\sigma}(\theta)\big)h_a(a,\theta)\,\mathrm{d}\sigma(\theta) - c'(a)\right) \\
&+ \gamma_{(a,\sigma)}\left(\int_{\Theta}\Big[u_H^{\sigma}(\theta)h(a,\theta) + u_L^{\sigma}(\theta)(1-h(a,\theta))\Big]\,\mathrm{d}\sigma(\theta) - c(a)\right),
\end{aligned}
$$

where $\lambda_{(a,\sigma)}$ and $\gamma_{(a,\sigma)}$ are the Lagrangian multipliers associated with IC and IR (under the signal σ), respectively.

Just as in the standard moral hazard problem, we can maximize pointwise over the wage. The first-order condition implies, for each state θ,

$$
\omega'\big(u_H^{\sigma}(\theta)\big) = \lambda_{(a,\sigma)} l_H(a,\theta) + \gamma_{(a,\sigma)}, \tag{1}
$$

$$
\omega'\big(u_L^{\sigma}(\theta)\big) = \lambda_{(a,\sigma)} l_L(a,\theta) + \gamma_{(a,\sigma)}. \tag{2}
$$

By the monotonicity of l_y, the wage for success is increasing in θ, while that for failure is decreasing; thus, the wage gap is increasing in θ. While this widening of the wage gap better aggregates incentives and thereby lowers the expected wage paid in higher states, it also exposes the agent to greater risk for higher states. Rearranging equations (1) and (2) gives the following equation, which shows how the principal balances incentive provision and risk sharing across states:

$$
\lambda_{(a,\sigma)} = \frac{\omega'\big(u_H^{\sigma}(\theta)\big) - \omega'\big(u_L^{\sigma}(\theta)\big)}{l_H(a,\theta) - l_L(a,\theta)}. \tag{3}
$$

Essentially, the numerator here measures the risk exposure across states, while the denominator gives the relative informativeness of an action given a state (the difference in likelihood ratio between the two outputs). Under the optimal contract, this ratio remains constant across all states in the support of σ.

Moreover, since the agent's aggregated IC and IR are both binding under the optimal contract, $\lambda_{(a,\sigma)}$ and $\gamma_{(a,\sigma)}$ are the solutions to these binding constraints. In the next step, the principal maximizes her expected payoff over the effort space, solving

$$
\max_a \int_{\Theta} \Big[\big(v - \omega(u_H^{\sigma}(\theta;\lambda_{(a,\sigma)},\gamma_{(a,\sigma)}))\big)h(a,\theta) - \omega\big(u_L^{\sigma}(\theta;\lambda_{(a,\sigma)},\gamma_{(a,\sigma)})\big)(1-h(a,\theta))\Big]\,\mathrm{d}\sigma(\theta).
$$

Figure 3 describes the wage schemes under both no information and complete information. Although the wage gaps in both cases are increasing with respect to θ, the underlying

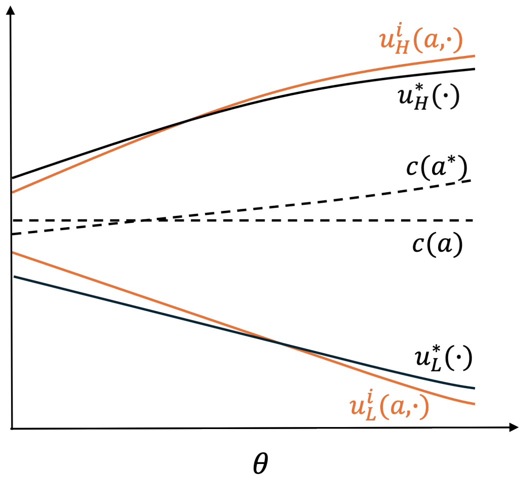


Figure 3: No information vs. complete information. Here, u_y^i denotes the wage scheme in the no-information environment.

economics in each case is different. Under incomplete information, the increase in the wage gap arises from the trade-off between providing insurance and providing incentives. In other words, when the contract is complete while information is not, the principal separates the tasks of risk sharing and incentive provision, assigning them to different states: she uses high states (where the output is more informative about the agent's effort) to provide incentive and low states (where the output is less informative about the agent's effort) to provide insurance.[10] Under complete information, however, the purpose of widening the wage gap is to incentivize the agent to follow the efficient pattern and exert more effort.

6. Contractibility vs. Information

As detailed in Sections 4 and 5, providing more information enhances efficiency, as it lets the principal to design incentive scheme to better match the optimal effort to the state, while providing less information saves costs, as it lets her smooth the incentive–risk trade-off across states. In this section, we discuss the relationship between these two forces.

We make the following assumption on the likelihood ratio.

Assumption 1. *For any two actions a and a', and for any θ, there exists an open neighborhood of θ such that $\frac{l_H(a,\theta)-l_L(a,\theta)}{l_H(a',\theta)-l_L(a',\theta)} \neq \frac{l_H(a,\theta')-l_L(a,\theta')}{l_H(a',\theta')-l_L(a',\theta')}$ for any θ' in the neighborhood of θ.*

Under Assumption 1, and given the differentiability of h, we have $\frac{\partial^2 \log(l_H - l_L)}{\partial a \partial \theta} \neq 0$ almost

[10] At a high level, this strategy is reminiscent of decision-linking, as used in the mechanism design literature (Jackson and Sonnenschein, 2007).

everywhere. This implies that the effort a and the state θ influence the likelihood ratio difference in a non-separable way. Assumption 1 is not necessary for our main theorem to be true, but we adopt it to simplify the analysis.

Theorem 1. *Suppose the contract is complete and Assumption 1 holds. Then providing perfect information over any subinterval of the state space is suboptimal. This implies that there exists a discrete (coarse) information structure giving the principal strictly higher payoff than the complete-information benchmark.*

Before providing a sketch of proof, we first elaborate the underlying intuition of this result. When the information is incomplete, the optimal wage scheme under any non-degenerate posterior distribution $\sigma(\theta)$ must satisfy equation (3): the optimal level of risk exposure divided by the likelihood ratio difference (risk-informativeness ratio) remains a constant across all states in the support of σ (all realizable states under σ share the same Lagrangian multiplier). That is, when those states are linked together by the principal to sustain one interim IC, the risk-informativeness ratio is smoothed out across states. In contrast, when the information is complete, the risk-informativeness ratio given the efficient effort $a^*(\theta)$, $\frac{\omega'(u_H^*(\theta))-\omega'(u_L^*(\theta))}{l_H(a^*(\theta),\theta)-l_L(a^*(\theta),\theta)}$, is independent across different states, since each ratio corresponds to the Lagrangian multiplier for each state. Therefore, if we cross-check the optimal wage schemes in the complete information benchmark across states, then the agent is generally over-insured in some states and under-insured in others.

Consider two nearby states $\theta < \theta'$, and a and a' are the corresponding efficient effort, respectively. The following expression determines whether the agent is over-insured at the higher state θ'.

$$\frac{\frac{d}{d\theta}\big(\omega'(u_H^*(\theta'))-\omega'(u_L^*(\theta'))\big)}{\omega'(u_H^*(\theta'))-\omega'(u_L^*(\theta'))}-\frac{\frac{\partial}{\partial\theta}\big(l_H(a',\theta')-l_L(a',\theta')\big)}{l_H(a',\theta')-l_L(a',\theta')}.$$

The first term represents the percentage change in the risk exposure across state, while the second term represents the percentage change in informativeness across states. If the difference between these two terms is negative, then the agent is over-insured at the higher state. In this case, the principal can benefit by pooling these two states and reallocating incentives across states to induce the same effort a' for both states. Specifically, she can increase the wage gaps at the higher state θ' to provide stronger incentives, while decrease the wage gaps at the lower state θ to maintain necessary insurance provision. Conversely, if this quantity is positive, then the agent is over-insured at the lower state θ, and so the principal can enhance incentive provision at θ and reallocate insurance to higher state θ' to smooth the risk-informativeness ratio. Assumption 1 ensures this quantity is non-zero for any open subinterval of the state space.

Now we provide a sketch of proof. Consider the following operation. Starting with the optimal contract under complete information, $\{u_H^*(\theta), u_L^*(\theta)\}$, pick two nearby states θ and θ', where $\theta < \theta'$, with associated efficient efforts a and a', respectively. Suppose the principal now pools θ and θ' together and tries to induce the higher action a' by adjusting the contract as follows. At θ', she increases $u_H^*(\theta')$ by $\overline{\varepsilon}_H$ and decreases $u_L^*(\theta')$ by $\overline{\varepsilon}_L$ so as to maintain the agent's IR at state θ'. This requires

$$\overline{\varepsilon}_H \, h(a', \theta') = \overline{\varepsilon}_L \left(1 - h(a', \theta')\right), \tag{4}$$

since the agent's IR for exerting effort a' is binding at state θ' under the original contract.

Next, at θ, the principal decreases $u_H^*(\theta)$ by $\underline{\varepsilon}_H$ and increases $u_L^*(\theta)$ by $\underline{\varepsilon}_L$ so as to satisfy the agent's IR at state θ, which requires

$$\left(u_H^*(\theta) - \underline{\varepsilon}_H\right) h(a', \theta) + \left(u_L^*(\theta') + \underline{\varepsilon}_L\right)\left(1 - h(a', \theta)\right) = c(a'), \tag{5}$$

where $c(a') = u_H(a', \theta) h(a', \theta) + u_L(a', \theta)\left(1 - h(a', \theta)\right)$. The treatment above ensures that the agent's IR is satisfied state-wise. Since the major gain from pooling states is incentive-smoothing, we consider the aggregated IC for both states,

$$\begin{aligned} & f(\theta)\left(u_H^*(\theta) - u_L^*(\theta) - \underline{\varepsilon}_H + \underline{\varepsilon}_L\right) h_a(a', \theta) + f(\theta')\left(u_H^*(\theta') - u_L^*(\theta') + \overline{\varepsilon}_H - \overline{\varepsilon}_L\right) h_a(a', \theta') \\ =& c'(a')\left(f(\theta) + f(\theta')\right). \end{aligned} \tag{6}$$

Substituting (5) and (4) into (6), we obtain the following relation between $\underline{\varepsilon}_L$ and $\overline{\varepsilon}_L$:

$$f(\theta') l_H(a', \theta') \overline{\varepsilon}_L = f(\theta) l_H(a', \theta)\left(u_L^*(\theta) - u_L(a', \theta)\right) + f(\theta) l_H(a', \theta) \underline{\varepsilon}_L. \tag{7}$$

The principal's payoff gain PG from this treatment is given by

$$\begin{aligned} PG =& f(\theta)\left(\Delta a \, h_a(a, \theta)\left[v - \left(\omega(u_H^*(\theta)) - \omega(u_L^*(\theta))\right)\right]\right) \\ & + f(\theta)\left(\omega'(u_H^*(\theta)) h(a', \theta) \underline{\varepsilon}_H - \omega'(u_L^*(\theta))\left(1 - h(a', \theta)\right) \underline{\varepsilon}_L\right) \\ & + f(\theta')\left(-\omega'(u_H^*(\theta')) h(a', \theta') \overline{\varepsilon}_H + \omega'(u_L^*(\theta'))\left(1 - h(a', \theta')\right) \overline{\varepsilon}_L\right), \end{aligned} \tag{8}$$

where $\Delta a \equiv a' - a = a^*(\theta') - a^*(\theta)$. The first term of PG represents the gain from inducing a higher action at state θ, the second term represents the expected gain from shrinking the wage gap at θ, and the third term represents the expected loss from widening the wage gap at θ'.

Note that the optimal effort a at state θ is determined by the principal's first-order condition,

$$h_a(a,\theta)\big[v-\big(\omega(u_H^*(\theta))-\omega(u_L^*(\theta))\big)\big]=h(a,\theta)\omega'(u_H^*(\theta))\frac{\partial u_H(a,\theta)}{\partial a}+(1-h(a,\theta))\omega'(u_L^*(\theta))\frac{\partial u_L(a,\theta)}{\partial a}. \tag{9}$$

By substituting equations (5), (4), (7), and (9) into (8), we can express the principal's payoff gain as follows:

$$\begin{aligned}&-f(\theta)\left[\omega'(u_H^*(\theta))\frac{\partial u_H}{\partial a}-\omega'(u_L^*(\theta))\frac{\partial u_L}{\partial a}\right]h_a(a,\theta)(\Delta a)^2\\&+f(\theta')l_H(a',\theta')\overline{\varepsilon}_L\Bigg[\underbrace{\frac{\omega'(u_H^*(\theta))-\omega'(u_L^*(\theta))}{l_H(a,\theta)-l_L(a,\theta)}\frac{l_H(a,\theta)-l_L(a,\theta)}{l_H(a',\theta)-l_L(a',\theta)}-\frac{\omega'(u_H^*(\theta'))-\omega'(u_L^*(\theta'))}{l_H(a',\theta')-l_L(a',\theta')}}_{\text{incentive-smoothing}}\Bigg].\end{aligned}$$

The first term corresponds to the efficiency loss from the mismatch between efforts and states; it is always negative. The second term corresponds to the cost savings from incentive-smoothing. The sign of its coefficient is ambiguous. For example, since the efficient effort is increasing in state and the likelihood ratio l_H-l_L is decreasing in a, then $\frac{l_H(a,\theta)-l_L(a,\theta)}{l_H(a',\theta)-l_L(a',\theta)}>1$. Therefore, if the risk–informativeness ratio under complete information, i.e., $\frac{\omega'(u_H^*(\theta))-\omega'(u_L^*(\theta))}{l_H(a^*(\theta),\theta)-l_L(a^*(\theta),\theta)}$, is decreasing in θ, then the incentive-smoothing coefficient is positive. This implies that under complete information, the principal overcompensates insurance in the higher states, so she may benefit from pooling those states and reallocate the insurance provision to lower states and use higher states to provide stronger incentives. On the other hand, if the risk–informativeness ratio is increasing and the increment is sufficient to compensate the change in informativeness, then the coefficient becomes negative. In this case, the principal overcompensates incentive in the higher states, so she may benefit from pooling those states and reallocate the incentive provision to lower states and use higher states to provide better insurance, i.e., she can choose a negative $\overline{\varepsilon}_L$. Assumption 1 implies that this coefficient does not equal to zero in any subinterval of the state space.[11]

Since the incentive-smoothing term is of the same order as Δa, and since $\overline{\varepsilon}_L$ is independent of Δa, we can always choose $\overline{\varepsilon}_L$ to be small enough, yet sufficiently large relative to Δa, to ensure that the gains from incentive-smoothing dominate the efficiency losses from effort

[11]If the coefficient is exactly 0 for all neighboring states, then $\frac{l_H(a,\theta)-l_L(a,\theta)}{l_H(a',\theta)-l_L(a',\theta)}$ is independent of θ. To see this, take three arbitrary neighboring states θ, θ', and θ''. The zero coefficient implies $\lambda^*_{\theta''}=\frac{l_H(a,\theta)-l_L(a,\theta)}{l_H(a'',\theta)-l_L(a'',\theta)}\lambda^*_\theta=\frac{l_H(a,\theta)-l_L(a,\theta)}{l_H(a',\theta)-l_L(a',\theta)}\frac{l_H(a',\theta')-l_L(a',\theta')}{l_H(a'',\theta')-l_L(a'',\theta')}\lambda^*_\theta$, where $\lambda^*_\theta=\frac{\omega'(u_H^*(\theta))-\omega'(u_L^*(\theta))}{l_H(a^*(\theta),\theta)-l_L(a^*(\theta),\theta)}$ is the optimal Lagrangian multiply for state θ under the optimal action. This implies $\frac{l_H(a',\theta)-l_L(a',\theta)}{l_H(a'',\theta)-l_L(a'',\theta)}=\frac{l_H(a',\theta')-l_L(a',\theta')}{l_H(a'',\theta')-l_L(a'',\theta')}$. Since θ' and θ'' are chosen arbitrarily, this essentially means $\frac{l_H(a',\theta)-l_L(a',\theta)}{l_H(a'',\theta)-l_L(a'',\theta)}$ is independent of θ, contradicting Assumption 1.

mismatching. That is, by pooling neighboring states, the principal can obtain first-order gains from incentive-smoothing, without incurring first-order efficiency losses. In the appendix, we provide a complete proof over two neighboring intervals; the intuition there is analogous to that of the pairwise argument above.

Crucially, this result establishing the suboptimality of providing complete information hinges on the contractibility of the ex-post realization of the state. Without contract completeness, providing complete information ex ante may still be optimal. In the next section, we explore incomplete contracts.

7. Optimality of Complete Information under Incomplete Contracts

We now consider a setting with *incomplete contracts*, in which the ex-post state θ is not contractible. That is, the wage scheme cannot be contingent on the realized state, only on the realization of the signal σ (see Aghion and Bolton (1992)). By a slight abuse of notation, given some signal realization σ, an incomplete wage scheme is simply a pair $(u_H^\sigma, u_L^\sigma) \subseteq \mathbb{R}^2$. Formally, given a specific posterior belief σ, the principal's program now becomes

$$\max_{a} \quad \left\{ \max_{u_H^\sigma, u_L^\sigma} \quad \int_\Theta \big(v - \omega(u_H^\sigma)\big)h(a,\theta) - \omega(u_L^\sigma)(1-h(a,\theta))\mathrm{d}\sigma(\theta) \right\} \tag{P_σ^I}$$

$$\text{s.t.} \qquad \int_\Theta \big(u_H^\sigma - u_L^\sigma\big)h_a(a,\theta)\mathrm{d}\sigma(\theta) = c'(a), \tag{IC_σ^I}$$

$$\int_\Theta \big[u_H^\sigma h(a,\theta) + u_L^\sigma(1-h(a,\theta))\big]\mathrm{d}\sigma(\theta) \geq c(a). \tag{IR_σ^I}$$

We denote by $\Pi^I(\sigma)$ the principal's profit under the optimal wage scheme and the optimal effort. The next result states that when the principal's profit function under complete information $\Pi^*(\theta)$ is convex, it is always optimal for her to provide complete information.

Proposition 2. *Suppose* $h_{a\theta\theta} < 0$. *If* $\Pi^*(\theta)$ *is convex, then* $\int_\Theta \Pi^*(\theta)\mathrm{d}\sigma(\theta) > \Pi^I(\sigma)$ *for any non-degenerate* σ.

Proof. First, suppose the wage scheme (u_H^σ, u_L^σ) is the optimal wage scheme that induces effort a under the distribution $\sigma(\theta)$ (both IC_σ^I and IR_σ^I are binding for effort a); then it induces some strictly higher effort $\tilde{a}$ under the Dirac measure on the expectation of σ.

To see this, note that by the concavity of h_a and $h_{a\theta}$, we have

$$c'(a) = (u_H^\sigma - u_L^\sigma)\int_\Theta h_a(a,\theta)\mathrm{d}\sigma(\theta) < (u_H^\sigma - u_L^\sigma)h_a(a, \mathbb{E}_\sigma[\theta]),$$

where the first equality follows from the premise (IC_σ^I). Then, by the convexity of c and the concavity of h, there exists some $\tilde{a} > a$ such that

$$c'(\alpha) \le (u_H^\sigma - u_L^\sigma) h_a\big(\alpha, \mathbb{E}_\sigma[\theta]\big)$$

for all $\alpha \in [a, \tilde{a}]$, with equality holding exactly at $\tilde{a}$.

In other words, $\tilde{a}$ satisfies IC under the wage scheme $\{u_H^\sigma, u_L^\sigma\}$ at state $\mathbb{E}_\sigma[\theta]$. Next we show that it also satisfies IR at state $\mathbb{E}_\sigma[\theta]$. By the binding ($\text{IR}_\sigma^I$),

$$c(a) = u_L^\sigma + \int_\Theta h(a,\theta) \mathrm{d}\sigma(\theta)\big(u_H^\sigma - u_L^\sigma\big) < u_L^\sigma + h\big(a, \mathbb{E}_\sigma[\theta]\big)\big(u_H^\sigma - u_L^\sigma\big),$$

where the second inequality comes from the concavity of h. Consequently,

$$\begin{aligned} c(a) + \int_a^{\tilde{a}} c'(\alpha)\mathrm{d}\alpha &\le u_L^\sigma + \big(u_H^\sigma - u_L^\sigma\big)\Big[h\big(a, \mathbb{E}_\sigma[\theta]\big) + \int_a^{\tilde{a}} h_a(\alpha, \mathbb{E}_\sigma[\theta])\mathrm{d}\alpha\Big] \\ \Rightarrow c(\tilde{a}) &\le u_L^\sigma + h\big(\tilde{a}, \mathbb{E}_\sigma[\theta]\big)\big(u_H^\sigma - u_L^\sigma\big). \end{aligned}$$

Hence, $\tilde{a}$ indeed can be implemented at state $\mathbb{E}_\sigma[\theta]$ under (u_H^σ, u_L^σ). Now we can establish the suboptimality of pooling information under σ. Suppose a and (u_H^σ, u_L^σ) are the optimal solutions under σ. Then

$$\begin{aligned} \Pi^I(\sigma) &= \int_\Theta h(a,\theta)\big[v - \omega(u_H^\sigma)\big] + \big(1 - h(a^\sigma, \theta)\big)\big[-\omega(u_L^\sigma)\big]\mathrm{d}\sigma(\theta) \\ &= \int_\Theta h(a,\theta)\big[v - \big(\omega(u_H^\sigma) - \omega(u_L^\sigma)\big)\big]\mathrm{d}\sigma(\theta) - \omega(u_L^\sigma) \\ &< h\big(a, \mathbb{E}_\sigma[\theta]\big)\big[v - \big(\omega(u_H^\sigma) - \omega(u_L^\sigma)\big)\big] - \omega(u_L^\sigma) \\ &< h\big(\tilde{a}, \mathbb{E}_\sigma[\theta]\big)\big[v - \big(\omega(u_H^\sigma) - \omega(u_L^\sigma)\big)\big] - \omega(u_L^\sigma) \\ &\le \Pi^*\big(\mathbb{E}_\sigma[\theta]\big) \\ &\le \int_\Theta \Pi^*(\theta)\,\mathrm{d}\sigma(\theta). \end{aligned}$$

Here the first inequality comes from the concavity of $h(\cdot,\theta)$. The third inequality is due to the agent's revealed preference as $\tilde{a}$ is an inducible effort. The last inequality follows from the convexity of Π^*. □

The lemma below provides a sufficient condition for the principal's indirect profit function $\Pi^*(\theta)$ to be convex.

Lemma 2. *Suppose the determinant of the Hessian matrix of* $\Pi(a,\theta)$ *is negative at* $(a^*(\theta),\theta)$ *for* *all θ. Then $\Pi^*(\theta)$ is convex.*

The contrast between the implications of Theorem 1 and Proposition 2 suggests that the suboptimality of complete information arises primarily from the ex-post contractibility of the state. Specifically, there are two channels by which the principal may use information about the state: she may use it ex ante to guide the agent's action, and she may use it ex post to smooth the agent's incentives. When the contract is incomplete, the effect of the second channel is minimal, as the principal cannot gain from the informativeness principle. In this case, if ex-ante information strictly improves efficiency (e.g., if $\Pi^*(\theta)$ is convex), then the first channel is dominant. On the other hand, this argument does not rule out the possibility that pooling is optimal when the state-wise optimal effort is inelastic. For example, the principal may prefer to induce the same effort in all states. This situation is discussed in Jehiel (2015), where pooling is still strictly optimal under incomplete contracts: for a fixed output, it is less costly for the principal to offer a constant wage scheme across states than to induce the same action using state-wise incentives.

8. Discussion

In this paper, we investigate the interplay between strategic information provision and ex-post contractibility within the framework of a moral hazard problem. Our analysis focuses on settings in which the state of the world is inherently uncertain, but the principal can learn and disclose information about it before the contracting stage. We emphasize the trade-off faced by the principal: she can provide detailed information to fine-tune the agent's level of effort, or she can provide coarser information and offer a more cost-effective contract. Our results show that when the ex-post state of the world is contractible, the optimal mechanism features endogenous informational incompleteness. In addition, extending our model to settings in which the ex-post state is not contractible, we identify the conditions under which complete information provision becomes optimal.

These results give rise to a testable prediction: in environments where institutional frameworks enable the contractibility of production factors, less information will be generated. Our insights are relevant to the broader study of moral hazard and information asymmetry, bridging the literature on incomplete contracts and complete contracts.

Furthermore, there is an alternative interpretation of the ex-post contractibility of the state. Instead of the actual state being revealed at time 2, we can consider that an ex-post contractible signal is generated at time 2, which may provide information about the state. In this context, our complete contract model represents an extreme case where the signal perfectly reveals the state, whereas our incomplete contract model represents the opposite extreme, where the signal is completely uninformative about the state. Our main result suggests a broader conjecture: when the ex-post signal is more informative about the state, the principal tends to disclose less information ex-ante.

References

Aghion, Philippe and Patrick Bolton, "An incomplete contracts approach to financial contracting," *Review of Economic Studies*, 1992, *59* (3), 473–494.

Beaudry, Paul, "Why an informed principal may leave rents to an agent," *International Economic Review*, 1994, *35* (4), 821–832.

Bedard, Nicholas Charles, "Contracts in informed-principal problems with moral hazard," *Economic Theory Bulletin*, 2017, *5* (1), 21–34.

Bergemann, Dirk, Tibor Heumann, and Stephen Morris, "Screening with persuasion," 2022. Cowles Foundation Discussion Paper No. 2338R.

Chaigneau, Pierre, Alex Edmans, and Daniel Gottlieb, "Does improved information improve incentives?," *Journal of Financial Economics*, 2018, *130* (2), 291–307.

_ , _ , and _ , "How should performance signals affect contracts?," *Review of Financial Studies*, 2022, *35* (1), 168–206.

Clark, Daniel, "Contracting with private information, moral hazard, and limited commitment," 2023. Working paper, 13 Feb 2023.

Corrao, Roberto, Joel P. Flynn, and Karthik A. Sastry, "Optimally coarse contracts," 2023. Working paper, 2 Nov 2023.

Gjesdal, Frøystein, "Information and incentives," *Review of Economic Studies*, 1982, *49* (3), 373–390.

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

Harris, Milton and Artur Raviv, "Optimal incentive contracts with imperfect information," *Journal of Economic Theory*, 1979, *20* (2), 231–259.

Hart, Oliver and Bengt Holmström, "The theory of contracts," in Truman Fassett Bewley, ed., *Advances in Economic Theory*, Vol. 1, Cambridge: Cambridge University Press, 1987, chapter 3, pp. 71–156.

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

_ , "Pay for performance and beyond," *American Economic Review*, 2017, *107* (7), 1753–1777.

Inderst, Roman, "Incentive schemes as a signaling device," *Journal of Economic Behavior & Organization*, 2001, *44* (4), 455–465.

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

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

Jehiel, Philippe, "On transparency in organizations," *Review of Economic Studies*, 2015, *82* (2), 736–761.

Jewitt, Ian, Ohad Kadan, and Jeroen M. Swinkels, "Moral hazard with bounded payments," *Journal of Economic Theory*, 2008, *143* (1), 59–82.

Jung, Jin Yong, Son Ku Kim, and Seung Joo Lee, "Ignorance is bliss," 2022. Working paper, 29 Oct 2022.

Karlin, Samuel and Yosef Rinott, "Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions," *Journal of Multivariate Analysis*, 1980, *10* (4), 467–498.

Kim, Son Ku, "Efficiency of an information system in an agency model," *Econometrica*, 1995, *63* (1), 89–102.

Kwak, Changhyun, "Optimal information disclosure with moral hazard," 2022. Working paper, July 2022.

Lewis, Tracy R. and David E. M. Sappington, "Information management in incentive problems," *Journal of Political Economy*, 1997, *105* (4), 796–821.

Mekonnen, Teddy, "Informed principal, moral hazard, and limited liability," *Economic Theory Bulletin*, 2021, *9* (1), 119–142.

Mirrlees, James A., "Notes on welfare economics, information and uncertainty," in Michael Balch, Daniel McFadden, and Shih yen Wu, eds., *Essays on economic behavior under uncertainty*, Amsterdam: North-Holland, 1974, chapter 9, pp. 243–258.

— , "The theory of moral hazard and unobservable behaviour," 1975. Working paper, Nuffield College, Oxford.

— , "The optimal structure of incentives and authority within an organization," *The Bell Journal of Economics*, 1976, *7* (1), 105–131.

—, "The theory of moral hazard and unobservable behaviour," *Review of Economic Studies*, 1999, *66* (1), 3–21.

Raith, Michael, "Specific knowledge and performance measurement," *RAND Journal of Economics*, 2008, *39* (4), 543–559.

Rantakari, Heikki V., "On the role of uncertainty in the risk-incentives tradeoff," *The B.E. Journal of Theoretical Economics (Topics)*, 2008, *8* (1), Article 10.

Ross, Stephen A., "The economic theory of agency," *American Economic Review: Papers & Proceedings*, 1973, *63* (2), 134–139.

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

Sobel, Joel, "Information control in the principal-agent problem," *International Economic Review*, 1993, *34* (2), 259–269.

Spence, Michael and Richard Zeckhauser, "Insurance, information, and individual action," *American Economic Review: Papers & Proceedings*, 1971, *61* (2), 380–387.

Tirole, Jean, "Incomplete contracts," *Econometrica*, 1999, *67* (4), 741–781.

Wagner, Christoph, Tymofiy Mylovanov, and Thomas Tröger, "Informed-principal problem with moral hazard, risk neutrality, and no limited liability," *Journal of Economic Theory*, 2015, *159* (A), 280–289.

Wilson, Robert, "The structure of incentives for decentralization under uncertainty," in "La décision" 1969.

Appendix A

Proof of Lemma 1.

$$\frac{\partial u_L^*(a,\theta)}{\partial \theta} = \frac{c''(a)\frac{\partial l_H}{\partial \theta}}{l_H^2} > 0, \quad \frac{\partial u_H^*(a,\theta)}{\partial \theta} = \frac{c''(a)\frac{\partial l_L}{\partial \theta}}{l_L^2} < 0$$

$$\frac{\partial u_L^*(a,\theta)}{\partial a} = \frac{\frac{h_{aa}}{h_a}c'(a) - c''(a)}{l_H} < 0, \quad \frac{\partial u_H^*(a,\theta)}{\partial a} = \frac{\frac{h_{aa}}{h_a}c'(a) - c''(a)}{l_L} > 0$$

□

Proof of Proposition 1. It suffices to prove the indirect profit function $\Pi(a,\theta)$ is supermodular. In particular,

$$\begin{aligned}
\frac{\partial^2 \Pi}{\partial a \partial \theta} &= \frac{\partial^2 \mathbb{E}[y|a,\theta]}{\partial a \partial \theta} - \frac{\partial^2 \mathbb{E}\Big[\omega\Big(u_y^*(a,\theta)\Big)\Big| a,\theta\Big]}{\partial a \partial \theta} \\
=&\Big[\big(H - \omega(u_H^*)\big) - \big(L - \omega(u_L^*)\big)\Big]h_{a\theta} + \Big[-\omega''(u_H^*)\frac{\partial u_H^*}{\partial \theta}\frac{\partial u_H^*}{\partial a}h - \omega''(u_L^*)\frac{\partial u_L^*}{\partial \theta}\frac{\partial u_L^*}{\partial a}(1-h)\Big] \\
&-\Bigg[\omega'(u_H^*)\Bigg(\frac{\partial^2\big(h\cdot u_H^*\big)}{\partial a \partial \theta} - u_H^* h_{a\theta}\Bigg) + \omega'(u_L^*)\Bigg(\frac{\partial^2\big((1-h)\cdot u_L^*\big)}{\partial a \partial \theta} + u_L^* h_{a\theta}\Bigg)\Bigg]
\end{aligned}$$

The first term in the first line is positive as the principal's profit is higher when the output is realized to be high. The second term is also positive, which comes from the convexity of ω and the fact that $\frac{\partial u_y^*}{\partial a}\frac{\partial u_y^*}{\partial \theta} < 0$ for either output. It remains to show the term in the square bracket in the second line is negative. By (IR_θ), we have:

$$\frac{\partial^2\big(h\cdot u_H^*\big)}{\partial a \partial \theta} = -\frac{\partial^2\big((1-h)\cdot u_L^*\big)}{\partial a \partial \theta} < 0$$

Since $u_H^* > u_L^*$, it follows immediately from the fact that h is supermodular and the convexity of ω. □

Proof of Theorem 1. Given some interval $[\underline{\theta}, \overline{\theta}] \subseteq [0,1]$, pick an arbitrary $\hat{t} \in (\underline{\theta}, \overline{\theta})$ and consider its neighborhood $(\hat{t} - \delta, \hat{t} + \delta) \subset [\underline{\theta}, \overline{\theta}]$. Our objective is to argue that there exists some δ such that the expected profit under full information on $(\hat{t} - \delta, \hat{t} + \delta)$ is lower than that inducing the pooling effort $\hat{a} := a^*(\hat{t})$ on the same region. In particular, consider the following perturbed incentive scheme $\tilde{u}_y^*$ that implements $\hat{a}$:

$$\tilde{u}_H^*(\theta) = \begin{cases} u_H^*(\theta) - \underline{\varepsilon_H}, & \text{if } \theta \in (\hat{t}-\delta, \hat{t}); \\ u_H^*(\theta) + \overline{\varepsilon_H}, & \text{if } \theta \in (\hat{t}, \hat{t}+\delta). \end{cases}$$

$$\tilde{u}_L^*(\theta) = \begin{cases} u_L^*(\theta) + \underline{\varepsilon_L}, & \text{if } \theta \in (\hat{t}-\delta, \hat{t}); \\ u_L^*(\theta) - \overline{\varepsilon_L}, & \text{if } \theta \in (\hat{t}, \hat{t}+\delta). \end{cases}$$

such that $\hat{a}$ is incentive-compatible and IR holds for both $(\hat{t}-\delta, \hat{t})$ and $(\hat{t}, \hat{t}+\delta)$. That is, we pick $(\overline{\varepsilon_H}, \overline{\varepsilon_L}, \underline{\varepsilon_H}, \underline{\varepsilon_L})$ such that the following three equalities are satisfied and eventually take the limit of the variable of choice to zero from above.

$$\int_{\hat{t}}^{\hat{t}+\delta} \left(\overline{\varepsilon_H}+\overline{\varepsilon_L}\right) h_a(\hat{a},\theta)\mathrm{d}F(\theta) - \int_{\hat{t}-\delta}^{\hat{t}} \left(\underline{\varepsilon_H}+\underline{\varepsilon_L}\right) h_a(\hat{a},\theta)\mathrm{d}F(\theta) = \int_{\hat{t}-\delta}^{\hat{t}+\delta} c'(\hat{a}) - \left(u_H^*(\theta) - u_L^*(\theta) h_a(\hat{a},\theta)\right)\mathrm{d}F(\theta) \tag{$\tilde{\text{IC}}$}$$

$$\int_{\hat{t}}^{\hat{t}+\delta} \left[\overline{\varepsilon_H} h(\hat{a},\theta) - \overline{\varepsilon_L}(1-h(\hat{a},\theta))\right]\mathrm{d}F(\theta) = \int_{\hat{t}}^{\hat{t}+\delta} c(\hat{a}) - \left[h(\hat{a},\theta)u_H^*(\theta) + (1-h(\hat{a},\theta))u_L^*(\theta)\right]\mathrm{d}F(\theta) \tag{$\overline{\text{IR}}$}$$

$$\int_{\hat{t}-\delta}^{\hat{t}} \left[\underline{\varepsilon_L}(1-h(\hat{a},\theta)) - \underline{\varepsilon_H} h(\hat{a},\theta)\right]\mathrm{d}F(\theta) = \int_{\hat{t}-\delta}^{\hat{t}} c(\hat{a}) - \left[h(\hat{a},\theta)u_H^*(\theta) + (1-h(\hat{a},\theta))u_L^*(\theta)\right]\mathrm{d}F(\theta). \tag{$\underline{\text{IR}}$}$$

Substituting ($\overline{\text{IR}}$) and ($\underline{\text{IR}}$) into ($\tilde{\text{IC}}$) yields

$$\begin{aligned} &\int_{\hat{t}-\delta}^{\hat{t}+\delta} h_a(\hat{a},\theta)\Big[\left(u_H^*(\hat{a},\theta) - u_L^*(\hat{a},\theta)\right) - \left(u_H^*(\theta) - u_L^*(\theta)\right)\Big]\mathrm{d}F(\theta) \\ =&l_h^+ \int_{\hat{t}}^{\hat{t}+\delta} \Big[\overline{\varepsilon_L} + h(\hat{a},\theta)\left(u_H^*(\hat{a},\theta) - u_H^*(\theta)\right) + (1-h(\hat{a},\theta))\left(u_L^*(\hat{a},\theta) - u_L^*(\theta)\right)\Big]\mathrm{d}F(\theta) \\ -&l_h^- \int_{\hat{t}-\delta}^{\hat{t}} \Big[\underline{\varepsilon_L} - h(\hat{a},\theta)\left(u_H^*(\hat{a},\theta) - u_H^*(\theta)\right) + (1-h(\hat{a},\theta))\left(u_L^*(\hat{a},\theta) - u_L^*(\theta)\right)\Big]\mathrm{d}F(\theta) \end{aligned}$$

where $l_h^+ = \frac{\int_{\hat{t}}^{\hat{t}+\delta} h_a(\hat{a},\theta)\mathrm{d}F(\theta)}{\int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)\mathrm{d}F(\theta)}$ and $l_h^- = \frac{\int_{\hat{t}-\delta}^{\hat{t}} h_a(\hat{a},\theta)\mathrm{d}F(\theta)}{\int_{\hat{t}-\delta}^{\hat{t}} h(\hat{a},\theta)\mathrm{d}F(\theta)}$.

The expression can be further rearranged as

$$
\begin{aligned}
l_h^+ \int_{\hat{t}}^{\hat{t}+\delta} \mathrm{d}F(\theta)\overline{\varepsilon_L} = l_h^- \int_{\hat{t}-\delta}^{\hat{t}} \mathrm{d}F(\theta)\underline{\varepsilon_L} \\
&- l_H^+ \int_{\hat{t}}^{\hat{t}+\delta} \big(u_L^*(\hat{a},\theta) - u_L^*(\theta)\big)\mathrm{d}F(\theta) - l_H^- \int_{\hat{t}-\delta}^{\hat{t}} \big(u_L^*(\hat{a},\theta) - u_L^*(\theta)\big)\mathrm{d}F(\theta) \\
&+ \int_{\hat{t}}^{\hat{t}+\delta} \big(h_a(\hat{a},\theta) - l_h^+ h(\hat{a},\theta)\big)\Big[\big(u_H^*(\hat{a},\theta) - u_L^*(\hat{a},\theta)\big) - \big(u_H^*(\theta) - u_L^*(\theta)\big)\Big]\mathrm{d}F(\theta) \\
&+ \int_{\hat{t}-\delta}^{\hat{t}} \big(h_a(\hat{a},\theta) - l_h^- h(\hat{a},\theta)\big)\Big[\big(u_H^*(\hat{a},\theta) - u_L^*(\hat{a},\theta)\big) - \big(u_H^*(\theta) - u_L^*(\theta)\big)\Big]\mathrm{d}F(\theta)
\end{aligned}
$$

Lemma 3. *Let $D(\theta) = u_H^*(\theta) - u_L^*(\theta) - \big(u_H^*(\hat{a},\theta) - u_L^*(\hat{a},\theta)\big)$, then*

$$
\int_{\hat{t}}^{\hat{t}+\delta} h_a(\hat{a},\theta)D(\theta)\mathrm{d}F(\theta) \int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)\mathrm{d}F(\theta) \geq \int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)D(\theta)\mathrm{d}F(\theta) \int_{\hat{t}}^{\hat{t}+\delta} h_a(\hat{a},\theta)\mathrm{d}F(\theta); \tag{10}
$$

$$
\int_{\hat{t}-\delta}^{\hat{t}} h_a(\hat{a},\theta)D(\theta)\mathrm{d}F(\theta) \int_{\hat{t}-\delta}^{\hat{t}} h(\hat{a},\theta)\mathrm{d}F(\theta) \geq \int_{\hat{t}-\delta}^{\hat{t}} h(\hat{a},\theta)D(\theta)\mathrm{d}F(\theta) \int_{\hat{t}-\delta}^{\hat{t}} h_a(\hat{a},\theta)\mathrm{d}F(\theta). \tag{11}
$$

Proof. First notice that by monotonicity of $a^*(\cdot)$ and $u_H^*(\cdot,\theta) - u_L^*(\cdot,\theta)$, $D(\theta)$ is strictly increasing, non-positive on $[\hat{t}-\delta, \hat{t}]$, and non-negative on $[\hat{t}, \hat{t}+\delta]$.

To prove (10), observe that h_a, h and D are all nonnegative, it suffices to show that for any $x, y \in [\hat{t}, \hat{t}+\delta]$,

$$
h(\hat{a}, x \wedge y)h_a(\hat{a}, x \vee y)D(x \vee y) \geq h_a(\hat{a}, x)h(\hat{a}, y)D(y).
$$

If $x \geq y$, then $LHS = h(\hat{a}, y)h_a(\hat{a}, x)D(x) \geq h_a(\hat{a}, x)h(\hat{a}, y)D(y) = RHS$ by monotonicity of D. If $x \leq y$, then $LHS = h(\hat{a}, x)h_a(\hat{a}, y)D(y) \geq h_a(\hat{a}, x)h(\hat{a}, y)D(y) = RHS$ by monotonicity of $l_h(\hat{a}, \cdot)$.

Therefore we can apply Theorem 2.1 in Karlin and Rinott (1980) and conclude (10) must hold.

Similarly for (11), it is equivalent to prove for any $x, y \in [\hat{t}-\delta, \hat{t}]$,

$$
-D(x)h_a(\hat{a}, x)h(\hat{a}, y) \leq -D(x \wedge y)h(\hat{a}, x \wedge y)h_a(\hat{a}, x \vee y).
$$

If $x \geq y$, then $LHS = -D(x)h_a(\hat{a},x)h(\hat{a},y) \leq -D(y)h(\hat{a},y)h_a(\hat{a},x)$ by monotonicity of D. If $x \leq y$, $LHS = -D(x)h_a(\hat{a},x)h(\hat{a},y) \leq -D(x)h(\hat{a},x)h_a(\hat{a},y)$, again by monotonicity of $l_h(\hat{a},\cdot)$. □

As a consequence of Lemma 3, we have

$$\int_{\hat{t}}^{\hat{t}+\delta} \big(\overline{\varepsilon_L} + u_L^*(\hat{a},\theta) - u_L^*(\theta)\big)\mathrm{d}F(\theta) \leq \frac{l_h^-}{l_h^+}\int_{\hat{t}-\delta}^{\hat{t}} \big[\underline{\varepsilon_L} - \big(u_L^*(\hat{a},\theta) - u_L^*(\theta)\big)\big]\mathrm{d}F(\theta). \tag{12}$$

Now we are ready to compare the profit between the pooling policy and the perfect information policy on the neighborhood. In particular, the profit difference, $\text{PD} := \Pi\big(\hat{a}, F|_{(\hat{t}+\delta,\hat{t}-\delta)}\big) - \int_{\hat{t}-\delta}^{\hat{t}+\delta} \Pi^*(\theta)dF(\theta)$, can be expressed as [12] :

$$\begin{aligned}
\text{PD} = &\underbrace{\int_{\hat{t}-\delta}^{\hat{t}+\delta} \big[h(\hat{a},\theta) - h(a^*(\theta),\theta)\big]P(\theta)\mathrm{d}F(\theta)}_{\text{Surplus Loss}} \\
&+ \underbrace{\int_{\hat{t}}^{\hat{t}+\delta} \big[\omega'(u_L^*(\theta))\overline{\varepsilon_L}(1-h(\hat{a},\theta)) - \omega'(u_H^*(\theta))\overline{\varepsilon_H}h(\hat{a},\theta)\big]\mathrm{d}F(\theta)}_{\text{Incentive Smoothing}^+} \\
&+ \underbrace{\int_{\hat{t}-\delta}^{\hat{t}} \big[\omega'(u_H^*(\theta))\underline{\varepsilon_H}h(\hat{a},\theta) - \omega'(u_L^*(\theta))\underline{\varepsilon_L}(1-h(\hat{a},\theta))\big]\mathrm{d}F(\theta)}_{\text{Incentive Smoothing}^-}
\end{aligned}$$

where $P(\theta) := H - \omega(u_H^*(\theta)) - \big(L - \omega(u_L^*(\theta))\big)$ is the net profit under perfect information and optimal contract at state θ.

Let $\Delta a(\theta) := a^{*\prime}(\theta)(\hat{t}-\theta)$. The first line of PD captures the expected profit loss due to the inflexibility of action $\hat{a}$, which can be further simplified as [13] :

$$\begin{aligned}
\text{Surplus Loss} &= \int_{\hat{t}-\delta}^{\hat{t}+\delta} \big[h_a(a^*(\theta),\theta)\Delta a(\theta)\big]P(\theta)\mathrm{d}F(\theta) \\
&= \int_{\hat{t}-\delta}^{\hat{t}+\delta} \big[\omega'(u_H^*(\theta))h(a^*(\theta),\theta)\frac{\partial u_H^*(a^*(\theta),\theta)}{\partial a} + \omega'(u_L^*(\theta))(1-h(a^*(\theta),\theta))\frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a}\big]\Delta a(\theta)\mathrm{d}F(\theta)
\end{aligned}$$

where the second equality comes from the optimality (first order condition) of $a^*(\theta)$ under

[12] This is up to the terms of the order $o(\varepsilon)$ in the integrand, which we ignore here and in what follows.

[13] Again, we omit all terms of the order $o(\delta)$ in the integrand here and in the remaining of the paper.

perfect information.

The second and third line of PD represent the net cost-saving through redistribute incentives across different states, which can be further expressed by:

$$\text{Incentive Smoothing}^{+} = \int_{\hat{t}}^{\hat{t}+\delta} \Big[\omega'\big(u_L^*(\theta)\big)\big(1-h(\hat{a},\theta)\big) - \omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\frac{\int_{\hat{t}}^{\hat{t}+\delta} 1-h(\hat{a},\theta)\mathrm{d}F}{\int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)\mathrm{d}F}\Big]\mathrm{d}F(\theta)\overline{\varepsilon_L}$$
$$-\underbrace{\frac{\int_{\hat{t}}^{\hat{t}+\delta}\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\mathrm{d}F(\theta)}{\int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)\mathrm{d}F(\theta)}}_{=\mathbb{E}\left[\omega'(u_H^*(\theta))\,\middle|\,H,+\right]}\int_{\hat{t}}^{\hat{t}+\delta}\big[h(\hat{a},\theta)\frac{\partial u_H^*(a^*(\theta),\theta)}{\partial a} + (1-h(\hat{a},\theta))\frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a}\big]\Delta a(\theta)\mathrm{d}F(\theta)$$

$$\text{Incentive Smoothing}^{-} = \int_{\hat{t}-\delta}^{\hat{t}} \Big[\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\frac{\int_{\hat{t}-\delta}^{\hat{t}} 1-h(\hat{a},\theta)\mathrm{d}F}{\int_{\hat{t}-\delta}^{\hat{t}} h(\hat{a},\theta)\mathrm{d}F} - \omega'\big(u_L^*(\theta)\big)\big(1-h(\hat{a},\theta)\big)\Big]\mathrm{d}F(\theta)\underline{\varepsilon_L}$$
$$-\underbrace{\frac{\int_{\hat{t}-\delta}^{\hat{t}}\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\mathrm{d}F(\theta)}{\int_{\hat{t}-\delta}^{\hat{t}} h(\hat{a},\theta)\mathrm{d}F(\theta)}}_{\mathbb{E}\left[\omega'(u_H^*(\theta))\,\middle|\,H,-\right]}\int_{\hat{t}-\delta}^{\hat{t}}\big[h(\hat{a},\theta)\frac{\partial u_H^*(a^*(\theta),\theta)}{\partial a} + (1-h(\hat{a},\theta))\frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a}\big]\Delta a(\theta)\mathrm{d}F(\theta)$$

where $\mathbb{E}\left[\omega'(u_H^*(\theta))\,\middle|\,H,+\right]$ and $\mathbb{E}\left[\omega'(u_H^*(\theta))\,\middle|\,H,-\right]$ are the expectations of u_H^* conditional on state is in one-sided neighborhood and high output being realized.

Observe that the coefficient of $\overline{\varepsilon_L}$ is negative. By convexity of ω' and monotonicity of u_H^* and $h(\hat{a},\cdot)$, we have

$$h(\hat{a},x\vee y)\omega'(u_H^*(x\vee y))(1-h(\hat{a},x\wedge y)) \geq h(\hat{a},y)\omega'(u_L^*(x))(1-h(\hat{a},x))$$

Again by Theorem 2.1 in Karlin and Rinott (1980),

$$\int_{\hat{t}}^{\hat{t}+\delta}\omega'\big(u_L^*(\theta)\big)\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta) \leq \frac{\int_{\hat{t}}^{\hat{t}+\delta} 1-h(\hat{a},\theta)\mathrm{d}F}{\int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)\mathrm{d}F}\int_{\hat{t}}^{\hat{t}+\delta}\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\mathrm{d}F(\theta).$$

Combining three parts, together with (12), yields

$$
\begin{aligned}
\text{PD} \geq \Bigg[& \Bigg(\frac{\int_{\hat{t}}^{\hat{t}+\delta} \omega'\big(u_L^*(\theta)\big)\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}{\int_{\hat{t}}^{\hat{t}+\delta}\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)} - \frac{\int_{\hat{t}}^{\hat{t}+\delta} \omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\mathrm{d}F(\theta)}{\int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)\mathrm{d}F(\theta)} \Bigg) \frac{l_H^-}{l_H^+} \frac{\int_{\hat{t}}^{\hat{t}+\delta}\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}{\int_{\hat{t}}^{\hat{t}+\delta} \mathrm{d}F(\theta)} \\
& + \Bigg(\frac{\int_{\hat{t}-\delta}^{\hat{t}} \omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\mathrm{d}F(\theta)}{\int_{\hat{t}-\delta}^{\hat{t}} h(\hat{a},\theta)\mathrm{d}F(\theta)} - \frac{\int_{\hat{t}-\delta}^{\hat{t}} \omega'\big(u_L^*(\theta)\big)\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}{\int_{\hat{t}-\delta}^{\hat{t}}\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)} \Bigg) \frac{\int_{\hat{t}-\delta}^{\hat{t}}\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}{\int_{\hat{t}-\delta}^{\hat{t}} \mathrm{d}F(\theta)} \Bigg] \int_{\hat{t}-\delta}^{\hat{t}} \mathrm{d}F(\theta) \underline{\varepsilon_L} \\
& + \int_{\hat{t}}^{\hat{t}+\delta} \Big[\omega'(u_H^*(\theta))h(a^*(\theta),\theta) - \mathbb{E}\big[\omega'(u_H^*(\theta)) \,\big|\, H,+\big]h(\hat{a},\theta) \Big] \frac{\partial u_H^*(a^*(\theta),\theta)}{\partial a} \Delta a(\theta)\mathrm{d}F(\theta) \\
& + \int_{\hat{t}-\delta}^{\hat{t}} \Big[\omega'(u_H^*(\theta))h(a^*(\theta),\theta) - \mathbb{E}\big[\omega'(u_H^*(\theta)) \,\big|\, H,-\big]h(\hat{a},\theta) \Big] \frac{\partial u_H^*(a^*(\theta),\theta)}{\partial a} \Delta a(\theta)\mathrm{d}F(\theta) \\
& + \int_{\hat{t}}^{\hat{t}+\delta} \Big[\omega'(u_L^*(\theta))\big(1-h(a^*(\theta),\theta)\big) - \mathbb{E}\big[\omega'(u_H^*(\theta)) \,\big|\, H,+\big]\big(1-h(\hat{a},\theta)\big) \Big] \frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a} \Delta a(\theta)\mathrm{d}F(\theta) \\
& + \int_{\hat{t}-\delta}^{\hat{t}} \Big[\omega'(u_L^*(\theta))\big(1-h(a^*(\theta),\theta)\big) - \mathbb{E}\big[\omega'(u_H^*(\theta)) \,\big|\, H,-\big]\big(1-h(\hat{a},\theta)\big) \Big] \frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a} \Delta a(\theta)\mathrm{d}F(\theta) \\
& - \int_{\hat{t}}^{\hat{t}+\delta} \Bigg[\omega'\big(u_L^*(\theta)\big)\big(1-h(\hat{a},\theta)\big) - \omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta) \frac{\int_{\hat{t}}^{\hat{t}+\delta} 1-h(\hat{a},\theta)\mathrm{d}F}{\int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)\mathrm{d}F} \Bigg] \frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a} \Delta a(\theta)\mathrm{d}F(\theta) \\
& - \int_{\hat{t}-\delta}^{\hat{t}} \frac{l_H^-}{l_H^+} \Bigg[\omega'\big(u_L^*(\theta)\big)\big(1-h(\hat{a},\theta)\big) - \omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta) \frac{\int_{\hat{t}}^{\hat{t}+\delta} 1-h(\hat{a},\theta)\mathrm{d}F}{\int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)\mathrm{d}F} \Bigg] \frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a} \Delta a(\theta)\mathrm{d}F(\theta)
\end{aligned}
$$

The integrands in third and fourth line are obviously $o(\delta)$.

We can combine the fifth and seventh line as

$$
\begin{aligned}
& \int_{\hat{t}}^{\hat{t}+\delta} \Bigg(\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta) \frac{\int_{\hat{t}}^{\hat{t}+\delta} 1-h(\hat{a},\theta)\mathrm{d}F}{\int_{\hat{t}}^{\hat{t}+\delta} h(\hat{a},\theta)\mathrm{d}F} - \mathbb{E}\big[\omega'(u_H^*(\theta)) \,\big|\, H,+\big]\big(1-h(\hat{a},\theta)\big) \Bigg) \frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a} \Delta a(\theta)\mathrm{d}F(\theta) \\
= & \int_{\hat{t}}^{\hat{t}+\delta} \underbrace{\Bigg(\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta) \frac{\mathbb{E}[1-h(\hat{a},\theta)\,|\,H,+]}{\mathbb{E}[h(\hat{a},\theta)\,|\,H,+]} - \mathbb{E}\big[\omega'(u_H^*(\theta)) \,\big|\, H,+\big]\big(1-h(\hat{a},\theta)\big) \Bigg)}_{\to 0 \text{ as } \delta \to 0} \frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a} \Delta a(\theta)\mathrm{d}F(\theta)
\end{aligned}
$$

which makes the integrand of order $o(\delta)$.

Similarly, we can combine the sixth and the eighth line as:

$$\int_{\hat{t}-\delta}^{\hat{t}}\Big(\underbrace{\frac{l_H^-}{l_H^+}}_{\to 1 \text{ as } \delta\to 0}\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\frac{\int_{\hat{t}}^{\hat{t}+\delta}1-h(\hat{a},\theta)\mathrm{d}F}{\int_{\hat{t}}^{\hat{t}+\delta}h(\hat{a},\theta)\mathrm{d}F}-\mathbb{E}\big[\omega'(u_H^*(\theta))\,\big|\,H,-\big]\big(1-h(\hat{a},\theta)\big)\Big)\frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a}\Delta a(\theta)\mathrm{d}F($$

$$=\int_{\hat{t}}^{\hat{t}+\delta}\underbrace{\Big(\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\frac{\mathbb{E}[1-h(\hat{a},\theta)\,|\,H,+]}{\mathbb{E}[h(\hat{a},\theta)\,|\,H,-]}-\mathbb{E}\big[\omega'(u_H^*(\theta))\,\big|\,H,+\big]\big(1-h(\hat{a},\theta)\big)\Big)}_{\to 0 \text{ as } \delta\to 0}\frac{\partial u_L^*(a^*(\theta),\theta)}{\partial a}\Delta a(\theta)\mathrm{d}F($$

Thus, its integrand also is of order $o(\delta)$.

In other words,

$$\mathrm{PD}\geq\Big[\Big(\frac{\int_{\hat{t}}^{\hat{t}+\delta}\omega'\big(u_L^*(\theta)\big)\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}{\int_{\hat{t}}^{\hat{t}+\delta}\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}-\frac{\int_{\hat{t}}^{\hat{t}+\delta}\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\mathrm{d}F(\theta)}{\int_{\hat{t}}^{\hat{t}+\delta}h(\hat{a},\theta)\mathrm{d}F(\theta)}\Big)\frac{l_H^-}{l_H^+}\frac{\int_{\hat{t}}^{\hat{t}+\delta}\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}{\int_{\hat{t}}^{\hat{t}+\delta}\mathrm{d}F(\theta)}$$

$$+\Big(\frac{\int_{\hat{t}-\delta}^{\hat{t}}\omega'\big(u_H^*(\theta)\big)h(\hat{a},\theta)\mathrm{d}F(\theta)}{\int_{\hat{t}-\delta}^{\hat{t}}h(\hat{a},\theta)\mathrm{d}F(\theta)}-\frac{\int_{\hat{t}-\delta}^{\hat{t}}\omega'\big(u_L^*(\theta)\big)\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}{\int_{\hat{t}-\delta}^{\hat{t}}\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}\Big)\frac{\int_{\hat{t}-\delta}^{\hat{t}}\big(1-h(\hat{a},\theta)\big)\mathrm{d}F(\theta)}{\int_{\hat{t}-\delta}^{\hat{t}}\mathrm{d}F(\theta)}\Big]\int_{\hat{t}-\delta}^{\hat{t}}\mathrm{d}F(\theta)\underline{\varepsilon_L}$$

So long as the coefficient of $\underline{\varepsilon_L}$ is not equal to 0 at all points in the small neighborhood of $\hat{t}$, then there exists a profitable deviation, rendering perfect information provision suboptimal.

□

Proof of Lemma 2. Take total differentiation with respect to θ, we obtain

$$\frac{\mathrm{d}\Pi^*(\theta)}{\mathrm{d}\theta}=\frac{\partial\Pi}{\partial a}\frac{\mathrm{d}a^*(\theta)}{\mathrm{d}\theta}+\frac{\partial\Pi}{\partial\theta}.$$

The second order differentiation further leads to,

$$\begin{aligned}\frac{\mathrm{d}^2\Pi^*}{\mathrm{d}\theta^2}&=\Big(\frac{\partial^2\Pi}{\partial a^2}\frac{\mathrm{d}a^*}{\mathrm{d}\theta}+\frac{\partial^2\Pi}{\partial a\partial\theta}\Big)\frac{\mathrm{d}a^*}{\mathrm{d}\theta}+\frac{\partial\Pi}{\partial a}\frac{\mathrm{d}^2a^*}{\mathrm{d}\theta^2}+\frac{\partial^2\Pi}{\partial a\partial\theta}\frac{\mathrm{d}a^*}{\mathrm{d}\theta}+\frac{\partial^2\Pi}{\partial\theta^2}\\&=\frac{\partial^2\Pi}{\partial a^2}\Big(\frac{\mathrm{d}a^*}{\mathrm{d}\theta}\Big)^2+2\frac{\partial^2\Pi}{\partial a\partial\theta}\frac{\mathrm{d}a^*}{\mathrm{d}\theta}+\frac{\partial^2\Pi}{\partial\theta^2}\\&=\frac{\partial^2\Pi}{\partial a^2}\Big(-\frac{\frac{\partial^2\Pi}{\partial\theta^2}}{\frac{\partial^2\Pi}{\partial a\partial\theta}}\Big)^2+2\frac{\partial^2\Pi}{\partial a\partial\theta}\Big(-\frac{\frac{\partial^2\Pi}{\partial\theta^2}}{\frac{\partial^2\Pi}{\partial a\partial\theta}}\Big)+\frac{\partial^2\Pi}{\partial\theta^2}\\&=\frac{\frac{\partial^2\Pi}{\partial\theta^2}}{\Big(\frac{\partial^2\Pi}{\partial a\partial\theta}\Big)^2}\Big[\frac{\partial^2\Pi}{\partial a^2}\frac{\partial^2\Pi}{\partial\theta^2}-\Big(\frac{\partial^2\Pi}{\partial a\partial\theta}\Big)^2\Big],\end{aligned}$$

which is positive if $\frac{\partial^2\Pi}{\partial a^2}\frac{\partial^2\Pi}{\partial \theta^2} < \left(\frac{\partial^2\Pi}{\partial a\partial\theta}\right)^2$. □

Appendix B: Risk-Neutral Agent

In this appendix, we provide the analysis to a risk-neutral agent. With risk-averse agent, we impose limited liability to the agent, and we further impose bounded payment such that $w_y^\sigma \in [0, \bar{w}]$ as introduced in Jewitt et al. (2008).

The production function, information environment, and the contract scheme are the same as in the main text. Except now, the agent is risk-neutral over wealth. His net utility is $w - c(a)$. The output y accrues directly to the principal. The principal is risk-neutral and maximizes her expected net profit $y - w$.

The principal's problem is the same but with additional constraints on payment.

$$\max_{\mathcal{E}, \{w_L^\sigma(\cdot), w_H^\sigma(\cdot)\}_{\sigma\in\Sigma}, a_\sigma} \int_\Sigma \int_\Theta \Big[(H - w_H^\sigma(\theta))h(a,\theta) + (L - w_L^\sigma(\theta))(1 - h(a,\theta))\Big] d\sigma(\theta) d\Pr(\sigma)$$

$$\text{s.t.} \int_\Theta \left(w_H^\sigma(\theta) - w_L^\sigma(\theta)\right) h_a(a_\sigma, \theta) d\sigma(\theta) = c'(a_\sigma), \quad \forall \sigma \in \Sigma \tag{IC$_\sigma$}$$

$$\int_\Theta \Big(w_H^\sigma(\theta) h(a_\sigma, \theta) + w_L^\sigma(\theta)(1 - h(a_\sigma, \theta))\Big) d\sigma(\theta) \geq c(a_\sigma), \quad \forall \sigma \in \Sigma \tag{IR$_\sigma$}$$

$$0 \leq w_L^\sigma \leq \bar{w};\ 0 \leq w_H^\sigma \leq \bar{w} \tag{M}$$

Claim 1. (Innes (1990)). *$w_L^\sigma = 0$ and IR$_\sigma$ is slack for all $\sigma \in \Sigma$.*

Without ambiguity, we adopt the same notation as in the illustration example.

Proposition 3. ***(Complete information solution)***. *Suppose the principal provides complete information to the agent, then the optimal wage scheme is the following: $w_H^*(\theta) = w_H(a^*(\theta), \theta)$ and $w_L^*(\theta) = 0$. Moreover, if $\frac{h_{aa}}{h_a}$ is increasing in state, then $a^*(\theta)$ is increasing in state.*

Proof. One can verify $h \cdot w_H^*$ is submodular, hence the profit is supermodular. □

Now consider another extreme scenario where the principal provides no information to the agent, i.e., all states map into a single signal σ which equals the prior distribution. In this case, the principal's problem becomes the following.

$$\max_{w_H^i(\theta), a} \int_\Theta \Big[\big(H - w_H^i(\theta)\big)h(a,\theta) + L\big(1 - h(a,\theta)\big)\Big] dF(\theta)$$

$$\text{s.t.} \int_{\Theta} w_H^i(\theta)\, h_a(a,\theta) dF = c'(a) \tag{IC$_F$}$$

$$0 \leq w_H^i \leq \bar{w} \tag{M}$$

Lemma 4. ***(Cost-minimizing wage for null information)*** *Suppose the principal provides no information to the agent. Then the optimal wage scheme for an fixed effort a is the following: $w_L^i(\theta) = 0$ for all $\theta \in [0,1]$ and*

$$w_H^i(\theta) = \begin{cases} \bar{w} & \text{if } \theta \in [t(a), 1] \\ 0 & \text{if } \theta \in [0, t), \end{cases}$$

where $t(a)$ is the solution to

$$\int_t^1 \bar{w} h_a(a,\theta) dF = c'(a). \tag{13}$$

In addition, the optimal action for no information is the solution to $\max_a \int_{\Theta} (H-L) h(a,\theta) dF(\theta) - \int_{t(a)}^1 \bar{w} h(a,\theta) dF(\theta)$.

The next result states that the complete-information benchmark is indeed suboptimal. More specifically, there exists a simple finite deterministic coarse information environment that dominates complete information.

Theorem 2. *Suppose $w_H^*(\theta) < \bar{w}$ for all $\theta \in \Theta$. Then there exists a finite number $N \geq 1$, such that dividing the state space into N equal width sub-environments dominates the complete information.*

Proof. Pick an arbitrary $\hat{\theta} \in \Theta$ and a small number $\delta > 0$. Consider a subset of the state space, $I_\delta \equiv [\hat{\theta} - \delta, \hat{\theta} + \delta]$. Consider the following two information environments. First, the principal pools the state $[\hat{\theta} - \delta, \hat{\theta} + \delta]$ into a single signal and designs the wage scheme as in Lemma 4 to induce the agent to take the action $\hat{a} \equiv a^*(\hat{\theta})$. Second, the principal provides perfect information to the agent on this subset of the state space. The profit difference can be decomposed as the sum of $IS(I_\delta)$ and $EL(I_\delta)$, where

$$IS(I_\delta) := \int_{\hat{\theta}-\delta}^{\hat{\theta}+\delta} w_H(\hat{a},\theta) h(\hat{a},\theta) dF - \int_{t(\hat{a},\delta)}^{\hat{\theta}+\delta} \bar{w} h(\hat{a},\theta) dF,$$

$$EL(I_\delta) := \int_{\hat{\theta}-\delta}^{\hat{\theta}+\delta} \Big(\underbrace{(H - L - w_H(\hat{a},\theta)) h(\hat{a},\theta)}_{\Pi(\hat{a},\theta)} - \underbrace{\big(H - L - w_H^*(\theta)\big) h(a^*(\theta),\theta)}_{\Pi(a^*(\theta),\theta)} \Big) dF,$$

are unconditional expectations. We take first-order and second-order derivatives for $IS(I_\delta)$ and

$EL(I_\delta)$ with respect to δ respectively.

$$\frac{dEL(I_\delta)}{d\delta} = f(\hat\theta+\delta)\big(\Pi(\hat a,\hat\theta+\delta)-\Pi(a^*(\hat\theta+\delta),\hat\theta+\delta)\big) + f(\hat\theta-\delta)\big(\Pi(\hat a,\hat\theta-\delta)-\Pi(a^*(\hat\theta-\delta),\hat\theta-\delta)\big)$$

$$\begin{aligned}\frac{d^2EL(I_\delta)}{d\delta^2} =& f'(\hat\theta+\delta)\big(\Pi(\hat a,\hat\theta+\delta)-\Pi(a^*(\hat\theta+\delta),\hat\theta+\delta)\big)\\ &- f'(\hat\theta-\delta)\big(\Pi(\hat a,\hat\theta-\delta)-\Pi(a^*(\hat\theta-\delta),\hat\theta-\delta)\big)\\ &+ f(\hat\theta+\delta)\bigg(\frac{\partial\Pi(\hat a,\hat\theta+\delta)}{\partial\theta}-\frac{\partial\Pi(a^*(\hat\theta+\delta),\hat\theta+\delta)}{\partial\theta}-\frac{\partial\Pi(a^*(\hat\theta+\delta),\hat\theta+\delta)}{\partial a}\frac{da^*(\hat\theta+\delta)}{d\theta}\bigg)\\ &- f(\hat\theta-\delta)\bigg(\frac{\partial\Pi(\hat a,\hat\theta-\delta)}{\partial\theta}-\frac{\partial\Pi(a^*(\hat\theta-\delta),\hat\theta-\delta)}{\partial\theta}-\frac{\partial\Pi(a^*(\hat\theta-\delta),\hat\theta-\delta)}{\partial a}\frac{da^*(\hat\theta-\delta)}{d\theta}\bigg)\end{aligned}$$

It is easy to verify that $\frac{dEL(I_\delta)}{d\delta}=0$ when δ converges to 0. Moreover, since $a^*(\theta)$ is determined by the first order condition, $\frac{da^*(\hat\theta)}{d\theta}=0$, hence, $\frac{d^2EL(I_\delta)}{d\delta^2}=0$. Now let us look at the incentive smoothing part. To simplify notation, we use t_δ to represent the solution of t given $(\hat a,\delta)$. By implicit function theorem,

$$\frac{dt_\delta}{d\delta}=\bigg(h_a(\hat a,\hat\theta+\delta)f(\hat\theta+\delta)-\frac{c'(\hat a)}{\bar w}\big(f(\hat\theta+\delta)+f(\hat\theta-\delta)\big)\bigg)\bigg/(h_a(\hat a,t_\delta)f(t_\delta)).$$

Hence,

$$\begin{aligned}\frac{dIS(I_\delta)}{d\delta} =& f(\hat\theta+\delta)\big(w_H^*(\hat a,\hat\theta+\delta)-\bar w\big)h(\hat a,\hat\theta+\delta)+\frac{dt_\delta}{d\delta}\bar w h(\hat a,t_\delta)f(t_\delta)+f(\hat\theta-\delta)w_H^*(\hat a,\hat\theta-\delta)h(\hat a,\hat\theta-\delta)\\ =& f(\hat\theta+\delta)\big(\bar w h_a(\hat a,\hat\theta+\delta)-c'(\hat a)\big)\bigg(\frac{h(\hat a,t_\delta)}{h_a(\hat a,t_\delta)}-\frac{h(\hat a,\hat\theta+\delta)}{h_a(\hat a,\hat\theta+\delta)}\bigg)+f(\hat\theta-\delta)c'(\hat a)\bigg(\frac{h(\hat a,\hat\theta-\delta)}{h_a(\hat a,\hat\theta-\delta)}-\frac{h(\hat a,t_\delta)}{h_a(\hat a,t_\delta)}\bigg)\end{aligned}$$

which is strictly positive if $\delta>0$ and equals 0 if δ goes to 0. Since both terms $\frac{h(\hat a,\hat\theta-\delta)}{h_a(\hat a,\hat\theta-\delta)}-\frac{h(\hat a,t_\delta)}{h_a(\hat a,t_\delta)}$ and $\frac{h(\hat a,t_\delta)}{h_a(\hat a,t_\delta)}-\frac{h(\hat a,\hat\theta+\delta)}{h_a(\hat a,\hat\theta+\delta)}$ converges to zero when δ goes to 0. Therefore,

$$\begin{aligned}\frac{d^2IS(I_\delta)}{d\delta}\bigg|_{\delta\to0} =& f(\hat\theta)\big(\bar w h_a(\hat a,\hat\theta)-c'(\hat a)\big)\bigg[\bigg(\frac{1}{l_H(\hat a,\hat\theta)}\bigg)'\frac{dt_\delta}{d\delta}-\bigg(\frac{1}{l_H(\hat a,\hat\theta)}\bigg)'\bigg]+f(\hat\theta)c'(\hat a)\bigg[-\bigg(\frac{1}{l_H(\hat a,\hat\theta)}\bigg)'-\bigg(\frac{1}{l_H(\hat a,\hat\theta)}\bigg)'\frac{dt_\delta}{d\delta}\bigg]\\ =& -f(\hat\theta)\bigg(\frac{1}{l_H(\hat a,\hat\theta)}\bigg)'\bigg(2c'(\hat a)\bigg(1+\frac{dt_\delta}{d\delta}\bigg)\bigg)\\ =& f(\hat\theta)\bigg(\frac{1}{l_H(\hat a,\hat\theta)}\bigg)'4c'(\hat a)\bigg(\frac{c'(\hat a)}{\bar w h_a(\hat a,\hat\theta)}-1\bigg).\end{aligned}$$

The final output is bigger than 0 because the state-wise compensation for action $\hat a$ is below $\bar w$. Moreover, since the optimal action in the complete-information benchmark is bounded below from 0,[14] and $\Big(\frac{1}{l_H(\hat a,\hat\theta)}\Big)'<0$, there exists a $\underline{\delta}>0$ such that $\frac{d^2IS(I_\delta)}{d\delta}+\frac{d^2EL(I_\delta)}{d\delta}>0$ for every $\delta<\underline{\delta}$ and for every $\hat\theta\in\Theta$. Hence, N is the smallest integer that is larger than $1/\underline{\delta}$. □

The reason that Incentive smoothing dominates efficiency gain for small interval of states is following. When states are close by, the optimal action incurs the first-order loss, which almost

[14]Since $h_a(a,0)\gg0$, $a^*(0)\gg0$.

equals to zero. On the other hand, incentive smoothing has a larger first-order gain because to induce the same action using state-wise incentive $w_H(\hat{a},\theta)$ is far away from providing the optimal wage incentive using the interim incentive.

Proposition 4. *Suppose there are two intervals I_1 and I_2 where $\max\{I_1\} < \min\{I_2\}$. If $\frac{h_{aa}}{h_a}$ is increasing in θ, then the optimal action for I_1 is smaller than that for I_2.*

Proof. For an interval $I = [\underline{I}, \overline{I}]$, the profit for inducing an action a is the following

$$\Pi(a,I) := \int_I \frac{(H-L)h(a,\theta)}{F(\overline{I})-F(\underline{I})}dF(\theta) - \int_{t_I(a)}^1 \frac{\bar{w}h(a,\theta)}{F(\overline{I})-F(\underline{I})}dF(\theta),$$

where $t_I(a)$ is the solution to $\int_t^{\overline{I}} \bar{w}h_a(a,\theta)dF = \int_I c'(a)dF$. We want to show $\int_I (H-L)h(a,\theta)dF(\theta)$ is supermodular in a and I and $\int_{t_I(a)}^1 \bar{w}h(a,\theta)dF(\theta)$ is submodular in a and I. The former one is easy to verify. We show the second one below. The first-order derivative of the expected wage with respect to a can be simplified to,[15]

$$\frac{d}{da}\left(\int_{t_I(a)}^1 \frac{\bar{w}h(a,\theta)}{F(\overline{I})-F(\underline{I})}dF(\theta)\right) = \left(\frac{1}{l_H(a,t_I(a))}\right)\left(c'(a)\frac{\int_{t_I(a)}^{\overline{I}} -h_{aa}dF}{\int_{t_I(a)}^{\overline{I}} h_a dF} + c''(a)\right) + c'(a).$$

Given that $-h_{aa}/h_a$ decreases in θ, then $\frac{\int_{t_I(a)}^{\overline{I}} -h_{aa}dF}{\int_{t_I(a)}^{\overline{I}} h_a dF}$ decreases if I increases in strong order set. In addition, $l_H(a,t_I(a))$ also increases if I increases in strong order set as $t_I(a)$ is interior in I. Hence, the right hand side is decreasing in I. Therefore, $\Pi(a,I)$ is supermodular and we then conclude the optimal action increases in the index k of the intervals. □

Intuitively, when the upper bound on wage $\bar{w}$ is larger, the principal should have larger incentive to pool on the information. The next result confirms this intuition.

Proposition 5. *N is decreasing $\bar{w}$.*

Proof. For an arbitrary $\hat{\theta} \in \Theta$ and $\delta > 0$, take implicit differentiation of t_δ,

$$\frac{\mathrm{d}t_\delta}{\mathrm{d}\bar{w}} = \frac{\int_{t_\delta}^{\hat{\theta}+\delta} h_a(a,\theta)dF}{\bar{w}h_a(a,t_\delta)f(t_\delta)}.$$

[15] Note that $\frac{dt_I(a)}{da} = \frac{\int_{t_I(a)}^{\overline{I}} \bar{w}h_{aa}dF - c''(a)\int_I dF}{\bar{w}h_a(a,t_I(a))f(t_I(a))}$.

Substitute it into $\frac{dIS(I_\delta)}{d\bar{w}}$,

$$\frac{dIS(I_\delta)}{\mathrm{d}\bar{w}} = \int_{t_\delta}^{\hat{\theta}+\delta} \left(\frac{l_H(a,\theta)}{l_H(a,t_\delta)} - 1 \right) h(a,\theta) \mathrm{d}F > 0,$$

because $\frac{l_H(a,\theta)}{l_H(a,t_\delta)} > 1$ for $\theta > t_\delta$. Hence, for each interval $[\hat{\theta} - \delta, \hat{\theta} + \delta]$, the gain from incentive smoothing is uniformly larger. Moreover, since the efficiency gain is independent of $\bar{w}$, this further implies that $\underline{\delta}$ derived in the proof of Theorem 2 is larger and thereby N is smaller. □