Two-Dimensional Comparison of Information Systems in Principal-Agent Models *

Jia Xie†

Initial version: June 06 2008

This version: January 29 2009

Abstract

This paper extends the comparison of information systems in contract theory to a general context with an unobservable outcome and a risk-averse principal. The model examines information systems on two dimensions: (1) incentive-providing and (2) risk-sharing. We first identify the most efficient information systems. Then we develop an extended informativeness criterion for comparison of inclusive information systems, and a MPS-SOSD criterion and a MPS-FOSD criterion for comparison of scalar information systems.

Keywords: Principal-agent model, moral hazard, information systems.

JEL Numbers: D8.

*The author is indebted to Professor Michael Rauh, Professor James Walker, and Professor Frank Page for pertinent comments and suggestions. Needless to say, Any mistakes are mine.

†Contact Information: Department of Economics, Indiana University, Bloomington, IN 47405; email: jiaxie@indiana.edu.

1 Introduction

The standard principal-agent model under moral hazard has the following structure: A principal (she) offers a contract to an agent (he) to induce him to carry out an action on behalf of the principal. If the contract is accepted, the agent then chooses an action which is unobservable to the principal and which stochastically generates an observable and verifiable information system: a monetary outcome and a monitoring signal. The monetary outcome is then shared between the principal and the agent as specified by the contract, which also denotes the end of the contract relation.

The standard model assumes that the decision process on information systems is outside the scope of the model. That is, the selection and modification of the information system is treated as outside the principal's control. Often, however, it is possible for the principal to further improve the optimal contracting if we endogenizing these controls. The literature diverges in two main approaches. The first approach assumes that the principal can improve the accuracy of the information system by devoting resources to that objective (monitoring). The optimal contracting should balance the marginal benefit and the marginal cost of monitoring. Since this approach is not the focus of this paper, we avoid the discussion here. Interesting readers can refer to Chapter 7 in Milgrom and Roberts (1995) and Demougin and Fluet (2001).

This paper belongs to the literature on the second approach. The second approach assumes that instead of just one information system, the principal has access to multiple information systems. Before offering a contract to the agent, however, the principal has to decide which information system to be written in the contract. A natural question raised within this approach is how the principal compares information systems with each other so as to choose the one that maximizes her own welfare.

This question was first raised and analyzed by Holstrom (1979) in Section 5 of that article. Holmstrom (1979, 1981) discovered the so-called informativeness criterion for comparison of inclusive information systems[1]: the information system with the additional signals is strictly preferred by the principal if and only if the additional signals are marginally informative about the agent's action. Holmstrom proved that the informativeness criterion is the necessary and sufficient condition for comparison of inclusive information systems, and suggested that informativeness should be related to the variability of likelihood ratios.

Kim (1995) formalized this intuition by proving that the Mean Preserving Spread (MPS) criterion is sufficient for comparing arbitrary information systems. The criterion states that for two information systems X and Y, X is more efficient than Y if the likelihood ratio distribution of X is a MPS of that of Y. Kim (1995) compared his MPS creation to Holmstrom's informativeness criterion and Blackwell's statistical sufficient criteria developed in the decision-theoretic context.[2] He found that Holmstrom's criterion is equivalent to the MPS criterion when applied to comparison of inclusive information systems, and that Blackwell's criteria is not only unnecessary for ranking information systems in the agency model, but it is also excessively restrictive as a sufficient condition. Moreover, Kim emphasized that there is an essential difference between a decision-theoretic problem and an agency problem. In the former the decision maker attempts to estimate some unverifiable variable, while in the later the principal attempts to control the agent's hidden action. Jewitt (1997 & 2007) proved that the MPS criterion is not only sufficient but also a necessary condition.

Based on Kim's work, in an agency model with both parties being risk neutral, Demougin and Fluet (1998) defined the notion of "mechanism sufficiency" and distinguished it from the concept of "statistical sufficiency" in Blackwell's theorem: "a signal is said to be statistically sufficient for an other random variable if no informational content is lost", while a signal is "to be 'mechanism sufficient' for an other random variable if none of the information relevant to the principal is lost." Later, Demougin and Fluet (2001) re-examined the difference between the two sufficiency concepts in the standard agency model with risk averse agent. More

[1]Two information systems are inclusive if one has more sources of signals than the other so that the later is a proper subset of the other.

[2]See Blackwell (1951, 1953).

importantly, they presented the integral criterion, which states that an information system is more efficient than another if its distribution function is more sensitive to the agents effort, and proved that the integral criterion is equivalent to Kim's MPS criterion.

One common characteristic of the above models is that either the principal is risk-neutral or the monetary outcome is observable and verifiable, so that the information system performs the incentive-providing function only, and thus all the existing criteria compare information systems one-dimensionally on their efficiency in incentive-providing. Indeed, Kim (1995) and Demougin and Fluet (1998, 2001) analyzed the standard agency model with a risk-neutral principal, whose expected utility is not affected by risk; in general, the information system providing more incentives at lower cost is more efficient. On the other hand, Holmstrom (1979, 1981) assumed that the monetary outcome is observable and verifiable. Because the monetary outcome, itself as an information system, is most efficient in sharing the risk inherent in the monetary outcome, any additional information system is valuable if and only if it is marginally more efficient in incentive-providing than the monetary outcome.

Thus the current literature leaves uncovered the case where both parties are risk-averse and the outcome is unobservable (at least at the time when the contract relation ends) so that information systems actually perform two functions: rewarding the agent for productive work and sharing between the parties the uncertainty inherent in the stochastic outcome. Real life situations often seem to be in this ignored case: often it is the quality outcome that the principal cares about, and the quality outcome is likely to be unobservable. For example, an risk-averse individual (the principal) drives her car to a service station (the agent) for repair, and she cares about the quality of the service that she receives. Because she has to pay the station right after the work is done, the quality is not perfectly observable at the time when she has to pay.

This paper intends to fill this gap. In this paper, we extend the literature to a general principal-agent model where both parties are risk-averse and where the final outcome is unobservable. We examine information systems on two dimensions: (1) incentive-providing and (2) risk-sharing. We first identify the most efficient information systems. Then we develop an extended informativeness criterion for comparison of inclusive information systems, and a MPS-SOSD criterion and a MPS-FOSD criterion for comparison of scalar information systems.

The rest of this paper is organized as follows. In section 2, we describe the basic framework of the model. We present the main results in Section 3: the most efficient information systems in Subsection 3.1; the extended informativeness criterion in Subsection 3.2; and the MPS-SOSD criterion and the MPS-FOSD criterion in Subsection 3.3. Section 4 concludes.

2 The model

A risk-averse principal faces a set of information systems, Γ. Once the principal chooses an information system, $X \in \Gamma$, she makes a take-it-or-leave-it offer, $s(x)$, to a single risk-averse agent with outside reservation utility of 0. If the contract is accepted, the agent then chooses a hidden action (or effort), $a \in A$, which has a stochastic effect on the information system X, and stochastically generates a outcome, which the principal cares about and which has a monetary representation, B.[3] We refer to B as the monetary outcome henceforce. After observing the realization of the information system, the principal pays the agent as specified by the contract and the contract relation ends. The time line of events are displayed in Figure 1.

We use the terms "information system" and "signal" interchangeably in this paper. The notion of information system can be made precise as follows.

Definition 1. *An information system, X, is an observable and verifiable random variable (or vector) whose distribution is affected by the agent's action.*

[3] B can be interpreted as the principal's willingness-to-pay for the outcome, if the outcome itself, such as the quality outcome, is not monetary.

This term "verifiable" is used to indicate that enforceable contracts can be written on the information systems.

On the other hand, we assume that the monetary outcome is unobservable at the time when the agent has to be paid, thus $B \notin \Gamma$ and no enforceable contract can be written on the monetary outcome.

The joint conditional probability distribution of the outcome and the information system, given agent's action is common knowledge and is denoted by $G_X(b, x|a)$ with $b \in \mathbb{B}$, $x \in \mathbb{X}$, and $a \in \mathbb{A}$, where $\mathbb{B}$ and $\mathbb{X}$ are the sample spaces of B and X respectively. The corresponding conditional density function is $g_X(b, x|a)$. The conditional marginal probability distribution function of the information system X, given the agent's action is $F_X(x|a)$, and the corresponding marginal density function is $f_X(x|a)$. The principal's utility function $U(w)$ is defined over wealth only, while the agent's utility function $H(w, a)$ is defined over wealth and action.

We make the following standard assumptions to simplify out model.

Assumption 1.

$$U' > 0, \quad U'' < 0; \quad H(w,a) = u(w) - c(a), \quad \textit{with} \quad u' > 0, \quad u'' < 0, \quad c' > 0, \quad \textit{and} \quad c'' > 0.$$

Assumption 2. *For any $f_X(x|a)$, the support $\mathbb{X}$ is independent of the agent's action choice.*

Assumption 3. *The "first order approach" to the principal's problem is valid.*[4]

The principal solves for $s(x)$ and a in the following program *P1*

$$\textit{P1:} \qquad \mathscr{U}^*(X) = \max_{s(x) \geq k, a} \int_{\mathbb{X}} \int_{\mathbb{B}} U(b - s(x)) g(b, x; a) db\, dx \tag{1}$$

$$s.t. \quad \int_{\mathbb{X}} u(s(x)) f(x; a) dx - c(a) \geq 0, \tag{2}$$

$$\int_{\mathbb{X}} u(s(x)) f_a(x; \hat{a}) dx - c_a(a) = 0. \tag{3}$$

[4]This is a standard assumption in the literature. See Holmstrom (1979), Kim (1995), and Demougin and Fluet (2001).

Here, k is the lower bound of the agent's compensation, which is introduced to avoid a possible nonexistence problem.[5] In the program *P1*, Constraint (2) is the Individual Rationality (IR) constraint, saying that the agent's expected utility must be no less than his outside reserve utility. Constraint (3) is the relaxed Incentive Compatibility (IC) constraint, saying that given the payment schedule $s(x)$, a maximizes the agent's expected utility. *P1* is referred to as *the relaxed program* (see Rogerson (1985)), because the original IC constraint is relaxed by the first-order necessary condition. Let λ^* be the multiplier for Constraint (2) and μ^* the multiplier for Constraint (3), then pointwise optimization of the Lagrangian with respect to $s(x)$ yields the following characterization of an optimal compensation schedule, which is referred to as the *Borch rule*, named after Borch (1962) (see also Bolton and Dewatripont (2005), Holmstrom (1979) and Kim (1995)).

$$\frac{\int U'(b - s(x))g_X(b|x,a)db}{u'(s(x))} = \lambda + \mu \frac{f_{X_a}}{f_X}(x|a). \tag{4}$$

We denote the solution to *P1* by $\mathscr{S}^*(X) \equiv \big(a^*_X, s^*_X(x), \mathscr{U}^*(X)\big)$.

3 The Main Results

Before we proceed with any criterion that compares information systems in terms of their relative efficiencies, the notion of efficiency should be made precise.

Definition 2. *Given two information systems, X and Y. X is more efficient than Y if $\mathscr{U}^*(X) \geq \mathscr{U}^*(Y)$ in any agency problem P1 satisfying Assumption 1. X is as efficient as Y if $\mathscr{U}^*(X) = \mathscr{U}^*(Y)$. X is strictly more efficient than Y if $\mathscr{U}^*(X) > \mathscr{U}^*(Y)$.*

3.1 The Most Efficient Information Systems

Proposition 3.1. *If $X = (a, B)$, then $\mathscr{U}^*(X) \geq \mathscr{U}^*(Y)$, $\forall Y \in \Gamma$.*

[5] See Mirlees (1974) and Kim (1995).

Proof. It suffices for us to prove that $\mathscr{U}^*(X) \geq \mathscr{U}^*((a,Y))$, $\forall Y \in \Gamma$, as it is trivial that $\mathscr{U}^*((a,Y)) \geq \mathscr{U}^*(Y)$.

Note that with information system $X = (a, B)$, the agent's action, a, is observable and verifiable, thus the principal solves the following program with only the IR constraint:

$$\mathscr{U}^*\big((a,B)\big) = \max_{s(a,b)\geq k,a} \int U(b - s(a,b))g(b;a)db \tag{5}$$

$$s.t. \quad \int u(s(a,b))f(b;a)db - c(a) \geq 0. \tag{6}$$

Similarly, with information system (a, Y), the principal solves the following problem

$$\mathscr{U}^*\big((a,Y)\big) = \max_{s(a,y)\geq k,a} \iint U(b - s(a,y))g(b,y;a)db\,dy \tag{7}$$

$$s.t. \quad \int u(s(a,y))f(y;a)dy - c(a) \geq 0. \tag{8}$$

Denote the solution to the program (7)-(8) by $\big(a_Y^*, s_Y^*(a,y)\big)$. Then we construct a compensation schedule $s(a_Y^*, b)$ such that

$$U(b - s(a_Y^*, b)) = \int U(b - s_Y^*(a_Y^*, y))g(y|b, a_Y^*)dy, \quad \forall b \in \mathbb{B}. \tag{9}$$

Then by Jensen's inequality, we have

$$b - s(a_Y^*, b) \leq \int \big[b - s_Y^*(a_Y^*, y)\big]g(y|b, a_Y^*)dy,$$

$$s(a_Y^*, b) \geq \int s_Y^*(a_Y^*, y)g(y|b, a_Y^*)dy,$$

$$u(s(a_Y^*, b)) \geq u(\int s_Y^*(a_Y^*, y)g(y|b, a_Y^*)dy) \geq \int u(s_Y^*(a_Y^*, y))g(y|b, a_Y^*)dy,$$

$$\int u(s(a_Y^*, b))g(b|a_Y^*)db - c(a_Y^*) \geq \int u(s_Y^*(a_Y^*, y))g(y|a_Y^*)dy - c(a_Y^*) \geq 0.$$

Thus there must exists a constant $\epsilon \geq 0$ such that $\int u(s(a_Y^*, b) - \epsilon)g(b|a_Y^*)db - c(a_Y^*) \geq 0$. Thus $\big(a_Y^*, s(a_Y^*, b) - \epsilon\big)$ satisfies the IP constraint in, and is a potential solution to, the program (5)-(6). Thus we have

$$\begin{aligned}\mathscr{U}^*(X) &\geq \int U(b - s(a_Y^*, b) + \epsilon)g(b|a_Y^*)db \geq \int U(b - s(a_Y^*, b))g(b|a_Y^*)db \\ &= \iint U(b - s_Y^*(a_Y^*, y))g(b, y|a_Y^*)dy\, db = \mathscr{U}^*\big((a, Y)\big),\end{aligned}$$

where the first inequality follows from the fact that $\mathscr{U}^*(X)$ is the principal's maximum expected utility in program (5)-(6), and the first equality follows from (9), and the last equality follows from the fact that $\big(a_Y^*, s(a_Y^*, y)\big)$ is the solution to the program (7)-(8). □

The proposition says that as long as the agent's action and the monetary outcome are observable and verifiable, there is no need to acquire any other information systems. No information system can generate a higher expected utility for the principal than the agent's action and the monetary outcome do. In this sense, the agent's action and the monetary outcome, together as an information system, dominate any other information systems.

This result should not be surprising. As we have mentioned in the introduction, in our model an information system performances two functions: giving incentives to the agent for productive work and sharing between parties the uncertainty inherent in the stochastic outcome. The agent's action, as a signal, is most efficient in incentive-providing, while the monetary outcome, as a signal, is most efficient in risk-sharing. Altogether, (a, B) is most efficient among all information systems.

In the rest part of the paper, we devote ourselves to the study of comparison criteria in cases when neither the agent's action nor the monetary outcome is observable.

3.2 The Extended Informativeness Criterion

In this subsection, we analyze the comparison of inclusive information systems, which is defined as follows.

Definition 3. *Two information systems are inclusive if one has more sources of signals than the other so that the later is a proper subset of the other.*

This term is firstly used by Kim (1995) to describe situations where Holmstrom(1979)'s informativeness criterion is applicable. Suppose there are two inclusive information systems, X and (X, Y). Naturally, the principal's maximum expected utility is nondecreasing with the number of signals in an information system. Thus $\mathscr{U}^*\big((X, Y)\big) \geq \mathscr{U}^*(X)$. If there exists a $s^*(x, y) = s(x)$, then $s^*(x, y)$ is also a solution to the program *P1* with X, thus $\mathscr{U}^*(X) = \mathscr{U}^*\big((X, Y)\big)$. Conversely, if $\mathscr{U}^*(X) < \mathscr{U}^*\big((X, Y)\big)$ then $s^*(x)$ can not be a solution to the program $P1$ with (X, Y), thus exists no $s^*(x, y) = s(x)$. Therefore we have the following lemma.

Lemma 1. $\mathscr{U}^*(X) \leq \mathscr{U}^*\big((X, Y)\big)$. $\mathscr{U}^*(X) = \mathscr{U}^*\big((X, Y)\big)$ *if and only if* $\exists s^*(x, y) = s(x)$.

Proposition 3.2. *(Extended Informativeness Criterion)*

$\mathscr{U}^*(X) = \mathscr{U}^*\big((X, Y)\big)$ *if and only if*

$$g(b, x, y|a) = \phi(x, y)g(b, x|a), \quad \forall a \quad \textit{and for almost all} \quad (x, y) \in \mathbb{X} \times \mathbb{Y}. \tag{10}$$

Proof. First, we prove that Condition (10) implies Holmstrom's informativeness criterion that

$$f(x, y|a) = \phi(x, y)f(x|a). \tag{11}$$

From (10) we have

$$g(b, x, y|a) = g(y|b, x, a)g(b, x|a) = \phi(x, y)g(b, x|a), \quad \forall a \quad \text{and for almost all} \quad (x, y).$$

Thus

$$g(y|b, x, a) = \phi(x, y), \text{ for all } a \text{ and almost all } (x, y). \tag{12}$$

Taking expectation on the LHS of (12) with respect to $f(b|a,x)$, we get

$$\int g(y|b,x,a)f(b|a,x)db = \int g(b,y|a,x)db = g(y|a,x).$$

Taking expectation on the RHS of (12) with respect to $f(b|a,x)$, we get

$$\int \phi(x,y)f(b|a,x)db = \phi(x,y).$$

Taken together, $g(y|a,x) = \phi(x,y)$ and thus $f(x,y|a) = \phi(x,y)f(x|a)$.

Proof of sufficiency. Assume (10). We prove that for any $s(x,y)$, $\exists s(x)$ that weakly pareto dominates $s(x,y)$.

Define $s(x)$ as follows

$$\int u(s(x,y))\phi(x,y)dy = u(s(x)). \tag{13}$$

Then from (11) and (13), we get

$$\int u(s(x,y))f(x,y|a)dxdy = \int u(s(x))f(x|a)dx.$$

Thus $s(x)$ result in the same action and welfare for the agent.

By Jensen's inequality, (13) implies

$$\int s(x,y)\phi(x,y)dy \geq s(x) = \int s(x)\phi(x,y)dy$$
$$\int (b - s(x,y))\phi(x,y)dy \leq \int (b - s(x))\phi(x,y)dy$$
$$\int U(b - s(x,y))\phi(x,y)dy \leq \int U(b - s(x))\phi(x,y)dy$$
$$\iint [\int U(b - s(x,y))\phi(x,y)dy]g(b,x|a)dxdb \leq \iint [\int U(b - s(x))\phi(x,y)dy]g(b,x|a)dxdb$$
$$\iiint U(b - s(x,y))g(b,x,y|a)dbdxdy \leq \iint U(b - s(x))g(b,x|a)dbdx$$

The principal is at least as well off with $s(x)$ as with $s(x,y)$. Thus $s(x)$ weakly pareto dominates $s(x,y)$. This completes the proof for sufficiency.

Proof of necessity. The proof of necessity is completed in two steps. We first prove that if (11) is false, then for $\forall s(x)$, there exists a $s(x,y)$ that strictly pareto dominates $s(x)$. Secondly, we prove that if (11) is true, but (10) is false, then $\mathscr{U}^*(X,Y) > \mathscr{U}^*(X)$. Because (10) is a stronger condition than (11), these two steps prove the necessity of (10).

Step 1: Suppose with the compensation schedule $s(x)$, the agent's action is $\hat{a}$, the multipliers for the agent's participation constraint and incentive compatibility constraint is λ and μ, respectively. Fix x for a moment. The principal and the agent's marginal return $d\mathscr{U}^p$ and $d\mathscr{U}^a$-conditional on x, from an small additional variation $ds(x,y)$ in the sharing rule, are [6]

$$d\mathscr{U}^p = -\int U'(b-s(x))g(b|x,\hat{a})db \int ds(x,y)f(x,y;\hat{a})dy + \lambda u'(s(x)) \int ds(x,y)f(x,y;\hat{a})dy$$

$$+ \mu u'(s(x)) \int ds(x,y) f_a(x,y;\hat{a})dy. \tag{14}$$

$$d\mathscr{U}^a = u'(s(x)) \int ds(x,y)f(x,y;\hat{a})dy. \tag{15}$$

From (11) it follows that there exists a set Y in the range of y with positive probability measure, and correspondingly for the complement Y^c, such that

$$\frac{f_a}{f}(x,Y|\hat{a}) > \frac{f_a}{f}(x,Y^c|\hat{a}).$$

Because $\mu \neq 0$,[7] choose a variation $ds(x,y)$ such that $ds(x,y) > 0$ if $\mu > 0$ ($ds(x,y) < 0$ if $\mu < 0$) and

$$ds(x,Y)f(x,Y|\hat{a}) + ds(x,Y^c)f(x,Y^c|\hat{a}) = 0.$$

[6] For the mathematical technique used for deriving these formula, refer to proposition 9.6.1 in Luenberger (1969). See also Holmstrom (1979).

[7] This is because the first-best solution is unavailable.

Then from (14) and (15), it follows that

$$d\mathscr{U}^p = \mu u'(s(x))ds(x,y)f(x,Y|\hat{a})[\frac{f_a}{f}(x,Y|\hat{a}) - \frac{f_a}{f}(x,Y^c|\hat{a})] > 0.$$

$$d\mathscr{U}^a = 0.$$

Thus the principal is strictly better off with the additive variation $ds(x,y)$.

Step 2: In this step we show that if (11) is true, but (10) is false, then $\mathscr{U}^*(X,Y) > \mathscr{U}^*(X)$.

The Borch rule for (X,Y) is

$$\int \frac{U'(b - s(x,y))}{u'(s(x,y))} g(b|x,y,a)db = \lambda + \mu \frac{f_a(x,y|a)}{f(x,y|a)}. \tag{16}$$

If (11) is true, then the RHS of (16) is constant in y, thus the LHS of (16) must also be constant in y. However, if (10) is false, then $g(b|x,y,a) \neq g(b|x,a)$. Thus to make the LHS of (16) constant in y, it must be true that $\frac{U'(b-s(x,y))}{u'(s(x,y))}$ varies in y, which implies that $s(x,y) \neq s(x)$. Then according to Lemma 1, $\mathscr{U}^*(X) < \mathscr{U}^*(X,Y)$. This completes our proof of necessity. □

To interpret the extended informativeness criterion, note first that (10) implies the Holmstrom's informativeness condition that $f(x,y|a) = \phi(x,y)f(x|a)$. That is, x is a sufficient statistic for the pair (x,y) with respect to a, which means that x carries all the relevant information about a, and y adds nothing to the power of incentive-providing. The signal Y is redundant for incentive-providing when X is available, and could only possibly be used for the risk-sharing purpose.

Secondly, Condition (10) can also be proven to be equivalent to $g(b|a,x) = g(b|a,x,y)$. That is, the condition distribution of B, given action a and realized signal x, is independent with y, which means that a and x carry all the relevant information about B, and y adds nothing to power of inference of B. Keep in mind that agent's action is actually known at the equilibrium, thus the signal Y is redundant for risk-sharing when X is accessible. Since Y is redundant for both the incentive-providing and risk-sharing purposes, Y is valueless when (10) holds.

On the other hand, if (10) is false, y contains some information about either a or B beyond that conveyed by x, y should then be used in the contract to improve the principal's welfare.

3.3 The MPS-SOSD and MPS-FOSD Criteria

Our task in this subsection is to develop comparison criteria that can be applied to scalar information systems, which are defined as follows.

Definition 4. *An information system X is scalar if it can be represented by a random variable over the support $\mathbb{X} \subset \mathbb{R}$.*

We refer to the following program *P2* as *the doubly-relaxed program*

$$\text{P2:} \qquad \mathscr{U}^{**}(X) = \max_{s(x)\geq k, a} \int_{\mathbb{X}} \int_{\mathbb{B}} U(b - s(x)) g(b, x; a) db\, dx, \tag{17}$$

$$s.t. \quad \int_{\mathbb{X}} u(s(x)) f(x; a) dx - c(a) = 0, \tag{18}$$

$$\int_{\mathbb{X}} u(s(x)) f_a(x; a) dx - c_a(a) \geq 0. \tag{19}$$

The term "doubly-relaxed" indicates the fact that the IC constraint in the above program is weakened even further from the relaxed program. Let $\mathscr{S}^{**}(X) = \big(a^{**}, s^{**}(x), \mathscr{U}^{**}(X)\big)$ be the solution to *P2* with the information system X.

To get prepared for the main result in this subsection, we state and prove two lemmas. Our main result, Proposition 3.3, follows directly from these two lemmas.

The first lemma is the Lemma 1 in Xie (2009), which states that under certain regularity conditions, any solution to *P1* is also a solution to *P2*, and the optimal payment schedule increases in the realization of the signal. We denote the principal's and the agent's absolute risk-aversion by $R_A(\cdot)$ and $r_A(\cdot)$ respectively.

Lemma 2. *Suppose that*

$$R_A\big(b-s\big) \leq r_A\big(s\big), \quad \forall b \textit{ and } \forall s. \tag{20}$$

If a scalar information system Z satisfies the following conditions

$$\frac{\partial \frac{f_a}{f}(z|a)}{\partial z} \geq 0, \quad \forall z \textit{ and } \forall a \tag{21}$$

$$\frac{\partial \frac{g_z}{g}(b|z,a)}{\partial b} \geq 0, \quad \forall z,\ \forall b,\ \textit{and } \forall a \tag{22}$$

$$\frac{\partial \frac{g_a}{g}(b|z,a)}{\partial b} \geq 0, \quad \forall z,\ \forall b,\ \textit{and } \forall a \tag{23}$$

then

$$\mathscr{S}^*(Z) = \mathscr{S}^{**}(Z), \tag{24}$$

$$\mu^* > 0, \tag{25}$$

$$s_z^*(z) > 0, \quad \forall z \in \mathbb{Z}. \tag{26}$$

Some insight can be gained by inspecting the conditions laid out in Lemma 2. Condition (20) assumes that the principal is less risk averse than the agent. This assumption is reasonable, because the principal often is more tolerant of risk and better able to bear it than the agent.

Conditions (21)-(23) put restrictions on the stochastic relationship between the agent's effort, the monetary outcome, and the signal. Condition (21) is a Monotone Likelihood Ratio Condition (MLRC) in z, which has a fairly natural interpretation of more effort, larger value of the signal. (22) is a MLRC in b. It says that for a given effort level a, a larger monetary outcome implies a larger value of the signal. Condition (23) is another MLRC in b. It says that for a given value of the signal, the monetary outcome tends to increase with effort.

We introduce the concept of *likelihood ratio distribution function* of a signal, which will be used in Lemma 3.

Definition 5. *Let $L_X^a(z)$, where $z = \frac{f_{X_a}}{f_X}(x|a) \in \mathbb{R}$, denote the likelihood ratio distribution function of an information system X at action a. That is,*

$$L_X^a(z) = Prob\Big[\frac{f_{X_a}}{f_X}(x|a) \leq z\Big] = \int_{\frac{f_{X_a}}{f_X}(x|a)\leq z} f_X(x|a)dx.$$

We now state and prove the second lemma.

Lemma 3. *Given two scalar information systems, X and Y, each of which satisfies the following two conditions*

$$\mathscr{S}^*(Z) = \mathscr{S}^{**}(Z), \tag{27}$$

$$s_z^*(z) \geq 0, \forall z \in \mathbb{Z}. \tag{28}$$

Then $\mathscr{U}^*(X) \geq \mathscr{U}^*(Y)$ if

$$L_X^a(z) \quad \text{is a Mean Preserving Spread (MPS) of} \quad L_Y^a(z), \quad \forall a \in A, \tag{29}$$

and either of the following two conditions is satisfied:

$$G_X\Big(b|a, F_X^{-1}(z|a)\Big) \quad \text{Second Order Stochastically Dominates (SOSD)} \quad G_Y\Big(b|a, F_Y^{-1}(z|a)\Big),$$

$$\forall a \in A \quad \text{and a.e.} \quad z \in [0,1]. \tag{30}$$

$$F_{F_X(Y|a)}(z|a,b) \quad \text{First Order Stochastically Dominates (FOSD)} \quad F_{F_X(X|a)}(z|a,b),$$

$$\forall a \in A \quad \text{and a.e.} \quad b \in \mathbb{B}. \tag{31}$$

Proof. We introduce a standardization of information systems. Define $Z_Y \equiv F_Y(y|a_Y^*)$ and

$Z_X \equiv F_X(x|a_Y^*)$, where a_Y^* is the effort level in the solution to Program *P1* with Y. Because Z_X and Z_Y are monotonic transformations of X and Y, by condition (28), $\partial s_{Z_X}^*(z)/\partial z \geq 0$ and $\partial s_{Z_Y}^*(z)/\partial z \geq 0$. The main advantage of the standardized information system is that they are uniformly distributed on $[0,1]$ at a_Y^*. Without lose of generality, we assume that both systems are standardized at a_Y^*.

By standardization, the MPS condition (29) is equivalent to the following integral condition[8]

$$-\frac{\partial F_X(Z|a)}{\partial a} \geq -\frac{\partial F_Y(Z|a)}{\partial a}, \quad \forall a \in A \quad \text{and} \quad z \in [0,1]. \tag{32}$$

The SOSD condition (30) is equivalent to

$$G_X(b|a,z) \quad \text{SOSD} \quad G_Y(b|a,z), \quad \forall a \in A \quad \text{and a.e.} \quad z \in [0,1]. \tag{33}$$

The FOSD condition (31) is equivalent to

$$F_Y(y|a,b) \quad \text{FOSD} \quad F_X(x|a,b), \quad \forall a \in A \quad \text{and a.e.} \quad b \in \mathbb{B}, \tag{34}$$

The solution to Program *P1* with Y is denoted by $\left(a_Y^*, S_Y^*(z)\right)$. We first prove that $\left(a_Y^*, s_Y^*(z)\right)$ satisfies the IR and IC constraints in Program *P2* with information system X.

As to the IR constraint, because $f_X(z|\hat{a}) = f_Y(z|\hat{a}) = 1$ on $z \in [0,1]$, we get

$$\int_0^1 u(s_Y^*(z) f_X(z|a_Y^*) dz = \int_0^1 u(s_Y^*(z) f_Y(z|a_Y^*) dz \geq c(a_Y^*).$$

The inequality follows from the fact that $s_Y^*(z)$ satisfies the IR constraint in Program *P1* with Y.

[8]See the proof of Proposition 3 in Demougin and Fluet (2001).

The IC constraint under X is

$$
\begin{aligned}
&\int_0^1 u(s_Y^*(z))\frac{\partial f_X(z|a_Y^*)}{\partial a}dz - c_a(a_Y^*)\\
&= -\int_0^1 u'(s_Y^*(z)){s_Y^*}'(z)F_{Xa}(z|a_Y^*)dz - c_a(a_Y^*)\\
&\geq -\int_0^1 u'(s_Y^*(z)){s_Y^*}'(z)F_{Ya}(z|a_Y^*)dz - c_a(a_Y^*)\\
&= \int_0^1 u(s_Y^*(z))\frac{\partial f_Y(z|a_Y^*)}{\partial a}dz - c_a(a_Y^*) = 0
\end{aligned}
$$

The first and second equalities follow from integration by part. The inequality follows from (32) and the fact that ${s^*_Y}_z(z) \geq 0$.

Since $\{a_Y^*, s_Y^*(z)\}$ satisfies both the IR and IC constraints in Program *P2* with X, then by condition (27), $\left(a_Y^*, s_Y^*(z)\right)$ is a potential solution to Program *P1* with X. Let $\mathscr{U}^*(X|a_Y^*, s_Y^*(\cdot)) = \int_0^1 \int_{\mathbb{B}} U(b - s_Y^*(z))g_X(b, z|a_Y^*)db\,dz$ be the principal's expected utility if the agent exerts effort a_Y^*, and the compensation schedule is $s_Y^*(x)$ written on X.

Then, on one hand, if the SOSD condition (33) is satisfied, we have

$$
\begin{aligned}
\mathscr{U}^*(X) \geq \mathscr{U}^*(X|a_Y^*, s_Y^*(\cdot)) &= \iint U(b - s_Y^*(z))g_X(b, z|a_Y^*)db\,dz\\
&= \iint U(b - s_Y^*(z))g_X(b|z, a_Y^*)db f_X(z|a_Y^*)dz\\
&\geq \iint U(b - s_Y^*(z))g_Y(b|z, a_Y^*)db f_Y(z|a_Y^*)dz\\
&= \iint U(b - s_Y^*(z))g_Y(b, z|a_Y^*)db\,dz\\
&= \mathscr{U}^*(Y)
\end{aligned}
$$

The first inequality follows from the fact that $\mathscr{U}^*(X)$ is the principal's maximum expected utility in Program *P1* with X, and the second inequality follows from condition (33) and the fact that $f_X(z|a_Y^*) = f_Y(z|a_Y^*) = 1$ by standardization. This finishes the proof of the MPS-SOSD criterion.

On the other hand, if the FOSD condition (34) is satisfied, we have

$$\begin{aligned}
\mathscr{U}^*(X) \geq \mathscr{U}^*(X|a_Y^*, s_Y^*(\cdot)) &= \iint U(b - s_Y^*(z)) g_X(b, z|a_Y^*) db\, dz \\
&= \iint U(b - s_Y^*(z)) g_X(z|b, a_Y^*) dz\, g(b|a_Y^*) db \\
&\geq \iint U(b - s_Y^*(z)) g_Y(z|b, a_Y^*) dz\, g(b|a_Y^*) db \\
&= \iint U(b - s_Y^*(z)) g_Y(b, z|a_Y^*) db\, dz \\
&= \mathscr{U}^*(Y)
\end{aligned}$$

The second inequality follows from condition (34) and condition (28). This finishes the proof of the MPS-FOSD criterion. □

Combining Lemma 2 and Lemma 3, one immediately gets the following proposition.

Proposition 3.3. ***(The MPS-SOSD and the MPS-FOSD criteria)*** *Assume that the agent is more absolutely risk-averse than the principal. Given two scalar information systems, X and Y, each of which satisfies the regularity conditions* (21)-(23) *in Lemma 2. Then $\mathscr{U}^*(X) \geq \mathscr{U}^*(Y)$, if the following MPS-SOSD criterion is satisfied:*

$$\begin{aligned}
L_X^a(z) \quad &\text{is a MPS of} \quad L_Y^a(z), \quad \forall a \in A, \\
G_X(b|a, F_X^{-1}(z|a)) \quad &\text{SOSD} \quad G_Y(b|a, F_Y^{-1}(z|a)), \quad \forall a \in A \quad \text{and a.e. } z \in [0,1],
\end{aligned}$$

or if the following MPS-FOSD criterion is satisfied:

$$\begin{aligned}
L_X^a(z) \quad &\text{is a MPS of} \quad L_Y^a, \quad \forall a \in A \\
F_{F_Y(Y|a)}(z|a, b) \quad &\text{FOSD} \quad F_{F_X(X|a)}(z|a, b), \quad \forall a \in A \quad \text{and a.e.} \quad b \in \mathbb{B}.
\end{aligned}$$

The conditions in Lemma 2 play a critical role in Proposition 3.3. They guarantee that the optimal compensations,$s_X^*(z)$ and $s_Y^*(z)$, are increasing functions. If the compensation functions are non-increasing, the MPS condition no longer implies relative efficiency in incentive-providing. For instance, when the principal is risk-neutral, if the optimal compensation functions are *decreasing*, i.e., ${s^*_X}'(z) \leq 0$ and ${s^*_Y}'(z) \leq 0$, $\forall z \in [0, 1]$, then X is more efficient than Y if the likelihood ratio distribution of Y is a MPS of the likelihood ratio distribution of X. (see the proof of Proposition 1 in Demougin and Fluet (2001).)

Proposition 3.3 differs from Kim's MPS criterion and Demougin and Fluet's integral criterion in that it requires a SOSD condition on the conditional distribution function of B given a and z ($z = x, y$), or a FOSD condition on the conditional distribution functions of Z ($Z = X, Y$) given a and b. Again, this difference arises from the additional risk-sharing function the information system performs in our model.

Proposition 3.3 says that for two scalar information systems, X and Y, satisfying certain regularity conditions so that the optimal payment schedules are increasing functions, if the likelihood ratio distribution of X is a MPS of that of Y, then X is more efficient than Y on incentive-providing; if the conditional distribution of B given a and x SOSD the conditional distribution of B given a and y, or if the conditional distribution of X given a and b FOSD the conditional distribution of Y given a and b, then X is more efficient than Y on risk-sharing. Since X is more efficient than Y on both dimensions, X is overall more efficient than Y, and any risk-averse principal will prefer X to Y.

4 Conclusion

The choice of information system has been a fundamental issue in agency theory. In the current literature, all criteria one-dimensionally compare information systems on their efficiencies in incentive-providing. The main objective of this paper is to extend the analysis to a two-dimensional comparison of information systems on their efficiencies on incentive-providing and risk-sharing.

In this paper, we first identify the most efficient information systems. Then we develop an extended informativeness criterion for comparison of inclusive information systems, and a MPS-SOSD and a MPS-FOSD criteria for comparison of scalar information systems.

The author is unaware of any two-dimensional comparison criterion on arbitrary information systems that are neither inclusive nor scalar. Unlike the MPS (integral) criterion on scalar information systems, which can be easily generalized to cases with non-scalar information systems (see Demougin and Fluet (1998ab, 2001)), our MPS-SOSD and MPS-FOSD criteria on scalar information systems can not be readily extended to non-scalar information systems.[9] This difficulty of extension arises from an intrinsic difference between our model and the standard model with a risk-neutral principal: in the standard model, the likelihood ratio of an information system is a scalar sufficient statistic of the information system, while in our model, the likelihood ratio is not.

One should also note that no general criterion can compare information systems X and Y, when X is more efficient on one function, while Y is more efficient on the other. In this case, the comparison depends on more specific characteristics of the agency model.

5 References

Blackwell, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability,* ed. by J. Neyman. Berkeley; University of California Press.

Blackwell, D. (1953): "Equivalent Comparison of Experiments," *Annals of Mathematics and Statistics,* 24, 265-272.

Bolton, P. and M. Dewatripont (2005): "Contract Theory," Cambridge, MA: MIT Press.

Demougin, D. and C. Fluet (1998):"Mechanism Sufficient Statistic in the Risk-Neutral Agency Problem," *Journal of Institutional and Theoretical Economics,* 154(4), 622-639.

[9]For example, one can check that the MPS-SOSD and the MPS-FOSD criteria are not readily applicable to inclusive information systems: the extended informativeness criterion developed in Section 3.2 is neither equivalent to nor a special case of the MPS-SOSD or the MPS-FOSD criteria.

Demougin, D. and C. Fluet (2001): "Ranking of Information Systems in Agency Models: An Integral Condition," *Economic Theory,* 17, 489-496.

Dewatripont, M., Jewitt, I., and Tirole, J. (1999): "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies,* 66, 32 183-198.

Fagart, M.-C. and B. Sinclair-Desgagne. (2007): "Ranking Contingent Monitoring Systems," *Management Science*, 53, 1501-1509.

Laffont, J., and D. Martimort. (2001): "The Theory of Incentives: The Principal-Agent Model," Princeton University Press.

Fagart, M., and B. Sinclair-Desgagne. (2004): "Auditing Policies and Information Systems in a Principal-Agent Model," mimeo, LEI-CREST, Paris.

Grossman, S., and O. Hart (1983): "An Analysis of the Principal-Agent Problem," *Econometrica,* 51, 7-45.

Holmstromm, B. (1979): " Moral Hazard and Observability," *Bell Journal of Economics,* 10, 74-91.

Holmstrom, B. (1982): "Moral Hazard in Teams," *Bell Journal of Economics,* 13, 324-340.

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

Jewitt, I. (1997): "Information and Principal Agent Problems," University of Bristol Discussion Paper 97/414.

Jewitt, I. (2007): "Information Order in Decision and Agency Problems," Personal Manuscript.

Jewitt, I. and S. Mukerji (2007): "Ordering Ambiguous Acts," Working Paper, University of Oxford.

Kim, S. K. (1995): "Efficiency of an Information System in an Agency Problem," *Econometrica,* 63, 89-102.

Mirrlees, J. (1999), "The Theory of Moral Hazard and Unobservable Behavior: Part 1", *The Review of Economic Studies*, 66, 3C23.

Rodriguez, G.E. (2007): “On the Value of Information in the Presence of Moral Hazard,” *Review of Economic Design*, 10: 341-361.

Rogerson, W. (1985):“The First Order Approach to Principal-Agent Problems,” *Econometrica,* 53, 1357-1368.

Xie, Jia. (2008): “The First-Order Approach to Principal-Agent Problems with An Unverifiable Monetary Outcome,” Working Paper, Indiana University.