Ignorance is Bliss: Ex-Ante vs. Ex-Post Information Systems in an Agency Model*

Jin Yong Jung[†] Son Ku Kim[‡] Seung Joo Lee[§]

September 30, 2025

Abstract

This paper studies the value of ex-ante information in a principal-agent model where such information is about random variables that affect the agent's utility and the timing of information revelation is an issue. We show that the principal and the agent's commonly observing information on those random variables ex ante (i.e., *before* the agent takes an action) adds no value to their observing it ex post (i.e., *after* the agent has taken an action). We also uncover that there is a negative relationship between the amount of ex-ante information contained in an information system and its efficiency in the principal-agent relation.

Keywords: Agency, Ex-Ante Information, Ex-Post Information

*We would like to thank Nicolae Gârleanu and Chris Shannon for helpful discussions.

[†]Department of Economics, Kangnam University. Email: jyjung@kangnam.ac.kr
[‡]Department of Economics, Seoul National University. Email: sonkukim@snu.ac.kr
[§]Saïd Business School, Oxford University. Email: seung.lee@sbs.ox.ac.uk

1 Introduction

The moral-hazard problem has been one of the central topics in the economics of information over the past four decades. It arises when a principal has to delegate to an agent to act on her behalf but cannot observe the agent's action choice directly. Most principal-agent theories have assumed that the agent's utility depends on monetary income to be granted by the principal. Thus, to control the agent's hidden action choice, the principal designs an incentive contract which ties the agent's pay to the observables that are imperfectly correlated with the agent's action choice.[1]

However, it is natural that the agent's utility from a job depend not only on the monetary compensation that is paid from the principal but also some other factors that are usually beyond the principal's control. Information on such variables will affect workers' decision making if revealed to the workers in advance. We can think about the following examples:

Example 1 A worker can infer about his personal match to the new workplace (e.g., fitness to the culture, aptitude for the nature of the work, a sense of belonging to different peer groups, and a feeling of accomplishment from the job) through on-site visits and some initial induction programs offered by the company. The company's hiring manager is deciding whether and when the firm should hold those events: (i) before he signs on the contract; (ii) after he signs on the contract but before starts his job; and (iii) never offers those programs for new hires ex ante and let workers figure out. Eventually, both the hiring manger and the new hire get to know about the degree of their match, and it is used as a contractual variable.

Example 2 A company provides its workers with pension plans along with regular salary schemes, which will be effective upon retirement. Final payouts from those pension plans are affected not by decisions made by the firm and workers but by a separate pension manager's portfolio decisions and the state of market. The company's CEO knows the track records of the pension manager's performance during the past few years, and is considering whether and when she should reveal those data to her new employee: (i) before he signs on the contract; (ii) after he signs on the contract but before starts his job; and (iii) never offers the information to the new hire ex ante and let the worker figure out. Eventually, the

[1]For the standard moral-hazard model, see Ross (1973), Mirrlees (1974), Harris and Raviv (1979), Holmström (1979), Holmström (1982) and Grossman and Hart (1983) among others.

pension performance will be revealed and included as a contractual variable.

In this paper, we ask the following natural question: When should information on those variables that affect the worker's utility be revealed? Should information on those variables be revealed as precisely as possible? For instance, in the above pension example, should the CEO make portfolio managers reveal information on the prospect of the company's pension plans before they sign on the contracts with their workers, or after the contracts but before the workers make their decisions, or even after the workers make their decisions? How does timing of such information revelation affect the efficiency of principal-agent relationships? We try to answer these questions through the standard theoretical framework.

We formulate a principal-agent model where the agent's utility depends not only on his monetary wage from the principal but also on some other variables such as non-wage benefits or his personal matches to the working environment. We consider the case in which the principal implements one out of the following three different information systems: (i) *the pre-contract information system*, (ii) *the post-contract ex-ante information system*, and (iii) *the post-contract ex-post information system*. Under the pre-contract information system, perfect information on those random variables is publicly revealed before the principal and the agent sign on the contract. On the other hand, the post-contract ex-ante information system publicly reveals perfect information about such variables after the contract is signed but before the agent takes an action, whereas the post-contract ex-post information system publicly reveals such information after the agent has taken an action.

The principal prefers the post-contract ex-ante information system to the pre-contract information system. This is because when the true value of a variable that affects the agent's utility is available even before the principal and the agent sign on the contract (i.e., the pre-contract information system), the principal should design an incentive contract that satisfies the agent's participation constraint for every realization of the variable, whereas she designs a wage contract that satisfies the agent's participation constraint only *on average* under the post-contract ex-ante information system where the information is revealed after they sign on the contract, which would be easier.

We show that the principal also prefers the post-contract ex-post information system to the post-contract ex-ante information system, implying that the principal and the agent's being informed ex ante (i.e., before the agent takes action) is not as efficient as their being informed ex post (i.e., after the agent takes action). Under the ex-ante information system, the principal can make use of such information to induce the agent to take an action which is optimal according to the given information. However, the agent will use such information for his own interests rather than for the principal's interests. Therefore, the principal should design a contract that satisfies the agent's incentive constraint for every realization of the random variable. On the other hand, if the principal and the agent are informed ex post, then the principal has to induce the agent to take a certain action that is constant across different realizations of that variable, which is usually sub-optimal. However, the principal, in this case, needs to design a wage contract that satisfies the agent's incentive constraint only on expectation, which is less costly. Therefore, there is a trade-off between both parties' being informed ex ante and their being informed ex post. We illustrate that the cost of their being informed ex ante always exceeds its benefit in reasonable cases.

In general, the efficiency of an information system in principal-agent models is affected not only by when it reveals information on random variables that affect the agent's hidden action choice, but also by how much it can say about those random variables. Therefore, we investigate a relationship between the amount of information contained in the post-contract ex-ante information system and its efficiency. To this end, we consider the case in which the principal has to implement one of two imperfect post-contract ex-ante information systems, where one information system publicly reveals more precise ex-ante information on random variables that affect the agent's utility than the other, in the sense that its information partition is finer than the other (i.e., the other system is more coarse). To see the pure effects that the "amount" of ex-ante information contained in an information system bears on the efficiency of that information system, we assume that perfect information on those variables is always available ex-post[2]. We show that there is a positive relationship between coarseness of ex-ante information system and its efficiency, implying that the more precise ex-ante information an information system generates, the less efficient the information system becomes in the principal-agent relation.[3]

The relationship between the amount of information contained in an information system and its efficiency in the principal-agent model has been widely discussed in the literature. Holmström (1979) and Shavell (1979) show that, between two costless public information systems, the information system with additional signals is strictly preferred by the principal

[2]Therefore, a contract can be dependent upon the realized values of those variables ('t' in our paper) as well as realized output.

[3]Thus, in the pension plan example, we can infer that the plan managers should not be allowed to reveal any information on prospects of their pension plans before the workers make their decisions. In the worker's match example, the company's manager should not hold on-site visits and induction sessions where new hires could get a sense of compatibility with the company. Note that we do not consider the productivity-boosting effect of those induction sessions.

if and only if at least one of those additional signals is informative about the agent's hidden action choice.[4] Gjesdal (1982) and Grossman and Hart (1983) apply Blackwell's notion of information sufficiency to the agency model, and showed that if two different information systems satisfy the Blackwell's information sufficiency conditions, then they can be ranked in terms of efficiency in canonical principal-agent models. Kim (1995) shows that Blackwell's information sufficiency conditions are unnecessarily strong for ranking information systems in agency models, and provides a weaker criterion, i.e., if one information system entails a likelihood ratio distribution which is a mean preserving spread of that of the other information system, then the former is more efficient than the latter in the agency model.[5]

Those previous works consider the efficiency of an information system which generates information on random variables that are affected by the agent's hidden action, e.g., outputs, annual earnings, and stock prices. Furthermore, information on those variables is available after the agent has taken an action (i.e., ex-post). In contrast, in this paper, we consider the efficiency of an information system which generates signals on random variables that are not affected by but affect the agent's action choice, and primarily focus on the relationship between the efficiency of the information system and its *timing* and *amount* of information revelation.

One related work to note is Sobel (1993), who compares three information systems in agency models: the zero information system, the pre-contract information system, and the post-contract information system (i.e., the post-contract ex-ante information system in our paper). He shows that the principal will always prefer the post-contract information system to the pre-contract information system,[6] but, depending on the underlying model specification, the principal may or may not prefer the pre-contract and post-contract information systems to the zero information system. However, there are several differences between his paper and ours. First, information is revealed privately to the agent in his paper, whereas in ours it is revealed both to the principal and the agent at the end of the period. Second, information is about the agent's cost function in his paper, whereas it is about the agent's utility function in ours. But, most of all, Sobel (1993) did not consider the post-contract ex-post information system, whereas comparing the post-contract ex-ante information system with the post-contract ex-post information system is the main result of our paper.

More recently, Silvers (2012) considers the case where, after the principal and the agent

[4]Their notion of informativeness is defined in terms of a sufficient statistic.

[5]For proper interpretations, see Kim (1995). Kim (1995) calls it 'MPS (Mean Preserving Spread) criterion' based on Rothsscild and Stiglitz (1970).

[6]The same result is derived in our paper in Lemma 1.

sign on contract, the principal receives a signal (private or public)[7] related to the technology. When the signal is private, then the principal announces the signal value to the agent possibly not truthfully. Related to our main results, he concludes that if the principal wants to implement the same action for any signal ('t' in our paper), then the public signal (before agent takes an action) lowers the efficiency, as the agent takes advantage of this public signal in pursuit of his own interests. In contrast to his paper, we do not consider adverse selection issues along with the truth-telling mechanism, and instead focus on the timing and amount of the public signal's revelation. Also, our signal is about the agent's marginal utility rather than technology. Banker et al. (2019) studies the value of information in the case in which the principal does not know some relevant traits of the agent ('t' in our paper) and uses the set (menu) of contracts to solve both adverse selection and moral hazard problems. We instead focus on the moral hazard side and how the timing and precision of information revelation affects the efficiency of agency relations.

Following the information design literature originated from Kamenica and Gentzkow (2011), Boleslavsky and Kim (2020) consider the world where the sender designs a signal about output, affected by the agent's hidden action, and develop a general solution method of characterizing the optimal "signal structure" when the moral hazard problem is at stake. We in contrast focus on the timing and coarseness of the relevant "exogenous" information relevant to the agent's utility, and its implication for the efficiency of agency relations.

Layout Section 2 formulates our basic model. In Section 3, we compare the efficiency of three different information systems. In Section 4, we compare the relative efficiency of two imperfect post-contract ex-ante information systems, one of which publicly reveals more precise ex-ante information than the other. Section 5 provides illustrative examples.[8] Concluding remarks are given in Section 6, and all the proofs of the lemmas and propositions in this paper are provided in Appendix A and Appendix B.

[7] The public signal means that the agent as well as the principal gets to know the signal's exact value. And contracts can be dependent upon both this signal value and the realized output level as in our specification.

[8] Examples in Section 5 allow us to directly decompose the agency cost in a closed form and illustrate the reason that the ex-ante information revelation hurts the efficiency of the agency relation, compared with the ex-post system.

2 The Basic Model

Consider a single-period principal-agent model in which a risk-averse agent works for a risk-neutral principal by providing an action (e.g., effort) $a \in A = [0, \infty)$. Output $x \in X = [\underline{x}, \overline{x}] \subset \mathbb{R}$ is determined not only by the agent's effort but also by the state of nature, θ, i.e., $x = X(a, \theta)$, and will be publicly observable at the end of the period.[9] We assume that production function $X(a, \theta)$ is additively separable. Thus, without loss of generality, $X(a, \theta)$ can be simply represented by

$$X(a, \theta) = a + \theta, \quad E(\theta) = 0.\text{[10]} \tag{1}$$

To suppress θ, we denote $f(x|a)$ as the output density function conditional on agent's effort choice, which is assumed to be twice differentiable with respect to a. After x is realized, the principal pays s to the agent as a monetary wage.

The agent has an additively separable utility function such as

$$U(s, t, a) = u(s, t) - v(a), \tag{2}$$

where $u(\cdot)$ denotes the agent's utility and $v(\cdot)$ denotes his disutility of exerting effort a. An important difference of the agent's utility function described in (2) from that of the standard principal-agent setting is that the agent's utility now depends not only on his monetary wage s but also on another random factor t. In the above equation, $t \in T = [\underline{t}, \overline{t}] \subset R$ represents all random factors that affect the agent's marginal utility. For example, it can be understood as variables that indicate the agent's matching characteristics to his working environments or non-wage benefits granted from the principal at the end of the period such as benefits from a pension plan, the exact amount of which is generally beyond the principal's control.

[9] $\mathbb{R}$ denotes the set of all real numbers. $\underline{x}$ can be $-\infty$ and $\overline{x}$ can also be $+\infty$.

[10] The general form of additively separable production function is

$$X(a, \theta) = \alpha(a) + \theta, \ \alpha' > 0.$$

However, one can easily rewrite the production function as

$$X(\alpha, \theta) = \alpha + \theta,$$

where α now can be considered as agent's effort level. In this case, the agent's cost of effort $v(a)$ can be rewritten as

$$c(\alpha) \equiv c(\alpha(a)) = v(a).$$

Information on variable t is eventually available to both the principal and the agent. However, *when* such information becomes available to each party depends on the information structure that is implemented by the principal at the outset. In this paper, we consider three different information systems: a (i) *pre-contract information system*, (ii) a *post-contract ex-ante information system*, and (iii) a *post-contract ex-post information system*. All these information systems publicly reveal perfect information on t but they differ in the timing of information revelation. Under the pre-contract information system, the principal and the agent can observe the true value of t before they sign on the contract. On the other hand, the post-contract ex-ante information system reveals the true value of t just after the contract and thereby the principal and the agent can observe it before the agent takes a, whereas the post-contract ex-post information system reveals it after the agent has taken a.

The timeline of our model is summarized as follows:

Stage 1: The principal implements one of the following information systems: (i) the pre-contract information system, (ii) the post-contract ex-ante information system, and (iii) the post-contract ex-post information system.

Stage 2: The principal and the agent agree upon the contract based on the information system chosen in **Stage 1**, i.e., $s^p(x,t), s^*(x,t)$ or $s^o(x,t)$, where $s^p(x,t)$ denotes the agent's optimal contract when the pre-contract information system is chosen, whereas $s^*(x,t)$ and $s^o(x,t)$ denote the agent's optimal wage contracts under the post-contract ex-ante and ex-post information systems, respectively.

Stage 3: The agent chooses his effort level a. Under the post-contract ex-post information system, the agent takes an effort before a true value of t is revealed. However, under either the pre-contract information system or the post-contract ex-ante information system, the agent takes an effort based on the true value of t.

Stage 4: The output level, x, is realized and the principal pays s to the agent according to the wage contract determined in Stage 2.

For analytical simplicity, we make the following assumptions.

Assumption 1. $u_s > 0,\ u_{ss} < 0,\ v' > 0,\ v'' > 0.$

Assumption 2. $u_{st} \neq 0$ *for some* (x,t).

In Assumptions 1 and 2, each subscript or prime indicates that derivatives are taken at the corresponding orders, i.e., $u_{ss} \equiv \frac{\partial^2}{(\partial s)^2}u$. Assumption 1 states that the agent is both risk and effort averse. Assumption 2 specifies the relationship between monetary wage s and random variable t in the agent's preference. It simply states that the agent's marginal utility with respect to s is affected by t.

3 Three Information Systems

3.1 Pre-Contract Information System

When pre-contract information system is implemented, a true value of t becomes available both to the principal and the agent *before* they agree upon the contract. Thus, the principal's optimization program for any given t[11] can be written as:

$$\begin{aligned}
&\max_{a_t \in A, s_t(x) \in S_t} \int_x (x - s_t(x)) f(x|a_t) dx \quad \text{s.t.} \\
&(i) \quad \int_x u(s_t(x), t) f(x|a_t) dx - v(a_t) \geq \overline{U}, \\
&(ii) \;\; a_t \in \arg\max_{a'} \int_x u(s_t(x), t) f(x|a') dx - v(a'), \; \forall a' \in A, \\
&(iii) \;\; s_t(x) \geq 0, \;\; \forall x \in X,
\end{aligned} \tag{3}$$

where $\overline{U}$ denotes the agent's reservation utility level that he obtains if he is employed elsewhere. In the above program, S_t denotes the set of admissible contracts given t, satisfying:

$$S_t \equiv \{s : X \to R | \; s(\cdot) \text{ is Lebesgue measurable}\}. \tag{4}$$

The first constraint of (3) is the agent's participation constraint and the second constraint is the agent's incentive constraint, reflecting the fact that both the principal and the agent can observe the true value of t even before they sign on the contract. Also, the third constraint represents the agent's limited liability, which indicates that the agent's subsistence level of monetary wage which is normalized by 0 must be guaranteed for any x given any realized t. This limited liability constraint on the agent's side is given to guarantee the existence of an optimal wage contract given t.[12]

[11] The optimally induced effort a_t and the optimal wage contract $s_t(\cdot)$ both depend upon the realized value of t.

[12] For details about this 'unpleasantness' arising from possibly unbounded likelihood ratios, see Mirrlees (1974) and Jewitt et al. (2008).

Since the contract, $s_t(x)$, as well as the agent's optimal effort choice, a_t, can depend on the realization of t, the principal's above optimization program can be equivalently written as:

$$\begin{aligned}
&\max_{a(t)\in A, s(x,t)\in S} \int_t \int_x (x - s(x,t)) f(x|a(t))h(t)dxdt \quad \text{s.t:} \\
&(i) \quad \int_x u(s(x,t),t) f(x|a(t))dx - v(a(t)) \geq \overline{U}, \qquad \forall t \in T, \\
&(ii) \quad a(t) \in \arg\max_{a'} \int_x u(s(x,t),t) f(x|a')dx - v(a'), \ \forall a' \in A, \ \forall t \in T, \\
&(iii) \quad s(x,t) \geq 0, \ \ \forall (x,t) \in X \times T,
\end{aligned} \tag{5}$$

where $h(t)$ is the probability density function of t which is common knowledge to both the principal and the agent. As before, S in program (5) denotes the set of admissible contracts satisfying:

$$S \equiv \{s : X \times T \to R | \ s(\cdot) \text{ is Lebesgue measurable}\}. \tag{6}$$

Assuming the first-order approach is valid,[13] the above optimization program reduces to:

[13] Following Jewitt (1988), it can be shown that the first-order approach is valid if the following conditions are satisfied:
(J1) Define $u_s^{-1}(\frac{1}{z},t) \equiv u_s^{-1}(u_s(s,t),t) \equiv s$ for given t. Then, $u(u_s^{-1}(\frac{1}{z},t),t) \equiv w_t(z)$, is concave in $z > 0$ for any given t.
(J2) The output density function $f(x|a)$ satisfies
(2.1) $\int_{-\infty}^{y} F(x|a)dx$ is non-increasing and convex in a for $\forall y$, where $F(x|a)$ denotes the cumulative distribution function of technology $f(x|a)$,
(2.2) $\int xf(x|a)dx$ is non-decreasing and concave in a, and
(2.3) $\frac{f_a}{f}(x|a)$ is non-decreasing and concave in x for each value of a.
More precisely, let $s^F(x,t)$ be the optimal contract obtained using the first-order approach. Then, conditions in (J1) and (2.3) of (J2) sufficiently guarantee that $u(s^F(x,t),t)$ is concave in x for any given t and, (2.1) and (2.2) of (J2) also sufficiently guarantee that

$$\int_x u(s^F(x,t),t) f(x|a)dx \equiv \phi_t(a) \tag{7}$$

is also concave in a for any given t. Consequently, the conditions in (J1) and (J2) sufficiently justify the use of the first-order approach in solving (5). As shown in Jewitt (1988), (J1) and (J2) are weaker than the usual MLRP (i.e., Monotone Likelihood Ratio Property) and CDFC (i.e., Convexity of Distribution Function Condition) conditions derived by Grossman and Hart (1983) and Rogerson (1985). For example, $u(s,t) = \frac{1}{r}(ts)^r$ or $u(s,t) = \frac{1}{r}(s+t)^r$ satisfies (J1) if $r \leq \frac{1}{2}$, and many familiar distribution families such as exponential, Chi-squared, and Poisson distributions satisfy (J2). For more details, see Jewitt (1988), and for more recent treatments of this issue, see Sinclair-Desgagné (1994), Conlon (2009), Jung and Kim (2015) among others.

$$
\begin{aligned}
&\max_{a(t)\in A, s(x,t)\in S} \int_t \int_x (x - s(x,t))\, f(x|a(t))h(t)dxdt \quad \text{s.t.}\\
&(i) \quad \int_x u(s(x,t),t) f(x|a(t))dx - v(a(t)) \geq \overline{U}, \quad \forall t \in T,\\
&(ii) \quad \int_x u(s(x,t),t) f_a(x|a(t))dx - v'(a(t)) = 0, \ \forall t \in T,\\
&(iii) \ \ s(x,t) \geq 0, \ \ \forall (x,t) \in X \times T,
\end{aligned}
\tag{8}
$$

where f_a denotes the first derivative of f with respect to a.

Let $(a^p(t), s^p(x,t))$ be the optimal solution for the above program. Then, solving Euler equation of the above program gives that $s^p(x,t)$ must satisfy:

$$
\frac{1}{u_s(s^p(x,t),t)} = \lambda^p(t) + \mu^p(t)\frac{f_a}{f}(x|a^p(t)), \tag{9}
$$

for almost every (x,t) where (9) has a solution $s^p(x,t) \geq 0$ and otherwise $s^p(x,t) = 0$. In equation (9), $\lambda^p(t)$ denotes the optimized Lagrangian multiplier of the agent's participation constraint given t and $\mu^p(t)$ is the optimized Lagrangian multiplier of the agent's incentive constraint given t.

We define the principal's optimized benefits under the pre-contract information system as:

$$
PW^p \equiv \int_t \int_x (x - s^p(x,t))\, f(x|a^p(t))h(t)dxdt. \tag{10}
$$

3.2 Post-Contract Ex-Post Information System

When the post-contract ex-post information system is implemented, t level will be publicly revealed to both the principal and the agent at the end of the contracting period. Therefore, it can be used as a contractual variable, and thus the agent's wage contract, s, is based on (x,t), i.e., $s = s(x,t)$. However, his effort choice, a, must be constant across different realizations of t because t will be available *after* the agent has taken action a.

Therefore principal's maximization program in this case is:[14]

[14]Below, [EP] corresponds to the post-contract ex-post contract, standing for 'Ex-Post'. Later, we use [EA] to denote post-contract ex-ante system.

$$
\begin{aligned}
&\text{[EP]} \max_{a\in A, s(x,t)\in S} \int_t \int_x (x - s(x,t))\, f(x|a)h(t)dxdt \quad \text{ s.t.} \\
&\quad (i) \quad \int_t \int_x u(s(x,t),t) f(x|a)h(t)dxdt - v(a) \geq \overline{U}, \\
&\quad (ii) \;\; a \in \arg\max_{a'} \int_t \int_x u(s(x,t),t) f(x|a')h(t)dxdt - v(a'),\; \forall a' \in A, \\
&\quad (iii) \;\; s(x,t) \geq 0, \;\; \forall (x,t) \in X \times T.
\end{aligned} \tag{11}
$$

The first constraint is agent's participation constraint, reflecting that the principal and the agent do not know the true value of t at the time of contracting. The second constraint is agent's incentive constraint, reflecting that the optimizing agent chooses his effort a *before* t is realized.

Assuming the first-order approach is valid as before, the above maximization program reduces to:

$$
\begin{aligned}
&\text{[EP]} \max_{a\in A, s(x,t)\in S} \int_t \int_x (x - s(x,t))\, f(x|a(t))h(t)dxdt \quad \text{ s.t.} \\
&\quad (i) \quad \int_t \int_x u(s(x,t),t) f(x|a)h(t)dxdt - v(a) \geq \overline{U}, \\
&\quad (ii) \quad \int_t \int_x u(s(x,t),t) f_a(x|a)h(t)dxdt - v'(a) = 0, \\
&\quad (iii) \;\; s(x,t) \geq 0, \;\; \forall (x,t) \in X \times T.
\end{aligned} \tag{12}
$$

Let $(a^o, s^o(x,t))$ be the optimal solution for program (12). Then, solving Euler equation of (12) gives that $s^o(x,t)$ must satisfy:

$$
\frac{1}{u_s(s^o(x,t),t)} = \lambda^o + \mu^o \frac{f_a}{f}(x|a^o) \equiv z^o(x), \tag{13}
$$

for almost every (x,t) where (13) has a solution $s^o(x,t) \geq 0$ and otherwise $s^o(x,t) = 0$. In (13), λ^o denotes the optimized Lagrangian multiplier of the agent's participation constraint and μ^o denotes the optimized Lagrangian multiplier of the agent's incentive constraint, where both λ^o and μ^o are independent of t.

One thing to note from (13) is that t does not appear on its right-hand side. Thus, equation (13) requires that the agent's marginal utility to be constant across different realizations of t when $s^o(x,t)$ is offered, as long as the agent's limited liability is not binding. By doing so, the principal can insure the agent from a risk stemming from t and improve the efficiency of her relation with the agent, as t does not affect the agent's effort a. Of course, if the agent's limited liability constraint is binding, then $s^o(x,t)$ will be characterized by the limited liability constraint, i.e., $s^o(x,t) = 0$.

We define the principal's optimized benefits under the post-contract ex-post information system as:

$$PW^o \equiv \int_t \int_x (x - s^o(x,t))\, f(x|a^o)h(t)dxdt. \tag{14}$$

3.3 Post-Contract Ex-Ante Information System

We now consider the case in which the principal chooses the post-contract ex-ante information system. Since the post-contract ex-ante information system publicly reveals information on t after the contracting but before the agent takes a, the principal's maximization program in this case is given by:

$$\begin{aligned}
&[\text{EA}] \max_{a(t)\in A, s(x,t)\in S} \int_t \int_x (x - s(x,t))\, f(x|a(t))h(t)dxdt \quad \text{s.t.} \\
&\quad (i) \;\; \int_t \left(\int_x u(s(x,t),t) f(x|a(t))dx - v(a(t)) \right) h(t)dt \geq \overline{U}, \\
&\quad (ii) \;\; a(t) \in \arg\max_{a'} \int_x u(s(x,t),t) f(x|a')dx - v(a'),\; \forall t \in T,\; \forall a' \in A, \\
&\quad (iii) \;\; s(x,t) \geq 0,\;\; \forall (x,t) \in X \times T.
\end{aligned} \tag{15}$$

The first constraint denotes the agent's participation constraint, reflecting that the principal and the agent do not know the true value of t at the time of contracting. The second constraint corresponds to the agent's incentive constraint, reflecting that the principal and the agent commonly know the true t before the agent takes effort a. The above maximization program differs from the maximization program under post-contract ex-post information system in two ways. First, the agent's incentive constraint must be satisfied for every realization of t, and second, the agent's effort choice is generally a non-trivial function of t. Actually, both differences arise from the fact that the principal and the agent can commonly observe true t before the agent takes his effort under the post-contract ex-ante information system.

Again assuming that using the first-order approach is valid, the principal's maximization program under post-contract ex-ante information system can be written as:

$$\begin{aligned}
&\text{[EA]} \max_{a(t)\in A, s(x,t)\in S} \int_t \int_x (x - s(x,t))\, f(x|a(t))h(t)dxdt \quad \text{s.t.} \\
&(i) \quad \int_t \left(\int_x u(s(x,t),t) f(x|a(t))dx - v(a(t)) \right) h(t)dt \geq \overline{U}, \\
&(ii) \quad \int_x u(s(x,t),t) f_a(x|a(t))dx - v'(a(t)) = 0, \ \forall t \in T, \\
&(iii) \quad s(x,t) \geq 0, \ \ \forall (x,t) \in X \times T.
\end{aligned} \tag{16}$$

Let $(a^*(t), s^*(x,t))$ be the optimal solution for the above program. Then, solving Euler equation of (16) gives that the optimal wage contract, $s^*(x,t)$, must satisfy:

$$\frac{1}{u_s(s^*(x,t),t)} = \lambda^* + \mu^*(t)\frac{f_a}{f}(x|a^*(t)) \equiv z^*(x,t), \tag{17}$$

for almost every (x,t) where (17) has a solution $s^*(x,t) \geq 0$ and otherwise $s^*(x,t) = 0$. λ^* denotes the optimized Lagragian multiplier attached to the agent's participation constraint, (i), and $\mu^*(t)$ is the optimized Lagrangian multiplier of his incentive constraint for a realized t, (ii). Since the agent's incentive constraint must be satisfied for every realized t, $\mu^*(t)$ must be a non-trivial function of t.

We define the principal's optimized benefits under the post-contract ex-ante information system as:

$$PW^* \equiv \int_t \int_x (x - s^*(x,t))\, f(x|a^*(t))h(t)dxdt. \tag{18}$$

3.4 Ranking of Three Information Systems

We first compare the efficiency of the pre-contract information system with the post-contract ex-ante information system.

Lemma 1.

$$PW^p \leq PW^*. \tag{19}$$

Lemma 1 states that the principal weakly prefers the post-contract ex-ante information system to the pre-contract information system. This is reasonable because the principal, when designing the optimal contract, considers the agent's participation constraint for every t under the pre-contract information system, whereas she considers it only in terms of its average across t under the post-contract ex-ante information system, which is less costly for the principal.[15]

Now, to compare the efficiency of the post-contract ex-ante information system, PW^*, with that of the post-contract ex-post information system, PW^o, we further make the following assumption.

Assumption 3. $v'''(a) \geq 0$.

Assumption 3 states that the agent's marginal disutility of effort is increasing in an increasing manner. That is, the agent's disutility function of taking effort, $v(a)$, is sufficiently convex.

Lemma 2. *Given that Assumptions 1, 2, 3 hold, we have*

$$PW^* \leq PW^o. \tag{20}$$

Note that random variable t in our model affects neither the production function nor the agent's cost function. Thus, the first-best effort level, a^{FB}, which satisfies

$$v'(a^{FB}) \equiv \frac{\partial E[X(a^{FB}, \theta)]}{\partial a} = 1, \tag{21}$$

does not vary with t. This indicates that knowing the true value of t ex-ante (i.e., before the agent takes a under the post-contract ex-ante information system) has no information value comparing with knowing t ex-post (i.e., after the agent has taken a under the post-contract ex-post information system) in the first-best scenario where the principal directly observes the agent's effort choice. This is because in the first-best scenario, the principal can always designs a *forcing contract* that enforces the agent to take a^{FB} under any circumstance and provides him the reservation utility, $\overline{U}$.[16]

[15] The same result was also derived in Sobel (1993) in a different context.

[16] For example, if the agent's cost of effort varies with t, i.e., $v = v(a, t)$, then the first-best effort level also varies with t such that

$$v'(a^{FB}(t), t) = 1. \tag{22}$$

In this case, both parties' knowing t ex-ante has some additional information value comparing with their knowing t ex-post even in the first-best world.

However, if t is known ex-ante, the effort level that the principal wishes to induce from the agent generally varies with t in the second-best world where she cannot directly observe the agent's effort choice. It is because t affects the agent's marginal utility, implying that knowing the true value of t ex-ante has some information value in the second-best scenario. Therefore, when the principal knows the agent observes the true value of t before he takes an effort, the principal makes use of this information to better insure the agent and induce an effort level which is better for given t from her standpoint. However, the agent, knowing a true t value ex-ante, uses this information to adjust effort level $a(t)$ for his own interest. Thus, the principal must design a wage contract that satisfies the agent's incentive constraint for every t under the post-contract ex-ante information system, which is more difficult than designing a wage contract under the post-contract ex-post information system which just satisfies the agent's incentive constraint in terms of the average across t. As a result, there is a trade-off between the agent's observing t ex ante and his observing it ex post. Lemma 2 states that the cost of observing t ex ante always dominates its benefit.

To understand a core idea of the proof of Lemma 2 more intuitively, consider the case in which the principal induces the agent to take $a^m \equiv \int_t a^*(t)h(t)dt$ under the post-contract ex-post information system. In that case, a^m under the post-contract ex-post information system produces the same expected output as $a^*(t)$ would produce under the post-contract ex-ante information system. However, inducing a^m under the post-contract ex-post information system is less costly than inducing $a^*(t)$ under the post-contract ex-ante information system. It is because, as shown in Appendix, the convexity of $v(\cdot)$ makes the participation constraint for inducing $a^*(t)$ under the post-contract ex-ante information system harder to satisfy, compared with the same constraint for inducing a^m under the post-contract ex-post information system, and the convexity of $v'(\cdot)$ makes the incentive constraint for inducing $a^*(t)$ under the post-contract ex-ante information system harder to satisfy, compared with the same constraint for inducing a^m under the post-contract ex-post information system.[17]

In sum, inducing a^m under the post-contract ex-post information system is not as costly as inducing $a^*(t)$ under the post-contract ex-ante information system. Since a^m may not be the optimal effort level under the post-contract ex-post information system (i.e., a^m may not be equal to a^o), the post-contract ex-post information system under which a^o is optimally induced cannot be less efficient than the post-contract ex-ante information system under which $a^*(t)$ is optimally induced for any given t. (i.e., $PW^* \leq PW^o$). This intuition will

[17]Roles of each assumption (i.e., convexity of $v(\cdot)$ and $v'(\cdot)$) in proving Lemma 2 will be illustrated more in detail in Section 5 with simple illustrative examples.

be better illustrated in Section 5 where we use specific types of the agent's utility function to illustrate this agency cost comparison between information systems.

Based on the results in Lemmas 1 and 2, we derive the following Proposition 1.

Proposition 1. *Given that Assumptions 1, 2, 3 hold, we have*

$$PW^p \leq PW^* \leq PW^o. \tag{23}$$

Given that the principal "weakly" prefers the post-contract ex-post information system to the post-contract ex-ante information system and the post-contract ex-ante information system to the pre-contract information system, it will be interesting to recognize conditions under which she becomes indifferent among those three information systems. To derive a sufficient condition, we start with the following Lemma 3.

Lemma 3. *If the agent's utility function satisfies $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$, and when the agent's limited liability constraint is not binding for any (x, t), then*

$$\frac{\partial^2}{\partial x \partial t} s^o(x, t) = 0. \tag{24}$$

Lemma 3 states: if the agent's utility function satisfies $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$ and the limited liability constraint is not binding for any (x, t), then the optimal wage contract $s^o(x, t)$ under the post-contract ex-post information system has the same pay-for-performance sensitivity across different realizations of t.

Lemma 4. *If the agent's utility function satisfies $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$, and when the agent's limited liability constraint is not binding for any (x, t), then*

$$\frac{d}{dt} \int_x u(s^o(x, t), t) f_a(x|a^o) dx = 0. \tag{25}$$

When t is not available until the agent chooses his effort level under the post-contract ex-post information system, the principal induces the agent to take effort a^o based on expectation with respect to t. Thus, in designing $s^o(x, t)$ under the post-contract ex-post information system, which induces a^o from the agent, the principal has to consider how to allocate effort incentives across different realizations of t. Note that $\int_t \int_x u(s^o(x,t),t) f_a(x|a^o) h(t) dx dt$ denotes the expected amount of incentive contained in the optimal wage contract, $s^o(x,t)$. Thus, $\int_x u(s^o(x,t),t) f_a(x|a^o) dx$ denotes the amount of incentive contained in $s^o(x,t)$ for given t. Lemma 4 states that, if $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$ and his limited liability constraint is not binding for any (x,t), then the optimal contract under the post-contract ex-post information system, $s^o(x,t)$, must be designed in a way that the same amount of incentive is assigned to every t. Intuitively, as shown in Lemma 3, if $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$, and if the agent's limited liability constraint is not binding for any (x,t), the agent's wage contract under the post-contract ex-post information system must have the same pay-for-performance sensitivity across different realizations of t, which, in turn, implies that the optimal wage contract, $s^o(x,t)$, must assign the same amount of effort incentive to every t.

Proposition 2. *If $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$, and if the agent's limited liability constraint is not binding for any (x,t), then the principal is indifferent between the post-contract ex-ante information system and the post-contract ex-post information system, i.e.,*

$$PW^p \leq PW^* = PW^o. \tag{26}$$

Furthermore, in this case,

$$a^*(t) = a^o, \quad \forall t \in T. \tag{27}$$

Proposition 2 derives a sufficient condition under which the effort level optimally induced from the agent under the post-contract ex-ante information system becomes constant across different realizations of t and thereby, the principal becomes indifferent between the post-contract ex-ante information system and the post-contract ex-post information system. As explained above, the cost of both parties' observing t ex ante compared with their observing t ex post arises from the fact that the principal has to design a wage contract which satisfies the agent's incentive constraint for every t, whereas the benefit comes from the fact that the principal can induce different effort levels from the agent for different realizations of t. However, as shown in Lemma 4, if the agent's utility function satisfies $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$ and the limited liability is not binding for any (x,t), the optimal contract under the post-contract ex-post information system, $s^o(x,t)$, would also satisfy the agent's incentive constraint for $\forall t$ because $s^o(x,t)$ must contain the same amount of effort incentive for every t. It implies that the cost of both parties' observing t ex ante compared with their observing t ex post disappears. On the other hand, as the effort level optimally induced from the agent under the post-contract ex-ante information system is constant across different realizations of t i.e., $a^*(t) = a^o, \ \forall t$, the benefit of both parties' observing t ex ante also disappears.

The condition that the agent's limited liability constraint is not binding for any (x, t) is satisfied if the agent's reservation utility level, $\overline{U}$, is sufficiently high and the gap between $\underline{t}$ and $\bar{t}$ is sufficiently small. On the other hand, the condition $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$ implies that $u_s(s, t)$ and $u_{ss}(s, t)$ have the same contour map on (s, t)-space. This will be satisfied if the agent's utility function has the following additively separable form between s and t:

$$u(s, t) = u(s + k(t)) + l(t). \tag{28}$$

In other words, it is satisfied if the agent's utility function $u(s, t)$ is composed of two additively separable parts: the part in which t affects the agent's marginal utility with respect to s, i.e., $u_s(s, t)$ in which s and t are perfect substitutes, and the part in which t does not affect the agent's marginal utility with respect to s.

As shown in equation (13), under the post-contract ex-post information system where both parties can only observe t ex post, the only reason t is included in the optimal contract, $s^o(x, t)$, is to neutralize the effect of t on the agent's marginal utility with respect to s. By doing so, the principal can minimize the amount of risk imposed on the agent arising from t. However, when s and t are perfect substitutes in the agent's marginal utility with respect to s, the way in which $s^o(x, t)$ can neutralize the effect of t on $u_s(s, t)$ is to make $s^o(x, t) = s^o(x) - k(t)$. Then, the agent's indirect utility $u(s^o(x, t), t)$ becomes $u(s^o(x)) + l(t)$, which makes $\int u(s^o(x, t), t) f_a(x|a^o) dx$ independent of t and gives the agent the same amount of effort incentive for every realization of t. However, even under the post-contract ex-ante information system where both parties can observe t ex ante, the principal can replicate the same result by designing $s^*(x, t) = s^o(x) - k(t)$. Therefore, both parties' observing t ex ante in this case will not affect the agent's incentive problem depending on t.

We have derived a sufficient condition under which the principal becomes indifferent between the post-contract ex-post information system and the post-contract ex-ante information system, while she still weakly prefers the two systems to the pre-contract information system. We now derive a new sufficient condition under which the principal becomes indifferent among those three information systems.

Proposition 3. *If* $\frac{u_s}{u_{ss}} = \frac{u_t}{u_{st}}$ *and the agent's limited liability constraint is not binding for* *any (x,t), then the principal is indifferent among the three information systems, i.e.,*

$$PW^p = PW^* = PW^o. \tag{29}$$

Proposition 3 derives a sufficient conditions under which the principal becomes indifferent among all of the three different information systems. Comparing with Proposition 2, we need a sufficient condition on the agent's utility function: $\frac{u_s}{u_{ss}} = \frac{u_t}{u_{st}}$, which is stronger than $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$ in Proposition 2. This new condition implies that $u(s,t)$ and $u_s(s,t)$ have the same contour maps on (s,t)-space, and it will be satisfied if

$$u(s,t) = u(s + k(t)). \tag{30}$$

In other words, it is satisfied if the agent's monetary wage s and his non-monetary benefits t are perfect substitutes in the agent's utility function.

Note that if s and t are perfect substitutes in the agent's utility, then they are also perfect substitutes in the agent's marginal utility with respect to s. Thus, as shown in Proposition 2, the case in which the principal and the agent observe t before the effort is taken, i.e., post-contract ex-ante, will not be different in terms of the agent's incentive problem compared with the case in which they observe it ex post (i.e., after the agent chooses a under the post-contract ex-post system). Furthermore, both parties' observing t before the contract is signed will not affect the agent's participation constraint, comparing with their observing it after the contract is signed, since $s^p(x,t) = s^*(x,t) = s^o(x,t) = s^o(x) - k(t)$.

4 Value of Coarse Ex-Ante Information

In Lemma 2, we already proved that the *post-contract ex-ante information system* which reveals perfect information on t ex ante (i.e., before the agent takes his effort) is less effi-

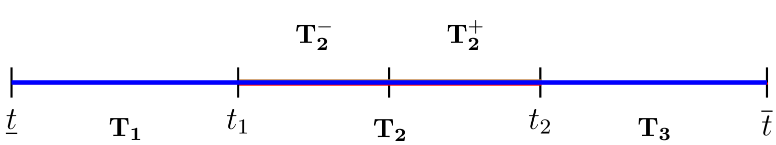

Figure 1: An Example of Information Systems N and N^+

cient than the *post-contract ex-post information system* which reveals t-value ex post. This result can be understood in a way that, given that the perfect information on random factors that affect the agent's marginal utility is always available and contractible ex post, both parties' observing such 'perfect' information ex ante is less efficient than their observing 'no' information ex ante at all. However, it is too early to conclude from this result that the same efficiency ordering can hold true among the post-contract information systems that are providing 'imperfect' ex-ante information.

In this section, we investigate whether the amount of imperfect ex-ante information (before the agent chooses his effort level) on random factors t that affect the agent's marginal utility affects the equilibrium payoff of the principal. To capture the pure effect the amount of ex-ante information on those random variables bears on the efficiency of the agency relation, we assume true values of those variables become publicly observable at the end of the period as in Section 3. Under this assumption, we compare two post-contract information systems, among which one partition of information is more coarse (i.e., it gives less precise imperfect information on those variables t) than the partition of the other information system. The finer partition generates more precise ex-ante information on the variable t.

To be specific, we consider two post-contract information systems: information system N and information system N^+. Information system N is represented by a set of partitions on $T = [\underline{t}, \overline{t}]$, given by $\{T_1, T_2, ..., T_j, ..., T_N\}$: it informs both the principal and the agent ex ante of a specific partition, say $T_{1\leq i\leq N}$, to which the true value of t belongs. We assume without loss of generality that those partitions are in order (i.e., $\forall t_1 \in T_1 \leq \forall t_2 \in T_2.... < \forall t_N \in T_N$).

On the other hand, information system N^+ is represented by $\{T_1, T_2, ..., T_j^-, T_j^+, ..., T_N\}$ where $T_j^- \cup T_j^+ = T_j$. Thus, information system N^+ provides more precise ex-ante information on t than information system N in the sense that it contains two sub-partitions, T_j^-, T_j^+, for T_j compared with information system N. The Figure 1 illustrates one example of N and N^+.

The principal's maximization program under information system N is given by:

$$\max_{\substack{\{a(T_i)\in A\}_{1\le i\le N}\\ s(x,t)\in S}} \sum_{i=1}^{N} p_i \left(\int_{t\in T_i} \int_x \left(x - s(x,t)\right) f\left(x|a(T_i)\right) h(t|T_i) dx dt \right) \quad \text{s.t.}$$

$$(i)\ \sum_{i=1}^{N} p_i \left(\int_{t\in T_i} \int_x u(s(x,t),t) f\left(x|a(T_i)\right) h(t|T_i) dx dt - v(a(T_i)) \right) \ge \overline{U}, \tag{31}$$

$$(ii)\ \int_{t\in T_i} \int_x u(s(x,t),t) f_a\left(x|a(T_i)\right) h(t|T_i) dx dt = v'(a(T_i)),\ \forall T_i \in T,$$

$$(iii)\ s(x,t) \ge 0,\ \forall (x,t) \in X \times T,$$

where $p_i \equiv \int_{t\in T_i} h(t)dt$ and $h(t|T_i)$ is the conditional density function of t given $t \in T_i$. Note that the wage contract when $t \in T_i$, i.e., $s(x, t \in T_i)$, is actually a function of t rather than T_i because the true value of t is publicly available at the end of the period. However, the agent's effort choice, $a(T_i)$, is a function of T_i because only T_i is available when the agent chooses his effort level.

Let $\left(a^N(T_i), s^N(x,t)\right)$ be the optimal solution for the above program (31). Then, the principal's expected benefits under information system N are

$$PW^N \equiv \sum_{i=1}^{N} p_i \left(\int_{t\in T_i} \int_x \left(x - s^N(x,t)\right) f\left(x|a^N(T_i)\right) h(t|T_i) dx dt \right). \tag{32}$$

Note that the principal's maximization program under information system N^+ is the same as (31) except that $\sum_{i=1}^{N}$ now becomes $\sum_{i=1}^{N^+}$ that covers the $N+1$ partitions

$$\{T_1, T_2, ..., T_j^-, T_j^+, ..., T_N\}.$$

Let $(a^{N^+}(T_i), s^{N^+}(x,t))$ be the optimal solution in this case. The principal's expected benefits under information system N^+ are then given by[18]

$$PW^{N^+} \equiv \sum_{i=1}^{N^+} p_i \left(\int_{t\in T_i} \int_x \left(x - s^{N^+}(x,t)\right) f\left(x|a^{N^+}(T_i)\right) h(t|T_i) dx dt \right). \tag{33}$$

Then, we obtain the following Proposition 4.

[18]For simplicity, we abuse the notations so p_j denotes the probability of partition j under the two information system N and N^+.

Proposition 4.

$$PW^{N^+} \leq PW^N. \tag{34}$$

Since it is assumed that a true value of t is available at the end of the period, the two imperfect post-contract ex-ante information systems do not differ in terms of the amount of ex-post information. However, they differ in the amount of ex-ante information. Proposition 4 states that the finer the ex-ante information in terms of partitions an information system provides, the less efficient the principal-agent relation becomes when the true value of t is available ex post. This implies that there is actually a positive relation between the coarseness of the ex-ante information system and its efficiency in the agency framework.

Implication We now return to our pension plan example and infer the policy implication. When t denotes a (random) final payout from a pension plan, s and t are perfect substitutes in the agent's utility function, and if the variation of t is sufficiently small relative to that of s (i.e., the final payout from the pension plan is negligible compared with s in the agent's utility function), then it is likely that the agent's limited liability is never binding given the optimal wage contract under the post-contract ex-post information system, $s^o(x,t)$. Thus, if the final payout from the pension plan t takes a sufficiently smaller portion than s in the agent's utility, then it is more likely that when and how much information on the prospects of the pension plan is revealed will not affect the efficiency of the principal-agent relationship. However, as the final payout from the pension plan takes up a more crucial portion in the agent's utility function, the principal prefers to regulate the plan manager not to reveal any information on those prospects ex ante. Even in the case where s and t are not perfect substitutes, the principal rather prefers to regulate the revelation of relevant information or reduce the amount of available information to the agent before the actual action is taken.

5 Illustrative Examples

We now derive our main results of Section 3.4 based on simple utility functions, to illustrate the underlying economic intuition that drives our results. Specifically, we use two different utility functions[19]: (i) *multiplicative* utility $(u(s,t) = 2t\sqrt{s}, t > 0)$ and (ii) *additive* utility

[19] The agent's square-root utility function, i.e., $u(s) = 2\sqrt{s}$ was used in Kim and Suh (1991) in characterization of the value of information in canonical agency frameworks. Later it was used in Jewitt et al. (2008) to characterize the agency cost in a closed form.

$(u(s,t) = 2\sqrt{s + l(t)}$, where $l(\cdot)$ is C^1 function of t). The first class of utility represents a case in which t affects the agent's marginal utility in a multiplicative way, e.g., the rate of return on workers' pension fund. The second class potentially features the random monetary allowance of the agent or hidden non-monetary satisfaction from workplaces that adds up with monetary payments, s, before entering the utility function $u(\cdot)$.

These examples are useful as they lead to a closed-form expression of the agency cost, and illustrate why ex-ante information revelation is inefficient from the principal's perspective. Here we assume $\overline{U}$ is large enough that limited liability, $s(x,t) \geq 0$, never binds. All detailed derivations are provided in the Appendix.[20]

5.1 Multiplicative Utility Case

We compare the *post-contract ex-ante system* and the *post-contract ex-post system* of Section 3. For the post-contract ex-ante system, let us first fix the set of actions $\{a^*(t)\}$ induced by the principal, and see how the principal designs contracts to minimize the agency cost. The principal solves the following optimization, where [EA]$_M$ stands for the multiplicative utility case under ex-ante systems.

$$\begin{aligned} AC^*_M(\{a^*(t)\}) \equiv \text{[EA]}_M \min_{s(x,t)\in S} & \int_t \int_x s(x,t) f(x|a^*(t)) h(t) dx dt \quad \text{s.t.} \\ & (i) \quad \int_t \left(\int_x 2t\sqrt{s(x,t)} f(x|a^*(t)) dx - v(a^*(t)) \right) h(t) dt \geq \overline{U}, \\ & (ii) \quad \int_x 2t\sqrt{s(x,t)} f_a(x|a^*(t)) dx - v'(a^*(t)) = 0, \ \forall t \in T, \\ & (iii) \quad s(x,t) \geq 0, \ \ \forall (x,t) \in X \times T. \end{aligned} \tag{35}$$

$AC^*_M(\{a^*(t)\})$ is defined as the agency cost associated with inducing the action $a^*(t)$ for each value of t in the post-contract ex-ante system. It turns out that we can express the

[20]We omit derivation of Section 5.2. If we fix the value of t attached to the utility at 1 in the derivation of the optimal contract in Section 5.1 and replace $s^*(x,t)$ and $s^o(x,t)$ with $s^*(x,t) + l(t)$ and $s^o(x,t) + l(t)$, then we can easily obtain expressions provided in Section 5.2.

agency cost in a closed-form as[21]

$$AC^*_M(\{a^*(t)\}) = \underbrace{\frac{\left(\int_t v(a^*(t))h(t)dt + \overline{U}\right)^2}{4\int_t t^2 h(t)dt}}_{\equiv AC^*_{M,\text{IR}}(\{a^*(t)\})} + \underbrace{\frac{1}{4\text{Var}\left(\frac{f_a}{f}\right)}\left(\int_t \frac{1}{t^2} v'(a^*(t))^2 h(t)dt\right)}_{\equiv AC^*_{M,\text{IC}}(\{a^*(t)\})}. \tag{36}$$

The first term $AC^*_{M,\text{IR}}(\{a^*(t)\})$ is interpreted as the cost of ensuring the agent's individual rationality, as the agent can always get $\overline{U}$ amount of utility in other jobs. The second term $AC^*_{M,\text{IC}}(\{a^*(t)\})$ is interpreted as the cost of ensuring that the agent's incentive compatibility constraint is satisfied for $\forall t$. As the expected output is $\int_t \int_x xf(x|a^*(t))h(t)dxdt = \int_t a^*(t)h(t)dt$, the principal chooses a set of actions $\{a^*(t)\}$ to induce from the agent by solving the following optimization program:

$$\begin{aligned} PW^*_M &= \max_{\{a^*(t)\}} \int_t a^*(t)h(t)dt - AC^*_M(\{a^*(t)\}) \\ &= \int_t a^*(t)h(t)dt - \frac{\left(\int_t v(a^*(t))h(t)dt + \overline{U}\right)^2}{4\int_t t^2 h(t)dt} - \frac{1}{4\text{Var}\left(\frac{f_a}{f}\right)}\left(\int_t \frac{1}{t^2} v'(a^*(t))^2 h(t)dt\right). \end{aligned} \tag{37}$$

The first-order conditions for $a^*(t)$ for each t are given as

$$1 = v'(a^*(t))\left(\underbrace{\frac{\int_t v(a^*(t))h(t)dt + \overline{U}}{2\int_t t^2 h(t)dt}}_{\text{Constant for } \forall t} + \frac{v''(a^*(t))}{2\text{Var}\left(\frac{f_a}{f}\right)t^2}\right), \quad \forall t. \tag{38}$$

Due to our assumptions that $v'(\cdot)$ and $v''(\cdot)$ are increasing, we get the unique solution $\{a^*(t)\}$ that is strictly increasing in t. As higher t is revealed ex ante before the agent takes an effort, the principal respects the fact that higher t corresponds to higher marginal utility of the agent, and thus the agent is induced to put more effort as t and s are complementary. In sum, we get a strictly dispersed, optimally induced action set $\{a^*(t)\}$.

[21] Due to our assumed structure about the technology $x = a + \theta$, Var $\left(\frac{f_a}{f}(x|a)\right)$'s value does not depend on the level of effort a.

For the post-contract ex-post system, we also fix the action a^o induced by the principal, and see how the principal minimizes the agency cost. In this case, the principal solves the following optimization:

$$
\begin{aligned}
AC_M^o(a^o) \equiv \text{[EP]}_M \min_{s(x,t)\in S} & \int_t \int_x s(x,t) f(x|a^o) h(t) dx dt \quad \text{s.t.} \\
& (i) \quad \int_t \int_x 2t\sqrt{s(x,t)} f(x|a^o) h(t) dx dt - v(a^o) \geq \overline{U}, \\
& (ii) \quad \int_t \int_x 2t\sqrt{s(x,t)} f_a(x|a^o) h(t) dx dt - v'(a^o) = 0, \\
& (iii) \;\; s(x,t) \geq 0, \;\; \forall (x,t) \in X \times T.
\end{aligned} \tag{39}
$$

$AC_M^o(a^o)$ is an agency cost associated with inducing the given action a^o for $\forall t$ in the post-contract ex-post system. We can express the agency cost in a closed form as

$$
AC_M^o(a^o) = \underbrace{\frac{\left(v(a^o) + \overline{U}\right)^2}{4\displaystyle\int_t t^2 h(t) dt}}_{\equiv AC^o_{M,\text{IR}}(a^o)} + \underbrace{\frac{v'(a^o)^2}{4\text{Var}\left(\dfrac{f_a}{f}\right)} \frac{1}{\displaystyle\int_t t^2 h(t) dt}}_{\equiv AC^o_{M,\text{IC}}(a^o)}, \tag{40}
$$

where the first term $AC^o_{M,\text{IR}}(a^o)$ is a cost of ensuring the agent's individual rationality, as the agent can always get $\overline{U}$ amount of utility in other jobs. The second term $AC^o_{M,\text{IC}}(a^o)$ is interpreted as the cost of satisfying the agent's incentive compatibility. As expected output is $\int_t \int_x x f(x|a^o) h(t) dx dt = a^o$, the principal chooses an action a^o to induce by solving the following optimization:

$$
PW_M^o = \max_{a^o} a^o - AC_M^o(a^o) = a^o - \frac{\left(v(a^o) + \overline{U}\right)^2}{4\displaystyle\int_t t^2 h(t) dt} - \frac{v'(a^o)^2}{4\text{Var}\left(\dfrac{f_a}{f}\right)} \frac{1}{\displaystyle\int_t t^2 h(t) dt}. \tag{41}
$$

The following Lemma 5 allows us to directly compare the relative sizes of each component of the agency costs of the two systems. Before we proceed, we define the average action $a^m \equiv \int_t a^*(t) h(t) dt$ in the post-contract ex-ante system.

Lemma 5. $AC^o_{M,IR}(a^m) < AC^*_{M,IR}(\{a^*(t)\})$ *and* $AC^o_{M,IC}(a^m) < AC^*_{M,IC}(\{a^*(t)\})$.

Lemma 5 implies inducing a^m, the mean of $\{a^*(t)\}$, in the post-contract ex-post system, costs less to the principal than inducing $\{a^*(t)\}$ in the post-contract ex-ante system. Since inducing a^m is inferior to inducing the optimal a^o, we get $PW^o_M > PW^*_M$[22]. In the ex-ante system, the principal understands that the agent observes a revealed level of t before taking effort, and thus induces different effort levels $\{a^*(t)\}$ for different realizations of t, since t affects the agent's marginal utility.

The resulting dispersion in $\{a^*(t)\}$ does not benefit the principal in terms of the output level compared to the case where the agent regardless of realized t takes the same action a^m. Furthermore, our assumption that $v(\cdot)$ is convex guarantees that $AC^o_{M,IR}(a^m)$ is lower than $AC^*_{M,IR}(\{a^*(t)\})$, meaning that if effort levels for each t are dispersed, to make sure that the agent stays on the contract, the principal on average must pay more to the agent, since the average cost of effort for the agent increases. Assumption 3 that $v'(\cdot)$ is convex guarantees that inducing a dispersed set of effort $\{a(t)\}$ under the ex-ante system is less efficient to the principal than inducing its average under the ex-post system, since the incentive cost inevitably rises on average. This agency cost decomposition (i.e., $AC^o_M(\cdot)$ into $AC^o_{M,IR}(\cdot)$ that governs the agent's participation constraint and $AC^o_{M,IC}(\cdot)$ that governs his incentive compatibility) clarifies underlying mechanisms for our main result, Lemma 2.

In the next Section 5.2, we study the case of additive utility where t enters into the utility in an additive way. We show that in this case, the principal can wisely design a contract to perfectly insure the agent from the t risk, and the timing of the information revelation does not affect the efficiency of the agency relation.

[22] Our utility specification does not satisfy the premises of Proposition 2. Here we have :

$$\frac{u_{ss}}{u_{sss}} = -\frac{2}{3}s \neq \frac{u_{ts}}{u_{tss}} = -2s. \tag{42}$$

5.2 Additive Utility Case

Similarly, we first fix a set of effort levels $\{a^*(t)\}$ induced by the principal, and minimize the agency cost. Under the post-contract ex-ante system, the principal solves:

$$
\begin{aligned}
AC^*_{Add}(\{a^*(t)\}) \equiv \text{[EA]}_{Add} \min_{s(x,t)\in S} & \int_t \int_x s(x,t) f(x|a^*(t))h(t)dxdt \quad \text{s.t.} \\
& (i) \quad \int_t \left(\int_x 2\sqrt{s(x,t)+l(t)} f(x|a^*(t))dx - v(a^*(t)) \right) h(t)dt \geq \overline{U}, \\
& (ii) \quad \int_x 2\sqrt{s(x,t)+l(t)} f_a(x|a^*(t))dx - v'(a^*(t)) = 0, \ \forall t \in T, \\
& (iii) \ \ s(x,t) \geq 0, \ \ \forall (x,t) \in X \times T,
\end{aligned}
\tag{43}
$$

where $AC^*_{Add}(\{a(t)\})$ is defined as the agency cost associated with inducing effort $a^*(t)$ for each t in the post-contract ex-ante system. We can similarly express the agency cost in a closed form as:[23]

$$
AC^*_{Add}(\{a(t)\}) = \underbrace{\frac{\left(\int_t v(a^*(t))h(t)dt + \overline{U}\right)^2}{4}}_{\equiv AC^*_{Add,\text{IR}}(\{a^*(t)\})} + \underbrace{\frac{1}{4\text{Var}\left(\frac{f_a}{f}\right)} \int_t v'(a^*(t))^2 h(t)dt}_{\equiv AC^*_{Add,\text{IC}}(\{a^*(t)\})} - \underbrace{\int_t l(t)h(t)dt}_{=\mathbb{E}(l(t))} . \tag{44}
$$

The first term, $AC^*_{Add,\text{IR}}(\{a^*(t)\})$, is the cost of satisfying the agent's individual rationality constraint, as the agent can always secure $\overline{U}$. The second term, $AC^*_{Add,\text{IC}}(\{a^*(t)\})$, is the cost of satisfying the agent's incentive compatibility constraint for $\forall t$. As expected output is $\int_t \int_x x f(x|a^*(t))dxh(t)dt = \int_t a^*(t)h(t)dt$, the principal chooses a set of actions $\{a^*(t)\}$ to induce by solving:

$$
\begin{aligned}
PW^*_{Add} &= \max_{\{a^*(t)\}} \int_t a^*(t)h(t)dt - AC^*_{Add}(\{a^*(t)\}) \\
&= \int_t a^*(t)h(t)dt - \frac{\left(\int_t v(a^*(t))h(t)dt + \overline{U}\right)^2}{4} \\
&\quad - \frac{1}{4\text{Var}\left(\frac{f_a}{f}\right)} \int_t v'(a^*(t))^2 h(t)dt + \int_t l(t)h(t)dt
\end{aligned}
\tag{45}
$$

[23] Derivations are very similar to those of Section 5.1.

The first-order condition for $a^*(t)$ yields

$$1 = v'(a^*(t)) \left(\underbrace{\frac{\int_t v(a^*(t))h(t)dt + \overline{U}}{2}}_{\text{Constant for } \forall t} + \frac{v''(a^*(t))}{2\mathrm{Var}\left(\frac{f_a}{f}\right)} \right), \quad \forall t. \tag{46}$$

As $v'(\cdot), v''(\cdot)$ are increasing, we get $a^*(t) = a^*$ for $\forall t \in T$.

Here, t-variable appears in the agent's utility in an additive way. Therefore, the principal can adjust contractual forms $s(\cdot,\cdot)$ in a way that it perfectly insures the agent from the risk around t and thus induce the same effort across different realizations of t, as a dispersion in action $\{a^*(t)\}$ costs more to the principal.

If we plug $a^*(t) = a^*$ for all t into (45), we get the following PW^*_{Add}:[24]

$$PW^*_{Add} = \max_{a^*} a^* - \frac{\left(\overline{U} + v(a^*)\right)^2}{4} - \frac{v'(a^*)^2}{4\mathrm{Var}\left(\frac{f_a}{f}\right)} + \int_t l(t)h(t)dt. \tag{48}$$

For the post-contract ex-post information system, we also first fix the action a^o induced by the principal, and see how the principal minimizes the agency cost. The principal solves

$$\begin{aligned}
AC^o_{Add}(a^o) \equiv [\mathrm{EP}]_{Add} \min_{s(x,t)\in S} & \int_t \int_x s(x,t) f(x|a^o)h(t)dxdt \quad \text{s.t.} \\
& (i) \quad \int_t \int_x 2\sqrt{s(x,t) + l(t)} f(x|a^o)h(t)dxdt - v(a^o) \geq \overline{U}, \\
& (ii) \quad \int_t \int_x 2\sqrt{s(x,t) + l(t)} f_a(x|a^o)h(t)dxdt - v'(a^o) = 0, \\
& (iii) \quad s(x,t) \geq 0, \ \forall (x,t) \in X \times T,
\end{aligned} \tag{49}$$

where $AC^o_{Add}(a^o)$ is the agency cost associated with inducing the effort a^o for all t in the post-contract ex-post system. In this case, the agency cost and the principal's welfare can be given in closed forms as follows:

[24] The optimal action a^* satisfies

$$1 = \left(\frac{\overline{U} + v(a^*)}{2} + \frac{v''(a^*)}{2\mathrm{Var}\left(\frac{f_a}{f}\right)} \right) v'(a^*) \tag{47}$$

$$
\begin{aligned}
AC^o_{Add}(a^o) &= \underbrace{\frac{\left(v(a^o)+\overline{U}\right)^2}{4}}_{\equiv AC^o_{Add,\mathrm{IR}}(a^o)} + \underbrace{\frac{v'(a^o)^2}{4\mathrm{Var}\left(\frac{f_a}{f}\right)}}_{\equiv AC^o_{Add,\mathrm{IC}}(a^o)} - \int_t l(t)h(t)dt, \\
PW^o_{Add} &= \min_{a^o} a^o - \frac{\left(v(a^o)+\overline{U}\right)^2}{4} - \frac{v'(a^o)^2}{4\mathrm{Var}\left(\frac{f_a}{f}\right)} + \int_t l(t)h(t)dt.
\end{aligned} \tag{50}
$$

From equations (48) and (50), we see $PW^o_{Add} = PW^*_{Add}$ holds.[25] For optimal $s^o(x,t)$, $s^o(x,t) + l(t)$ does not depend on t, so the agent is perfectly hedged against t-risk, and thus the individual rationality constraint is satisfied for $\forall t \in T$, which means $PW^*_{Add} = PW^o_{Add} = PW^p_{Add}$.

6 Conclusion

The main objective of this paper is to analyze how the efficiency of an information system is affected not only by the amount of information it contains but also by the time it reveals information.

We compare three different information systems and show that the principal prefers the post-contract ex-post information system, in which perfect information on random variables affecting the agent's marginal utility is publicly revealed after the agent decides his action choice, to the post-contract ex-ante information system, in which the same information is publicly revealed before the agent decides his action choice, and the post-contract ex-ante information system to the pre-contract information system, in which the same information is publicly revealed even before the principal and the agent agree upon the contract. In addition, we show that there is a negative relationship between the amount of ex-ante information an information system contains and its efficiency in the principal-agent relation, implying that the principal and the agent's observing more precise ex-ante information on random variables that affect the agent's marginal utility reduces the efficiency compared

[25] Actually our additive utility satisfies the assumptions of Proposition 3:

$$
\frac{u_s}{u_{ss}} = -2(s + l(t)) = \frac{u_t}{u_{ts}}. \tag{51}
$$

with their observing less precise ex-ante information.

Before closing, two remarks need to be in order. First, information systems considered in this paper reveal information not only to the principal but also to the agent. In contrast, if the principal acquires such information privately[26], then the principal's information revealing strategy must be involved and thereby our ranking of the three information systems may be affected. Second, the random factors in this paper are assumed to affect only the agent's marginal utility. Thus, the principal and the agent's observing information on those variables ex ante has no additional information value, compared with their observing it ex post in the first-best situation in which the principal can directly observe the agent's action choice. However, in a situation where the random factors affect either the agent's production function (e.g., Silvers (2012)) or his disutility (i.e., cost) function (e.g., Sobel (1993)), the principal and the agent's observing information on those variables ex ante will have some additional information value compared with their observing it ex-post even when the principal can directly observe the agent's action choice (i.e., the first-best case). Therefore in those cases, in comparing the efficiency of the post-contract ex-ante information system with that of the post-contract ex-post information system, we also need to consider such additional information value of ex-ante information. Also, the ex-ante revelation of relevant information about the technology could yield a better matching between the principal and the agent, and improves efficiency. Consequently, in those cases, we cannot generally obtain the same conclusion that the post-contract ex-ante information system is always less efficient than the post-contract ex-post information system under the agency relationship.

[26] As we mentioned before, Silvers (2012) considers the environment where the principal gets a private or public signal about technology. In those cases, the principal might reveal the signal to the agent not truthfully, and the optimal contract is distorted to account for this information revealing strategy.

References

Banker, Rajiv D., Masako Darrough, Shaopeng Li, and Lucas Threinen, "The Value of Precontract Information About an Agent's Ability in the Presence of Moral Hazard and Adverse Selection," *Journal of Accounting Research*, 2019, *57* (5), 1201–1245.

Boleslavsky, Ralph and Teddy Kim, "Bayesian Persuasion and Moral Hazard," *Working Paper*, 2020.

Conlon, John R., "Two New Conditions Supporting the First-Order Approach to Multisignal Principal-Agent Problems," *Econometrica*, 2009, *77* (1), 249–278.

Gjesdal, Frøystein, "Information and Incentives: The Agency Information Problem," *The Review of Economic Studies*, 1982, *49* (3), 373–390.

Grossman, Sanford J. and Oliver Hart, "An Analysis of the Principal-Agent Problem," *Econometrica*, 1983, *51*, 7–45.

Harris, Milton and Artur Raviv, "Optimal Incentive Contracts with Imperfect Information," *Journal of Economic Theory*, 1979, *20* (2), 231–259.

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

_ , "Moral Hazard in Teams," *The Bell Journal of Economics*, 1982, *13*, 324–340.

Jewitt, Ian, "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica*, 1988, *56* (5), 1177–1190.

_ , "Information and Principal Agent Problems," *Working Paper*, 1997.

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

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

Kamenica, Emir and Matthew Gentzkow, "Bayesian Persuasion," *American Economic Review*, 2011, *101*, 2590–2615.

Kim, Son Ku, "Efficiency of an Information System in an Agency Model," *Econometrica*, 1995, *63* (1), 89–102.

— **and Yoon Suh**, "Ranking of Accounting Information Systems for Management Control," *Journal of Accounting Research*, 1991, *29* (2), 386–396.

Luenberger, D. G., *Optimization by Vector Space Methods*, John Wiley & Sons, Inc. 27, 28, 30, 1969.

Mirrlees, James A., "Notes on Welfare Economics, Information, and Uncertainty," *Essays on Economic Behavior Under Uncertainty: E. Balch, D. McFadden, and H. Wu, eds., North-Holland, Amsterdam.*, 1974.

Rogerson, William, "The First Order Approach to Principal-Agent Problem," *Econometrica*, 1985, *53*, 1357–1368.

Ross, Steve, "The Economic Theory of Agency: The Principal's Problem," *The American Economic Review*, 1973, *63*, 134–139.

Rothsscild, Michael and Joseph E. Stiglitz, "Increasing Risk I: A Definition," *Journal of Economic Theory*, 1970, *2* (3), 225–243.

Shavell, Steven, "Risk Sharing and Incentives in the Principal and Agent Relationship," *The Bell Journal of Economics*, 1979, *10* (1), 55–73.

Silvers, Randy, "The value of information in a principal–agent model with moral hazard: The ex post contracting case," *Games and Economic Behavior*, 2012, *74* (1), 352–365.

Sinclair-Desgagné, Bernanrd, "The First-Order Approach to Multi-Signal Principal-Agent Problems," *Econometrica*, 1994, *62* (2), 459–465.

Sobel, Joel, "Information Control in the Principal-Agent Problem," *International Economic Review*, 1993, *34*, 259–269.

Appendix A. Proofs

Proof of Lemma 1: The optimal solution for the pre-contract information system, $(s^p(x,t), a^p(t))$, satisfies both the participation constraint and the incentive constraint in [EA], indicating that $(s^p(x,t), a^p(t))$ is at least feasible under the post-contract ex-ante information system. Thus, we directly have

$$PW^p \leq PW^*. \tag{I.1}$$

Proof of Lemma 2: First, let us consider the case where the principal has chosen the *post-contract ex-post* information system but induces the agent to take a^m instead of a^o, where

$$a^m \equiv \int_t a^*(t)h(t)dt. \tag{I.2}$$

Thus, a^m is the average (across t) effort level of $a^*(t)$. Based on the first-order approach, the principal's optimization program in this case is :

$$\begin{aligned}
&\max_{s(x,t)\in S} \int_t \int_x \left(x - s(x,t)\right) f(x|a^m)h(t)dxdt \quad \text{s.t.}\\
&(i) \quad \int_t \int_x u(s(x,t),t) f(x|a^m)h(t)dxdt - v(a^m) \geq \overline{U},\\
&(ii) \quad \int_t \int_x u(s(x,t),t) f_a(x|a^m)h(t)dxdt - v'(a^m) = 0,\\
&(iii) \quad s(x,t) \geq 0, \;\; \forall (x,t) \in X \times T.
\end{aligned} \tag{I.3}$$

Let $s^m(x,t)$ be the optimal contract for the optimization program (I.3). Similar to (9), we obtain that $s^m(x,t)$ must satisfy:

$$\frac{1}{u_s(s^m(x,t),t)} = \lambda^m + \mu^m \frac{f_a}{f}(x|a^m) \equiv z^m(x), \tag{I.4}$$

when $s^m(x,t) \geq 0$ and $s^m(x,t) = 0$ elsewhere. In (I.4), λ^m is the optimized Lagrangian multiplier of the agent's participation constraint under the post-contract ex-post information system, whereas μ^m denotes the optimized Lagrangian multiplier of the agent's incentive constraint.

Based on (I.4), we transform the optimal contract, $s^m(x,t)$, which is defined on (x,t)-space to a wage contract that is defined on (z,t)-space, that is,

$$s^m(x,t) \equiv r(z^m(x),t). \tag{I.5}$$

Actually, $r(\cdot)$ function can be interpreted as:

$$r(z^m(x),t) \equiv u_s^{-1}(u_s(s^m(x,t),t),t) = u_s^{-1}\left(\frac{1}{z^m(x)},t\right). \tag{I.6}$$

From (I.6), we see that the functional form of $r(\cdot)$ is determined by $u_s(\cdot)$: it does not depend on the information systems. $r(\cdot)$ is increasing in z because $u_{ss} < 0$. For notational simplicity, we will drop x from $z^m(x)$ and denote $z^m \equiv z^m(x)$. Then, based on equation (I.4), $r(z^m,t)$ must satisfy:

$$u_s(r(z^m,t),t)z^m = \begin{cases} 1, & \text{if } r(z^m,t) > 0, \\ u_s(0,t)z^m, & \text{if } r(z^m,t) = 0. \end{cases} \tag{I.7}$$

Accordingly, the principal's optimized benefits in this case can be written as

$$PW^m = \int_t \int_x (x - r(z^m,t))\, f(x|a^m)h(t)dxdt. \tag{I.8}$$

Now, consider the case where the principal chooses the post-contract ex-ante information system. Then, based on (17), we can rewrite

$$s^*(x,t) \equiv r(z^*(x,t),t). \tag{I.9}$$

Actually,

$$r(z^*(x,t),t) \equiv u_s^{-1}(u_s(s^*(x,t),t),t) = u_s^{-1}\left(\frac{1}{z^*(x,t)},t\right). \tag{I.10}$$

Again, for notational simplicity, we can drop (x,t) from $z^*(x,t)$ and denote $z^* \equiv z^*(x,t)$. Based on (17), $r(z^*,t)$ satisfies

$$u_s(r(z^*,t),t)z^* = \begin{cases} 1, & \text{if } r(z^*,t) > 0, \\ u_s(0,t)z^*, & \text{if } r(z^*,t) = 0. \end{cases} \tag{I.11}$$

Using $r(z^*,t)$, the principal's optimized benefits under the post-contract ex-ante information system can be rewritten as:

$$PW^* = \int_t \int_x (x - r(z^*,t))\, f(x|a^*(t))h(t)dxdt. \tag{I.12}$$

Since

$$\int_t \int_x x f(x|a^*(t))h(t)dxdt = \int_t a^*(t)h(t)dt = a^m = \int_t \int_x x f(x|a^m)h(t)dxdt, \tag{I.13}$$

by subtracting equation (I.12) from equation (I.8), we have

$$PW^m - PW^* = \int_t \int_x r(z^*, t) f(x|a^*(t)) h(t) dx dt - \int_t \int_x r(z^m, t) f(x|a^m) h(t) dx dt. \tag{I.14}$$

We next define the following function, which was extensively used in Kim (1995) and Jewitt et al. (2008):

$$\psi(z,t) \equiv r(z,t) - u(r(z,t),t)z. \tag{I.15}$$

Since

$$\psi_z(z,t) = r_z(z,t) - u_s(r(z,t),t) r_z(z,t) z - u(r(z,t),t), \tag{I.16}$$

by using $u_s(r(z,t),t)z = 1$ where $r(z,t) > 0$, we have

$$\psi_z(z,t) = \begin{cases} -u(r(z,t),t), & \text{if } r(z,t) > 0, \\ -u(0,t), & \text{if } r(z,t) = 0, \end{cases} \tag{I.17}$$

and

$$\psi_{zz}(z,t) = \begin{cases} -u_s(r(z,t),t) r_z(z,t), \leq 0 & \text{if } r(z,t) > 0, \\ 0, & \text{if } r(z,t) = 0, \end{cases} \tag{I.18}$$

Thus, it is easy to see that $\psi(z,t)$ is globally concave in z (and is a C^1 function).

From equation (I.15):

$$r(z,t) = \psi(z,t) + u(r(z,t),t)z, \tag{I.19}$$

and we can rewrite (I.14) as

$$\begin{aligned} PW^m - PW^* = & \int_t \int_x \psi(z^*, t) f(x|a^*(t)) h(t) dx dt - \int_t \int_x \psi(z^m, t) f(x|a^m) h(t) dx dt \\ & + \int_t \int_x u(r(z^*,t),t) z^* f(x|a^*(t)) h(t) dx dt - \int_t \int_x u(r(z^m,t),t) z^m f(x|a^m) h(t) dx dt. \end{aligned} \tag{I.20}$$

Due to the complementary slackness of the participation constraints in both programs (post-contract ex-post and post-contract ex-ante), we must have:

$$\begin{gathered} \lambda^* \left(\int_t \int_x u(r(z^*,t),t) f(x|a^*(t)) h(t) dx dt \right) = \lambda^* \left(\overline{U} + \int_t v(a^*(t)) h(t) dt \right) \\ \int_t \int_x u(r(z^*,t),t) f(x|a^*(t)) h(t) dx dt \geq \overline{U} + \int_t v(a^*(t)) h(t) dt, \end{gathered} \tag{I.21}$$

and

$$\lambda^m \left(\int_t \int_x u(r(z^m,t),t) f(x|a^m) h(t) dx dt\right) = \lambda^m \left(\overline{U} + v(a^m)\right)$$
$$\int_t \int_x u(r(z^m,t),t) f(x|a^m) h(t) dx dt \geq \overline{U} + v(a^m). \tag{I.22}$$

Also, the incentive constraints for the two programs must be satisfied, respectively:

$$\int_x u(r(z^*,t),t) f_a(x|a^*(t)) dx = v'(a^*(t)),\ \forall t,$$
$$\int_t \int_x u(r(z^m,t),t) f_a(x|a^m) h(t) dx dt = v'(a^m). \tag{I.23}$$

From equations (17), (I.21), and (I.23), we obtain

$$\begin{aligned}
&\int_t \int_x u(r(z^*,t),t) z^* f(x|a^*(t)) h(t) dx dt \\
&\quad = \int_t \int_x u(r(z^*,t),t) \left(\lambda^* + \mu^*(t) \frac{f_a}{f}(x|a^*(t))\right) f(x|a^*(t)) h(t) dx dt \\
&\quad = \lambda^* \int_t \int_x u(r(z^*,t),t) f(x|a^*(t)) h(t) dx dt + \int_t \mu^*(t) \left(\int_x u(r(z^*,t),t) f_a(x|a^*(t)) dx\right) h(t) dt \\
&\quad = \lambda^* \left(\overline{U} + \int_t v(a^*(t)) h(t) dt\right) + \int_t \mu^*(t) v'(a^*(t)) h(t) dt,
\end{aligned} \tag{I.24}$$

which is a function of $\lambda^*, \{\mu^*(t)\}, \{a^*(t)\}$. Similarly, by equations (I.4), (I.22), and (I.23), we similarly obtain

$$\begin{aligned}
&\int_t \int_x u(r(z^m,t),t) z^m f(x|a^m) h(t) dx dt \\
&\quad = \int_t \int_x u(r(z^m,t),t) \left(\lambda^m + \mu^m \frac{f_a}{f}(x|a^m)\right) f(x|a^m) h(t) dx dt \\
&\quad = \underbrace{\lambda^m \int_t \int_x u(r(z^m,t),t) f(x|a^m) h(t) dx dt}_{=\lambda^m \left(\overline{U} + v(a^m)\right)} \\
&\qquad + \mu^m \underbrace{\int_t \int_x u(r(z^m,t),t) f_a(x|a^m) h(t) dx dt}_{=v'(a^m)} \\
&\quad = \lambda^m \left(\overline{U} + v(a^m)\right) + \mu^m v'(a^m).
\end{aligned} \tag{I.25}$$

Thus, from the (I.24) and (I.25), we can rewrite (I.20) as[1]

$$
\begin{aligned}
PW^m - PW^* = & \int_t \int_x \psi(z^*, t) f(x|a^*(t)) h(t) dx dt - \int_t \int_x \psi(z^m, t) f(x|a^m) h(t) dx dt \\
& + \lambda^* \left(\overline{U} + \int_t v(a^*(t)) h(t) dt \right) - \lambda^m \left(\overline{U} + v(a^m) \right) + \left(\int_t \mu^*(t) v'(a^*(t)) h(t) dt \right) - \mu^m v'(a^m)
\end{aligned}
\tag{I.26}
$$

Now, following Kim (1995), we define

$$
z^h \equiv \lambda^m + \mu^m \frac{f_a}{f}(x|a^*(t)). \tag{I.27}
$$

Since $\psi(z, t)$ is globally concave in z,[2] we have

$$
\begin{aligned}
& \int_t \int_x (\psi(z^h, t) - \psi(z^*, t)) f(x|a^*(t)) h(t) dx dt \\
& \quad \leq \int_t \int_x \psi_z(z^*, t)(z^h - z^*) f(x|a^*(t)) h(t) dx dt \\
& \quad = - \int_t \int_x u(r(z^*, t), t) \left((\lambda^m - \lambda^*) f(x|a^*(t)) + (\mu^m - \mu^*(t)) f_a(x|a^*(t)) \right) h(t) dx dt \\
& \quad = \lambda^* \left(\int_t \int_x u(r(z^*, t), t) f(x|a^*(t)) h(t) dx dt \right) - \lambda^m \left(\int_t \int_x u(r(z^*, t), t) f(x|a^*(t)) h(t) dx dt \right) \\
& \qquad + \int_t (\mu^*(t) - \mu^m) \left(\int_x u(r(z^*, t), t) f_a(x|a^*(t)) dx \right) h(t) dt \\
& \quad \leq (\lambda^* - \lambda^m) \left(\overline{U} + \int_t v(a^*(t)) h(t) dt \right) + \int_t (\mu^*(t) - \mu^m) v'(a^*(t)) h(t) dt \\
& \quad \leq \lambda^* \left(\overline{U} + \int_t v(a^*(t)) h(t) dt \right) - \lambda^m \left(\overline{U} + v(a^m) \right) + \int_t (\mu^*(t) - \mu^m) v'(a^*(t)) h(t) dt \\
& \quad \leq \lambda^* \left(\overline{U} + \int_t v(a^*(t)) h(t) dt \right) - \lambda^m \left(\overline{U} + v(a^m) \right) + \int_t \mu^*(t) v'(a^*(t)) h(t) dt - \mu^m v'(a^m).
\end{aligned}
\tag{I.28}
$$

The first inequality holds from the concavity of $\psi(z, t)$ function in z. For the first equality, we use (17), (I.17), and (I.27), and for the second inequality, we use (I.21), (I.22), and (I.23). The third inequality is from $a^m = \int_t a^*(t) h(t) dt$ and that $v(\cdot)$ is convex, whereas the last inequality is from Assumption 3. By substituting equation (I.28) into equation (I.26),

[1] In the absence of the agent's limited liability constraint $s(x, t) \geq 0$, equation (I.26) can be derived using the Lagrange Duality theorem, as derived in Jewitt (1997) and Jewitt et al. (2008). Our derivation holds even in the presence of the limited liability constraint, i.e., $s(x, t) \geq 0$.

[2] See Kim (1995) and Jewitt et al. (2008) for this issue in the canonical principal-agent framework.

we can obtain

$$\begin{aligned} PW^m - PW^* &\geq \int_t\int_x \psi(z^h,t)f(x|a^*(t))h(t)dxdt - \int_t\int_x \psi(z^m,t)f(x|a^m)h(t)dxdt \\ &= \int_t\int_x \psi\left(\lambda^m + \mu^m\frac{f_a}{f}(x|a^*(t)),t\right)f(x|a^*(t))h(t)dxdt \\ &\quad - \int_t\int_x \psi\left(\lambda^m + \mu^m\frac{f_a}{f}(x|a^m),t\right)f(x|a^m)h(t)dxdt. \end{aligned} \tag{I.29}$$

As $x = a + \theta$, where θ is a random variable with probability density function $g(\cdot)$, we can write $f(x|a) = g(x-a)$. Then, we obtain $f_a(x|a) = -g'(x-a)$ and $\frac{f_a}{f}(x|a) = -\frac{g'(\theta)}{g(\theta)}$. Therefore, when $x \sim f(x|a)$, $\frac{f_a}{f}(x|a)$'s distribution does not depend on a. Accordingly, for any given t,[3]

$$\int_x \psi\left(\lambda^m + \mu^m\frac{f_a}{f}(x|a^*(t)),t\right)f(x|a^*(t))dx = \int_x \psi\left(\lambda^m + \mu^m\frac{f_a}{f}(x|a^m),t\right)f(x|a^m)dx. \tag{I.30}$$

Thus, from (I.29) and (I.30), we derive

$$PW^* \leq PW^m. \tag{I.31}$$

Since inducing a^m may not be optimal under the ex-post information system (i.e., a^m is inferior to a^o, as a^o is by definition the optimally induced effort), we obtain

$$PW^m \leq PW^o, \tag{I.32}$$

which leads to

$$PW^* \leq PW^o. \tag{I.33}$$

[3] Jewitt et al. (2008) show that if $F(\cdot|\cdot)$ satisfies *Concave Local Informativeness* (CLI), i.e.,

$$L'(a) \equiv F_{ee}(x|a) - \frac{f_a}{f}(x|a)F(x|a)$$

is increasing in a, i.e., $L(a)$ is convex in a, then $\psi(\lambda^m + \mu^m\frac{f_a}{f}(x|a),t)$ becomes convex in a, which proves (I.31) from (I.29). If $x = a+\theta$, $L'(a) = L''(a) = 0$ holds for any a, satisfying (I.30). Note that our problem of comparing two systems, the post-contract ex-ante system and the ex-post system is not a direct corollary of their result that the agency cost $C_S(a)$ in inducing effort a, in the canonical agency model à la Holmström (1979), is convex when $F(\cdot|\cdot)$ satisfies the *Concave Local Informativeness* (CLI), as (i) the two systems only differ in their incentive constraints while sharing the participation constraint (both are post-contract); (ii) we consider the agent's limited liability; and (iii) t affects the agent's marginal utility function and is contractible. Still, our interim result that $PW^m \geq PW^*$ is closely related to their result.

Proof of Lemma 3: As the agent's limited liability is not binding for any (x,t), we have $s^o(x,t) > 0$, implying that the optimal wage contract, $s^o(x,t)$, is characterized solely by equation (13) and we have

$$u_s(s^o(x,t),t) = \frac{1}{z^o(x)}, \quad \forall (x,t) \in X \times T. \tag{I.34}$$

Since $s^o(x,t)$ in equation (13) is differentiable, by taking a derivative with respect to t on both sides of the above equation, we obtain

$$u_{ss}s^o_t + u_{st} = 0. \tag{I.35}$$

Again, by taking a derivative with respect to x on both sides of equation (I.35), we obtain

$$(u_{sss}s^o_t + u_{sst})s^o_x + u_{ss}s^o_{xt} = 0. \tag{I.36}$$

Using (I.35), we can rewrite (I.36) as

$$\left(-\frac{u_{sss}}{u_{ss}}u_{st} + u_{sst}\right) s^o_x + u_{ss}s^o_{xt} = 0. \tag{I.37}$$

Since $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$ and since $u_{ss} < 0$, we have $s^o_{xt} = 0$.

Proof of Lemma 4: Since $u_s(s^o(x,t),t)$ is independent of t as shown in (13), we have

$$\frac{\partial^2}{\partial x \partial t}u(s^o(x,t),t) = \frac{\partial^2}{\partial t \partial x}u(s^o(x,t),t) = u_s s^o_{xt}. \tag{I.38}$$

Since $u_s > 0$ and $s^o_{xt} = 0$ by Lemma 3, we see that $\frac{\partial^2}{\partial x \partial t}u(s^o(x,t),t) = 0$, i.e., $\frac{\partial}{\partial t}u(s^o(x,t),t)$ is constant in x. By integration by parts, we have the following results:

$$\begin{aligned} \frac{d}{dt}\int_x u(s^o(x,t),t)f_a(x|a^o)dx &= \int_x \frac{\partial}{\partial t}u(s^o(x,t),t)f_a(x|a^o)dx \\ &= \underbrace{\frac{\partial}{\partial t}u(s^o(x,t),t)F_a(x|a^o)dx\Big|_{\underline{x}}^{\overline{x}}}_{=0} - \int F_a(x|a^o)\underbrace{\frac{\partial^2}{\partial t \partial x}u(s^o(x,t),t)}_{=0}dx = 0. \end{aligned} \tag{I.39}$$

Proof of Proposition 2: From Lemma 4, if $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$ and the agent's limited liability is not binding for any (x,t), then the optimal solution under the post-contract ex-post information system, $(s^o(x,t), a^o)$, also satisfies the following optimization program's constraints.

$$\max_{\substack{a\in A, s(x,t)\in S,\\ s(x,t)\geq 0}} \int_t\int_x (x-s(x,t))\, f(x|a)h(t)dxdt \quad \text{s.t.}$$
$$(i) \quad \int_t\int_x u(s(x,t),t)f(x|a)h(t)dxdt - v(a) \geq \overline{U}, \tag{I.40}$$
$$(ii) \quad \int_x u(s(x,t),t)f_a(x|a)dx - v'(a) = 0, \ \forall t \in T,$$

Usually, the only difference that the optimization program under the post-contract ex-post information system has from that under the post-contract ex-ante information system is: the principal is restricted to induce the same effort level from the agent across different realizations of t in the former case, whereas such a restriction is not imposed on the principal under the post-contract ex-ante information system. In this case, our agent does not have an incentive to deviate from a^o even when he knows the value of t. Trivially, we can directly have[4]

$$PW^o \leq PW^*. \tag{I.41}$$

However, since $PW^o \geq PW^*$ by Lemma 2, we finally have

$$PW^o = PW^*. \tag{I.42}$$

Now, assume to the contrary that $a^o \neq a^*(t)$ for some positive Borel measure of t. Then we see from the proof of Lemma 2 that $PW^o > PW^*$. So, there is a contradiction.

Proof of Proposition 3: Since $\frac{u_t}{u_s} = \frac{u_{st}}{u_{ss}}$, by taking a derivative with respect to s on both sides, we have

$$\frac{u_{st}u_s - u_t u_{ss}}{(u_s)^2} = \frac{u_{sst}u_{ss} - u_{st}u_{sss}}{(u_{ss})^2}. \tag{I.43}$$

Since $\frac{u_t}{u_s} = \frac{u_{st}}{u_{ss}}$, we have that the left-hand side of (I.43) equals zero, which also implies

$$\frac{u_{st}}{u_{ss}} = \frac{u_{sst}}{u_{sss}}. \tag{I.44}$$

Since the agent's limited liability constraint is not binding for any (x,t) and $\frac{u_{st}}{u_{ss}} = \frac{u_{sst}}{u_{sss}}$ from (I.44), by Proposition 2, we have

$$PW^p \leq PW^* = PW^o. \tag{I.45}$$

[4]As $(s^o(x,t), a^o)$ satisfies the constraints of the above optimization (I.40) and these constraints are those of the post-contract ex-ante optimization programs.

Since $s_t^o(x,t) = -\frac{u_{st}(s^o(x,t),t)}{u_{ss}(s^o(x,t),t)}$ from equation (I.35), and since $\frac{u_t}{u_s} = \frac{u_{st}}{u_{ss}}$, we have

$$\frac{\partial}{\partial t} u(s^o(x,t),t) = u_s(s^o(x,t),t)s_t^o + u_t(s^o(x,t),t) = 0, \tag{I.46}$$

implying that $u(s^o(x,t),t)$ is independent of t and, in turn, $\int_x u(s^o(x,t),t)f(x|a^o)dx - v(a^o)$ is independent of t. Thus,

$$\int_t \int_x u(s^o(x,t),t)f(x|a^o)h(t)dxdt - v(a^o) \geq \overline{U} \tag{I.47}$$

implies

$$\int_x u(s^o(x,t),t)f(x|a^o)dx - v(a^o) \geq \overline{U}, \ \ \forall t \in T. \tag{I.48}$$

Accordingly, $(s^o(x,t), a^o)$ satisfies the constraints of following optimization program, where the agent's participation and incentive constraints holds for every given t.

$$\begin{aligned}
&\max_{a \in A, s(x,t) \in S} \int_t \int_x (x - s(x,t))\, f(x|a)h(t)dxdt \quad \text{s.t.} \\
&(i) \quad \int_x u(s(x,t),t)f(x|a)dx - v(a) \geq \overline{U}, \ \ \forall t \in T, \\
&(ii) \quad \int_x u(s(x,t),t)f_a(x|a)dx - v'(a) = 0, \ \forall t \in T, \\
&(iii) \ \ s(x,t) \geq 0, \ \ \forall (x,t) \in X \times T.
\end{aligned} \tag{I.49}$$

The above (I.49) differs from the maximization program under the pre-contract information system only in that the principal is restricted to induce an effort level that is constant across realizations of t from the agent. Thus we have

$$PW^* = PW^o \leq PW^p. \tag{I.50}$$

Thus, from Proposition 2, we finally derive

$$PW^p = PW^* = PW^o. \tag{I.51}$$

Proof of Proposition 4: Given that $(a^{N^+}(T_i), s^{N^+}(x,t))$ is the solution for the optimization program in (31) under the information system N^+, let

$$K^{N^+}(T_i) \equiv \int_{t \in T_i} \int_x u\left(s^{N^+}(x,t),t\right) f\left(x|a^{N^+}(T_i)\right) h(t|T_i)dxdt - v\left(a^{N^+}(T_i)\right), \tag{I.52}$$

where $i = 1, 2, ..., j^-, j^+, ..., N$, and

$$\sum_{i=1}^{N^+} p_i K^{N^+}(T_i) \geq \overline{U}. \tag{I.53}$$

Then, $s^{N^+}(x,t)$ for $t \in T_i$, $i = 1, 2, ..., j^-, j^+, ..., N$, solves the following optimization:

$$\begin{aligned}
&\max_{\substack{s(x,t)\in S,\\ s(x,t)\geq 0}} \int_{t\in T_i}\int_x (x - s(x,t))\, f\left(x|a^{N^+}(T_i)\right) h(t|T_i)dxdt \quad \text{s.t.} \\
&(i)\ \int_{t\in T_i}\int_x u(s(x,t),t) f\left(x|a^{N^+}(T_i)\right) h(t|T_i)dxdt - v\left(a^{N^+}(T_i)\right) \geq K^{N^+}(T_i), \\
&(ii)\ \int_{t\in T_i}\int_x u(s(x,t),t) f_a\left(x|a^{N^+}(T_i)\right) h(t|T_i)dxdt = v'\left(a^{N^+}(T_i)\right).
\end{aligned} \tag{I.54}$$

Define the principal's expected benefits under information system N^+ given $t \in T_i$ as

$$PW^{N^+}(T_i) \equiv \int_{t\in T_i}\int_x \left(x - s^{N^+}(x,t)\right) f\left(x|a^{N^+}(T_i)\right) h(t|T_i)dxdt. \tag{I.55}$$

Now, let us go back to the system N and define $\left(a^m(T_i), s_m^N(x,t)\right)$ such that:[5]

$$a^m(T_i) = \begin{cases} a^{N^+}(T_i) & \text{for } i \neq j, \\ \dfrac{p_j^- a^{N^+}\left(T_j^-\right) + p_j^+ a^{N^+}\left(T_j^+\right)}{p_j^- + p_j^+} & \text{for } i = j, \end{cases} \tag{I.56}$$

where $p_j^- \equiv \int_{t\in T_j^-} h(t)dt$ and $p_j^+ \equiv \int_{t\in T_j^+} h(t)dt$.

We also define $s_m^N(x,t)$ for $t \in T_i$, $i = 1, 2, ..., j, ..., N$ as solving the following optimization program under the information system N:[6]

$$\begin{aligned}
&\max_{\substack{s(x,t)\in S,\\ s(x,t)\geq 0}} \int_{t\in T_i}\int_x (x - s(x,t))\, f\left(x|a^m(T_i)\right) h(t|T_i)dxdt \quad \text{s.t.} \\
&(i)\ \int_{t\in T_i}\int_x u(s(x,t),t) f\left(x|a^m(T_i)\right) h(t|T_i)dxdt - v\left(a^m(T_i)\right) \geq K^{m,N}(T_i), \\
&(ii)\ \int_{t\in T_i}\int_x u(s(x,t),t) f_a\left(x|a^m(T_i)\right) h(t|T_i)dxdt = v'\left(a^m(T_i)\right).
\end{aligned} \tag{I.57}$$

[5]It is similar to the way we defined a^m from $\{a(t)\}$ in equation (I.2).

[6]Therefore, the contract $s_m^N(x,t)$ induces the effort level $\{a^m(T_i)\}$ from the agent given $t \in T_i$, $i = 1, 2, ..., j, ..., N$.

where

$$K^{m,N}(T_i) = \begin{cases} K^{N^+}(T_i) & \text{for } i \neq j, \\ \dfrac{p_j^- K^{N^+}\left(T_j^-\right) + p_j^+ K^{N^+}\left(T_j^+\right)}{p_j^- + p_j^+} & \text{for } i = j. \end{cases} \tag{I.58}$$

Also, define the principal's expected benefits with $\left(a^m(T_i), s_m^N(x,t)\right)$ given $t \in T_i$ as

$$PW^m(T_i) \equiv \int_{t \in T_i} \int_x \left(x - s_m^N(x,t)\right) f\left(x|a^m(T_i)\right) h(t|T_i) dx dt. \tag{I.59}$$

Then, trivially we obtain when $i \neq j$:

$$PW^m(T_i) = PW^{N^+}(T_i) \;\text{ for }\; i \neq j, \tag{I.60}$$

and when $i = j$, we use the following Lemma 6, which we prove later.

Lemma 6. *For $i = j$, we have the following inequality.*

$$PW^m(T_j) \geq \frac{p_j^- PW^{N^+}\left(T_j^-\right) + p_j^+ PW^{N^+}\left(T_j^+\right)}{p_j^- + p_j^+}. \tag{I.61}$$

Therefore, we have the following:

$$\sum_{i=1}^{N} p_i PW^m(T_i) \geq PW^{N^+}. \tag{I.62}$$

Since $\sum_{i=1}^{N+} p_i K^{N^+}(T_i) \geq \overline{U}$, we can see that $(a^m(T_i), s_m^N(x,t))$ satisfies all constraints in (31) under the information system N given actions $\{a^m(T_i)\}$ such as

$$\sum_{i=1}^{N} p_i \left(\int_{t \in T_i} \int_x u\left(s_m^N(x,t), t\right) f\left(x|a^m(T_i)\right) h(t|T_i) dx dt - v(a^m(T_i)) \right) \geq \overline{U}, \tag{I.63}$$

$$\int_{t \in T_i} \int_x u\left(s_m^N(x,t), t\right) f_a\left(x|a^m(T_i)\right) h(t|T_i) dx dt = v'\left(a^m(T_i)\right), \; \forall T_i \in T, \tag{I.64}$$

and

$$s_m^N(x,t) \geq 0, \;\; \forall (x,t) \in X \times T. \tag{I.65}$$

This implies that $\left(a^m(T_i), s_m^N(x,t)\right)$ is in the feasible set of the optimization program under the information system N in (31). Therefore, we can conclude that

$$\sum_{i=1}^{N} p_i PW^m(T_i) \leq PW^N. \tag{I.66}$$

Consequently, we derive

$$PW^N \geq PW^{N^+}. \tag{I.67}$$

Now we have to prove the above Lemma 6.

Proof of Lemma 6: As $(a^{N^+}(T_i), s^{N^+}(x,t))$ is the optimal effort and contract under the system N^+, $s^{N^+}(x,t)$ solves the optimization in (I.54) for $i = j^-, j^+$. Let us define the agent's action as a function of t conditional on $t \in T_j \equiv T_j^- \cup T_j^+$ as

$$\text{for } \forall t \in T_j, \quad a_1(t) \equiv \begin{cases} a^{N^+}(T_j^-) & \text{for } \ t \in T_j^-, \\ a^{N^+}(T_j^+) & \text{for } \ t \in T_j^+. \end{cases} \tag{I.68}$$

Then we can express the variables defined above as[7]

$$a^m(T_j) = \frac{p_j^- a^{N^+}(T_j^-) + p_j^+ a^{N^+}(T_j^+)}{p_j^- + p_j^+} = \int_{t \in T_j} a_1(t) h(t|T_j) dt. \tag{I.70}$$

Then given $\{a_1(t)\}$, $s^{N^+}(x,t), t \in T_j = T_j^- \cup T_j^+$ solves the following optimization pro-

[7] We can easily obtain

$$\int_{t\in T_j}\int_x (x - s(x,t))\, f(x|a_1(t)) h(t|T_j) dx dt$$
$$= \frac{p_j^-}{p_j^- + p_j^+} \underbrace{\int_{T_j^-}\int_x (x - s(x,t))\, f\left(x|a^{N^+}_{T_j^-}\right) h\left(t|T_j^-\right) dx dt}_{\equiv PW^{N^+}\left(T_j^-\right) \text{ if } s(\cdot) = s^{N^+}(\cdot)} + \frac{p_j^+}{p_j^- + p_j^+} \underbrace{\int_{T_j^+}\int_x (x - s(x,t))\, f\left(x|a^{N^+}_{T_j^+}\right) h\left(t|T_j^+\right) dx dt}_{\equiv PW^{N^+}\left(T_j^+\right) \text{ if } s(\cdot) = s^{N^+}(\cdot)},$$

and

$$\int_{t\in T_j}\int_x u(s(x,t),t) f(x|a_1(t)) h(t|T_j) dx dt$$
$$= \frac{p_j^-}{p_j^- + p_j^+} \int_{T_j^-}\int_x u(s(x,t),t) f\left(x|a^{N^+}_{T_j^-}\right) h\left(t|T_j^-\right) dx dt + \frac{p_j^+}{p_j^- + p_j^+} \int_{T_j^+}\int_x u(s(x,t),t) f\left(x|a^{N^+}_{T_j^+}\right) h\left(t|T_j^+\right) dx dt. \tag{I.69}$$

Since $s^{N^+}(x,t)$ solves (31) given $\{a^{N^+}(T_i)\}, i = 1, 2, ..., j^-, j^+, ..., N$, and it satisfies constraints of (I.71) for $t \in T_j = T_j^- \cup T_j^+$, it solves **[P1]** as $s^{N^+}(x,t)$ for $t \in T_i$ solves the optimization program in (I.54).

gram $[\mathbf{P1}]$.

$$
\begin{aligned}
&[\mathbf{P1}] \max_{s(x,t)\in S} \int_{t\in T_j}\int_x (x - s(x,t))\, f(x|a_1(t))h(t|T_j)dxdt \quad \text{s.t.}\\
&(i) \ \int_{t\in T_j}\left(\int_x u(s(x,t),t) f(x|a_1(t))dx - v(a_1(t))\right) h(t|T_j)dt\\
&\geq K^{m,N}(T_j) = \frac{p_j^- K^{N^+}\left(T_j^-\right) + p_j^+ K^{N^+}\left(T_j^+\right)}{p_j^- + p_j^+},\\
&(ii) \ \begin{cases} \int_{t\in T_j^-}\int_x u(s(x,t),t) f_a\left(x|a^{N^+}(T_j^-)\right) h(t|T_j^-)dxdt = v'\left(a^{N^+}(T_j^-)\right), \\ \int_{t\in T_j^+}\int_x u(s(x,t),t) f_a\left(x|a^{N^+}(T_j^+)\right) h(t|T_j^+)dxdt = v'\left(a^{N^+}(T_j^+)\right), \end{cases}\\
&(iii) \ s(x,t) \geq 0, \ \forall (x,t) \in X \times T_i.
\end{aligned}
\tag{I.71}
$$

Thus we have the following expression, where $PW([\mathbf{P1}])$ stands for the principal's optimized value in $[\mathbf{P1}]$, which is given by (7).

$$
PW([\mathbf{P1}]) = \frac{p_j^- PW^{N^+}\left(T_j^-\right) + p_j^+ PW^{N^+}\left(T_j^+\right)}{p_j^- + p_j^+}. \tag{I.72}
$$

Comparing $[\mathbf{P1}]$ with (I.57) with $a^m(T_j) = \int_{t\in T_j} a_1(t)h(t|T_j)dt$ in (I.70), we observe that the only difference between $[\mathbf{P1}]$ and (I.57) is that (I.57) induces the average action $a^m(T_j)$ from the agent across T_j^- and T_j^+, T_j's two partitions, whatever $t \in T_j$ is realized. In this case, Lemma 2 proves it gives a higher efficiency from principal's standpoint, thereby:

$$
PW^m(T_j) \geq PW([\mathbf{P1}]) = \frac{p_j^- PW^{N^+}(T_j^-) + p_j^+ PW^{N^+}(T_j^+)}{p_j^- + p_j^+}. \tag{I.73}
$$

which proves Lemma 6.

Appendix B. Derivation of Section 5.1

In this section, we obtain the closed-form solution of $AC^*_M(\{a(t)\}$ and $AC^o_M(a^o)$ and prove Lemma 5. Let λ^* and $\mu^*(t)h(t)$ be the Lagrange multiplier to the constraints (i) and (ii) respectively, of the following optimization (I.74):

$$
\begin{aligned}
AC^*_M(\{a^*(t)\}) \equiv [\text{EA}]_M \min_{s(x,t)\in S} & \int_t\int_x s(x,t)f(x|a^*(t))h(t)dxdt \quad \text{s.t.} \\
& (i) \quad \int_t \left(\int_x 2t\sqrt{s(x,t)}f(x|a^*(t))dx - v(a^*(t))\right) h(t)dt \geq \overline{U}, \\
& (ii) \quad \int_x 2t\sqrt{s(x,t)}f_a(x|a^*(t))dx - v'(a^*(t)) = 0, \ \forall t \in T, \\
& (iii) \ \ s(x,t) \geq 0, \ \ \forall(x,t) \in X \times T.
\end{aligned}
\tag{I.74}
$$

Solving Euler equation for (I.74), the solution $s^*(x,t)$ must be represented as follows, since we assume away the limited liability constraint (iii).

$$
\frac{1}{u_s(s^*(x,t))} = \frac{\sqrt{s^*(x,t)}}{t} = \lambda^* + \mu^*(t)\frac{f_a}{f}(x|a^*(t)), \ \ s(x,t) \geq 0, \tag{I.75}
$$

leading to

$$
s^*(x,t) = t^2\left(\lambda^* + \mu^*(t)\frac{f_a}{f}(x|a^*(t))\right)^2 \tag{I.76}
$$

and

$$
u(s^*(x,t)) = 2t\sqrt{s^*(x,t)} = 2t^2\left(\lambda^* + \mu^*(t)\frac{f_a}{f}(x|a^*(t))\right) \geq 0. \tag{I.77}
$$

We plug the above (I.77) into the above constraints and obtain

$$
\mu^*(t) = \frac{v'(a^*(t))}{2t^2}\frac{1}{\text{Var}\left(\frac{f_a}{f}\right)} \tag{I.78}
$$

and

$$
\lambda^* = \frac{\int_t v(a^*(t))h(t)dt + \overline{U}}{2\int_t t^2h(t)dt}. \tag{I.79}
$$

Finally we can write $AC^*_M(\{a^*(t)\})$ in (I.74) as

$$\begin{aligned}
AC^*_M(\{a(t)\}) &= \int_t \left(\int_x s^*(x,t) f(x|a^*(t))dx \right) h(t)dt \\
&= (\lambda^*)^2 \left(\int_t t^2 h(t)dt \right) + \mathrm{Var}\left(\frac{f_a}{f}\right) \left(\int_t t^2 \mu^*(t)^2 h(t)dt \right) \\
&= \underbrace{\frac{\left(\int_t v(a^*(t))h(t)dt + \overline{U} \right)^2}{4 \int_t t^2 h(t)dt}}_{\equiv AC^*_{M,\mathrm{IR}}(\{a^*(t)\})} + \underbrace{\frac{1}{4\mathrm{Var}\left(\frac{f_a}{f}\right)} \int_t \frac{1}{t^2} v'(a^*(t))^2 h(t)dt}_{\equiv AC^*_{M,\mathrm{IC}}(\{a^*(t)\})} .
\end{aligned} \tag{I.80}$$

For the post-contract ex-post system, we define λ^o and μ^o to be the multipliers to the constraints (i) and (ii) respectively of the following optimization:

$$\begin{aligned}
AC^o_M(a^o) \equiv \text{[EP]}_M \min_{s(x,t)\in S} & \int_t \int_x s(x,t) f(x|a^o) h(t) dx dt \quad \text{s.t.} \\
(i) \ & \int_t \int_x 2t\sqrt{s(x,t)} f(x|a^o) dx h(t) dt - v(a^o) \geq \overline{U}, \\
(ii) \ & \int_t \int_x 2t\sqrt{s(x,t)} f_a(x|a^o) dx h(t) dt - v'(a^o) = 0, \\
(iii) \ & s(x,t) \geq 0, \ \ \forall (x,t) \in X \times T.
\end{aligned} \tag{I.81}$$

Solving Euler equation for (I.81) assuming that the limited liability constraint (iii) does not bind, the solution $s^o(x,t)$ must be represented as

$$\frac{1}{u_s(s^o(x,t))} = \frac{\sqrt{s^o(x,t)}}{t} = \lambda^o + \mu^o \frac{f_a}{f}(x|a^o), \tag{I.82}$$

where the solution $(s^o(x,t))$ must satisfy:

$$s^o(x,t) = t^2 \left(\lambda^o + \mu^o \frac{f_a}{f}(x|a^o) \right)^2 \ \text{ and } \ u(s^o(x,t)) = 2t^2 \left(\lambda^o + \mu^o \frac{f_a}{f}(x|a^o) \right). \tag{I.83}$$

Plugging (I.83) into the constraints in (I.81), we obtain λ^o and μ^o.

$$\mu^o = \frac{v'(a^o)}{2\int_t t^2 h(t)dt} \cdot \frac{1}{\mathrm{Var}\left(\frac{f_a}{f}\right)}, \quad \lambda^o = \frac{v(a^o) + \overline{U}}{2\int_t t^2 h(t)dt}. \tag{I.84}$$

Finally, we obtain the value of the objective function in (I.81).

$$\begin{aligned}
AC_M^o(a^o) &= \int_t \Big(\int_x s^o(x,t) f(x|a^o) dx \Big) h(t) dt \\
&= (\lambda^o)^2 \left(\int_t t^2 h(t) dt \right) + \mathrm{Var}\left(\frac{f_a}{f} \right) (\mu^o)^2 \left(\int_t t^2 h(t) dt \right) \\
&= \underbrace{\frac{\big(v(a^o) + \overline{U}\big)^2}{4 \displaystyle\int_t t^2 h(t) dt}}_{\equiv AC_{M,\text{IR}}^o(a^o)} + \underbrace{\frac{v'(a^o)^2}{4\mathrm{Var}\left(\frac{f_a}{f} \right)} \frac{1}{\displaystyle\int_t t^2 h(t) dt}}_{\equiv AC_{M,\text{IC}}^o(a^o)}.
\end{aligned} \tag{I.85}$$

Proof of Lemma 5: $AC_{M,\text{IR}}^o(a^m) < AC_{M,\text{IR}}^*(\{a(t)\})$ holds since $v(\cdot)$ is convex as seen in

$$AC_{M,\text{IR}}^o(a^m) = \frac{\big(v(a^m) + \overline{U}\big)^2}{4 \displaystyle\int_t t^2 h(t) dt} < \frac{\left(\displaystyle\int_t v(a^*(t)) h(t) dt + \overline{U} \right)^2}{4 \displaystyle\int_t t^2 h(t) dt} = AC_{M,\text{IR}}^*(\{a(t)\}). \tag{I.86}$$

To show $AC_{M,\text{IC}}^o(a^m) < AC_{M,\text{IC}}^*(\{a(t)\})$, it is sufficient to prove

$$\int_t \frac{1}{t^2} v'(a^*(t))^2 h(t) dt > \frac{v'(a^m)^2}{\displaystyle\int_t t^2 h(t) dt}. \tag{I.87}$$

If define the random variables $X = \frac{1}{Z} v'(a^*(Z)), Y = Z, Z \sim h(Z)$, since $v'(\cdot)$ is convex, we get the following inequality, which proves our claim.[8]

$$\begin{aligned}
&\mathbb{E}(X^2)\mathbb{E}(Y^2) = \left(\int_t \frac{1}{t^2} v'(a^*(t))^2 h(t) dt \right) \left(\int_t t^2 h(t) dt \right) \\
&> (\mathbb{E}(XY))^2 = \left(\int_t v'(a^*(t)) h(t) dt \right)^2 \\
&> v'(a^m)^2.
\end{aligned} \tag{I.88}$$

[8] Due to the Cauchy-Schwarz inequality, as random variable X and Y are not linearly dependent, the first inequality must be strict.