Auditing Policies and Information*

Marie-Cécile Fagart and Bernard Sinclair-Desgagné

February 2003

We first point out that, using any of the current criteria for comparing information systems in principal-agent models with moral hazard (such as Kim (1994)'s MPS criterion), it is often impossible to contrast the value of information obtained from different policies of contingent audits that bear the same cost. Given two such policies A and B where, say, the lower cumulated frequencies of audits are always larger under B than under A, we show, however, that the likelihood ratio distribution associated with A dominates the one associated with B in the *third order*. A new, strictly finer, ranking of information systems then implies that the value of information is greater under A than under B when the agent's negative inverse utility function exhibits some *prudence*. The practical upshot is that the design of auditing policies involves somewhat more than the classical tradeoff between risk-sharing and incentives; it also requires to balance incentives and *downside risk*.

KEYWORDS: Principal-agent, moral hazard, value of information, likelihood ratio distributions, third-order stochastic dominance, prudence.

* Fagart: CREST-LEI, 28 Rue des Saints-Pères, 75007 Paris, France; fagart@ensae.fr. Sinclair-Desgagné (corresponding author): CIRANO and HEC, Université de Montréal, 3000 Côte-Ste-Catherine, Montréal (Québec), Canada H3T 2A7; bsd@hec.ca. We are grateful to Jacques Crémer, Glenn Ellison, Claude Fluet, Christian Gollier, Jean-Pierre Ponssard, Patrick Rey, François Salanié and, most of all, Louis Eeckhoudt for encouragements and valuable discussions on the topic of this paper. We also acknowledge useful comments from seminar participants at Nottingham, IDEI-Toulouse, HEC Montréal, and the University of Paris (ETAPE).

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; its introduction

[1] The MPS criterion says that (assuming the first-order approach to the considered principal-agent problem 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.

also constituted a radical improvement, for it allowed comparisons between information systems which are not necessarily nested.

One significant group (both in theory and practice) of information systems, however, largely eludes the MPS criterion: those which are induced by contingent auditing policies. An intuitive explanation of this fact would be the following. Several economically plausible auditing policies are, for instance, upper-tailed or lower-tailed (see, e.g., Baiman and Demski (1980), Dye (1986), Jewitt (1988), and Sinclair-Desgagné (1999)), i.e. they prescribe that audits be triggered only by the observation of respectively high or low signals.[2] A rational principal who seeks to bring about a given action by the agent would then typically have to 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 , L_Y are two independent likelihood ratios. Yet, the respective likelihood ratio distributions associated with A and B clearly have the same mean (since both L_X and L_Y have mean zero) and variance, so neither is a

[2] Baiman and Demski (1980) first found, assuming that the agent's utility function belongs to the HARA (hyperbolic absolute risk aversion) family, that optimal audits might often be upper-tailed or lower-tailed. In a more general setting, Dye (1986) and Jewitt (1988) have next characterized the situations where optimal audits are *lower-tailed* (Jewitt (1988) also provided sufficient assumptions for the first-order approach to be valid). More recently, Sinclair-Desgagné (1999) showed that *upper-tailed* contingent audits can help raising the power of incentives in a multitasking context.

Of course, optimal contingent audits do not need to be upper or lower-tailed. Lambert (1985) and Young (1986), for instance, provide examples where they are in fact *two-tailed* - i.e. triggered only by the observation of high or low signals. Our main results (Theorems 1 and 2) can also cope with such policies.

mean-preserving spread of the other.

The objective of this paper is now to develop a ranking which is consistent with the MPS criterion and allows to better select among audit-generated information systems.

The upcoming section lays out a standard principal-agent model with audits. Section 3 next contains our first key result: consider two auditing policies A and B that have the same expected frequency of audits but where the lower cumulated frequencies are always larger under B than under A, then the likelihood ratio distribution associated with A dominates the one associated with B in the *third* order. This conclusion means that implementing a new contingent auditing policy without changing the overall expected frequency of audits amounts to making mean *and* variance-preserving transformations of the actual information system (see Menezes et al. (1980)). It suggests, furthermore, that choosing among various auditing policies by comparing the resulting information systems should still be possible, provided a suitable generalization of the MPS criterion is made available. Such a generalization is introduced in section 4. It allows indeed to discriminate between A and B, if the sign of the third derivative of the agent's inverse utility function is constant. Further implications in more general settings - respectively where the cost of audits may vary and when monitoring and auditing signals can be correlated - are explored in section 5. All these developments suggest, finally, that in designing a contingent auditing policy one must not only weigh the agent's incentives and overall risk bearing, but also the agent's exposure to *downside risk*. Some conjectures arising from this practical remark are briefly stated in the concluding section 6.

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 (until section 5) 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 $f(x, a)$ and $g(y, a)$ that exhibit constant supports (noted Γ_X and Γ_Y) and are twice continuously differentiable in a for all x, y.

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 $L_Y(y, a) = \dfrac{g_a(y, a)}{g(y, a)}$.[3] A standard assumption is that these ratios share the *Monotone Likelihood Ratio Property* (MLRP), that is: $L_X(x, a)$ and $L_Y(y, a)$ increase in x and in y respectively, for every a. Clearly, L_X and L_Y are themselves random variables, and their respective distribution - called a *likelihood ratio distribution* - constitutes a formal representation of an *information system*.[4] 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.[5]

[3] Throughout this paper the subscript $_a$ refers to the partial derivative with respect to a.

[4] Actually, it is the density functions f and g themselves which are usually interpreted as information systems. But since there is a one-one relationship between these and their associated likelihood ratio distributions, we deem that also calling the latter an information system will not create confusion.

[5] The Fisher information index is well-known to statisticians and econometricians (see Gouriéroux and Monfort (1989), for example). Note that $E_X[(L_X)^2] = E_X[-\frac{\partial L_X}{\partial a}]$, so this index measures the sensitivity

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

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

The latter integral $M(a) = \int_{\Gamma_X} m(x) dF(x, a)$ gives the expected probability of an audit (or the *overall frequency*, or the *intensity* of audits) under a *policy* $m(X)$.

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 given by

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

of the likelihood ratio with respect to a (or the informational content of X about a). For a compelling illustration of the usefulness of this index in principal-agent analyses, see Dewatripont et al. (1999).

In the upcoming sections, we let $\varphi = u^{-1}$ denote the inverse of $u(\cdot)$.

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. Formally, this amounts to minimize (1), subject to participation and incentive compatibility constraints given respectively by

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

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

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:

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

We assume that this so-called "first-order approach" always yields a solution that constitutes an incentive compatible allocation (hence which solves the initial problem as well).[6]

[6]Suitable sufficient conditions for the validity of the first-order approach in the present context - conditions that do not put further *a priori* restrictions on the agent's utility function - can be found in Sinclair-Desgagné (1994).

3. SORTING INFORMATION SYSTEMS

If someone wants to distinguish among various auditing policies, at least two natural features come to mind. First, a policy exhibits an overall intensity M, which determines its total cost $K \cdot M$. Second, a contingent auditing policy m_A can be described as *relatively more upper-tailed* than another contingent policy m_B if the latter exhibits larger cumulated auditing frequencies (or larger *downside intensities*) than the former, that is: $\int_{Inf\Gamma_X}^{x} m_B dF \geq \int_{Inf\Gamma_X}^{x} m_A dF$ for all x. The question we now ask is whether those features can be used to infer some characteristics of the induced information systems.

Accordingly, consider a contingent auditing policy $m(X)$ of intensity M and associated likelihood ratio L^m. Clearly, the event $\{L^m \leq l\}$ is the same as[7]

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

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

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

where $\delta(z) = 1$ as long as $z \geq 0$, and $\delta(z) = 0$ otherwise. And the first and second

[7] The summation in the second subset comes immediately from the separability of the joint distribution of X and Y. For the likelihood ratio associated with this joint distribution is precisely

$$L_{X,Y}(x,y,a) = \frac{\partial[f(x,a)g(y,a)]/\partial a}{f(x,a)g(y,a)} = \frac{f_a(x,a)g(y,a)+f(x,a)g_a(y,a)}{f(x,a)g(y,a)} = L_X(x,a) + L_Y(y,a).$$

moments of this distribution are respectively[8]

$$E(L^m) = 0 \text{ and } Var(L^m) = I_X(a) + M(a)I_Y(a)\ . \tag{6}$$

A significant implication of (6) is that two contingent auditing policies (be they relatively upper-tailed, relatively lower-tailed, two-tailed, or totally random, for instance) that share the same intensity will generate likelihood ratio distributions with the same variance. It follows that Kim (1994)'s MPS criterion is often not helpful to differentiate the information systems correxponding to different policies of audits: two policies $m_A(X)$ and $m_B(X)$ where $M_A = M_B$ will have associated likelihood ratio distributions so that noone can yield the other via some mean-preserving spread of probability mass.

But what about mean *and* variance-preserving transformations? Recall that, while classifying probability distributions according to mean-preserving spreads amounts to making second-order stochastic dominance comparisons (Rothschild and Stiglitz (1970)), using mean and variance-preserving transformations relates to *third-order* stochastic dominance (Menezes et al. (1980)).[9] Let us now write $R \gtrsim_n S$ when the distribution of the

[8] To be sure, notice that

$$\begin{aligned}
E(L^m)^2 &= E[E[(L^m)^2 \mid X]] \\
&= E[(1 - m(X))(L_X)^2 + m(X)E[(L_X + L_Y)^2 \mid X]] \\
&= E[(1 - m(X))(L_X)^2 + m(X)E[(L_X)^2 + 2L_XL_Y + (L_Y)^2 \mid X]] \\
&= E[(1 - m(X))(L_X)^2 + m(X)((L_X)^2 + 2L_XE[L_Y \mid X] + E[(L_Y)^2 \mid X])] \\
&= E(L_X)^2 + 0 + E[m(X)E[(L_Y)^2 \mid X]] \\
&= I_X + MI_Y \quad .
\end{aligned}$$

[9] Let R and S be any two random variables with respective distribution functions $H(r)$ and $P(s)$, and densities $h(r)$ and $p(s)$ which are strictly positive on the open interval $(\underline{t}, \overline{t})$. Recall that the distribution

random variable R dominates the distribution of the random variable S in the n^{th} order. The following theorem spells out our current intuition.

Theorem 1. *Let $m_A(X)$ and $m_B(X)$ be some auditing policies with the same intensity. If for any $x \in \Gamma_X$ we have that $\int_{Inf\Gamma_X}^{x} m_B dF \geq \int_{Inf\Gamma_X}^{x} m_A dF$, the inequality being strict for a set of positive measure, then $L^A \gtrsim_3 L^B$.*

In other words, the information system induced by a contingent auditing policy m_A stochastically dominates in the third order the information system generated by an equally expensive policy m_B when the downside intensity of audits is lower under the former than under the latter, or equivalently when the former is relatively more upper-tailed than the latter. A proof of this statement can be found in the appendix.

Let $m_{UT}(X)$ denote an *upper-tailed* auditing policy, i.e. a policy such that $m_{UT}(x) = 1$ if $x > \overline{x}$ and $m_{UT}(x) = 0$ if $x \leq \overline{x}$; and similarly, let $m_{LT}(X)$ refer to a *lower-tailed* auditing policy, so $m_{LT}(x) = 0$ when $x > \underline{x}$ and $m_{LT}(x) = 1$ otherwise. Given an upper-tailed, a lower-tailed, and an arbitrary auditing policy $m(X)$ which all exhibit the same expected frequency, it can be checked that

$$\int_{Inf\Gamma_X}^{x} m_{LT} dF \geq \int_{Inf\Gamma_X}^{x} m dF \geq \int_{Inf\Gamma_X}^{x} m_{UT} dF \tag{7}$$

of R stochastically dominates the distribution of S in the n^{th} order, noted $R \gtrsim_n S$, if for all $t \in (\underline{t}, \overline{t}]$ we have that $\int_{\underline{t}}^{t} (t-z)^{n-1}\{h(z) - p(z)\}dz \leq 0$, the inequality being strict on a subset of $(\underline{t}, \overline{t}]$ of positive measure. It can be shown that $R \gtrsim_n S$ implies that $R \gtrsim_{n+1} S$, while the converse is not true. Third-order stochastic dominance thus provides a finer ordering than second and first-order dominance.

for all x. By the above theorem, one can now conclude that

$$L^{m_{UT}} \gtrsim_3 L^m \gtrsim_3 L^{m_{LT}} \quad . \tag{8}$$

That is: relative to the third-order stochastic dominance ordering, any set of information systems generated by cost-equivalent auditing policies is bounded above and below by the systems corresponding respectively to an upper and a lower-tailed auditing policy.

Theorem 1 finally clarifies what amending an existing auditing policy actually does to the associated information system. Note that the conditional expectation and variance of the likelihood ratio distribution associated with an auditing policy $m(X)$ are respectively

$$E(L^m \mid X = x) = L_X(x, a) \text{ so } Var(L^m \mid X = x) = m(x) I_Y(a) \quad . \tag{9}$$

Raising the probability $m(x)$ of performing an audit amounts therefore to increasing the local variance of L^m at $X = x$ Conversely, decreasing $m(x')$ contracts the distribution of L^m at $X = x'$. A combination of both transformations thereby involves a reallocation of *local* variance - leftward if $x < x'$ or rightward if $x' > x$ - within the information system.[10]

Now that we have a convenient classification of audit-generated information systems, the upcoming section will turn to decision-making and the principal's choice of policy.

[10] Equivalently, let $L_{X,Y}$ be the joint likelihood ratio of X and Y; it can be shown that the distribution of $L_{X,Y}$ is a mean-preserving spread of that of L_X (see Kim (1994), proposition 2). Raising (diminishing) the contingent probability $m(x)$ of performing an audit, by increasing (decreasing) the relative frequency of $L_{X,Y}$ and lowering (augmenting) that of L_X, amounts therefore to increasing (decreasing) the local dispersion of L^m at $X = x$.

4. CHOOSING AN INFORMATION SYSTEM

It is intuitive that the cost of auditing will first have an impact on the type of policy to be set by the principal. The following proposition clarifies this matter.

PROPOSITION 1. *An optimal auditing policy is such that auditing intensity is decreasing with respect to K, and $M \equiv 1$ when $K = 0$.*

A proof can be found in the Appendix. Note that the second part of this proposition constitutes an extension of Holmström (1979)'s celebrated "sufficient statistic" result: it says indeed that any informative signal about the agent's effort has positive value for the principal, *even* when gathering such a signal is an endogenous (i.e. strategic) decision.

Given a cost of auditing K, the relevant set of policies reduces therefore to those displaying the appropriate expected frequency. This result certainly provides some guidance for selecting an auditing policy, but it still leaves out a huge set of policies to choose from. In order to pursue further, let us now rewrite the principal-agent problem (taking stock from Grossman and Hart (1983)) as follows. Let

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

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 observing x, and $w_A(x)$, $\rho(x)$ denote respectively the "certainty equivalent" and the "risk premium" associated with this lottery. Expressions (1), (3) and (5) are then respectively the same as

$$EC = \int_{\Gamma_X} \{(1-m)\varphi(u_N) + m[\varphi(u_A) + \rho]\} dF + K \int_{\Gamma_X} m dF \tag{10}$$

$$EU = \int_{\Gamma_X} \{(1-m)u_N + m u_A\} dF - a \geq \underline{U} \tag{11}$$

$$EU_a = \int_{\Gamma_X} \{(1-m)u_N + m u_A\} dF_a + \int_{\Gamma_X} \int_{\Gamma_Y} m \omega dF dG_a - 1 \geq 0. \tag{12}$$

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

$$\rho(x) = E_Y[\varphi(u_A(x) + \omega(x, Y))] - \varphi(u_A(x)). \tag{13}$$

The current optimization 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. The optimality conditions imply, furthermore, that

$$\mu L_Y = \varphi'(u_A(x) + \omega(x, Y)) - \varphi'(u_N(x)), \tag{14}$$

where $\mu \geq 0$ is the Lagrange multiplier associated with constraint (12). The contingent lotteries $\omega(x, Y)$ must now 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).[11] This time the decision-maker exhibits precautionary motives, or *prudence*. 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)). Starting from the previous solution ($u_N(x) = u_A(x)$, and $\omega(x, Y)$ invariant with respect to x), she would thus set $m(x)$ larger when x is higher and $m(x)$ smaller when x is lower. This suggests than a

[11] 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 to 1/2. For concreteness, a complete treatment of the knife-edge case $u(t) = t^{1/2}$ has been put in the Appendix.

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.

preferred auditing policy would now be relatively more upper-tailed. Moreover, prudence together with (13) 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 such an alternative 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 (using Baiman and Demski's wording) that a better auditing policy would have $u_A(x)>u_N(x)$, thereby constituting a "carrot" rather than a "stick" for the agent.

This discussion draws attention to the function $\varphi(\cdot)$ and the sign of its third derivative as key ingredients of choice. Indeed, the general criterion we will now introduce, which allows to select among audit-generated information systems that bear the same cost, relies on a surrogate of φ.

Write $\Delta(w,\sigma)=u(w)\sigma-w$ and let $\Delta^*(\sigma)=Max_{w\in W}\{\Delta(w,\sigma)\}$. By the envelope theorem, $\Delta^{*\prime}(\sigma)=u(w(\sigma))$, where $w(\sigma)$ satisfies $u'(w)\sigma=1$ or equivalently $\varphi'(u(w))=\sigma$.

Hence, $\Delta^{*\prime}(\sigma) = \varphi'^{-1}(\sigma)$ and

$$\varphi'''(\cdot) > 0 \text{ if and only if } \Delta^{*'''}(\cdot) < 0 \; . \tag{15}$$

The second major result of this paper, which proof is in the Appendix, is now at hand.

THEOREM 2. *The principal prefers a signal R to a signal S to implement a given action a if $E_R[\Delta^*(\lambda_R+\mu_R L_R)] \geq E_S[\Delta^*(\lambda_R+\mu_R L_S)]$, where λ_R and μ_R are the multipliers of the participation and the incentive constraints which appear in the principal-agent problem with signal R.*

First note that the following assertion - a restatement of Kim (1995)'s proposition 1 - is a direct consequence of the above. Hence, theorem 2 encompasses the MPS criterion.[12]

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

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, for $0 \leq \alpha \leq 1$,

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

[12] Theorem 2 is actually implicit in Kim (1995)'s proposition 1, where the function $\psi(q)$ defined by expression number (4) in the proof corresponds to our function $\Delta^*(\sigma)$. Our current presentation simply brings up and exploits the potential of Δ^* (thereby contributing also a simpler proof of the MPS criterion).

so $\Delta^*(\cdot)$ is a convex function.[13] The statement now simply follows from the fact that $E_J[\Delta^*(\lambda_R + \mu_R L_J)] = E_{L_J}[\Delta^*(\lambda_R + \mu_R L_J)]$, for $J = T, Z$. ■

The next result will finally allow to select among various contingent auditing policies that bear the same cost.

COROLLARY 2: *Let* $\Delta^{*'''} > (<) 0$. *The information system from* R *is preferred by the principal to that from a signal* S *when* $L_R \gtrsim_3 L_S \quad (L_R \lesssim_3 L_S)$.

Proof. Recall from Whitmore (1970) that any risk-averse decision maker whose marginal utility function is strictly convex would prefer a lottery A over a lottery B having the same expectation when the former dominates the latter in the sense of third-order stochastic dominance. ■

Building on the classification of audit-generated information systems made available in section 3, Corollary 2 entails that (corroborating the discussion previous to theorem 2), if $\varphi''' < 0$, then the principal will adopt a contingent auditing policy $m_A(X)$ instead of an alternative policy $m_B(X)$ of equal intensity when the information system induced by the former exhibits less local variance at lower values of X (and consequently more local variance at higher values of X) than the one corresponding to the latter.

This remark supports (this time from an information-value perspective) Baiman and Demski (1980)'s first characterization of optimal auditing policies: considering the inequalities in (8), optimal audits are upper-tailed when $\varphi''' < 0$ and lower-tailed if $\varphi''' > 0$.

[13] 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)).

These developments also hint at a practical recipe for incrementally improving an auditing policy.

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

· When this is done, determine the convenient local intensity of audits at specific values of the signal X. This would be achieved through successive mean and variance-preserving transformations of the information system, and the relative strength of the agent's prudence (as captured by the sign of the third derivative of φ) would then indicate the appropriate reallocation (rightward or leftward) of local variance.

5. EXTENSIONS

The above analysis uses the framework which is standard in the auditing literature: namely, there is a unit cost per audit and the signals X and Y are conditionally independent. This section will now show that the approach developed in this paper can provide useful insights when those assumptions are relaxed.

5.1 *Convex Auditing Costs*

Assume that the cost of audit is strictly convex in the auditing probability, that is: the function $K : [0,1] \to [0,\infty)$ is such that $K(0) = 0$, $K'(\cdot) > 0$ and $K''(\cdot) > 0$. The principal' s expected cost when the agent delivers effort a is then given by

$$EC = \int_{\Gamma_X} \int_{\Gamma_Y} \{(1-m)w + ms\} fg dx dy + \int_{\Gamma_X} K(m) f dx \tag{16}$$

but the constraints (3) and (4) of the principal-agent problem remain the same.

Such a cost function indicates that the principal now dislikes variability in the probability of auditing. This has to be traded off against the informational returns from contingent audits. According to theorems 1 and 2, which are still valid in this context, the latter is higher for relatively upper-tailed (lower-tailed) audits when $\varphi''' < 0$ ($\varphi''' > 0$). It is therefore intuitive that an optimal auditing policy would be increasing (resp. decreasing; constant) with respect to X when $\varphi''' < 0$ (resp. $\varphi''' > 0$; $\varphi''' = 0$), and that it may now be *strictly random* for values of X located in the middle of Γ_X. The following statement formalizes this assertion; a proof can be found in the Appendix.

PROPOSITION 2: *If the cost of audits $K(\cdot)$ is a strictly convex function and $\varphi''' < 0$ ($\varphi''' > 0$; $\varphi''' = 0$), then an optimal auditing policy $m^*(X)$ is such that:*

- *either $m^*(X) \equiv 0$;*

- *or $m^*(X) \equiv 1$;*

- *or $m^*(\cdot)$ is continuous and increasing (decreasing; constant) in the values of X, and there exists a subinterval $(\underline{x}, \overline{x})$ of Γ_X such that $0 < m^*(X) < 1$ when $X \in (\underline{x}, \overline{x})$.*

5.2 *Correlated Signals*

What does an optimal auditing policy look like when the signals X and Y are dependent random variables (given the effort a)? To answer this question,[14] let $h(x, y, a)$ be the joint density function of (X, Y). In this context, $f(x, a)$ now denotes the marginal density of X, i.e. $f(x, a) = \int_{\Gamma_Y} h(x, y, a)dy$, and $g(x, y, a) = \dfrac{h(x, y, a)}{f(x, a)}$ stands for the density of Y conditional upon observing $X = x$. The likelihood ratio corresponding to an auditing policy $m(X)$ is then

$$\begin{aligned} L^m &= (1 - m(x))\frac{f_a(x, a)}{f(x, a)} + m(x)\frac{h_a(x, y, a)}{h(x, y, a)} = \frac{f_a(x, a)}{f(x, a)} + m(x)\frac{g_a(x, y, a)}{g(x, y, a)} \\ &= L_X(x, a) + m(x)L_Y(x, y, a). \end{aligned} \tag{17}$$

Since

$$E_{Y/X=x}[L_Y(x, Y, a)] = \int_{\Gamma_Y} \frac{g_a(x, y)}{g(x, y)} g(x, y)dy = 0, \tag{18}$$

$E_{Y/X}[L_X L_Y] = L_X[E_{Y/X}(L_Y)] = 0$. The conditional mean and variance of the associated likelihood ratio distribution are then respectively given by

$$E(L^m \mid X = x) = L_X(x, a) \ \text{ and } \ Var(L^m \mid X = x) = m(x)E_{Y/X}[(L_Y)^2] \ . \tag{19}$$

Comparing (19) and (9) reveals that the local variance now depends not only on the probability $m(x)$ but also on the informational content of the signal Y when $X = x$. If that content is monotone increasing (decreasing) in the value X may take, it is intuitive that this would then back up the principal's preference for upper-tailed (lower-tailed) audits

[14] A different but related question would be to ask for the value of an additional signal Y, as a function of the linear correlation between X and Y. This issue is addressed by Rajan et Sarath (1997), in a monitoring context (i.e. where $m(X) \equiv 1$) with binary random variables.

in the presence of a uniformly negative (positive) function $\varphi'''(\cdot)$. Our last proposition makes this statement rigorous. (The proof is in the Appendix.)

PROPOSITION 3: *If* $L_Y(x, Y, a) \gtrsim_2 L_Y(x', Y, a)$ *for all* $x > x'$ *and* $\varphi''' \leq 0$*, then the optimal auditing policy is relatively upper-tailed. If, on the other hand,* $L_Y(x', Y, a) \gtrsim_2 L_Y(x, Y, a)$ *for all* x *and* x' *such that* $x > x'$ *and* $\varphi''' \geq 0$*, then the optimal policy is relatively lower-tailed.*

6. CONCLUDING REMARKS

This paper first brings together two important streams of literature in principal-agent theory: that which started with Holmstrom (1979) on ordering information systems, and that which began with Baiman and Demski (1980) on auditing.

It contributes to the former by providing, through Theorem 2 and Corollary 1, a strict extension (and a simpler proof) of Kim (1995)'s MPS criterion. It offers new insights as well for the auditing and principal-agent literature, through Propositions 2 and 3 which deal with more general contexts that the ones previously studied, and through Theorem 1 which makes it clear that the design of auditing policies not only trades off risk sharing and incentives, but also incentives and *downside* risk. (This conclusion constitutes, furthermore, a new economic application of third-order stochastic dominance.)

The latter conclusion would now support some conjectures for further principal-agent research. On the positive side, empirical work seems to have found little relationship between risk and incentives (see, for instance, Prendergast (1999, 2002)); based on the latter conclusion, however, this may be because, in circumstances where it is harder to infer effort from output, firms can nevertheless introduce more incentives via compensated changes in downside risk. On the normative side, one way to set higher-powered incentives in more uncertain environments (in multitasking, for example) might be to harness the agent's precautionary motives and consider explicitly the configuration of local risks.[15]

APPENDIX

PROOF OF THEOREM 1: By definition (see Menezes et al. (1980)), a random variable Z dominates a random variable T to the third order whenever the following inequality holds for all real t, this inequality being strict on a set of values of t of positive measure:

$$E[Max(t - Z, 0)^2] \leq E[Max(t - T, 0)^2] \ .$$

Applying this to our problem, $L^A \gtrsim_3 L^B$ thus means that

$$\Omega(t) = E[Max(t - L^A, 0)^2] - E[Max(t - L^B, 0)^2] \leq 0$$

for any $t \in \Gamma$, the inequality being strict on a subset of Γ_L with positive measure.

[15] This assertion can actually be supported further, thanks to some recent results from Keenan and Snow (2002, p. 274-5): "(...) in simple problems of portfolio choice and labor supply, risk averse decision makers with constant absolute risk aversion increase their exposure to risk in response to compensated increases in downside risk, but would respond in the opposite manner to compensated increases in risk."

As

$$E[Max(t - L^A, 0)^2] = \int_{\Gamma_X}\int_{\Gamma_Y} (1 - m_A(x)) Max(t - L_X, 0)^2 dF + \int_{\Gamma_X}\int_{\Gamma_Y} m_A(x) Max(t - L_X - L_Y, 0)^2 dF dG, \tag{20}$$

we obtain that

$$\begin{aligned} \Omega(t) &= \int_{\Gamma_X} (m_B - m_A)[Max(t - L_X, 0)]^2 dF - \int_{\Gamma_X}\int_{\Gamma_Y} (m_B - m_A)[Max(t - L_X - L_Y, 0)]^2 dF dG \\ &= \int_{\Gamma_X} (m_A - m_B)\Psi(t - L_X) dF \ , \end{aligned} \tag{21}$$

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

$$\Psi(t) = E_Y[Max(t - L_Y, 0)^2] - Max(t, 0)^2. \tag{22}$$

Note that $\Psi(\cdot)$ is a differentiable function, since the derivative of $Max(C, 0)^2$ exists and is equal to $2Max(C, 0)$. Therefore, by Jensen's inequality,

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

so $\Psi(\cdot)$ is increasing on $]Inf \frac{g_a}{g}, Sup \frac{g_a}{g}[$.

The right-hand side of (21) can now be integrated by parts, which yields

$$\Omega(t) = \int_{\Gamma_X} \{\int_{Inf\Gamma_X}^{x} (m_A(z) - m_B(z))dF(z,a)\}\Psi'(t - L_X)(\frac{\partial L_X}{\partial x})dx \ . \tag{23}$$

We conclude that, if $\int_{Inf\Gamma_X}^{x} m_B(z)dF(z,a) \geq \int_{Inf\Gamma_X}^{x} m_A(z)dF(z,a)$ for any x, with strict inequality on a subset of positive measure, then $\Omega(t) \leq 0$, as claimed. ■

PROOF OF PROPOSITION 1:

Part I (Optimality conditions): Write $\Delta(w,\sigma) = u(w)\sigma - w$ and $\Delta^*(\sigma) = Max_{w \in W}\{\Delta(w,\sigma)\}$. And let Λ denote the Lagrangian function associated with the principal-agent problem, that is:

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

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, 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 $s(x,y) = Argmax_w \ \Delta(w, \lambda + \mu L_X(x,a) + \mu L_Y(y,a))$,

3. and for all x,

$$m(x) = \arg \max_{m \in [0,1]} m\{\int_{\Gamma_Y} [u(s)(\lambda + \mu L_X + \mu L_Y) - s]dG \quad (24)$$

$$- [u(w)(\lambda + \mu L_X) - w] - K\}.$$

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

$$u'(w)\{\lambda + \mu L_X\} = 1, \quad (25)$$

$$u'(s)\{\lambda + \mu L_X + \mu L_Y\} = 1 \quad (26)$$

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 (7) and (8). Condition 3 therefore says that $m(x)$ maximizes $m \cdot Q(L_X(x, a))$ on $[0, 1]$, where $Q(\cdot)$ is defined as

$$Q(z) = E_Y[\Delta^*(\lambda + \mu z + \mu L_Y)] - \Delta^*(\lambda + \mu z) - K. \quad (27)$$

Note that, together with the Monotone Likelihood Ratio Property, equations (25) and (26) entail that the optimal wages $w(x)$ and $s(x, y)$ are nondecreasing in x and y.

Part II (Comparative statics): 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 $ET(K') + M(K')K' \leq ET(K) + M(K')K$. Summing 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

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

(the inequality being strict at an interior solution), and so $Q(L_X(x, a))$ is always nonnegative when $K = 0$. ■

The constant relative risk aversion (CRRA) case: Suppose that the agent's risk preferences can be represented by a utility index of the form $u(t) = t^{1/2}$.

By equations (25) and (26), 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

$$EC^* = (\frac{\lambda}{2})^2 + (\frac{\mu}{2})^2\{I_X + MI_Y\} + KM = (a + \underline{U})^2 + \frac{1}{I_X + MI_Y} + KM. \qquad (28)$$

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).

PROOF OF THEOREM 2: 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:

$$\int_{\Gamma_i} w(t)dH(t,a,i) \equiv \overline{EC}_i \tag{29}$$

$$\int_{\Gamma_i} u(w(t))dH(t,a,i) - a \geq \underline{U} \tag{30}$$

$$\int_{\Gamma_i} u(w(t))dH_a(t,a,i) \geq 1. \tag{31}$$

The Lagrangian function associated with this problem is

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

or equivalently

$$\Lambda_i = E_i[\Delta(w, \lambda_i + \mu_i L_i)] - \lambda_i(a + \underline{U}) - \mu_i \ . \tag{32}$$

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:

$$\lambda_i\{\int_{\Gamma_i} u(w_i(t))dH(t,a,i) - a - \underline{U}\} = \mu_i\{\int_{\Gamma_i} u(w_i(t))dH(t,a,i) - 1\} = 0 \ . \tag{33}$$

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

$$\Lambda^*_T - \Lambda^*_Z = \overline{EC}^*_Z - \overline{EC}^*_T \quad . \tag{34}$$

It follows that (Note that, in this model, the multiplier μ_T is strictly positive.)

$$\begin{aligned} \overline{EC}^*_Z - \overline{EC}^*_T &\geq E_T[\Delta^*(\lambda_T + \mu_T L_T)] - E_Z[\Delta(w_Z, \lambda_T + \mu_T L_Z)] \quad (35) \\ &\geq E_T[\Delta^*(\lambda_T + \mu_T L_T)] - E_Z[\Delta^*(\lambda_T + \mu_T L_Z)] \quad . \end{aligned}$$

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. ■

PROOF OF PROPOSITION 2: When the cost of auditing is given by the function $K(\cdot)$, the necessary optimality conditions (25) and (26) remain the same but (24) is replaced by:

$$\text{for all } x,\ m(x) = \arg\max_{m \in [0,1]} \ mQ(L_X(x,a)) - K(m)$$

$$\text{with } Q(z) = \{E_Y[\Delta(s, \lambda + \mu z + \mu L_Y(y,a))] - \Delta(w, \lambda + \mu z)\}.$$

Denote by $m^*(x)$ the solution of

$$Q(L_X(x,a)) = K'(m^*(x)). \tag{36}$$

$m^*(x)$ is a continuous function of x which, from Part I of the proof of proposition 1, increases (decreases; is constant) with x provided $\Delta^{*'''}(\cdot) > (<;=)\ 0$. Several cases may now arise.

If $K'(0) \geq Q(L_X(x,a))$ for all x, then the optimal auditing policy clearly is $m^*(X) \equiv 1$. And if $K'(1) \leq Q(L_X(x,a))$ for all x, then it is optimal to set $m^*(X) \equiv 0$.

When none of the latter inequalities is satisfied for all x, however, there exists at least one value $\widehat{x} \in \Gamma_X$ at which $m^*(\widehat{x})$ belongs to the open interval $(0,1)$. If $\Delta^{*'''}(\cdot)$ happens to be always 0 in this case, then $m^*(X)$ will be constant and equal to $m^*(\widehat{x})$; but if $\Delta^{*'''}(\cdot) >$ or < 0, then $m^*(x)$ will lie strictly between 0 and 1 as long as $Q(L_X(x,a)) < K'(1)$ and $Q(L_X(x,a)) > K'(0)$. ■

Proof of proposition 3: When X and Y are correlated, the necessary optimality conditions (25) and (26) still hold and (24) now becomes

$$\text{for all } x,\ m(x) = \arg\max_{m\in[0,1]}\ mQ(L_X(x,a),x)$$

$$\text{with } Q(z,x) = E_{Y/X=x}[\Delta^*(\lambda + \mu z + \mu L_Y(x,Y,a))] - \Delta^*(\lambda + \mu z) - K\}.$$

Let us write

$$Q(L_X(x,a),x) - Q(L_X(x',a),x') = A + B \tag{40}$$

where

$$\begin{aligned} A &= Q(L_X(x,a),x) - Q(L_X(x,a),x') \text{ and} \\ B &= Q(L_X(x,a),x') - Q(L_X(x',a),x'). \end{aligned}$$

Notice that: (1) $Q(z,x)$ increases with z provided that $E_{Y/X=x}[\Delta^{*\prime}(\lambda+\mu z+\mu L_Y(x,Y,a))] > \Delta^{*\prime}(\lambda+\mu z)$, and the latter occurs when $\Delta^{*\prime\prime\prime} > 0$; (2) $L_X(x,a)$ increases with x. Hence, when $x > x'$, B is positive (negative; equal to 0) if $\Delta^{*\prime\prime\prime}$ is positive (negative; equal to 0).

The term A, on the other hand, can be written as

$$\begin{aligned} A &= E_{Y/X=x}[\Delta^*(\lambda+\mu L_X(x,a)+\mu L_Y(x,Y,a))] \\ &\quad - E_{Y/X=x'}[\Delta^*(\lambda+\mu L_X(x,a)+\mu L_Y(x',Y,a))]. \end{aligned} \tag{41}$$

Recall that $L_Y(x,Y,a)$ and $L_Y(x',Y,a)$ have the same mean 0. Since $\Delta^*(\cdot)$ is a convex function, A will be positive (negative) if $L_Y(x,Y,a)$ dominates (is dominated by) $L_Y(x',Y,a)$ in the sense of second-order stochastic dominance.

It follows from the above that:

1. If $L_Y(x,Y,a) \succsim_2 L_Y(x',Y,a)$ for all $x > x'$ and $\Delta^{*\prime\prime\prime}(\cdot) \geq 0$, then $A+B$ is positive so $Q(L_X(x,a),x)$ increases with x. In this case, the optimal policy is thus relatively

upper-tailed.

2. If $L_Y(x', Y, a) \succsim_2 L_Y(x, Y, a)$ for all $x > x'$ and $\Delta^{*'''}(\cdot) \leq 0$, then $A + B$ is negative so $Q(L_X(x, a), x)$ decreases with x. In this case, the optimal policy is relatively lower-tailed. ■

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.

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.

Keenan, D. C., and A. Snow (2002): "Greater Downside Risk Aversion," *Journal of Risk and Uncertainty*, 24, 267-277.

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.

Prendergast, C. (2002): "The Tenuous Trade-off between Risk and Incentives," *Journal of Political Economy*, 110, 1071-1102.

Prendergast, C. (1999): "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37, 7-63.

Rajan, M. V., and B. Sarath (1997): "The Value of Correlated Signals in Agencies," *The Rand Journal of Economics*, 28, 150-167.

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.

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.

Whitmore, G. A. (1970): "Third-Degree Stochastic Dominance," *American Economic Review*, 60, 457-459.

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.