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

Jin Yong Jung† Kangnam University

Son Ku Kim‡ Seoul National University

Seung Joo Lee§ U.C. Berkeley

October 18, 2021

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 show 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.

†jyjung@kangnam.ac.kr

‡sonkukim@snu.ac.kr

§seungjoo@berkeley.edu

1 Introduction

The moral-hazard problem has been one of the central topics in the economics of information over the past three decades. It arises when a principal has to delegate to an agent to act on his 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 quite 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. For example, most firms in the U.S. provide their workers with various pension plans along with regular salary schemes which will be effective upon retirement. The final payouts from those pension plans are generally affected neither by the firms' decisions nor by the workers' decisions but by the outside pension managers' portfolio decisions and the state of market. However, information on such payouts will have an effect on the workers' decision-making if it is revealed to the workers in advance.

Other examples for such factors include each worker's personal matches to his workplace, his pride of being a member of a specific company, and feeling of accomplishment from the job. Those variables are random in the sense that they are controlled neither by the principal nor by the agent, and they are usually unknown both to the principal and the agent in the beginning. But, there may be an information system that generates some information on those variables, and information on such variables will affect the agent's action choice if it is revealed before the agent takes an action because the agent's utility depends on those variables. Thus, we ask the following questions: When should information on those variables be revealed? Should information on those variables be revealed as precisely as possible? For instance, in the above pension plan example, should CEOs make plan managers reveal information on prospects of companies' 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 have made their decisions? How does timing of such information revelation affect the efficiency of a principal-agent relationship? Here we try to answer those questions through the standard theoretical framework.

To answer those questions, we formulate a principal-agent model in which the agent's utility depends not only on his monetary wage but also on some other variables such as non-wage benefits or his personal matches to the working environments. We consider the case in which the principal can implement one out of three different information systems: a pre-contract information system, a

[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.

post-contract ex-ante information system, and a post-contract ex-post information system. The pre-contract information system publicly reveals perfect information on those random variables before principal and agent make an agreement on the contract. On the other hand, the post-contract ex-ante information system publicly reveals perfect information 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.

We first show that 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 principal and agent sign the contract (i.e., the pre-contract information system), the principal should design a wage contract that satisfies the agent's participation constraint for every realization of the variable, whereas the principal has to design a wage contract that satisfies the agent's participation constraint only on average under the post-contract ex-ante information system, which is easier.

We also show that the principal prefers the post-contract ex-post information system to the post-contract ex-ante information system, implying that the principal and agent's being informed ex-ante (i.e., before the agent takes an action) is not as efficient as their being informed ex-post (i.e., the agent has taken his action). If principal and agent are informed ex-ante, 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 the principal's interests. Thus, principal should design a wage 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. But, the principal, in this case, needs to design a wage contract that satisfies the agent's incentive constraint only on expectation, which is easier. Therefore, there is a trade-off between both parties' being informed ex-ante and their being informed ex-post. We show 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 agency 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 variables. Thus, we also investigate a relationship between the amount of information contained in a 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 (or the other system is more coarse). To see 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. Thus, it implies that in the pension plan example described above, plan managers should not be allowed to reveal any information on prospects of their pension plans ex-ante (i.e., before the workers make their decisions).

The relationship between the amount of information contained in an information system and its efficiency in an agency model has been widely discussed in the literature. For instance, Holmström (1979) and Shavell (1979) showed that, between two costless public information systems, the information system with additional signals is strictly preferred by the principal if and only if at least one of those additional signals is informative about the agent's hidden action choice. Their notion of informativeness is defined in terms of a sufficient statistic. Gjesdal (1982) and Grossman and Hart (1983) applied 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 general principal-agent models. Kim (1995) showed that Blackwell's information sufficiency conditions are unnecessarily strong for ranking information systems in agency models, and provided a weaker criterion. That is, if one information system has 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.[3]

Those papers consider the efficiency of an information system which generates information on random variables that are affected by the agent's hidden action, such as outputs, annual earnings. stock prices, and so on. Furthermore, information on those variables is available <u>after</u> the agent has taken an action (i.e., ex-post). In this paper, we consider the efficiency of an information system which generates information on random variables that are not affected by but affect the agent's action choice, and primarily focus on the relationship between efficiency of an information system and its timing of information revelation.

One related paper to ours is Sobel (1993). He compared three information systems in agency models: a zero information system, a pre-contract information system, and a post-contract information system (i.e., the post-contract ex-ante information system in our paper). He showed that the principal always prefers the post-contract information system to the pre-contract information system,[4] but, depending on the underlying model specification, the principal may or may not prefer

[2]Thus in this case that a contract can always be dependent upon the realized values of those variables ('t' in our paper) as well as realized output.

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

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

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 it is revealed both to the principal and the agent in ours. 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) considered the case where, after principal and the agent sign on contract, principal gets a signal (private or public[5]) related to the technology. If the signal is private, then principal announces the signal to the agent possibly not truthfully. Related to our main results, he concluded if the principal wants to implement the same action for any signal ('t' in our paper), then public signal (before agent takes an action) lowers the efficiency, as the agent can take advantage of this public signal in a pursuit of his own interest. In contrast to his paper, however, we abstract from the adverse selection issue and truth-telling mechanism and instead focus on the timing and amount of that public signal's revelation, and our signal is about the agent's marginal utility rather than the technology. Banker et al. (2019) analyzed a value of information in the case where principal doesn't know some relevant traits of the agent ('t' in our paper) and uses the set of contracts to tackle both the adverse selection and moral hazard issues. Our paper rather focuses on the moral hazard side and how the timing and precision of information revelation affects the efficiency of an agency relationship. À la information design literatures originated from Kamenica and Gentzkow (2011), recently Boleslavsky and Kim (2020) considered the world where sender designs a signal about the output, which is driven by the agent's hidden action, and developed the general method of characterizing the optimal signal structure when moral hazard issue exists. Our work in contrast focuses on the timing and coarseness of the relevant 'exogenous' information, which is relevant to the agent's marginal utility, instead of endogenous output level, and its implication for the agency efficiency, abstracting from the endogenous design of signal to the technology in the existence of moral hazard problems.

The rest of the paper is organized in the following way. Section 2 formulates our basic framework. In Section 3, we compare efficiencies of three different information systems. In Section 4, we compare relative efficiencies 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 an example that illustrates main propositions in Section 3. This example allows us to directly decompose an agency cost and shows in the most intuitive way why under our assumptions the ex-ante information revelation can hurt the efficiency of agency relation, compared with the ex-post system.

[5]Public signal means the agent as well as the principal gets to know the exact value of the signal. And contracts can be dependent upon both this signal value and (realized) output level as in our specification.

Concluding remarks are given in Section 6, and all the proofs of the lemmas and propositions in this paper are provided in the Appendix.

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 effort $a \in A = [0, \infty)$. Output $x \in X = [\underline{x}, \overline{x}] \subset R$ 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.[6] We assume that production function $X(a, \theta)$ is additively separable. Thus, without loss of generality, $X(a, \theta)$ can be simply represented by

$$X(a, \theta) = a + \theta, \quad E(\theta) = 0.^{7} \tag{1}$$

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 a. An important difference of the agent's utility function described in equation (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

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

[7] 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

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

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.

Information on those variables is available both to the principal and the agent. However, when such information becomes available depends on the information structure that is implemented by the principal at the outset. We assume that there are three different information systems: a pre-contract information system, a post-contract ex-ante information system, and 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 three different information systems: the pre-contract information system, the post-contract ex-ante information system, and 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 wage 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. When the post-contract ex-post information system is chosen by the principal, the agent takes an effort before a true value of t is known. However, under either 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 above assumptions, each subscript or prime indicates that derivatives are taken at the corresponding orders. For example, $u_{ss} \equiv \frac{\partial^2}{(\partial s)^2}u$. 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 utility. 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[8] can be written as:

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

where $\overline{U}$ denotes the agent's reservation utility level that he can obtain 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 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 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 for any given t. This limited liability constraint on the agent's side is given for the existence of an optimal wage contract given t.[9]

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

[8] Optimally induced action a_t and the optimal contract $s_t(\cdot)$ both depend upon the realized value of t.

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

$$\max_{a(t)\in A, s(x,t)\in S} \int_t \int_x [x - s(x,t)] f(x|a(t)) h(t) dx dt \quad \text{s.t:}$$

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

where $h(t)$ denotes the density function of t which is common knowledge to both principal and the agent. As before, S in the above program 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,[10] the above optimization program reduces to"

$$\max_{a(t)\in A, s(x,t)\in S} \int_t \int_x [x - s(x,t)] f(x|a(t)) h(t) dx dt \quad \text{s.t.}$$

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

[10]Following Jewitt (1988), it can be shown that using the first-order approcah is valid if 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 each 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 that is obtained by 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, conditions in (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 concave in a for any given t. Consequently, the conditions in (J1) and (J2) sufficiently justify the use of the first-order approach in the above optimization program. As shown in Jewitt (1988), the conditions in (J1) and (J2) are much weaker than the usual MLRP (Monotone Likelihood Ratio Property) and CDFC (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 families of distributions such as exponential, Chi-squared, and Poisson distributions satisfy (J2). For more details, see Jewitt (1988), and for more recent developments, see Sinclair-Desgagné (1994), Conlon (2009), Jung and Kim (2015) among others.

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 equation (9) has a solution $s^p(x,t) \geq 0$ and otherwise $s^p(x,t) = 0$. In the above equation, $\lambda^p(t)$ 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:

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

3.2 Post-Contract Ex-Post Information System

When the post-contract ex-post information system is implemented, t will be publicly known at the end of the period. Thus, it can be used as a contractual variable, and thus agent's wage contract, s, will be based on (x,t), i.e., $s = s(x,t)$. However, agent's effort choice, a, must be constant across different realizations of t because t will be available <u>after</u> the agent has taken a.

Therefore principal's maximization program in this case is:[11]

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

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 <u>before</u> t is realized.

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

[11] Below [EP] stands for the post-contract ex-post contract. Later we will use [EA] to denote post-contract ex-ante system.

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

Let $(a^o, s^o(x,t))$ be the optimal solution for the above program. Then, solving Euler equation of the above program 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 equation (13) has a solution $s^o(x,t) \geq 0$ and otherwise $s^o(x,t) = 0$. In the above equation, 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 equation (13) is that t does not appear on the right-hand side of the equation. Thus, equation (13) requires that agent's marginal utility on income to be constant across different realizations of t when $s^o(x,t)$ is offered, as long as the agent's limited liability constraint is not binding. By doing so, principal can insure the agent from a risk stemming from t and improve an efficiency of her relation with the agent, as t does not affect the agent's action 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:

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

3.3 Post-Contract Ex-Ante Information System

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

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

The first constraint denotes agent's participation constraint, reflecting that principal and agent do not know the true value of t at the time of contracting. The second constraint denotes the agent's incentive constraint, reflecting that the principal and the agent commonly know the true value of t before the agent takes 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, agent's effort choice is generally a non-trivial function of t. Actually, both differences arise from the fact that principal and agent can commonly observe the true value of t before the agent takes a 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:

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

Let $(a^*(t), s^*(x,t))$ be the optimal solution for the above program. Then, solving Euler equation of the above program 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 equation (17) has a solution $s^*(x,t) \geq 0$ and otherwise $s^*(x,t) = 0$. λ^* denotes the optimized Lagragian multiplier of agent's participation constraint and $\mu^*(t)$ is the optimized Lagrangian multiplier of agent's incentive constraint when t is realized. Since the agent's incentive constraint must be satisfied for every t in the above program, $\mu^*(t)$ must be a non-trivial function of t.

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

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

3.4 Ranking of Three Information Systems

We first compare the efficiency of the pre-contract information system with that of 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 wage contract, should consider the agent's participation constraint for every t under the pre-contract information system, whereas he should consider the agent's participation constraint only on average (across t) under post-contract ex-ante information system, which is easier for the principal.[12]

Now, to compare efficiency of the post-contract ex-ante information system with that of post-contract ex-post information system, 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

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

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

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 world where the principal directly observes the agent's effort choice. This is because in the first-best world, principal can just design a forcing contract that enforces the agent to take a^{FB} under any circumstance and guarantees him the reservation utility, $\overline{U}$.[13] 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 the principal cannot directly observe the agent's effort choice. This is because t affects the agent's marginal utility. It implies that knowing the true value of t ex-ante has some information value in the second-best world. Thus, when the principal knows the agent observes the true value of t before he takes an action, she can make 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 value of t ex-ante, will use this information to adjust effort level $a(t)$ for his own interest. Therefore, principal must design a wage contract that satisfies 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 based on expectation about t. As a result, there is a trade-off between 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 over its benefit.

To understand this result more precisely, 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. Then, a^m under the post-contract ex-post information system will produce the same expected output as $a^*(t)$ would produce under the post-contract ex-ante information system. However, inducing a^m under post-contract ex-post information system is less costly than inducing $a^*(t)$ under the post-contract ex-ante information system. This is because, as shown in the proof in the 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 comparing 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 infrmation system harder to satisfy comparing with the same constraint for inducing a^m under post-contract ex-post information system. Therefore, from the agency perspective, 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. However,

[13] 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.

since inducing a^m may not be optimal 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. (i.e., $PW^* \leq PW^o$). This intuition is better illustrated in Section 5 in which we use the specific type of 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 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 if 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 that if the agent's utility function satisfies $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$, and if limited liability is not binding for any (x, t), then the optimal wage contract under 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 if 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 post-contract ex-post information system, principal will induce the agent to take a certain effort level based on expectation with respect to t. Therefore, 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 incentive 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 the agent's utility function satisfies $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$, and if the agent's limited liability constraint is not binding for any (x,t), the optimal wage 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,

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

Proposition 2 derives sufficient conditions under which the effort level that will be optimally induced from the agent under the post-contract ex-ante information system is actually 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 mentioned earlier, 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 if limited liability is not binding for any (x,t), the optimal contract under post-contract ex-post information system, $s^o(x,t)$, would also satisfy the agent's incentive constraint for every t because $s^o(x,t)$ must contain the same amount of effort incentive for every t. This implies that the cost of both parties' observing t ex-ante compared with their observing t ex-post disappears. On the other hand, since the effort level optimally induced from the agent under 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 agent's limited liability constraint is not binding for any (x,t) will be 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 functions $u_s(s,t)$ 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 will be 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 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, principal can minimize the amount of risk imposed on the agent which stems 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, 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. But, even under the post-contract ex-ante information system where both parties can observe t ex-ante, 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 the sufficient conditions under which the principal becomes indifferent between post-contract ex-post information system and the post-contract ex-ante information system, while she still prefers the both post-contract information systems to the pre-contract information system. We now derive sufficient conditions under which a principal becomes indifferent among those three information systems.

Proposition 3 *If $\frac{u_s}{u_{ss}} = \frac{u_t}{u_{st}}$, and if 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 sufficient conditions under which the principal becomes indifferent among three different information systems. Comparing with Proposition 2, we need a stronger condition on the agent's utility function $\frac{u_s}{u_{ss}} = \frac{u_t}{u_{st}}$, which is stronger than $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$. This 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 will be 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 the perfect substitutes in agent's marginal utility with respect to s. Thus, if s and t are perfect substitutes in the agent's utility, then as shown in Proposition 2, the case that both parties observe t before the action is taken will not be different in terms of the agent's incentive problem compared to the case they observe it ex-post (i.e., after the agent has taken a). 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 since it will be the case $s^p(x,t) = s^*(x,t) = s^o(x,t) = s^o(x) - k(t)$.

4 Value of Coarse Ex-Ante Information

As shown in Lemma 2, we have already proved that the post-contract ex-ante information system which reveals perfect information on t ex-ante is less efficient than the post-contract ex-post information system which reveals perfect information on t 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 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.

Thus in this section, we investigate whether the amount of imperfect ex-ante information (before an action is taken) on random factors t that affect agent's marginal utility matters to the equilibrium payoff of 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 that true values of those variables become publicly observable at the end of the period as before. 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 hidden variables) than the partition of the other system. The finer partition generates more precise ex-ante information on those variables.

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

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 has two sub-partitions for T_j compared with system N.

Then, the principal's maximization program under information system N is:

$$\begin{aligned}
&\max_{\{a(T_i)\in A\}_{1\leq i\leq N}, s(x,t)\in S} \sum_{i=1}^{N} p_i \int_{t\in T_i}\int_x [x - s(x,t)] f(x|a(T_i)) h(t|T_i) dx dt \quad \text{s.t.} \\
&(i) \ \sum_{i=1}^{N} p_i \Big(\int_{t\in T_i}\int_x u(s(x,t),t) f(x|a(T_i)) h(t|T_i) dx dt - v(a(T_i)) \Big) \geq \overline{U}, \\
&(ii) \ \int_{t\in T_i}\int_x u(s(x,t),t) f_a(x|a(T_i)) h(t|T_i) dx dt = v'(a(T_i)), \ \forall T_i \in T, \\
&(iii) \ s(x,t) \geq 0, \ \ \forall (x,t) \in X \times T,
\end{aligned} \tag{31}$$

where $p_i \equiv \int_{t\in T_i} h(t)dt$ and $h(t|T_i)$ denotes the conditional density function of t given $t \in T_i$. Note that the agent's 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 $(a^N(T_i), s^N(x,t))$ be the optimal solution for the above program. Then, the principal's expected benefits under information system N are

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

Note that the principal's maximization program under information system N^+ is the same as equation (31) except that the summation $\sum_{i=1}^{N}$ now becomes $\sum_{i=1}^{N^+}$ that covers $N+1$ 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. Then, the principal's expected benefits under information system N^+ are

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

Then, we obtain the following Proposition 4:

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 do differ in the amount of ex-ante information. Proposition 4 states that the finer ex-ante information in terms of partition an information system provides, the less efficient it makes the principal-agent relation when the true value of t is available ex-post. This implies that there is actually a positive correlation between the coarseness of ex-ante information system and its efficiency in the agency framework.

We now return to our pension plan example and based on the results derived above infer the following policy implication. When t denotes the final payout from the pension plan, s and t are perfect substitutes in the agent's utility, 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), then it is more likely that the agent's limited liability is not 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 relation. However, as the final payout from the pension plan takes up a more crucial portion in the agent's utility, 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 during the agency relationship or reduce the amount of available information to the agent before the actual action is taken.

5 Illustrative Examples

In this section we provide examples with simple utility functions to illustrate the points made in previous sections and underlying economic intuitions that drive results. For that purpose, we use two different types of utility function[14]: (i) multiplicative utility ($u(s,t) = 2t\sqrt{s}$, $t > 0$) and (ii) additive utility ($u(s,t) = 2\sqrt{s + l(t)}$, where $l(\cdot)$ is C^1 function of t). The first class of utility represents cases in which t is multiplicative factor that affects the agent's marginal utility, for example, the rate of return on workers' pension fund. The second class may feature the workers' hidden non-monetary satisfaction from workplaces that adds up with monetary payments, s, before entering utility function. These examples turn out to allow us to get the closed form expression of agency costs, and illustrate why ex-ante information revelation is inefficient from the principal's perspective. Here we assume $\overline{U}$ is large enough that a limited liability ($s(x,t) \geq 0$) never binds. All

[14]An agent's square-root utility function ($u(s) = 2\sqrt{s}$) was used in Kim and Suh (1991) in characterization of value of information in canonical agency frameworks.

detailed derivations are provided in the Appendix.[15]

5.1 Multiplicative Utility Case

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

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

$AC_M^*(\{a^*(t)\})$ is an 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 agency cost in a closed-form as[16]

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

The first term ($AC_{M,\text{IR}}^*(\{a^*(t)\})$) is the cost of ensuring 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 an expected output is $\int_t \int_x x f(x|a^*(t)) dx h(t) dt = \int_t a^*(t)h(t)dt$, principal chooses set of actions $\{a^*(t)\}$ to induce by solving such an optimization as:

[15] We omit derivation of Section 5.2 in Appendix. 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 get expressions in Section 5.2.

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

$$PW_M^* = \max_{\{a^*(t)\}} \int_t a^*(t)h(t)dt - AC_M^*(\{a^*(t)\})$$
$$= \int_t a^*(t)h(t)dt - \frac{\left(\int_t v(a^*(t))h(t)dt + \overline{U}\right)^2}{4\int_t t^2 h(t)dt} - \frac{1}{4\mathrm{Var}\left(\frac{f_a}{f}\right)} \int_t \frac{1}{t^2} v'(a^*(t))^2 h(t)dt. \tag{37}$$

First-order conditions for $a^*(t)$ for each t tells us how a set of optimally induced $\{a^*(t)\}$ depends on the realization of t as shown in

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

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 a higher t is revealed ex-ante before agent takes an action, principal respects the fact that a higher t leads to a higher marginal utility of the agent, and thus the agent wants to put more effort as t and s are complementary. In sum, we get a strictly dispersed, optimally induced action set $\{a^*(t)\}$.

For the post-contract ex-post system, let us also first fix the action a^o induced by the principal, and see how the principal minimizes the agency cost. Principal solves the following optimization:

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

$AC_M^o(a^o)$ is an agency cost associated with inducing the given action a^o for $\forall t$ in the post-contract ex-post system. It turns out that we can express the agency cost in a closed-form in the following way.

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

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 ($AC^o_{M,\mathrm{IC}}(a^o)$) is interpreted as the cost of satisfying agent's incentive compatibility. As expected output is $\int_t \int_x x f(x|a^o)h(t)dxdt = a^o$, the principal chooses an action a^o to induce by solving the following optimization:

$$PW^o_M = \max_{a^o} a^o - AC^o_M(a^o) = a^o - \frac{(v(a^o) + \overline{U})^2}{4\int_t t^2 h(t)dt} - \frac{v'(a^o)^2}{4\mathrm{Var}\left(\frac{f_a}{f}\right)} \frac{1}{\int_t t^2 h(t)dt}. \tag{41}$$

The following Lemma 5 allows us to directly compare the relative size of each component of two systems' agency costs. Before we proceed, let us 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 tells us that inducing a^m, a mean of $\{a^*(t)\}$ under the post-contract ex-post system, costs less to the principal than inducing $\{a^*(t)\}$ under the post-contract ex-ante system. As inducing a^m is inferior (from the principal's point of view) to inducing the optimal a^o in the ex-post system, we get $PW^o_M > PW^*_M$[17]. In the ex-ante system, principal respects the fact that the agent observes a revealed t before he takes an action, and thus induces different actions $\{a^*(t)\}$ for different realizations of t, since each different t value affects the agent's marginal utility and the degree of incentive compatibility, and the principal takes advantage of it. This 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 actions for each t are dispersed, to make sure the agent sign on the contract, principal on average must pay more to the agent, since average cost of actions for the agent goes up. Assumption 3 that $v'(\cdot)$ is convex guarantees that inducing a dispersed set of action $\{a(t)\}$ under the ex-ante system is less efficient to the principal than inducing its average in ex-post system, since incentive cost inevitably rises on average. This decomposition of the agency cost ($AC^o_M(\cdot)$) into the component ($AC^o_{M,IR}(\cdot)$) that governs the agent's participation constraint and the other one ($AC^o_{M,IC}(\cdot)$) 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. It turns out that in the case where t enters into the utility in an additive way, principal can wisely design a contract to perfectly insure

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

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

the agent from the t risk, and the timing of information revelation does not affect the efficiency of the agency relation.

5.2 Additive Utility Case

For post-contract ex-ante system, we follow the similar strategy we applied in the above Section 5.1. We first fix the set of actions $\{a^*(t)\}$ induced by the principal, and see how principal designs contract to minimize the agency cost. Principal solves the following optimization in equation (43). Here $[\text{EA}]_{Add}$ stands for the additive utility case under ex-ante system.

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

where $AC^*_{Add}(\{a(t)\})$ is defined as the agency cost[18] associated with inducing action $a(t)$ for each value of t in the post-contract ex-ante system. It turns out that we can express the agency cost in a closed-form as:

$$AC^*_{Add}(\{a(t)\}) = \underbrace{\frac{\left(\int_t v(a^*(t))h(t)dt + \overline{U}\right)^2}{4}}_{\equiv AC^*_{Add,\text{IR}}(\{a^*(t)\})} + \underbrace{\frac{1}{4\text{Var}\left(\frac{f_a}{f}\right)} \int_t v'(a^*(t))^2 h(t)dt}_{\equiv AC^*_{Add,\text{IC}}(\{a^*(t)\})} - \underbrace{\int_t l(t)h(t)dt}_{=\mathbb{E}(l(t))}. \tag{45}$$

The first term ($AC^*_{Add,\text{IR}}(\{a^*(t)\})$) is the cost of ensuring agent's individual rationality, as the agent can always get $\overline{U}$ amount of utility in other jobs. The second term ($AC^*_{Add,\text{IC}}(\{a^*(t)\})$) is interpreted as the cost of satisfying agent's incentive compatibility constraint for $\forall t$ as before. Here $\mathbb{E}(l(t))$ and $AC^*_{Add,\text{IR}}(\{a^*(t)\})$ are perfect substitutes, as s and $l(t)$ are perfect substitutes. As an expected output is $\int_t \int_x x f(x|a^*(t))dxh(t)dt = \int_t a^*(t)h(t)dt$, principal chooses a set of actions $\{a^*(t)\}$ to induce by

[18] If λ^* and $\mu^*(t)h(t)$ are multipliers for the participation and incentive constraint (for specific t), respectively, these variables are characterized as

$$\lambda^* = \frac{\overline{U} + \int_t v(a^*(t))h(t)dt}{2}, \quad \mu^*(t) = \frac{v'(a^*(t))}{2\text{Var}\left(\frac{f_a}{f}\right)}. \tag{44}$$

solving the following optimization.

$$
\begin{aligned}
PW^*_{Add} &= \max_{\{a^*(t)\}} \int_t a^*(t)h(t)dt - AC^*_{Add}(\{a^*(t)\}) \\
&= \int_t a^*(t)h(t)dt - \frac{\left(\int_t v(a^*(t))h(t)dt + \overline{U}\right)^2}{4} - \frac{1}{4\text{Var}\left(\frac{f_a}{f}\right)} \int_t v'(a^*(t))^2 h(t)dt + \int_t l(t)h(t)dt
\end{aligned} \tag{46}
$$

First-order condition tells us how the optimally induced $a^*(t)$ depends on the realization of t, as shown in

$$
1 = v'(a^*(t))\Big[\underbrace{\frac{\int_t v(a^*(t))h(t)dt + \overline{U}}{2}}_{\text{Constant for } \forall t} + \frac{v''(a^*(t))}{2\text{Var}\left(\frac{f_a}{f}\right)}\Big], \quad \forall t. \tag{47}
$$

As $v'(\cdot), v''(\cdot)$ are increasing, we get $a^*(t) = a^*$ for $\forall t \in T$. Here t appears in the agent's utility in an additive manner. 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 action across different realizations of t, as a dispersion in action $\{a^*(t)\}$ costs more to the principal, as we observed in the above Section 5.1.

If we plug $a^*(t) = a^*$ for all t into equation (46), we get the following PW^*_{Add}:

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

For the post-contract ex-post system, let us also first fix the action a^o induced by the principal, and see how the principal minimizes the agency cost. Principal solves the following optimization.

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

where $AC^o_{Add}(a^o)$ is an agency cost associated with inducing the action a^o for $\forall t$ in the post-contract ex-post system. It turns out that we can express the agency cost and principal's welfare in a closed-form as

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

We can see from equation (48) and equation (50) that $PW^{o}_{Add} = PW^{*}_{Add}$ holds.[19] For the 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.

First, we show that the principal prefers the post-contract ex-post information system in which perfect information on random variables that affect the agent's marginal utility is publicly revealed after the agent has made his action choice to the post-contract ex-ante information system in which the same information is publicly revealed before the agent makes 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. Second, 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 with their observing less precise ex-ante information.

Before closing, two remarks need to be in order. First, information systems considered in this paper are assumed to reveal information not only to principal but also to agent. In contrast, if the principal acquires such information privately[20], then the principal's information revealing strat-

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

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

[20] As we mentioned before, Silvers (2012) considers the environment where the principal gets a private or public

egy 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, principal and agent's observing information on those variables ex-ante has no additional information value comparing with their observing it ex-post in the first-best situation where 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 (i.e., Silvers (2012)) or his disutility (i.e., cost) function (i.e., 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 (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, ex-ante revelation of relevant information about the technology might yield the better matching between principal and 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 agency relation.

signal about the technology. In this case principal might announce the signal to the agent not truthfully and optimal contract is distorted to take into account this information revealing strategy.

References

Rajiv D. Banker, 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*, 57(5):1201–1245, 2019.

Ralph Boleslavsky and Teddy Kim. Bayesian persuasion and moral hazard. *Working Paper*, 2020.

John R. Conlon. Two new conditions supporting the first-order approach to multisignal principal-agent problems. *Econometrica*, 77(1):249–278, 2009.

Frøystein Gjesdal. Information and incentives: The agency information problem. *The Review of Economic Studies*, 49(3):373–390, 1982.

Sanford J. Grossman and Oliver Hart. An analysis of the principal-agent problem. *Econometrica*, 51: 7–45, 1983.

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

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

Bengt Holmström. Moral hazard in teams. *The Bell Journal of Economics*, 13:324–340, 1982.

Ian Jewitt. Justifying the first-order approach to principal-agent problems. *Econometrica*, 56(5): 1177–1190, 1988.

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

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

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

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

Son Ku Kim and Yoon Suh. Ranking of accounting information systems for management control. *Journal of Accounting Research*, 29(2):386–396, 1991.

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

James A. Mirrlees. 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.

William Rogerson. The first order approach to principal-agent problem. *Econometrica*, 53:1357–1368, 1985.

Steve Ross. The economic theory of agency: The principal's problem. *The American Economic Review*, 63:134–139, 1973.

Michael Rothsscild and Joseph E. Stiglitz. Increasing risk i: A definition. *Journal of Economic Theory*, 2(3):225–243, 1970.

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

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

Bernanrd Sinclair-Desgagné. The first-order approach to multi-signal principal-agent problems. *Econometrica*, 62(2):459–465, 1994.

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

Appendix. Derivations and 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, consider the case in which 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 $a^*(t)$.
By using the first-order approach, the principal's optimization program in this case is :

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

Let $s^m(x,t)$ be the optimal contract for the above program. Then, similar to equation (8), 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. Here λ^m denotes 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 in that case.
Based on equation (I.4), we transform the optimal wage contract, $s^m(x,t)$, that 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 follows.

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

It is worth to note the functional form of $r(\cdot)$ is governed solely by $u_s(\cdot)$ and does not depend on information systems. And, we know that $r(\cdot)$ 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

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

Now consider the case in which the principal has chosen the post-contract ex-ante information system. Then, based on equation (17), we can rewrite

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

Actually,

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

Again, for notational simplicity, we will drop (x,t) from $z^*(x,t)$ and denote $z^* \equiv z^*(x,t)$. Then, based on equation (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 $r(z^*,t)$, principal's optimized benefits under the post-contract ex-ante information system can be rewritten as:

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

Since

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

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

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

Let us define the following function, which was extensively used in Kim (1995):

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

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

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

Thus, it is easy to see that $\psi(z,t)$ is globally concave in z.
Since from equation (I.15)

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

we can rewrite equation (I.14) as

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

Due to the complementary slackness of participation constraints in both programs, we have:

$$\begin{aligned} \lambda^* \Big(\int_t \int_x u(r(z^*,t),t) f(x|a^*(t))h(t)dxdt \Big) &= \lambda^* \Big(\overline{U} + \int_t v(a^*(t))h(t)dt \Big) \\ \int_t \int_x u(r(z^*,t),t) f(x|a^*(t))h(t)dxdt &\ge \overline{U} + \int_t v(a^*(t))h(t)dt, \end{aligned} \tag{I.21}$$

and

$$\begin{aligned} \lambda^m \Big(\int_t \int_x u(r(z^m,t),t) f(x|a^m)h(t)dxdt \Big) &= \lambda^m \Big(\overline{U} + v(a^m) \Big) \\ \int_t \int_x u(r(z^m,t),t) f(x|a^m)h(t)dxdt &\ge \overline{U} + v(a^m). \end{aligned} \tag{I.22}$$

Also, incentive constraints for each programs are respectively:

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

By using equation (17), equation (I.21), and equation (I.23), we have

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

Also, by using equation (I.4), equation (I.22), and equation (I.23), we have

$$\begin{aligned} &\int_t \int_x u(r(z^m,t),t)z^m f(x|a^m)h(t)dxdt = \int_t \int_x u(r(z^m,t),t)[\lambda^m + \mu^m \frac{f_a}{f}(x|a^m)] f(x|a^m)h(t)dxdt \\ &\quad = \lambda^m \int_t \int_x u(r(z^m,t),t) f(x|a^m)h(t)dxdt + \mu^m \int_t \int_x u(r(z^m,t),t) f_a(x|a^m)h(t)dxdt \\ &\quad = \lambda^m [\overline{U} + v(a^m)] + \mu^m v'(a^m). \end{aligned} \tag{I.25}$$

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

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

Now, define (similarly to Kim (1995))

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

Since $\psi(z,t)$ is globally concave in z,[1] we have

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

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

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

Since we assume $x = a + \theta$, where θ is random variable with the probability density $g(\cdot)$, we can write $f(x|a) = g(x-a)$. then $f_a(x|a) = -g'(x-a)$ and $\frac{f_a}{f}(x|a) = -\frac{g'(x-a)}{g(x-a)} = -\frac{g'(\theta)}{g(\theta)}$. Thus, when $x \sim f(x|a)$,

[1] See Kim (1995) for this issue in a more canonical principal-agent framework.

$\frac{f_a}{f}(x|a)$'s distribution does not depend on the value of a.[2] Accordingly, for any given t,

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

Thus, from equation (I.29) and equation (I.30), we finally 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 action), we have

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

Consequently, we derive

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

Proof of Lemma 3: Since the agent's limited liability constraint is not binding for any (x, t), we have $s^o(x,t) > 0$. It implies 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 <u>with respect to x</u> on the 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 equation (I.35), we can rewrite equation (I.36) as

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

Since $\frac{u_{ss}}{u_{sss}} = \frac{u_{st}}{u_{sst}}$ and since $u_{ss} < 0$ (by Assumption 1), 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 equation (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 $u_s > 0$ (by Assumption 1) 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.,

[2] $f(x|a + \triangle a)$ is a shift of $f(x|a)$ by $\triangle a$. Therefore, $f(x^0|a^*(t)) = f(x^1|a^m)$ and $f_a(x^0|a^*(t)) = f_a(x^1|a^m)$ for any x^1 and x^0 satisfying $x^1 - a^m = x^0 - a^*(t)$, and we have $\frac{f_a}{f}(x^0|a^*(t)) = \frac{f_a}{f}(x^1|a^m)$ for any x^1 and x^0 satisfying $x^1 - a^m = x^0 - a^*(t)$.

$\frac{\partial}{\partial t}u(s^o(x,t),t)$ 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, we know that, 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.

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

Usually, the only difference that the optimization program with the post-contract ex-post information system has from that with the post-contract ex-ante information system is that 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 in the optimization program with the post-contract ex-ante information system. In this case agent does not have an incentive to deviate from a^o even when he knows the value of t. Trivially, we can directly have[3]

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

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 easily 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

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

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

Since $\frac{u_t}{u_s} = \frac{u_{st}}{u_{ss}}$, we have the left-hand side of equation (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 we have $\frac{u_{st}}{u_{ss}} = \frac{u_{sst}}{u_{sss}}$ from equation (I.44), by Proposition 2, we have

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

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

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

Accordingly, $(s^o(x,t),a^o)$ satisfies the constraints of following optimization program.

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

The above program differs from the maximization program under a 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 equation (31) under information system N^+, let

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

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

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

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

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

Now, let us go back to the system N and consider $(a^m(T_i), s_m^N(x,t))$ such that:

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

where $p_j^- \equiv \int_{t\in T_j^-} h(t)dt$ and $p_j^+ \equiv \int_{t\in T_j^+} h(t)dt$ and $s_m^N(x,t)$ for $t \in T_i$, $i = 1, 2, ..., j, ..., N$ is defined as solving the following optimization program under information system N.[4]

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

where

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

[4] Thus the contract $s_m^N(x,t)$ induces the action $\{a^m(T_i)\}$ from the agent given $t \in T_i$.

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

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

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.*

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

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 easily see that $(a^m(T_i), s_m^N(x,t))$ satisfies all the constraints in equation (31) under information system N given actions $\{a^m(T_i)\}$ such as

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

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

and

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

This implies that $(a^m(T_i), s_m^N(x,t))$ is actually in the feasible set of the optimization program under information system N in equation (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: Given that $(a^{N^+}(T_i), s^{N^+}(x,t))$ is the solution for equation (31), we know that $s^{N^+}(x,t)$ solves the optimization in equation (I.54) for $i = j^-, j^+$. Let us define some variables 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[5]

$$a^m(T_j) = \frac{p_j^- a^{N^+}(T_j^-) + p_j^+ a^{N^+}(T_j^+)}{p_j^- + p_j^+} = \int_{t \in T_j} a_1(t) h(t|T_j) dt. \tag{I.71}$$

Then given $\{a_1(t)\}$, $s^{N^+}(x,t), t \in T_j = T_j^- \cup T_j^+$ solves the following optimization program [**P1**].

$$\begin{aligned}
&[\mathbf{P1}] \max_{s(x,t) \in S} \int_{t \in T_j} \int_x [x - s(x,t)] f(x|a_1(t)) h(t|T_j) dxdt \quad \text{s.t.} \\
&\quad (i) \;\; \int_{t \in T_j} \Big(\int_x u(s(x,t),t) f(x|a_1(t)) dx - v(a_1(t)) \Big) h(t|T_j) dt \geq K^{m,N}(T_j) = \frac{p_j^- K^{N^+}(T_j^-) + p_j^+ K^{N^+}(T_j^+)}{p_j^- + p_j^+}, \\
&\quad (ii) \;\; \begin{cases} \displaystyle\int_{t \in T_j^-} \int_x u(s(x,t),t) f_a(x|a^{N^+}(T_j^-)) dx h(t|T_j^-) dt = v'(a^{N^+}(T_j^-)), \\ \displaystyle\int_{t \in T_j^+} \int_x u(s(x,t),t) f_a(x|a^{N^+}(T_j^+)) dx h(t|T_j^+) dt = v'(a^{N^+}(T_j^+)), \end{cases} \\
&\quad (iii) \;\; s(x,t) \geq 0, \;\; \forall (x,t) \in X \times T_i.
\end{aligned} \tag{I.72}$$

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

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

---

[5] Since

$$\begin{aligned}
&\int_{t \in T_j} \int_x [x - s(x,t)] f(x|a_1(t)) h(t|T_j) dxdt \\
&= \frac{p_j^-}{p_j^- + p_j^+} \underbrace{\int_{t \in T_j^-} \int_x [x - s(x,t)] f(x|a^{N^+}(T_j^-)) h(t|T_j^-) dxdt}_{\equiv PW^{N^+}(T_j^-) \text{ if } s(\cdot) = s^{N^+}(\cdot)} + \frac{p_j^+}{p_j^- + p_j^+} \underbrace{\int_{t \in T_j^+} \int_x [x - s(x,t)] f(x|a^{N^+}(T_j^+)) h(t|T_j^+) dxdt}_{\equiv PW^{N^+}(T_j^+) \text{ if } s(\cdot) = s^{N^+}(\cdot)},
\end{aligned} \tag{I.69}$$

and

$$\begin{aligned}
&\int_{t \in T_j} \int_x u(s(x,t),t) f(x|a_1(t)) h(t|T_j) dxdt \\
&= \frac{p_j^-}{p_j^- + p_j^+} \int_{t \in T_j^-} \int_x u(s(x,t),t) f(x|a^{N^+}(T_j^-)) h(t|T_j^-) dxdt + \frac{p_j^+}{p_j^- + p_j^+} \int_{t \in T_j^+} \int_x u(s(x,t),t) f(x|a^{N^+}(T_j^+)) h(t|T_j^+) dxdt,
\end{aligned} \tag{I.70}$$

we know $s^{N^+}(x,t)$ solves equation (31) given $a^{N^+}(T_i), i = 1, 2, ..., j^-, j^+, ..., N$, and it satisfies the constraints of equation (I.72) for $t \in T_j = T_j^- \cup T_j^+$. Thus it solves **[P1]** as we know $s^{N^+}(x,t)$ for $t \in T_i (i \neq j^-, j^+)$ solves equation (I.54).

Finally we consider the optimization [**P2**] with $a^m(T_j) = \int_{t\in T_j} a_1(t)h(t|T_j)dt$, which is the same as equation (I.57).

$$\begin{aligned}
&[\mathbf{P2}] \max_{s(x,t)\in S} \int_{t\in T_j}\int_x [x - s(x,t)] f(x|a^m(T_j))h(t|T_j)dxdt \quad \text{s.t.}\\
&\quad (i) \ \int_{t\in T_j}\int_x u(s(x,t),t)f(x|a^m(T_j))dxh(t|T_j)dt - v(a^m(T_j)) \geq K^{m,N}(T_j),\\
&\quad (ii) \ \int_{t\in T_j}\int_x u(s(x,t),t)f_a(x|a^m(T_j))dxh(t|T_j)dt = v'(a^m(T_j)),\\
&\quad (iii) \ s(x,t)\geq 0, \ \forall (x,t)\in X\times T_i.
\end{aligned} \tag{I.74}$$

Due to the Lemma 2[6,7], $PW([\mathbf{P1}]) \leq PW([\mathbf{P2}])$ holds. As $PW([\mathbf{P2}]) = PW^m(T_j)$, we finally derive the claim that $PW([\mathbf{P2}]) = PW^m(T_j) \geq PW([\mathbf{P1}])$.Thus we get

$$PW^m(T_j) \geq \frac{p_j^- PW^{N^+}(T_j^-) + p_j^+ PW^{N^+}(T_j^+)}{p_j^- + p_j^+}. \tag{I.76}$$

Derivation of Section 5.1 : Multiplicative Utility

First, we get the closed-form solution of $AC_M^*(\{a(t)\}$ and $AC_M^o(a^o)$ and prove the Lemma 5. Let λ^* and $\mu^*(t)h(t)$ be the Lagrange multiplier to the constraints (i) and (ii) respectively, of the following equation (I.77).

$$\begin{aligned}
AC_M^*(\{a^*(t)\}) \equiv [\text{EA}]_M &\min_{s(x,t)\in S} \int_t\int_x s(x,t)f(x|a^*(t))h(t)dxdt \quad \text{s.t.}\\
&(i) \ \int_t \Big[\int_x 2t\sqrt{s(x,t)}f(x|a^*(t))dx - v(a^*(t))\Big]h(t)dt \geq \overline{U},\\
&(ii) \ \int_x 2t\sqrt{s(x,t)}f_a(x|a^*(t))dx - v'(a^*(t)) = 0, \ \forall t\in T,\\
&(iii) \ s(x,t)\geq 0, \ \forall (x,t)\in X\times T.
\end{aligned} \tag{I.77}$$

Solving Euler equation for equation (I.77), we get the solution $s^*(x,t)$ must be represented as follows,

---

[6] Note that the only difference between [**P1**] and [**P2**] is that [**P2**] induces the average action $a^m(T_j)$ across t's partitions (T_j^- and T_j^+) whatever $t\in T_j$ is realized, and Lemma 2 proves it gives a higher efficiency from principal's standpoint.

[7] Actually equation (I.24) in this case become

$$\begin{aligned}
&\int_{t\in T_j}\int_x u(r(z^*,t),t)z^* f(x|a^*(t))h(t|T_j)dxdt\\
&\quad = \int_{t\in T_j}\int_x u(r(z^*,t),t)\Big(\lambda^* + \mu^*(t)\frac{f_a}{f}(x|a^*(t))\Big)f(x|a^*(t))h(t|T_j)dxdt\\
&\quad = \lambda^*\int_{t\in T_j}\int_x u(r(z^*,t),t)f(x|a^*(t))h(t|T_j)dxdt + \int_{t\in T_j}\mu^*(t)\int_x u(r(z^*,t),t)f_a(x|a^*(t))dxh(t|T_j)dt\\
&\quad = \lambda^*\Big[K^{m,N}(T_j) + \frac{p_j^- v(a^{N^+}(T_j^-)) + p_j^+ v(a^{N^+}(T_j^+))}{p_j^- + p_j^+}\Big] + \frac{p_j^-\mu^*(T_j^-)v'(a^{N^+}(T_j^-)) + p_j^+\mu^*(T_j^+)v(a^{N^+}(T_j^+))}{p_j^- + p_j^+}.
\end{aligned} \tag{I.75}$$

since we assume away the limited liability constraint (iii).

$$\frac{1}{u_s(s^*(x,t))} = \frac{\sqrt{s^*(x,t)}}{t} = \lambda^* + \mu^*(t)\frac{f_a}{f}(x|a^*(t)),\;\; s(x,t) \geq 0. \tag{I.78}$$

Thus we easily obtain

$$\begin{aligned} s^*(x,t) &= t^2\Big(\lambda^* + \mu^*(t)\frac{f_a}{f}(x|a^*(t))\Big)^2 \\ &= t^2\Big((\lambda^*)^2 + 2\lambda^*\mu^*(t)\frac{f_a}{f}(x|a^*(t)) + \mu^*(t)^2\Big(\frac{f_a}{f}\Big)^2(x|a^*(t))\Big), \\ u(s^*(x,t)) &= 2t\sqrt{s^*(x,t)} = 2t^2\Big(\lambda^* + \mu^*(t)\frac{f_a}{f}(x|a^*(t))\Big) \geq 0. \end{aligned} \tag{I.79}$$

We plug the above $u(s^*(x,t))$ into the constraint (i) and (ii) and write λ^* and $\mu^*(t)$ as

$$\mu^*(t) = \frac{v'(a^*(t))}{2t^2}\frac{1}{\mathrm{Var}\Big(\frac{f_a}{f}\Big)},\;\; \lambda^* = \frac{\int_t v(a^*(t))h(t)dt + \overline{U}}{2\int_t t^2h(t)dt}. \tag{I.80}$$

Finally we can write the explicit formula for the value of the objective function in equation (I.77) as

$$\begin{aligned} AC_M^*(\{a(t)\}) &= \int_t\Big(\int_x s^*(x,t)f(x|a^*(t))dx\Big)h(t)dt = (\lambda^*)^2\Big(\int_t t^2h(t)dt\Big) + \mathrm{Var}\Big(\frac{f_a}{f}\Big)\Big(\int_t t^2\mu^*(t)^2h(t)dt\Big) \\ &= \underbrace{\frac{\Big(\int_t v(a^*(t))h(t)dt + \overline{U}\Big)^2}{4\int_t t^2h(t)dt}}_{\equiv AC^*_{M,\mathrm{IR}}(\{a^*(t)\})} + \underbrace{\frac{1}{4\mathrm{Var}\Big(\frac{f_a}{f}\Big)}\int_t \frac{1}{t^2}v'(a^*(t))^2h(t)dt}_{\equiv AC^*_{M,\mathrm{IC}}(\{a^*(t)\})}. \end{aligned} \tag{I.81}$$

For the post-contract ex-post system, let λ^o and μ^o be the multipliers to the constraints (i) and (ii) respectively of the following optimization in equation (I.82):

$$\begin{aligned} AC_M^o(a^o) \equiv [\mathrm{EP}]_M \min_{s(x,t)\in S} &\int_t\int_x s(x,t)f(x|a^o)h(t)dxdt \quad \text{s.t.} \\ &(i)\;\; \int_t\int_x 2t\sqrt{s(x,t)}f(x|a^o)dxh(t)dt - v(a^o) \geq \overline{U}, \\ &(ii)\;\; \int_t\int_x 2t\sqrt{s(x,t)}f_a(x|a^o)dxh(t)dt - v'(a^o) = 0, \\ &(iii)\;\; s(x,t) \geq 0,\;\; \forall(x,t) \in X \times T. \end{aligned} \tag{I.82}$$

Solving Euler equation for equation (I.82), the solution $s^o(x,t)$ must be represented as follows, since we assume away the limited liability constraint (iii):

$$\frac{1}{u_s(s^o(x,t))} = \frac{\sqrt{s^o(x,t)}}{t} = \lambda^o + \mu^o \frac{f_a}{f}(x|a^o), \tag{I.83}$$

where the solution $(s^o(x,t))$ must be as follows, since we assume away the limited liability constraint (iii):

$$\begin{aligned} s^o(x,t) &= t^2\Big(\lambda^o + \mu^o \frac{f_a}{f}(x|a^o)\Big)^2 \\ &= t^2\Big((\lambda^o)^2 + 2\lambda^o\mu^o \frac{f_a}{f}(x|a^o) + (\mu^o)^2\Big(\frac{f_a}{f}\Big)^2(x|a^o)\Big), \\ u(s^o(x,t)) &= 2t^2\Big(\lambda^o + \mu^o \frac{f_a}{f}(x|a^o)\Big). \end{aligned} \tag{I.84}$$

Plugging the above $u(s^o(x,t))$ into the constraint (i) and (ii), we obtain λ^o and μ^o as

$$\mu^o = \frac{v'(a^o)}{2\int_t t^2 h(t)dt} \cdot \frac{1}{\mathrm{Var}\Big(\frac{f_a}{f}\Big)}, \;\; \lambda^o = \frac{v(a^o) + \overline{U}}{2\int_t t^2 h(t)dt}. \tag{I.85}$$

Finally we can get the value of the objective function in equation (I.82).

$$\begin{aligned} AC^o_M(a^o) &= \int_t \Big(\int_x s^o(x,t) f(x|a^o)dx\Big)h(t)dt = (\lambda^o)^2\Big(\int_t t^2 h(t)dt\Big) + \mathrm{Var}\Big(\frac{f_a}{f}\Big)(\mu^o)^2\Big(\int_t t^2 h(t)dt\Big) \\ &= \underbrace{\frac{(v(a^o)+\overline{U})^2}{4\int_t t^2 h(t)dt}}_{\equiv AC^o_{M,\mathrm{IR}}(a^o)} + \underbrace{\frac{v'(a^o)^2}{4\mathrm{Var}\Big(\frac{f_a}{f}\Big)} \frac{1}{\int_t t^2 h(t)dt}}_{\equiv AC^o_{M,\mathrm{IC}}(a^o)}. \end{aligned} \tag{I.86}$$

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

$$AC^o_{M,\mathrm{IR}}(a^m) = \frac{(v(a^m)+\overline{U})^2}{4\int_t t^2 h(t)dt} < \frac{(\int_t v(a^*(t))h(t)dt + \overline{U})^2}{4\int_t t^2 h(t)dt} = AC^*_{M,\mathrm{IR}}(\{a(t)\}). \tag{I.87}$$

To show $AC^o_{M,\mathrm{IC}}(a^m) < AC^*_{M,\mathrm{IC}}(\{a(t)\})$, it is sufficient to prove

$$\int_t \frac{1}{t^2} v'(a^*(t))^2 h(t)dt > \frac{v'(a^m)^2}{\int_t t^2 h(t)dt}. \tag{I.88}$$

Let us 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]

$$\mathbb{E}(X^2)\mathbb{E}(Y^2) = \Big(\int_t \frac{1}{t^2} v'(a^*(t))^2 h(t)dt\Big)\Big(\int_t t^2 h(t)dt\Big) > \mathbb{E}(XY)^2 = \Big(\int_t v'(a^*(t))h(t)dt\Big)^2 \tag{I.89}$$
$$> v'(a^m)^2.$$

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