Auditing policies and information systems in a principal-agent model

By Marie-Cécile Fagart and Bernard Sinclair-Desgagné*

January 2002

This paper considers the information systems induced by auditing policies in a principal-agent model with moral hazard, assuming the first-order approach is valid. We point out that in this context two information systems A and B are seldom comparable using the customary mean-preserving spread relation between their respective likelihood ratio distributions. We offer a general extension of this criterion, however, and we use it to show that, provided the sign of the third derivative of the agent's inverse utility function is constant, it is yet often possible to rank A against B because one of the corresponding likelihood ratio distributions dominates the other in the third order.

KEYWORDS: Principal-agent, moral hazard, audits, likelihood ratio distribution, stochastic dominance, prudence.

---

* We are grateful to Claude Fluet, Dominique Henriet and, most of all, Louis Eeckhoudt for valuable discussions on the topic of this paper. The authors respectively thank CIRANO and the Laboratoire d'Économie Industrielle (LEI-CREST, Paris) for their hospitality and support while this paper was being written.

1. Introduction

An important topic in the analysis of principal-agent relationships is the comparison of information systems that imperfectly correlate some common observables with the agent's hidden actions. Any classification should first lead to identify and discard information systems under which the principal achieves a relatively lower expected payoff. A "practical" (i.e. robust) ranking criterion, however, would also rely as little as possible on specific features of the current relationship, such as the agent's utility function.

Starting with the seminal contribution of Holmström (1979), some orderings have successively been studied by Gjesdal (1982), Grossman and Hart (1983), Kim (1995), Jewitt (1997), and Demougin and Fluet (2000). One shortcoming of the suggested rankings is that they hardly convey the actual costs of gathering and communicating the prescribed observables (see Baker (1992)). A second weakness, which most of the literature primarily addresses, is that they are incomplete and may not allow to decide in some contexts between relevant information systems.

Among the available orderings, the "MPS criterion" introduced by Kim (1995) - which classifies information systems according to the mean-preserving spread relation between their respective likelihood ratio distributions - is now the one that best deals with the latter criticism.[1] This criterion embodies those that were proposed earlier and constituted

[1] To be precise, the MPS criterion says that (assuming the first-order approach to principal-agent problems is valid) an information system A yields a higher expected payoff to the principal than an information system B if the likelihood ratio distribution associated with A is a mean-preserving spread of the one associated with B, or in other words if the latter dominates the former in the sense of second-order stochastic dominance. Alternative criteria were recently introduced and discussed by Jewitt (1997) and Demougin and Fluet (2000), who show that these are actually equivalent to the MPS criterion.

indeed a radical improvement, for it allows comparisons between information systems that are not necessarily nested.

While studying the information systems induced by auditing policies, however, we found out that this significant group largely eluded the MPS criterion. An intuitive explanation of this fact would be the following. Previous work by Baiman and Demski (1980), Dye (1986), Sinclair-Desgagné (1999) and others have shown that optimal (and economically plausible) audits are often either upper-tailed or lower-tailed, i.e. triggered by the observation of respectively good or bad signals. In selecting an auditing policy to bring about a given action by the agent, a rational principal would thus typically discriminate between compound information systems of the form

A (upper-tailed policy): use $L_X + L_Y$ if signal $X \geq x'$, and L_X otherwise;

versus

B (lower-tailed policy): use $L_X + L_Y$ if signal $X \leq x''$, and L_X otherwise;

where $\text{Prob}\{X \geq x'\} = \text{Prob}\{X \leq x''\}$, i.e. the two policies entail the same frequency of audits (hence the same cost), and L_X and L_Y are two independent likelihood ratio distributions. Yet, A and B clearly have the same mean zero (since both L_X and L_Y have mean zero) and the same variance, so neither is a mean-preserving spread of the other.

The objective of this paper is thus to develop a ranking criterion which supplements the MPS criterion and allows to make comparisons between information systems that commonly occur in the analysis of auditing policies.

The upcoming section lays out the basic principal-agent model and its main assumptions, which are meant in particular to guarantee the validity of the first-order approach.[2] Section 3 discusses optimal audits and formalizes the above intuitive explanation. Proposition 1 first establishes that the frequency of optimal audits decreases with their cost, this frequency being equal to 1 when the cost is 0. Proposition 2 points out next that, except for peculiar instances of the agent's utility function, optimal audits are not just characterized by an appropriate frequency but they are also contingent upon observing the level of some predefined signal.[3] When we compute the variance of the likelihood ratio distribution associated with an auditing policy, however, we find that it only depends on the frequency of audits and on the Fisher information indices associated with the underlying likelihood ratio distributions; hence, the MPS criterion cannot distinguish between two contingent auditing policies that bear the same frequency.

Section 4 is then devoted to developing a finer ranking that would still be based on the comparison of likelihood ratio distributions. A convenient criterion is presented through proposition 3. According to it, a rational principal would again prefer an information system which likelihood ratio distribution is dominated in the sense of second-order stochastic dominance (corollary 1); this new criterion thus coincides with the MPS criterion. Provided the sign of the third derivative of the agent's inverse utility remains constant,

[2] The first-order approach could have been justified in the present context using the assumptions made in Jewitt (1988). But since it is important here not to restrict a priori the range of possible agent's utility functions, our assumptions are finally adapted from Sinclair-Desgagné (1994).

[3] For instance, the principal would rather use upper-tailed audits when the agent's coefficient of absolute prudence (as defined in Kimball (1990)) is larger that three times his coefficient of absolute risk aversion. This key result, and other related ones, will be discussed below.

furthermore, the criterion yields a finer ordering of information systems based on third-order stochastic dominance (corollary 2). This feature is used in section 5, where we finally show that it allows a ranking of upper-tailed and lower-tailed audits which portrays the principal's choices (proposition 4). Section 6 contains some concluding remarks.

2. The Model

Consider a one-period relationship between a principal and an agent. An amount of effort $a \in [0, \infty)$ is expected from the latter. This effort, however, is only imperfectly observable through some random variables X and Y. We assume that X and Y are conditionally independent, so for a given effort a the realizations x and y of the random variables obey the conditional distributions $F(x, a)$ and $G(y, a)$ respectively. Those distributions have respective densities noted $f(x, a)$ and $g(y, a)$ that exhibit constant support (denoted by Γ_X and Γ_Y) with respect to a and are twice continuously differentiable in a for every x and y. Throughout this paper the subscript $_a$ refers to the partial derivative with respect to a.

The likelihood ratios associated with X and Y will now be respectively denoted $L_X(x, a) = \dfrac{f_a(x, a)}{f(x, a)}$ and $L_Y(y, a) = \dfrac{g_a(y, a)}{g(y, a)}$. Clearly, these ratios are themselves random variables, and their respective distribution is called a "likelihood ratio distribution." It is well known that all likelihood ratio distributions have the same mean $E_X[L_X] = E_Y[L_Y] = 0$. The variance of, say, L_X is then given by $Var(L_X) = E[(L_X)^2]$; it is often denoted I_X and called the "Fisher information index" associated with X.[4]

[4]The Fisher information index is well-known to statisticians and econometricians (see Gouriéroux and

The risk neutral principal routinely observes the value of X. Based on this, she may either compensate the agent immediately according to a wage schedule $w(X)$, or she may audit the agent at a constant cost K - thereby also gathering signal Y - and pay him according to a sharing rule $s(X, Y)$. We suppose that the principal can commit to a probability $m(x)$ of making an audit upon observing $X = x$. Her expected cost when the agent delivers effort a is therefore given by

(1)

$$EC = \int_{\Gamma_X} \int_{\Gamma_Y} \{(1 - m(x))w(x) + m(x)s(x, y)\} dF(x, a) dG(y, a) + K \int_{\Gamma_X} m(x) dF(x, a).$$

The latter integral $M(a) = \int_{\Gamma_X} m(x) dF(x, a)$ yields the "frequency" (or the "intensity") of audits under a policy $m(X)$, when the agent expends an effort level a.

The agent's preferences are assumed to be additively separable in effort and wealth. The cost of effort is scaled so that its first-order derivative is equal to 1. The agent's attitude with respect to uncertain variations of his wealth exhibits risk aversion and is represented by a positive, strictly concave and three-times continuously differentiable Von Neumann-Morgenstern utility index $u(\cdot)$. The agent's expected utility after putting an effort a under a contract $[w, s, m]$ is then precisely

$$(2) \qquad EU = \int_{\Gamma_X} \int_{Y} \{(1 - m(x))u(w(x)) + m(x)u(s(x, y))\} dF(x, a) dG(y, a) - a.$$

Monfort (1989), for example) Note that $E_X[(L_X)^2] = E_X[-\frac{\partial L_X}{\partial a}]$, so this index provides a measurement of the sensitivity of the likelihood ratio with respect to a. For a recent account of the pervasiveness and usefulness of the Fisher information index in principal-agent analysis, see Dewatripont et al. (1999).

In the upcoming sections, we let $\varphi = u^{-1}$ denote the inverse of $u(\cdot)$, and the following transformations of the agent's utility index, $\Delta(w, \sigma) \equiv u(w)\sigma - w$ and $\Delta^*(\sigma) \equiv Max_{w \in W}\{\Delta(w, \sigma)\}$, will be quite useful.

A rational principal will select an auditing policy $m(X)$ and wage schedules $w(X)$ and $s(X, Y)$ that implement a given effort a at a minimal cost, provided the agent thereby achieves his reservation utility level $\underline{U}$ and is also willing to deliver the expected effort level. This amounts formally to minimize (1), subject to participation and incentive compatibility constraints given respectively by

(3) $$EU = \int_{\Gamma_X} \int_{\Gamma_Y} \{(1 - m)u(w) + mu(s)\} dF dG - a \geq \underline{U},$$

(4) $$a = \arg\max_e \int_{\Gamma_X} \int_{\Gamma_Y} \{(1 - m)u(w) + mu(s)\} dF(x, e) dG(y, e) - e.$$

The latter constraint involves a continuum of inequalities and is thus not generally tractable. In what follows we replace it by a friendlier one which requires that the effort level a be an interior stationary point of the agent's expected utility function, that is:

(5)
$$\int_{\Gamma_X} \int_{\Gamma_Y} [f_a(x, a)g(y, a) + f(x, a)g_a(y, a)]\{(1 - m(x))u(w(x) + m(x)u(s(x, y))\} dx dy - 1 \geq 0.$$

We want this so-called "first-order approach" to always yield a solution that constitutes an incentive compatible allocation (so that solves the initial problem as well). It can be shown that this will be the case if the following assumptions are met.

Monotone Likelihood Ratio Property (MLRP): *For all* a, *the likelihood ratios* $L_X(x,a) = \dfrac{f_a(x,a)}{f(x,a)}$ *and* $L_Y(y,a) = \dfrac{g_a(y,a)}{g(y,a)}$ *are nondecreasing in* x *and* y.

Convexity of the Distribution Function Condition (CDFC): *For all* x, y, *and* a, $F_{aa}(x) \geq 0$ *and* $G_{aa}(y) \geq 0$.

The first assumption is quite common in statistics and principal-agent analysis. It implies (for univariate distributions only) that $F_a(x) \leq 0$ and $G_a(y) \leq 0$ for all x, y and a, so larger realizations of X and Y make it more likely that the agent's effort was higher. The second assumption is often invoked in principal-agent analyses that use the first-order approach and it corresponds in turn to some (stochastic) decreasing returns to effort. Conditional distributions that satisfy those two assumptions can easily be constructed. For instance, take a pair of different continuous distribution functions $P(z)$ and $Q(z)$ with similar support such that $P(z) \leq Q(z)$ for all z, and a pair of functions $\alpha : [0,\infty) \to [0,1]$ and $\beta : [0,\infty) \to [0,1]$ increasing and concave such that $\alpha(0) = \beta(0) = 0$ and $\lim_{a\to\infty} \alpha(a) = \lim_{a\to\infty} \beta(a) = 1$. One may then let $F(x,a) = \alpha(a)P(x) + (1-\alpha(a))Q(x)$ and $G(y,a) = \beta(a)P(y) + (1-\beta(a))Q(y)$.

3. Optimal Auditing Policies

Let Λ denote the Lagrangian function associated with the current principal-agent problem. Using the notation of section 2, we have that

$$\begin{aligned}\Lambda &= -K\int_{\Gamma_X} mdF + \int_{\Gamma_X}(1-m)\Delta(w,\lambda+\mu L_X)dF \\ &\quad + \int_{\Gamma_X}\int_{\Gamma_Y} m\Delta(s,\lambda+\mu(L_X+L_Y))dFdG - \lambda(a+\underline{U}) - \mu,\end{aligned}$$

where λ and μ are the multipliers corresponding to the participation and the incentive constraints respectively. If $[w(X), s(X,Y), m(X)]$ solves the principal-agent problem and constitutes thereby an optimal contract, then the following conditions have to be satisfied for some $\lambda \geq 0$ and $\mu \geq 0$:

1. if $m(x) < 1$, then $w(x) = Argmax_w\ \Delta(w, \lambda + \mu L_X(x,a))$,

2. if $m(x) > 0$, then $s(x,y) = Argmax_w\ \Delta(w, \lambda + \mu L_X(x,a) + \mu L_Y(y,a))$, and

for all x, $m(x)$ maximizes $m(x) \cdot G(L_X(x,a))$ on $[0,1]$, where $G(.)$ is defined as

$$G(z) \equiv E_Y[\Delta^*(\lambda + \mu z + \mu L_Y)] - \Delta^*(\lambda + \mu z) - K. \tag{6}$$

Given the first two conditions the latter can be written as

$$m(x) = \arg\max_{m\in[0,1]} m\{\int_{\Gamma_Y}[u(s)(\lambda+\mu L_X+\mu L_Y)-s]dG - [u(w)(\lambda+\mu L_X)-w] - K\}. \tag{7}$$

And when the decision to audit is randomized, i.e. when $1 > m(x) > 0$ at some x, the first and second conditions can also be written respectively as

(8) $$u'(w)\{\lambda + \mu L_X\} = 1,$$

(9) $$u'(s)\{\lambda + \mu L_X + \mu L_Y\} = 1.$$

If $m(x) = 0$ or 1 at some signal x, however, there is a multiplicity of optimal contracts, since $s(x, Y)$ can be set arbitrarily at $m(x) = 0$ and any $w(x)$ is also a possible solution at $m(x) = 1$. In what follows, we shall suppose without losing generality that in this case $s(x, Y)$ and $w(x)$ still satisfy conditions 1 and 2, and so equations (8) and (9).

Together with the Monotone Likelihood Ratio Property, the latter two equations entail that the optimal wages $w(x)$ and $s(x, y)$ are nondecreasing in x and y. Let us now turn to auditing policies. The next statement first establishes that the frequency of audits is naturally related to their unit cost K.

PROPOSITION 1: *An optimal auditing policy is such that, for any given* a, $M(a)$ *is decreasing with respect to* K *and* $M(a) = 1$ *when* $K = 0$.

The proof can be found in the Appendix. Note that the second part of this proposition extends somewhat Holmström (1979)'s celebrated "sufficient statistic" result: it says that any informative signal about the agent's effort has positive value for the principal, *even* when gathering such a signal could be a strategic decision. The following example now brings up a situation which highlights further the relationship between audit cost and audit intensity.

EXAMPLE: Let the agent's preferences exhibit constant relative risk aversion (CRRA) equal to 1/2, so they can be represented by a utility index of the form $u(t) = t^{1/2}$. By equations (8) and (9), the wage schedules in this case are given by

$$w(X) = (\frac{\lambda + \mu L_X}{2})^2 \ \text{ and } s(X,Y) = (\frac{\lambda + \mu L_X + \mu L_Y}{2})^2.$$

Making substitutions in the participation constraint (3) and the incentive constraint (5) then yields the following relationships:

$$EU = \frac{\lambda}{2} - a = \underline{U}$$

and

$$EU_a = \frac{\mu}{2}\{\int_{\Gamma_X} (L_X)^2 dF + M \int_{\Gamma_Y} (L_Y)^2 dG)\} - 1 = \frac{\mu}{2}\{I_X + MI_Y\} - 1 = 0.$$

The principal's expected cost can thus be written as

$$(10) \qquad EC^* = (\frac{\lambda}{2})^2 + (\frac{\mu}{2})^2\{I_X + MI_Y\} + KM = (a + \underline{U})^2 + \frac{1}{I_X + MI_Y} + KM.$$

It appears therefore that this cost depends exclusively on the unit cost of an audit K and on the intensity $M(a)$ of the chosen auditing policy. The latter would actually be set so that

$$
\begin{aligned}
M(a) &= 1 \quad \text{when } K \leq \frac{I_Y}{(I_X + I_Y)^2} \ , \\
M(a) &= 0 \quad \text{when } K \geq \frac{I_Y}{(I_X)^2} \ , \quad \text{and} \\
M(a) &= \frac{1}{I_Y}\{(\frac{I_Y}{K})^{1/2} - I_X\} \ \text{ when } \frac{I_Y}{(I_X + I_Y)^2} < K < \frac{I_Y}{(I_X)^2} \ .
\end{aligned}
$$

Observe also that this policy exhibits the intuitive property that the agent would be audited less often under a signal X which is more informative (in the sense of Fisher).

In this example the principal is indifferent between auditing policies that would be contingent on the observed value of X, as long as such policies have the same intensity. This outcome is obviously rather peculiar. In their seminal article, for instance, Baiman and Demski (1981) have identified a significant range of situations where an optimal auditing policy would either be lower-tailed or upper-tailed. The next proposition is a generalization of their main theorem.

PROPOSITION 2: *When $\varphi'''(\cdot) = 0$, optimal auditing is a matter of setting the appropriate auditing intensity. When $\varphi'''(\cdot) < 0$ (resp. $\varphi'''(\cdot) > 0$), however, an optimal auditing policy prescribes that audits be also triggered - with probability equal to 1 - by the higher (resp. the lower) values of X.*[5]

The proof is also presented in the Appendix. A heuristic derivation of this statement

[5] Note that contingent - albeit *two-tailed* - audits are also optimal when either both $\varphi'''(\cdot) \leq 0$ and $\varphi'''(\cdot) \geq 0$ are untrue (Young (1986)) or the observables X and Y are correlated (Lambert (1985)).

might run as follows. First, define

$$\begin{aligned} u_N(x) &= u(w(x)), \\ u_A(x) &= E_Y[u(s(x,Y))] = u(w_A(x)), \\ u(s(x,y)) &= u_A(x) + \omega(x,y) \ \text{ with } E_Y[\omega(x,Y)] = 0, \\ \text{and} \quad \rho(x) &= E_Y[s(x,Y)] - w_A(x), \end{aligned}$$

so $\omega(x,Y)$ represents the contingent "lottery" (with prizes expressed in the units of the agent's utility function) associated with an audit that comes after an appraisal x, and $w_A(x)$, $\rho(x)$ denote respectively the "certainty equivalent" and the "risk premium" associated with such a lottery. A new formulation of the current optimization problem is now available, that is:

$$EC = \int_{\Gamma_X} \{(1-m)\varphi(u_N) + m[\varphi(u_A) + \rho]\} dF + K \int_{\Gamma_X} m dF \tag{11}$$

$$EU = \int_{\Gamma_X} \{(1-m)u_N + m u_A\} dF - a \geq \underline{U} \tag{12}$$

$$EU_a = \int_{\Gamma_X} \{(1-m)u_N + m u_A\} dF_a + \int_{\Gamma_X} \int_{\Gamma_Y} m\omega dF dG_a - 1 \geq 0. \tag{13}$$

Note that the risk premium ρ can in turn be written as

$$\rho(x) = E_Y[\varphi(u_A(x) + \omega(x,Y))] - \varphi(u_A(x)). \tag{14}$$

REMARK 1: The principal's problem is thereby equivalent to that of a Von-Neumann-Morgenstern decision-maker with utility index $-\varphi(\cdot)$ who must select feasible contributions $u_N(X)$ and $u_A(X)$ together with fair lotteries of the form $\omega(x, Y)$ and their contingent probabilities of occurrence $m(x)$.

If $\varphi'''(\cdot) \equiv 0$, then ρ is invariant with respect to u_A. In this case the decision-maker prefers to set $u_N(x) = u_A(x)$ whenever $0 < m(x) < 1$, because φ is a convex function. By equations (8) and (9), moreover,

$$\mu L_Y = \varphi'(u_A(x) + \omega(x, Y)) - \varphi'(u_N(x)), \tag{15}$$

so the contingent lotteries $\omega(x, Y)$ must be identical since φ' is a linear function. The decision-maker's problem amounts therefore to minimize

$$EC = \int_{\Gamma_X} \varphi(u_N(x))dF + M\rho + KM$$

subject to

$$\begin{aligned} EU &= \int_{\Gamma_X} u_N(x)dF - a \geq \underline{U} \\ EU_a &= \int_{\Gamma_X} u_N(x)dF_a + M\int_{\Gamma_Y} \omega dG_a - 1 \geq 0. \end{aligned}$$

Clearly, the only feature of audits that matters here is their intensity M.

Now, let φ''' be negative (the treatment of $\varphi''' > 0$ is symmetric).[6] This time the decision-maker exhibits *prudence* (resp is *non-prudent*). When having to face a mean-preserving additional risk, a prudent decision-maker prefers to see it attached to the best rather than the worst outcomes (see Eeckhoudt et al. (1995)). At the previous solution ($u_N(x) = u_A(x)$, and $\omega(x, Y)$ invariant with respect to x), she would thus rather go for $m(x)$ larger when x is higher and $m(x)$ smaller when x is lower. This suggests than an optimal audit might now be upper-tailed. Moreover, prudence together with (14) implies that the premium ρ must decrease with u_A (see Kimball (1990), and Hartwick (1999)), and that

$$E_Y[\varphi'(u_A(x) + \omega(x, Y))] - \varphi'(u_A(x)) < 0.$$

When being offered a slight increase in $u_A(x)$ that keeps $(1 - m)u_N + mu_A$ constant the decision-maker would therefore depart from any proposal in which $u_N(x) \geq u_A(x)$ and $0 < m(x) < 1$, for this alternative offer entails that

$$\begin{aligned} dEC(x) &= (1 - m)\varphi'(u_N)du_N + m[E_Y[\varphi'(u_A(x) + \omega(x, Y))]du_A \\ &= m\{E_Y[\varphi'(u_A(x) + \omega(x, Y))] - \varphi'(u_N)\}du_A < 0. \end{aligned}$$

This suggests in turn that one should see $u_A(x) > u_N(x)$ at the optimum, i.e. an audit

[6] The sign of φ is negative, positive or zero when, for instance, the agent's utility function shows constant relative risk aversion (CRRA) respectively lower than, greater than, or equal (as in the above example) to 1/2. More generally, $\varphi'''(.) < (> \text{ or } =)\ 0$ if and only if $P > (< \text{ or } =)\ 3R$, where $P = \dfrac{-u'''}{u''}$ is the agent's coefficient of absolute prudence, as defined and interpreted in Kimball (1990), and $R = \dfrac{-u''}{u'}$ that of absolute risk aversion.

would accordingly constitute a carrot rather than a stick from the agent's viewpoint (and conversely audit would be perceived as a stick when $\varphi''' > 0$).

Let us now consider the information system generated by an auditing policy $m(X)$ of intensity M. Let L^m denote the likelihood ratio associated with such a policy. Clearly, the event $\{L^m \leq l\}$ is the same as

$$\{L_X(X,a) \leq l \text{ and there is no audit}\} \cup \{L_X(X,a) + L_Y(Y,a) \leq l \text{ and an audit occurs}\} .$$

The cumulative distribution $\Phi_m(\cdot)$ of L^m is therefore given by

$$(16)\quad \Phi_m(l) = \Pr ob(L^m \leq l) = \int_{\Gamma_X} (1-m)\delta(l-L_X)dF + \int_{\Gamma_X}\int_{\Gamma_Y} m\delta(l-L_X-L_Y)dF dG \ ,$$

where $\delta(z) = 1$ as long as $z \geq 0$, and $\delta(z) = 0$ otherwise. The first and second moments of this distribution are respectively

$$(17)\qquad E(L^m) = 0 \text{ and } Var(L^m) = I_X(a) + MI_Y(a) \ ,$$

so the variance of the likelihood ratio distribution depends only on auditing intensity and the Fisher information indices associated with X and Y. This supports the following remark.

Remark 2: Two distinct contingent auditing policies that have the same intensity cannot be ranked according to the mean preserving spread (MPS) relation between their respective likelihood distributions.

The MPS criterion is thus rather ineffective when one deals with the information systems associated with various contingent auditing policies. The upcoming section will now develop a suitable refinement of this criterion.

4. A General Ranking Criterion

The previous discussion and the outcome arrived at in (17) suggest that the design of optimal contingent audits might finally come to performing some mean *and* variance-preserving transformations (MVPT) of likelihood ratio distributions. According to previous works on such manipulations of probability distributions (see Menezes et al. (1980), for instance), this would mean that in order to compare the obtained information systems, and thereby supplement the MPS criterion, one should now invoke stochastic dominance of the third order. The following subsection thus recalls briefly the notions of third and n^{th}-order stochastic dominance, their relationships and some useful implications. Subsection 4.2 and section 5 will next substantiate our current intuition.

4.1. *Stochastic Dominance Orderings*

Let X and Y be two random variables with corresponding distributions functions $F(x)$ and $G(y)$ and densities $f(x)$ and $g(y)$ which are strictly positive on the interval (a, b). These distributions are often compared according to some probability-weighted function of deviations below an arbitrary target. This approach yields to the following partial orderings of probability distributions.

DEFINITION: We say that X stochastically dominates Y in the n^{th} order, noted $X \gtrsim_n Y$, if for all $t \in [a, b]$ we have that $\int_a^t (t-z)^{n-1}\{f(z) - g(z)\}dz \leq 0$, the inequality being strict on a subset of (a, b) of positive measure.

Equivalent forms of this definition are often useful. In order to concisely state one of them, let us introduce some new notation. For Z a random variable taking values in (a, b) and having a distribution function H with density h, let us write $H^{(0)}(z) = h(z)$, $H^{(1)}(z) = H(z)$, and $H^{(n)}(z) = \int_a^z H^{(n-1)}(r)dr$. The following identity can now be derived (via straightforward integration by parts):

$$\int_a^t (t-z)^{n-1}\{f(z) - g(z)\}dz = (n-1)![F^{(n)}(t) - G^{(n)}(t)]. \tag{18}$$

It follows that $X \gtrsim_n Y$ if and only if $F^{(n)}(t) - G^{(n)}(t) \leq 0$, the inequality being strict on a subset of (a, b) with positive measure.

It can be shown that

$$X \gtrsim_n Y \quad \text{implies that} \quad X \gtrsim_{n+1} Y \quad , \tag{19}$$

where the converse is obviously not true. Hence, third-order stochastic dominance in particular provides a finer ordering than second and first-order stochastic dominance.

Furthermore, consider an individual with Von Neumann-Morgenstern utility index $u : [a, b] \to R$. By definition, she prefers strictly a lottery with prizes X to another one with prizes given by Y if

$$\int_a^b u(x)[f(z) - g(z)]dz > 0.$$

After integrating the left-hand side by parts successively three times, we get

$$\begin{aligned}
\int_a^b u(x)[f(x) - g(x)]dx &= -\int_a^b u'(x)[F(x) - G(x)]dx \\
&= -u'(b)[F^{(2)}(b) - G^{(2)}(b)] + \int_a^b u''(x)[F^{(2)}(x) - G^{(2)}(x)]dx \\
&= -u'(b)[F^{(2)}(b) - G^{(2)}(b)] + u''(b)[F^{(3)}(b) - G^{(3)}(b)] \\
&\quad - \int_a^b u'''(x)[F^{(3)}(x) - G^{(3)}(x)]dx\ .
\end{aligned}$$

The upcoming assertions which concern the relationship between stochastic dominance and decision-making are now a direct consequences of the above.[7]

REMARK 3: (i) If $u'(\cdot) > 0$ and $X \gtrsim_1 Y$, then X is strictly preferred to Y. (ii) If $u'(\cdot) > 0$, $u''(\cdot) < 0$, and $X \gtrsim_2 Y$, then X is strictly preferred to Y. (iii) If $E(X) = E(Y)$, $u''(\cdot) < 0$, and $X \gtrsim_2 Y$, then X is strictly preferred to Y. (iv) If $X \gtrsim_1 Y$, then $E(X) > E(Y)$. (v) If $X \gtrsim_2 Y$, then $E(X) \geq E(Y)$. (vi) If $u''(\cdot) \leq 0$, $u'''(\cdot) > 0$, $u'(b)[E(X) - E(Y)] \geq 0$, and $X \gtrsim_3 Y$, then X is strictly preferred to Y.

As we implicitly pointed out in the previous sections, there are also some important linkages between stochastic dominance and the notion of variance. By definition, the

[7] Notice that $E(X) = \int_a^b xf(x)dx = b - \int_a^b F(x)dx = b - F^{(2)}(b)$.

difference between the variance of X and that of Y is given by

$$V(X) - V(Y) = \int_a^b x^2[f(x) - g(x)]dx - E(X)^2 + E(Y)^2.$$

However, note that

$$\int_a^b x^2[f(x) - g(x)]dx = -2b[F^{(2)}(b) - G^{(2)}(b)] + 2[F^{(3)}(b) - G^{(3)}(b)].$$

Hence,

$$V(X) - V(Y) = 2[F^{(3)}(b) - G^{(3)}(b)] + [F^{(2)}(b) - G^{(2)}(b)][F^{(2)}(b) + G^{(2)}(b)] \quad . \tag{20}$$

One can now draw the following conclusions.

REMARK 4: (i) If $X \gtrsim_2 Y$, then $V(X) < V(Y)$. (ii) If $X \gtrsim_3 Y$ and $E(X) \geq E(Y)$, then $V(X) \leq V(Y)$. (iii) If $E(X) = E(Y)$, $V(X) = V(Y)$, $u'''(\cdot) > 0$, and $X \gtrsim_3 Y$ ($Y \gtrsim_3 X$), then X is strictly preferred to Y (Y is strictly preferred to X).

Finally, an appealing characteristic of the notions of second and third-order stochastic dominance is that they are "constructive" in the sense that, if two distributions can be compared using those rankings, then one can be obtained from the other through a finite number of straightforward manipulations involving mean-preserving spreads (i.e. transfers of the probability mass from the center to the tails without changing the mean) and mean-preserving contractions (i.e. transfers of the probability mass from the tails to the center without changing the mean). Our last remark constitute a formal statement of this practical feature.

REMARK 5: (i) (Rothschild and Stiglitz (1970)) If $E(X) = E(Y)$, then $X \gtrsim_2 Y$ if and only if $G(\cdot)$ can be obtained from $F(\cdot)$ via a mean-preserving spread (MPS). (ii) (Menezes et al. (1980)) If $E(X) = E(Y)$ and $V(X) = V(Y)$, then $X \gtrsim_3 Y$ if and only if $G(\cdot)$ can be obtained from $F(\cdot)$ via a mean and variance-preserving transformation (MVPT).

The first part is well-known and has been widely used throughout the economics and decision-theoretic literatures. Lesser known and apparently much less intuitive is the second part. According to Menezes et al. (1980), however, an MVPT is actually just a combination of mean-preserving spreads and mean-preserving contractions.

4.2 *Comparing Likelihood Ratio Distributions*

With the previous background and discussion, we are now ready to state and prove a general ranking criterion for the information systems arising generally in principal-agent problems. This criterion is also based on likelihood ratio distributions. The MPS criterion is thereafter derived as a special version of it.

PROPOSITION 3: *The principal prefers a signal T to a signal Z to implement a given action a if $E_T[\Delta^*(\lambda_T + \mu_T L_T)] \geq E_Z[\Delta^*(\lambda_T + \mu_T L_Z)]$, where λ_T and μ_T are the multipliers of the participation and the incentive constraints which appear in the principal-agent problem with signal T.*

Proof. Let Γ_i , $H(t,a,i)$, $h(t,a,i)$, and L_i denote the support, distribution function, density function, and likelihood ratio associated with signal $i = T, Z$. The corresponding objective, participation constraint, and incentive compatibility contraint of the principal-agent problem are now respectively written:

$$\overline{EC}_i = \int_{\Gamma_i} w(t)dH(t,a,i) \tag{21}$$

$$\int_{\Gamma_i} u(w(t))dH(t,a,i) - a \geq \underline{U} \tag{22}$$

$$\int_{\Gamma_i} u(w(t))dH_a(t,a,i) \geq 1. \tag{23}$$

The Lagrangian function associated with this problem is

$$\Lambda_i = -\int_{\Gamma_i} w(t)dH(t,a,i) + \lambda_i\{\int_{\Gamma_i} u(w(t))dH(t,a,i) - a - \underline{U}\} + \mu_i\{\int_{\Gamma_i} u(w(t))dH_a(t,a,i) - 1\},$$

or equivalently

$$\Lambda_i = E_i[\Delta(w, \lambda_i + \mu_i L_i)] - \lambda_i(a + \underline{U}) - \mu_i \ . \tag{24}$$

>From the necessary optimality conditions, we know that there exist some nonnegative multipliers λ_i and μ_i such that the wage schedule $w_i(\cdot)$ maximizes $\Delta(w, \lambda_i + \mu_i L_i)$ and

the following equations are satisfied:

$$\lambda_i\{\int_{\Gamma_i} u(w_i(t))dH(t,a,i) - a - \underline{U}\} = \mu_i\{\int_{\Gamma_i} u(w_i(t))dH(t,a,i) - 1\} = 0 \ . \tag{25}$$

The principal now prefers the information system generated by signal T to the one generated by signal Z if using the former is cheaper, that is if $\overline{EC}_Z^* - \overline{EC}_T^* \geq 0$. At an optimum, we have that

$$\Lambda_i^* = E_i[\Delta^*(\lambda_i + \mu_i L_i)] - \lambda_i(a + \underline{U}) - \mu_i \leq E_i[\Delta(w_i, \lambda + \mu L_i)] - \lambda(a + \underline{U}) - \mu$$

for any $\lambda \geq 0$ and μ, and

$$\Lambda_T^* - \Lambda_Z^* = \overline{EC}_Z^* - \overline{EC}_T^* \ . \tag{26}$$

It follows that

$$\begin{aligned} \overline{EC}_Z^* - \overline{EC}_T^* &\geq E_T[\Delta^*(\lambda_T + \mu_T L_T)] - E_Z[\Delta(w_Z, \lambda_T + \mu_T L_Z)] \\ &\geq E_T[\Delta^*(\lambda_T + \mu_T L_T)] - E_Z[\Delta^*(\lambda_T + \mu_T L_Z)] \ . \end{aligned} \tag{27}$$

Hence, the principal selects signal T over signal Z to implement an action a whenever $E_T[\Delta^*(\lambda_T + \mu_T L_T)] \geq E_Z[\Delta^*(\lambda_T + \mu_T L_Z)]$, as claimed. (Note that, in this model, the multiplier μ_T is strictly positive. ■

The following assertion, which is a restatement of Kim (1995)'s proposition 1, is now a direct consequence of the proposition.

COROLLARY 1 (MPS criterion): *The information system from a signal T is preferred by the principal to the one from a signal Z if the likelihood ratio distribution of T is a mean-preserving spread of the likelihood ratio distribution of Z, that is if $L_Z \gtrsim_2 L_T$.*

PROOF: By definition, $\Delta^*(\sigma) = \max_{w \in W} \Delta(w, \sigma)$ where $\Delta(w, \sigma)$ is a linear function of σ. As a consequence,

$$
\begin{aligned}
\Delta^*(\lambda\sigma_0 + (1-\lambda)\sigma_1) &= \lambda\Delta(w(\lambda\sigma_0 + (1-\lambda)\sigma_1), \sigma_0) + (1-\lambda)\Delta(w(\lambda\sigma_0 + (1-\lambda)\sigma_1), \sigma_1) \\
&\leq \lambda\Delta^*(\sigma_0) + (1-\lambda)\Delta^*(\sigma_1) \ ,
\end{aligned}
$$

so $\Delta^*(\cdot)$ is a convex function.[8] Since

$$E_K[\Delta^*(\lambda_T + \mu_T L_K)] = E_{L_K}[\Delta^*(\lambda_T + \mu_T L_K)], \text{ for } K = T, Z,$$

the statement follows from Remark 3(ii). ■

Using Remark 3(vi), furthermore, an additional ranking criterion is now also available, which is based on stochastic dominance of the third order.

COROLLARY 2: *Let $\Delta^{*'''} \geq (\leq) 0$. The information system from T dominates from the principal's viewpoint that from a signal Z when $L_T \gtrsim_3 L_Z \quad (L_Z \gtrsim_3 L_T)$.*

[8] The reader might have noticed that Δ^* is actually the mathematical conjugate of φ. And the conjugate function of a convex function is itself convex (Rockafellar (1970)).

In practice, the sign of the third derivative $\Delta^*(\cdot)$ could be inferred from the sign of the third derivative of $\varphi(\cdot)$. By the envelope theorem $\Delta^{*\prime}(\sigma) = u(w(\sigma))$, where $w(\sigma)$ satisfies $u'(w)\sigma = 1$ or equivalently $\varphi'(u(w)) = \sigma$. This entails that $\Delta^{*\prime}(\sigma) = \varphi'^{-1}(\sigma)$, and so

$$\varphi'''(\cdot) > 0 \Longleftrightarrow \Delta^{*'''}(\cdot) < 0 \quad . \tag{28}$$

The latter brings Corollary 2 closer to Proposition 2 of the preceding section, which pertains to optimal auditing policies. The actual linkage will be spelled out in the upcoming section.

5. Comparing audit-generated information systems

In order to use our general ranking criterion, we need a statement that precisely relates the design of auditing policies to some stochastic ordering of the implied information systems. This is the purpose of our last proposition, which proof can be found in the Appendix.

Proposition 4: *Let $m(X)$ and $\widehat{m}(X)$ be some auditing policies with the same intensity. If for any $x \in \Gamma_X$ we have that $\int_{Inf X}^{x} m dF \geq \int_{Inf X}^{x} \widehat{m} dF$ - the inequality being strict for a set of positive measure, then $L^{\widehat{m}} \gtrsim_3 L^m$.*

Thanks to this result, it is now possible to compare and rank any two contingent auditing policies that have the same frequency. Let $m_{UT}(X)$, $m_{LT}(X)$, and $m(X)$ represent respectively an upper-tailed, a lower-tailed, and a random auditing policy. The proposition says that

$$L^{m_{UT}} \gtrsim_3 L^m \gtrsim_3 L^{m_{LT}} \quad . \tag{29}$$

According to Corollary 2, the principal therefore prefers m_{UT} when $\varphi''' < 0$ and m_{LT} when $\varphi''' > 0$, which is consistent with Remark 4(iii) and corroborates Proposition 2.

Furthermore, starting from a given auditing policy, the principal can enhance the efficiency of her information system by auditing more intensively the highest (resp. the lowest) values of signal X and less intensively the lowest (resp. the highest) ones when $\varphi''' < 0$ (resp. $\varphi''' > 0$).

This result finally hints at a procedure for setting up an optimal auditing policy (given, of course, some auditing technology).

· First, determine the appropriate frequency of audits. This would involve standard considerations of risk sharing and incentives, taking into account the unsunk cost of auditing.

· Secondly, determine what instances of the signal X would trigger an audit. The agent's prudence (or the way his risk attitude changes when his wealth varies) would now be relevant, and via the construction pointed out in Remark 5(ii) the principal might then consider making the agent's compensation more sensitive to the observables when these are higher or lower.

6. Conclusion

The literature on moral hazard has so far exclusively emphasized the tradeoff between incentives and insurance. The above analysis indicates, however, that the design of optimal auditing policies must also take into account the agent's prudence, for it amounts finally to examine mean and variance-preserving transformations of given information systems.

This general insight could be useful in various context. In multi-tasking, for instance (provided the current one-dimensional analysis can be extended to this context), submitting hard-to-appraise activities to audits triggered by the observation of high performance on the more staightforward ones might alleviate incentive problems (Sinclair-Desgagné (1999)). Prudence and MVPT should also play a role in setting optimal contracts within repeated principal-agent relationships (Rogerson (1985)) or in agencies submitted to background risk (Gollier and Pratt (1995)).

Université de Rouen and CREST-LEI, Rue des Saints-Pères, 75007 Paris, France.

CIRANO and Institut d'Économie Appliquée, HEC Montréal, Canada H3T 2A7.

Appendix

Proof of proposition 1: For the sake of this proof, let us abuse notation and denote respectively $ET(K)$ and $M(K)$ the expected optimal transfer and the intensity of an optimal auditing policy at a given effort level a, when the unit cost of an audit is K. At different cost levels K and K', the principal's objective function would be such that $ET(K) + M(K)K \leq ET(K') + M(K)K'$. Similarly, reversing the respective roles of K by K' also gives $ET(K') + M(K')K' \leq ET(K) + M(K')K$. The sum of these two inequalities yields $(K - K')[M(K) - M(K')] \leq 0$. Accordingly, the intensity of an optimal audit must decrease with K.

To prove the second part of the proposition, observe that according to (6) the term within brackets in (7) is precisely

$$E_Y[\Delta^*(\lambda + \mu L_X + \mu L_Y)] - \Delta^*(\lambda + \mu L_X) - K.$$

So we have that

$$E_Y[\Delta^*(\lambda + \mu L_X + \mu L_Y)] \geq E_Y[\Delta(w(x), \lambda + \mu L_X + \mu L_Y)] = \Delta^*(\lambda + \mu L_X)$$

(the inequality being strict at an interior solution), and the bracketed term in (7) is always nonnegative when $K = 0$. ■

Proof of proposition 2: Let $u_N(x) = u(w(x))$ and $u_A(x, y) = u(s(x, y))$. First, equations (8) and (9) become respectively

$$\varphi'(u_N(x)) = \lambda + \mu L_X(x, a) \text{ and } \varphi'(u_A(x, y)) = \lambda + \mu L_X(x, a) + \mu L_Y(y, a),$$

which implies that

$$E_Y[\varphi'(u_A(x,Y))] = \lambda + \mu L_X(x,a) = \varphi'(u_N(x)). \tag{30}$$

On the other hand, applying the Envelope Theorem gives $\Delta^{*\prime}(\sigma) = u(w(\sigma))$ when $w(\sigma) = Arg\max_w \Delta(w,\sigma)$; and it follows that the function $G(\cdot)$ defined in (6) has a first-order derivative at $L_X(x,a)$ which is given by

$$G'(L_X(x,a)) = E_Y[u_A(x,Y)] - u_N(x). \tag{31}$$

The combination of (30) and (31) allows us now to conclude that:

- When $\varphi'''(.) > 0$ so φ' is convex, we have that $\varphi'(u_N(x)) = E_Y[\varphi'(u_A(x,Y))] > \varphi'(E_Y([u_A(x,Y)])$. Consequently, $u_N(x) > E_Y[u_A(x,Y)]$ and $G(\cdot)$ is a decreasing function, so the optimal auditing policy must be lower-tailed.
- When $\varphi'''(.) < 0$ so φ' is concave, conversely, $\varphi'(u_N(x)) = E_Y[\varphi'(u_A(x,Y))] < \varphi'(E_Y[u_A(x,Y)])$. In this case $u_N(x) < E_Y[u_A(x,Y)]$ and $G(z)$ increases with z, so the optimal auditing policy is upper-tailed.
- When $\varphi'''(.) = 0$ so φ' is a linear function, finally, then $\varphi'(u_N(x)) = E_Y[\varphi'(u_A(x,Y))] = \varphi'(E_Y[u_A(x,Y)])$. In this case, $u_N(x) = E_Y[u_A(x,Y)]$ and the agent is indifferent between being audited or not, so the principal will select any auditing policy that has the appropriate intensity. ■

Proof of proposition 4: Denote as $\Gamma_L = [Inf(\frac{f_a}{f}+\frac{g_a}{g}), Sup(\frac{f_a}{f}+\frac{g_a}{g})]$ the common support of the distributions Φ_m and $\Phi_{\widehat{m}}$ corresponding to the likelihood ratios L^m and $L^{\widehat{m}}$.

By definition, $L^{\widehat{m}} \gtrsim_3 L^m$ means that

$$(\upsilon) = \int_{\Gamma_L}\int_{\Gamma_L}[\Phi_m(z) - \Phi_{\widehat{m}}(z)]\delta(t-z)\delta(\upsilon - t)dzdt \geq 0$$

for any $\upsilon \in \Gamma$, the inequality being strict on a subset of Γ_L with positive measure. Applying Fubini's theorem and using a little algebra then yields

$$\begin{aligned}(\upsilon) &= \int_{\Gamma_L}[\Phi_m(z) - \Phi_{\widehat{m}}(z)]\{\int_{\Gamma_L}\delta(t-z)\delta(\upsilon - t)dt\}dz \\ &= \int_{\Gamma_L}[\Phi_m(z) - \Phi_{\widehat{m}}(z)]Max(\upsilon - z, 0)dz.\end{aligned}$$

Meanwhile, the expression for $\Phi_m(z)$ given by (16) gives:

$$\Phi_m(z) - \Phi_{\widehat{m}}(z) = \int_{\Gamma_X}(\widehat{m} - m)\delta(z - L_X)dF - \int_{\Gamma_X}\int_{\Gamma_Y}(\widehat{m} - m)\delta(z - L_X - L_Y)dFdG.$$

So we have that

$$\begin{aligned}(\upsilon) &= C - D\ , \quad \text{where} \\ C &= \int_{\Gamma_L}\int_{\Gamma_X} Max(\upsilon - z, 0)(\widehat{m} - m)\delta(z - L_X)dzdF \text{ and} \\ D &= \int_{\Gamma_L}\int_{\Gamma_X}\int_{\Gamma_Y} Max(\upsilon - z, 0)(\widehat{m} - m)\delta(z - L_X - L_Y)dzdFdG\ \ .\end{aligned}$$

Invoking Fubini's theorem again, the latter can be written as

$$
\begin{aligned}
C &= \int_{\Gamma_X} (\widehat{m} - m)\{\int_{\Gamma_L} Max(v - z, 0)\delta(z - L_X)dz\}dF \\
D &= \int_{\Gamma_X}\int_{\Gamma_Y} (\widehat{m} - m)\{\int_{\Gamma_L} Max(v - z, 0)\delta(z - L_X - L_Y)dz\}dFdG.
\end{aligned}
$$

And since $\int_{z \geq z} Max(v - z, 0)dz = (1/2)(Max(v - z, 0)^2$, (v) is finally written as

$$
\begin{aligned}
(v) &= (1/2)\{\int_{\Gamma_X} (\widehat{m} - m)[Max(v - L_X, 0)]^2 dF \\
&- \int_{\Gamma_X}\int_{\Gamma_Y} (\widehat{m} - m)[Max(v - L_X - L_Y, 0)]^2 dFdG\} \\
&= (1/2)\int_{\Gamma_X} (m - \widehat{m})\Psi(v - L_X)dF \quad , \qquad (32)
\end{aligned}
$$

where the function $\Psi(\cdot)$ is defined as

$$\Psi(t) = E_Y[Max(t - L_Y, 0)^2] - Max(t, 0)^2. \qquad (33)$$

Note that $\Psi(\cdot)$ is a differentiable function because the derivative of $Max(C, 0)^2$ exists and is equal to $2Max(C, 0)$. Therefore,

$$\Psi'(t) = 2E_Y[Max(t - L_Y, 0)] - 2Max(t, 0) \geq 0,$$

which entails that $\Psi(\cdot)$ is strictly increasing on $]Inf\frac{g_a}{g}, Sup\frac{g_a}{g}[$ and is constant elsewhere. In particular, $\Psi'(0) > 0$.

The right-hand of (32) can now be integrated by parts, which yields

$$(34) \qquad (\upsilon) = \int_{\Gamma_X} \{ \int_{InfX}^{x} (m(z) - \widehat{m}(z)) dF(z,a) \} \Psi'(\nu - L_X) (\frac{\partial L_X}{\partial x}) dx \ .$$

We conclude that: (i) if $\int_{InfX}^{x} m(z)dF(z,a) \geq \int_{InfX}^{x} \widehat{m}(z)dF(z,a)$ for any x, then $(\upsilon) \geq 0$, and (ii) if $\int_{InfX}^{x} m(z)dF(z,a) > \int_{InfX}^{x} \widehat{m}(z)dF(z,a)$ on some interval $[c,d]$, then letting $\upsilon = L_X(d,a)$ we get

$$(\upsilon) \geq \int_{c}^{d} \{ \int_{InfX}^{x} (m(z) - \widehat{m}(z)) dF(z,a) \} \Psi'(L_X(d,a) - L_X(x,a)) (\frac{\partial L_X}{\partial x}) dx > 0$$

in a neighbourhood of $L_X(d,a)$. In both cases, if $\int_{InfX}^{x} m(z)dF(z,a) \geq \int_{InfX}^{x} \widehat{m}(z)dF(z,a)$ with strict inequality on a subset of positive measure, then $(\upsilon) \geq 0$, as claimed. ∎

References

Baiman, S., and J. Demski (1980): "Economically Optimal Performance Evaluation and Control Systems," *Journal of Accounting Research*, 18, 184-220.

Baker, G. P. (1992): "Incentive Contracts and Performance Measurement," *Journal of Political Economy*, 100, 598-614.

Demougin, D., and C. Fluet (2000): "Ranking of information systems in agency models: an integral condition," *Economic Theory*, 15,

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

Dye, R. A. (1986): "Optimal monitoring Policies in Agencies," *The Rand Journal of Economics*, 17, 339-350.

Eeckhoudt, L., C. Gollier, and T. Schneider (1995): "Risk-Aversion, Prudence and Temperance: A Unified Approach," *Economics Letters*, 49, 331-336.

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

Gollier, C., and J. W. Pratt (1995): "Risk Vulnerability and the Tempering Effect of Background Risk," *Econometrica*, 1109-1123.

Gouriéroux, C., and A. Monfort (1989): *Statistiques et Modèles Économétriques.* Paris: Éditions Économica.

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

Hartwick, J. M. (1999): "Insuring and $u'''(y)$," *Economic Letters*, 65, 205-212.

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

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

Jewitt, I. (1997): "Information and Principal-Agent Problems," *Mimeo*, University of Bristol.

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

Kimball, M. S. (1990): “Precautionary Saving in the Small and in the Large,” *Econometrica*, 58, 53-73.

Lambert, R. A. (1985): “Variance Investigation in Agency Settings,” *Journal of Accounting Research*, 23, 633-647.

Menezes, C., C. Geiss, and J. Tressler (1980): “Increasing Downside Risk,” *American Economic Review*, 70, 921-932.

Rothschild, M. and J. Stiglitz (1970): “Increasing Risk: I. A Definition,” *Journal of Economic Theory*, 2, 225-243.

Rockafellar, R. T. (1970): *Convex Analysis.* Princeton, New Jersey: Princeton University Press.

Rogerson, W. (1985): “Repeated Moral Hazard,” *Econometrica*, 53, 69-76.

Sinclair-Desgagné B. (1999): “How to Restore Higher-Powered Incentives in Multitask Agencies,” *Journal of Law, Economics and Organization*, 15, 418-433.

Sinclair-Desgagné B. (1994): “The First-Order Approach to Multi-Signal Principal-Agent Problems,” *Econometrica*, 62, 459-465.

Young, R. A. (1986): “A Note on ”Economically Optimal Performance Evaluation and Control Systems”: The Optimality of Two-Tailed Investigations,” *Journal of Accounting Research*, 24, 231-240.