Australian School of Business

Working Paper

Never Stand Still

Australian School of Business

Australian School of Business Research Paper No. 2012 ECON 12

Optimal Contract Under Moral Hazard with Soft Information

Guillaume Roger

Optimal contract under moral hazard with soft information

Guillaume Roger *

The University of New South Wales

January 12, 2012

Abstract

I study a model of moral hazard with soft information: the agent *alone* observes the stochastic outcome of her action; hence the principal faces a problem of *ex post* adverse selection. With limited instruments the principal cannot solve these two problems independently; the *ex post* incentive for misreporting interacts with the *ex ante* incentives for effort. The optimal transfer is option-like, the contract leaves the agent with some *ex ante* rent and fails to elicit truthful revelation in all states. Audit and transfer co-vary positively, which likely is a forgotten component of many real-life contracts.

Keywords: moral hazard, asymmetric information, soft information, contract, mechanism, audit. JEL Classification: D82.

*School of Economics, UNSW. Email: g.roger@unsw.edu.au. I specially thank Guido Friebel, Claudio Mezzetti, Mike Peters, Ralph Winter, my colleagues Bill Schworm, Chris Bidner, Carlos Pimienta and Suraj Prasad, and Luis Vasconcelos for helpful comments. I benefited from remarks in seminars at UNSW, U. of Sydney, U. of Queensland, U. of Adelaide, Goethe Universität, Universidade Nova de Lisboa, U. of Melbourne and conference participants at the Australasian Theory Workshop 2010, the IIOC 2010, the 9th Journées L.-A. Gerard-Varet, the 2010 EEA Congress and the 7th ISI Conference (Delhi, 2011). Financial support from the Australian School of Business at UNSW is gratefully acknowledged.

1 Introduction

The standard solution of a moral hazard problem requires the observation of some informative signal of the agent's action. It is then possible to design a second-best contract, which is conditioned on that information instead of the actual action. While convenient to investigate questions such as the cost of moral hazard or explore properties of the solution, this model strongly relies on there being an observable signal of the agent's action. However performance may be difficult to observe or noisy. When the observed signal is not very informative it may be complemented. For example, Dye and Sridhar [10] suggests to gather additional information *ex post*, and thus to condition the contract on a broader set of data. Sometimes performance is not observed at all: an accounting report, for instance, is *not* a direct observation of the state of an firm. Rather it is a message that is sent by the same self-interested agent who generated that information in the first place. Then one may construct an audit mechanism (e.g. Kanodia [18] and Mookherjee and Png [26]). With enough instruments (as in [26]), the problems of *ex ante* effort provision and *ex post* information revelation can be treated separately. More precisely, separability allows *ex post* truthful revelation without any consequence on the incentive device used to solve the *ex ante* moral hazard problem. Their connection becomes moot and the moral hazard problem can be solved in standard fashion.

This paper explores exactly that connection. I present a model of moral hazard *cum* adverse selection, in which the principal cannot disentangle the *ex post* and the *ex ante* problems. I show this has significant consequences for the optimal scheme used to solve the moral hazard problem, which is quite different from the standard second-best. In addition, the agent may, optimally and rationally, not truthfully reveal her information, unlike in Mookherjee and Png [26] and Kanodia [18]. Applications of this model are broad-ranging. For example, after hiring the CEO, a board often asks of him (her) to report his (her) results while on the job; a regulated firm may be asked to reveal its production cost after investing in an uncertain technology.

The starting premise is that the real world does not accord with the results of Mookherjee and Png [26] or Kanodia [18]. Enron executives did not truthfully reveal their information, neither did Merrill Lynch's nor Lehman's.[1] As the Greek economy was imploding it was revealed that its national accounts were not reflective of its true state of affairs. More broadly, Reinhart and Rogoff [28] document that governments facing a sovereign debt crises tend to not disclose the true impact of their actions. Thus a model that systematically predicts truthful revelation has limited applicability. In addition, separability voids any interaction between *ex post* adverse selection and *ex ante* moral hazard. The analysis of such interaction has received scant attention in economics, possibly because the Revelation Principle (applied by Mookherjee and Png [26] and others) is too powerful in some sense. Indeed the accounting literature roots misreporting of information in some failure of the Revelation Principle (e.g. Arya, Glover and Sunder [2] or Demski and Frimor [9]). I suggest a different route that affirms and exploits the Revelation Principle.

Bar for the issue of observability, the model mirrors that of a standard moral hazard problem. A risk-neutral principal delegates production to a risk-averse agent. The agent's action a governs the distribution $F(\cdot|a)$ of a stochastic outcome θ, which she *alone* observes. That information must therefore be elicited *ex post*. Because the principal otherwise observes nothing, the contract must include an audit and some (exogenous) punishment. The model is *not* reliant on endogenous penalties nor rewards; that is, the principal possesses fewer instruments than in Kanodia [18] or Mookherjee and Png [26]. Furthermore, the model attempts to be faithful to audit as a sampling process, which is imperfect.[2] This paucity of instruments induces an inability to separate the problems. It introduces a fundamental tension between *ex ante* effort provision, which requires a state-contingent compensation, and *ex post* information revelation, which is best addressed with a constant transfer. The equilibrium is fully characterized and its properties are explored.

The audit function, optimal action and transfer schedule are all jointly determined. Three possible information revelation regimes may arise: completely truthful, partially truthful and

[1] Kedia and Philippon [20] develop and test a model of earnings management (a euphemism for fraudulent accounting). They document how pervasive the practice is.

[2] For example, financial audits are sampling processes.

never truthful. Which of these regimes prevails depends on the whole contract, not just the audit. An *ex ante* rent must be left to the agent (i.e. the participation constraint is slack) because the penalty for misreporting acts like an implicit limited liability constraint. As in Jewitt, Kadan and Swinkels [17] the transfer function is "option-like" (see Figure 3), which accords well with many real-life instances. Complete truth-telling can never be an equilibrium because the optimal transfer schedule is constant (optimally at zero) below a performance threshold θ_a. Thus for any realisation of the state beneath θ_a, the agent has nothing to lose by misreporting. So lack of observability combined with weak punishments require a peculiar contract in response, which in turn prevents complete truth-telling; and lack of truthful revelation induces further distortions. When truthful revelation is possible for at least some states, the agent misreports in the worse states, where the incentive is strongest and the cost is lowest.

Furthermore, the optimal audit and transfer co-vary positively. I suggest that together these two sets of results bring us a step closer to real life. For example with Enron, Merrill Lynch or Lehman Brothers (to name only a few), bankruptcy might not just have been a case of poor auditing but also (excessively) powerful incentives that can only lead agents to manipulate information. Indeed, the more powerful the *ex ante* incentives for high effort (i.e. the steeper the transfer function), the more attractive is the option to manipulate information *ex post*, especially when it is bad. Therefore the more accurate must the audit be.

The papers closest to this one are Kanodia [18] and Mookherjee and Png [26]. Both consider a combination of moral hazard and *ex post* adverse selection with no observability. Kanodia [18] renders both information revelation and moral hazard problems vacuous by assuming constant wages (Equation 13). Mookherjee and Png [26] combine a Grossman-Hart [12] model with an *ex post* revelation mechanism, where the principal may be a tax authority. The agent's message conditions a payment to the principal and the probability of a perfect audit. In equilibrium the principal offers rewards (not fines, by limited liability) for truthful revelation; these may be arbitrarily large.[3] These are enough instruments to elicit truth-telling without any bearing on the contract used to solve the moral hazard problem; their connection vanishes. Close to Mookherjee and Png [26], Border and Sobel [4] construct an audit mechanism with endogenous penalties as well. The optimal probability of audit is varying in the messages sent; truthful revelation obtains. In a similar vein, Reinganum and Wilde [27] show that a simple audit cut-off rule does at least as well as a random audit rule. Both [4, 27] ignore the agent's participation decision, as pointed out by Mookherjee and Png [26].[4] In all these papers, auditing is perfect but the principal controls the probability of running an audit. I depart from them in two ways. First, there are no endogenous penalties for misreporting nor rewards for truthful revelation; the principal thus must do with fewer instruments. Second, the audit is imperfect. That technology is closer to one of sampling, which is what most real audits do, and has been modeled by Bushman and Kanodia [5] or Demski and Dye [8].

Others combine moral hazard and adverse selection, however not with soft information. Gromb and Martimort [11] let (an) expert(s) search for some information by exerting some effort, who then has (have) to disclose it to the principal. The expert(s) receive(s) a soft signal, but whether a project is eventually successful is publicly observable. To overcome the moral hazard problem, the expert's incentive contract must be made state-dependent. Like in this paper, this very fact introduces adverse selection. However, a contract can be conditioned on the final outcome, unlike here. For the purpose of this discussion, Krähmer and Strausz [22] adopt a similar construct in the context of pre-project planning. Malcolmson [25] studies a problem where, as in Gromb and Martimort [11], the agent acquires soft information and the return to the principal is publicly observable. That soft information

[3] Mookherjee and Png's model yields a quirky byproduct: the agent strictly prefers being audited. This owes to the construction of the revelation constraint (2), which implicitly only allows *reward* to be offered for truth-telling.

[4] In Khalil [21] truthful revelation can be obtained through a standard direct revelation mechanism. Auditing relaxes the agent's incentive constraint; the principal trades-off the audit cost with the information rent.

may be used by the agent to make a decision yielding the verifiable outcome. The principal may have incentives to distort the decision rule away from the first-best to foster information acquisition. In all these papers, information is *exogenously* given although *ex ante* unknown to the agent. Here the private information emerges endogenously. Levitt and Snyder [24] develop a contracting model in which the agent receives an early (soft) signal about the likely success of the project, however the eventual outcome is fully observed by the principal, hence contractible. With appropriate early information, the principal can decide whether to shut-down or continue. To obtain this information, the principal must commit to shut-down less frequently than the unconstrained solution prescribes.

After introducing the model, Section 3 deals with the *ex post* information revelation problem. Next I characterize the optimal contract; Sections 5 explores some properties. Section 6 offers a discussion of modeling choices (penalties, participation fees, commitment, audit). The proofs and some of the technical material are relegated to the Appendix.

2 Model

A principal delegates a task to an agent. She undertakes an action $a \in \mathcal{A}$, which is a compact subset of $\mathbb{R}_+$. The action's cost $c(a)$ is increasing and convex, and yields a stochastic outcome $\theta \in \left[\underline{\theta}, \overline{\theta}\right] \equiv \Theta \subset \mathbb{R}$ with conditional distribution $F(\theta|a)$ and corresponding density $f(\theta|a) > 0$. The density $f(\theta|a)$ satisfies the MLRP: f_a/f is non-decreasing, concave in θ; therefore $F(\theta|a')$ stochastically dominates $F(\theta|a)$ in a first-order sense when $a' > a$. The agent *alone* observes the outcome θ and reports a message $\omega \in \Omega$ to the principal, whereupon she receives a transfer t. Her net utility is given by $u(t, a) = v(t) - c(a)$, where $v : \mathbb{R} \mapsto \mathbb{R}$ is a continuous, increasing, concave function with $v(0) = 0$. The agent's reservation value is 0 and I do not exogenously impose a limited liability constraint (but I purposefully disregard forcing contracts). The principal receives a net payoff $S(t, \theta) = \theta - t$.[5] If the true state θ

[5]That θ is not observed does not prevent maximising $\mathbb{E}[S(t, \theta)]$ or any other monotone transformation $\mathbb{E}[S(t, g(\theta))]$. See also Grossman and Hart [12], Remark 4.

were observable by the principal, the model would collapse to the textbook moral hazard problem. Throughout I make the essential assumption that the principal can commit to the contract.

At the stage of information revelation, effort is sunk so all that matters is the utility $v(t)$ from the transfer t, which can only be conditioned on the message ω. Given the monotonicity of $v(t)$, either all types pool to the same message if $t(\omega)$ is increasing, or have no effort incentive at all if it is constant. Auditing restores a measure of *ex post* observability; it breaks the monotonicity of $v(t)$. It has zero marginal cost (and therefore always run). However it is imperfect and uncovers misreporting with probability $p(\omega - \theta; \alpha)$, where $p : \mathbb{R} \mapsto [0, 1]$ is a continuous, differentiable function in both arguments and $p(0; \alpha) = p(\cdot; 0) = 0$.[6] The technology $p(\cdot; \alpha)$ is costly to acquire; it is drawn from a family of functions parametrized by an investment α at cost $k(\alpha)$, increasing and convex. The parameter α affects the slope of $p(\cdot; \alpha)$ at 0, that is, the precision of the audit. I presume that $\forall \alpha,\ \partial p(0|\alpha)/\partial z < \infty$ (where $z = \omega - \theta$), so that auditing remains imperfect. If discovered the agent receives nothing.[7] With this construction the expected utility function of an agent at the revelation stage is $U = v(t(\omega)) [1 - p(\omega - \theta; \alpha)]$. Hence, taking α fixed,

$$\frac{\partial U}{\partial t} = v' [1 - p] \geq 0; \ \frac{\partial^2 U}{\partial t \partial \theta} = v' p' \tag{2.1}$$

is a sorting condition on the *ex post* expected utility of the agent, akin to the Spence-Mirrlees condition. Further discussion of the properties of $p(\cdot; \alpha)$ is postponed to the next section. The timing is almost standard:

1. The principal offers a contract $\mathcal{C} = \langle \Omega, t(\omega), p(\omega - \theta; \alpha) \rangle$ made of a message space, a transfer function and an audit technology;

2. The agent accepts or rejects the contract. If accepting, she also chooses an action a;

3. Action a generates an outcome $\theta \in \Theta$ observed by the agent only;

[6] This is akin to a sampling process, as in Bushman and Kanodia [5].

[7] By application of the Maximal Punishment Principle, this penalty is optimal absent any harsher one (Baron and Besanko [3]). See also the discussion in Section 6.

4. The agent reports a message $\omega \in \Omega$;

5. Audit occurs;

6. Transfers are implemented and payoffs are realized.

That payoffs are realized needs not imply that they are observed by the principal, as in the accounting example. Alternatively one can interpret this as the agent being gone by then, so that the *ex post* information embedded in the principal's payoff cannot be used.

3 Degrees of Information Revelation

I start by showing that truthful revelation in any arbitrary state θ amounts to a condition relating the transfer function $t(\cdot)$ to the probability $p(\cdot|\alpha)$. This defines three regimes: complete, partial or no information revelation. To do so I exploit two results contained in a companion paper (Roger [29]); *(i)* a direct mechanism where $\Omega = \Theta$ induces a measure of pooling, which is bad for incentives and *(ii)* there is no loss of generality in restricting attention to a simple *separating* mechanism, in which $\Omega = \widehat{\mathcal{M}}$ and $\Theta \subset \widehat{\mathcal{M}} \subset \mathbb{R}$.

Fix the contract $\langle t, p \rangle$ and consider the agent's problem after the action a has been sunk. She sends a message $\widehat{m}$ such that $\max_{\widehat{m} \in \widehat{\mathcal{M}}} v(t(\widehat{m}))\,[1 - p(\widehat{m} - \theta)]$. Her best reply $m(\theta)$ solves:[8]

$$v't'(m)[1 - p(m - \theta)] - v(t(m))p'(m - \theta) = 0 \tag{3.1}$$

Let $\mathcal{M} \equiv \left\{ m \in \widehat{\mathcal{M}} | m \textit{ solves } (3.1) \right\}$ – this is the set of optimal messages. For a mechanism to be truthful, $v(t(\theta)) \geq v(t(m(\theta)))\,[1 - p]$, that is, truth-telling corresponds to a maximum: $v(t(\theta)) = \max_{\widehat{m} \in \widehat{\mathcal{M}}} v(t(\widehat{m}))\,[1 - p(\widehat{m} - \theta)]$. Using (3.1), this implies

$$v't'(\theta) = v(t(\theta))p'(0; \alpha) \tag{3.2}$$

at some θ. Because the solution to (3.1) is unique, (3.2) is sufficient at θ for truthful revelation. Given that $t' \geq 0\ \forall \widehat{m}$, and strictly for at least some $\widehat{m}$, is necessary to induce

[8] For a validation of this differentiable approach see Laffont and Martimort [23].

a non-trivial action, this equation can hold only if $p'(0;\alpha) \geq 0$ (and strictly for at least some values). Define $\mathcal{P}$ as the set of audit technologies satisfying this minimum condition. Equation (3.2) embodies a requirement on the precision of the audit at 0; that is, it defines a subset $\mathcal{P}_0(t) \subseteq \mathcal{P}$ of audit functions that can elicit truthful revelation for at least some values of θ, given the transfer t. That Condition (3.2) holds at some θ does not mean it does for all values. There may be three cases of interest.

Case 1: Truthful revelation. This occurs when Condition (3.2) is satisfied for all values of the private information θ; more precisely, $\forall\theta,\ v't'(\theta) \leq v(t(\theta))p'(0;\alpha)$. That is, jointly with the transfer, the audit technology $p(\cdot;\alpha)$ is sufficiently precise to induce truthful revelation: $\forall\theta,\ m(\theta) = \theta$.

Case 2: Partial truthful revelation. This corresponds to condition $v't'(\tilde{\theta}) = v(t(\tilde{\theta}))p'(0;\alpha)$ for some value $\tilde{\theta} \in \left(\underline{\theta}, \overline{\theta}\right)$. If $v(t(\cdot))$ is concave, then $v't'|_{\theta\geq\tilde{\theta}} \leq v(t(\tilde{\theta}))p'(0;\alpha)$ and truthtelling obtains above $\tilde{\theta}$ (so $m(\theta) = \theta$). Similarly, $v't'|_{\theta<\tilde{\theta}} > v(t(\tilde{\theta}))p'(0;\alpha)$ and truthtelling is out of reach below $\tilde{\theta}$ (where $m(\theta) > \theta$). The converse is true for $v(t(\cdot))$ convex. Figure 1 (left panel) depicts an interior example of $\tilde{\theta}$ when $v(t(\cdot))$ is a concave function.

Case 3: No truthful revelation. Here Condition (3.2) fails to hold anywhere on the range Θ, i.e. $\forall\theta \in \Theta,\ v't'(\theta) > v(t(\theta))p'(0;\alpha)$. This is shown on the right panel of Figure 1.

In Cases 2 and 3 an agent who is induced to exert any effort necessarily misreports her private information with positive probability. This owes to the fundamental tension between *ex ante* effort incentives, which require a state-contingent transfer schedule, and *ex post* information revelation that is best addressed with state-independent transfers. This rich array of outcomes obtains because *(i)* the audit technology is allowed to be imperfect, unlike much of the audit literature; and *(ii)* the principal possesses limited instruments.

One last remark is in order. There may exist many kinds of contracts satisfying $t' \geq 0$: some may include jumps, there may be intervals on which $t' = 0$ and so on, with implications

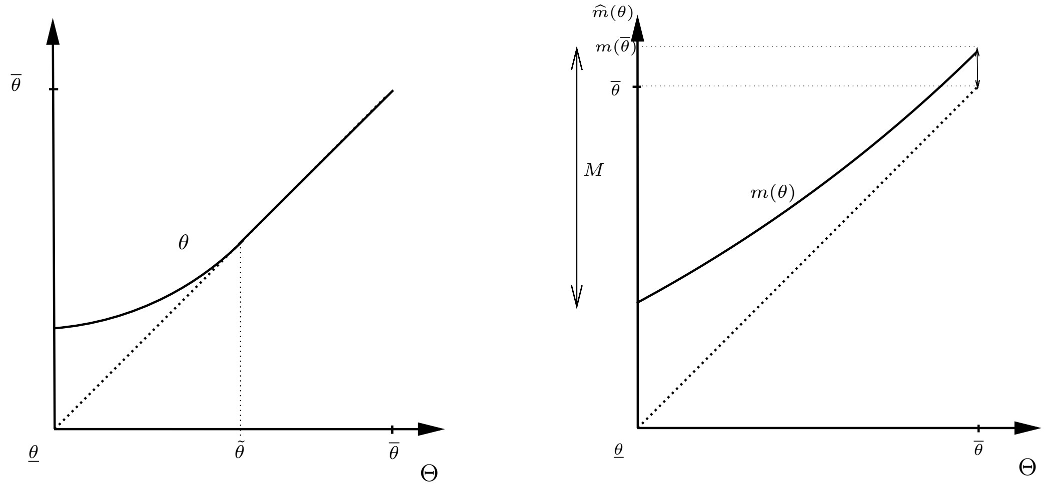

Figure1: Optimal messages above and below $\tilde{\theta}$ (left); with extended message space (right)

for the message $m(\theta)$. It is not obvious that $m(\theta)$ must be continuous, as it is depicted in Figure 1. To see why, consider a transfer scheme $t(\cdot)$ that is flat on some range, say, on $\Theta_f \equiv [\theta_1, \theta_2]$. If $\tilde{\theta} \geq \theta_2$ the agent misreports her information on Θ_f as anywhere else below $\tilde{\theta}$. If $\tilde{\theta} \leq \theta_1$, she may face the conditions $v't'(\theta_1) \leq v(t(\theta_1))p'(0;\alpha)$ but $v't'(\theta_2) \geq v(t(\theta_2))p'(0;\alpha)$, i.e. $t(\cdot)$ may be steeper at θ_2 than at θ_1 and (3.2) is reversed. Then one moves from truthful revelation above $\tilde{\theta}$ and below θ_1 to misreporting from θ_2 on, i.e. there is a jump in the optimal message (because $v(t(\tilde{\theta})) \geq (1-p)v(t(m(\tilde{\theta})))$ at $\tilde{\theta}$ but $v(t(\theta_2)) < (1-p)v(t(m(\theta_2)))$. This is shown on Figure 2, where the left panel is the transfer offered and the right one the agent's optimal message.

4 Characterising the contract

To proceed, I first seek to understand the behaviour of the contract for some fixed audit technology $p(\cdot;\alpha)$. Then I endogenize α, to which all other endogenous variables also respond, and optimize fully over the whole set of instruments t, a, α. I use the first-order approach.[9]

[9]See Jewitt [16], Araujo and Moreira [1] or Conlon [6] for validations; [16] specifically for sufficient conditions.

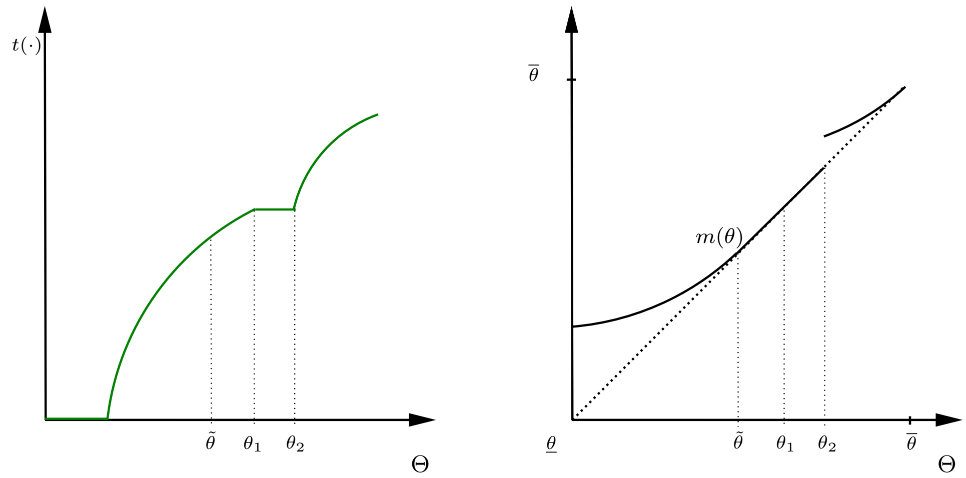

Figure 2: contract (left) may induce jump(s) in the optimal message (right)

The ensuing analysis may be problematic in that the agent's utility

$$U = \begin{cases} v(t(\theta)), & \theta \geq \tilde{\theta}; \\ (1 - p(m - \theta))v(t(m)), & \theta < \tilde{\theta}. \end{cases}$$

may not be smooth, nor even continuous, at $\tilde{\theta}$. It turns out that it must be both (see Lemma 5 in the Appendix). From this it follows that the optimal message is also a smooth function of θ at $\tilde{\theta}$ by the Theorem of the Maximum (see Figure 1), so the regime change at $\tilde{\theta}$ is "smooth" and the first-order approach can safely be applied. The other potential source of discomfort is that highlighted in Figure 2, i.e. a jump away from truth-telling above $\tilde{\theta}$; this is addressed later. Defining t over $\widehat{\mathcal{M}}$, i.e. $t : \widehat{\mathcal{M}} \mapsto \mathbb{R}$, the principal's program is

Problem 1

$$\max_{\alpha,t,a} \int_{\underline{\theta}}^{\tilde{\theta}} [x - (1 - p(m(x) - x; \alpha))t(m(x))] \, dF(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} [x - t(x)] \, dF(x|a) - k(\alpha)$$

s.t.

$$m(\theta) = \arg\max_{\widehat{m} \in \widehat{\mathcal{M}}} (1 - p(\widehat{m} - \theta))v(t(\widehat{m})) \tag{4.1}$$

$$\int_{\underline{\theta}}^{\tilde{\theta}} v(t(m(x)))[1 - p(m(x) - x)]dF(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t(x))dF(x|a) - c(a) \geq 0 \tag{4.2}$$

$$\int_{\underline{\theta}}^{\tilde{\theta}} v(t(m(x)))[1 - p(m(x) - x)]dF_a(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t(x))dF_a(x|a) = c'(a) \tag{4.3}$$

where $\tilde{\theta} \equiv \tilde{\theta}(p(\cdot;\alpha), t, a)$. The *ex post* message may be entirely truthful (only drawn from Θ), not at all (and only drawn from $\mathcal{M}$) or some of both depending on where $\tilde{\theta}$ lies.[10] From an *ex ante* standpoint the principal must account for any of these possibilities, which the objective function and the constraints reflect. Condition (4.1) is the agent's information revelation constraint – the novelty in this paper. Let λ be the Lagrange multiplier of the moral hazard constraint (4.3) and μ that of the participation constraint (4.2).

4.1 Form of the contract

In this model the *ex post* problem interacts with the provision of *ex ante* incentives. This affects the form of the contract in possibly several ways. First, the ability to inflate one's performance may alter the expected cost to the principal, as well as a the choice of action by the agent. The principal responds by distorting the transfer schedule, as shown below.

Lemma 1 *Fix a and α. The first-order conditions of Problem 1 are given by:-*

$$\frac{1}{v'(t^O(m(\theta)))} = \mu + \lambda \frac{f_a}{f}; \tag{4.4}$$

for $\theta < \tilde{\theta}$; and

$$\frac{1}{v'(t^O(\theta))} = \mu + \lambda \frac{f_a}{f}; \tag{4.5}$$

for $\theta \geq \tilde{\theta}$, where $m(\theta)$ is determined by (4.1) and $\mu, \lambda \geq 0$.

The case of complete information revelation (Case 1) is obtained by extending $\tilde{\theta}$ to $\underline{\theta}$. Then the first-order condition is standard; (4.5) holds over Θ. Case 3 corresponds to $\tilde{\theta} \geq \overline{\theta}$. Conditions (4.4) and (4.5) closely resemble one another, bar for the exact argument of $t^O(\cdot)$.

[10] More comprehensively the program allows for jumps as described in Section 3; the principal's objective is then $\max_{\alpha,t,a} \int_{\underline{\theta}}^{\tilde{\theta}} [x - (1 - p(m(x) - x;\alpha))t(m(x))] \, dF(x|a) + \int_{\tilde{\theta}}^{\theta_2} [x - t(x)] \, dF(x|a) + \int_{\theta_2}^{\hat{\theta}} [x - (1 - p(m(x) - x;\alpha))t(m(x))] \, dF(x|a) + + \int_{\hat{\theta}}^{\overline{\theta}} [x - t(x)] \, dF(x|a) - k(\alpha)$, with a jump at θ_2 and two thresholds $\tilde{\theta}, \hat{\theta}$–and the agent's utility is similarly modified. The analysis extends immediately. Note that although the problem does not specify a distribution over the message space $\mathcal{M}$, $F(\theta|a)$ is still the relevant distribution because $m(\theta)$ is injective. For details, see Roger [29].

When the agent can report $m(\theta) > \theta$, she is being paid "too much" given a, which the principal anticipates.

The second issue speaks to the nature of the constraints of Problem 1. In the standard problem the principal presents the agent with a transfer function of the form

$$\frac{1}{v'(t^S)} = \mu^S + \lambda^S \frac{f_a}{f} \tag{4.6}$$

for some action a^S, and where μ^S, λ^S are both strictly positive (see [16]). Two observations must be made. Firstly, it is immediate from (3.2) that no truthful revelation can be compatible with a binding participation constraint ($\mu^S > 0$). To see that, suppose truthful revelation obtains in equilibrium (i.e $p(\cdot; \alpha) = 0$ and $\tilde{\theta} = \underline{\theta}$), then the first-order condition of Problem 1 is exactly (4.6). Now (3.2) at $\underline{\theta}$ implies that $v(t(\underline{\theta})) \geq 0$. By monotonicity of $t(\cdot)$ therefore $\int_\Theta v(t)dF > 0$ for any action a. So the agent could accept any contract $\langle \tilde{t}, \tilde{a}, \tilde{p} \rangle$, $\tilde{a} > \min a$ such that $\int_\Theta v(\tilde{t})dF(\cdot|\tilde{a}) - c(\tilde{a}) = 0$, select $a = \min a$ at cost $c(a) = 0$ and receive an *ex ante* rent. Secondly, this reasoning holds for any revelation (truthful or otherwise). Given the (zero) penalty specified, the principal's reliance on the agent's messages to condition compensation implies that the transfers actually implemented in equilibrium can only be strictly positive. Indeed, any transfer schedule must contain at least some positive elements to induce participation with costly effort provision, as in the standard problem (see Holmström [13], Rogerson [30], Jewitt [16]), but also some negative ones for the participation constraint to bind everywhere (Rogerson [30], Jewitt [16]). Here the agent can *always* do better than accepting a negative transfer. As Condition (3.1) states, she can simply take the lottery $\{p, 1-p\}$ over 0 and some positive $v(t(m))$ by exaggerating her message. No message resulting in a negative transfer will ever be sent, and no negative transfer will ever be implemented. That is, the *ex post* adverse selection problem (together with the choice of punishment) introduces an implicit and endogenous limited liability constraint in the moral hazard problem.

I draw on the work of Jewitt, Kadan and Swinkel [17], who consider exogenous bounds on payments, to characterize the transfer function. Because the ratio f_a/f is monotonic and $\mathbb{E}_\Theta [f_a/f] = 0$, for some action a there exists some θ_a such that $f_a(\theta_a|a)/f(\theta_a|a) = 0$. Then

Proposition 1 *Fix a and α, the optimal transfer t^O takes the form*

$$\frac{1}{v'(t^O)} = \begin{cases} \kappa, & \forall\ m(\theta) \leq \theta_a; \\ \kappa + \lambda \frac{f_a}{f}, & \forall\ m(\theta) > \theta_a. \end{cases}$$

where $\kappa \geq 0$, $\kappa \neq \mu$ and $m(\theta)$ solves (4.1).

The next result furthers the characterization of the optimal transfer schedule.

Lemma 2 *The multiplier λ of the moral hazard constraint (4.3) is strictly positive.*

Therefore the optimal transfer function $t^O(\cdot)$ is fully described by Lemma 1 and Proposition 1, and it behaves according to the ratio f_a/f. Next I complete the description of the transfer schedule.

Proposition 2 *The optimal transfer function t^O solving Problem 1 is continuous and non-decreasing over Θ; in particular, it is:-*

- *continuous but with a kink at θ_a;*
- *non-decreasing concave for all θ above θ_a; and*
- *continuous and differentiable at $\tilde{\theta}$.*

The second part of Proposition 2 is trivially true when $\tilde{\theta} = \underline{\theta}$ or $\tilde{\theta} = \overline{\theta}$, for then either (4.4) or (4.5) prevails over the whole range Θ. When $\tilde{\theta}$ is interior, t^O is still continuous at $\tilde{\theta}$. The reason is that $m(\theta)$ smoothly converges to θ at $\tilde{\theta}$ because the function U is smooth. (See the left panel of Figure 1). Proving the first part is simple; to understand it, recall the informational value of the ratio f_a/f at θ_a. This is the point where F_a is the most negative, that is, where effort has the highest marginal effect. Thus a signal θ_a is indicative of an effort level that is the most valuable for the principal, who offers the agent the steepest incentives at that point (the transfer t is concave, increasing from θ_a on). This feature of the contract accords well with practice, where boni may be observed when a hurdle is passed.[11] A technical but re-assuring Corollary follows from the collection of the previous results.

[11] Jewitt, Kadan and Swinkels [17] call this kind of scheme "option contracts".

Corollary 1 *The optimal message $m(\theta)$ is everywhere continuous on Θ; i.e. there are no jumps.*

This follows from the fact that the optimal transfer function t^O is monotone concave from θ_a on. There can be no pair $\theta_1 < \theta_2$ such that $v't'(\theta_2) > v't'(\theta_1)$; thus Condition (3.2) cannot be simultaneously holding at θ_1 but reversed at θ_2. Consequently the conditions outlined by Jewitt [16] are also sufficient here. Furthermore, there can be only at most one threshold $\tilde{\theta}$, and the three simple regimes described in Section 3 are exhaustive.

4.2 Optimal contract

As part of the optimal contract the principal selects his audit technology $p(\cdot;\alpha) \in \mathcal{P}$ by choice of α. This may have two effects. First, fixing $t(\cdot)$ and a, it may alter the degree of information revelation, i.e. the cutoff $\tilde{\theta}$ (Cases 1 to 3). Second, $t(\cdot)$ and a are endogenous variables, so they too adjust to a change in α. The optimal contract balances all these effects.

Proposition 3 *The optimal contract is characterised by:-*

1. *a continuous transfer scheme* $t^O = \begin{cases} t^O(m(\theta)), & \theta < \tilde{\theta}; \\ t^O(\theta), & \theta \geq \tilde{\theta}. \end{cases}$ *determined by Proposition 1, and Conditions (4.4) and (4.5) on the relevant ranges;*

2. *an action a^O solving the first-order condition*

$$\begin{aligned} &\int_{\underline{\theta}}^{\tilde{\theta}} [x - t(m(x))(1-p)]dF_a + \int_{\tilde{\theta}}^{\bar{\theta}} [x - t(x)]dF_a \\ &+\lambda \left[\int_{\underline{\theta}}^{\tilde{\theta}} v(t(m(x)))(1-p)dF_{aa} + \int_{\tilde{\theta}}^{\bar{\theta}} v(t(x))dF_{aa} - c''(a) \right] = 0 \end{aligned} \tag{4.7}$$

3. *and an audit investment $\alpha^O = \alpha_1^O + \alpha_2^O$, where α_1^O solves*

$$v't'(\underline{\theta}) = v(t(\underline{\theta}))p'(0;\alpha_1^O) \tag{4.8}$$

and $\alpha^O \geq \alpha_1^O$ solves

$$\int_{\underline{\theta}}^{\tilde{\theta}} t(m)p_\alpha dF(x|a) + \lambda \int_{\underline{\theta}}^{\tilde{\theta}} v(t(m))p_\alpha dF_a(x|a) = k'(\alpha) \tag{4.9}$$

The cut-off $\tilde{\theta} \in [\underline{\theta}, \overline{\theta}]$ *is determined by (3.2) given* t^O, a^O, α^O*.*

The threshold $\tilde{\theta}$ is free to lie at either boundary or to be interior; it is endogenous to the contract and so is the regime one operates under. The first two items of Proposition 3 resemble standard ones. The last one determines the level of investment in the audit technology. It allows for α_2^O to be zero, that is, $\tilde{\theta} = \underline{\theta}$. If so, the technology is sufficiently inexpensive (or equivalently, precise) for Condition (3.2) to hold at $\underline{\theta}$. Condition (4.8) thus pins down the smallest investment necessary for truthful revelation. In that case, the transfer is determined by (4.5) and (4.7) collapses to the standard expression; the *pair* t^O, α^O, together with the 0 penalty, are such that they compel truthful revelation. If α_1^O is not sufficient, the investment may be increased from α_1^O to α^O (i.e. by α_2^O), and this entails a trade-off given by (4.9); that is, further distortions arise. The total marginal benefit (LHS) includes saving on undue transfers, as well as *tightening* of the moral hazard constraint, which is discussed later. When truthful revelation is impossible, the transfer is determined solely by (4.4) and (4.7) is modified by extending $\tilde{\theta}$ to $\overline{\theta}$. Importantly, truth-telling cannot be guaranteed (as in Mookherjee and Png [26]), because t^O, α^O are *jointly* determined. Whether truthful revelation obtains does not just depends on the audit procedure because the problems of moral hazard (*ex ante*) and adverse selection (*ex post*) are meshed; there is no independent instrument such as fines to solve the information revelation problem separately from the moral hazard problem. Figure 3 shows the optimal transfer scheme.

The optimal transfer function is so different from the standard one that it is difficult to perform cross-model comparisons and comment on the level of transfers and the level of the action a^O. Indeed, consider this: from Proposition 3 it may be tempting to assert that when $\alpha^O = \alpha_1^O$, truth-telling obtains and no distortion ensues, as in Mookherjee and Png [26]. But Proposition 1 still applies; so it is really the *property* of separability, rather than the *outcome* of truthful revelation, that is key.

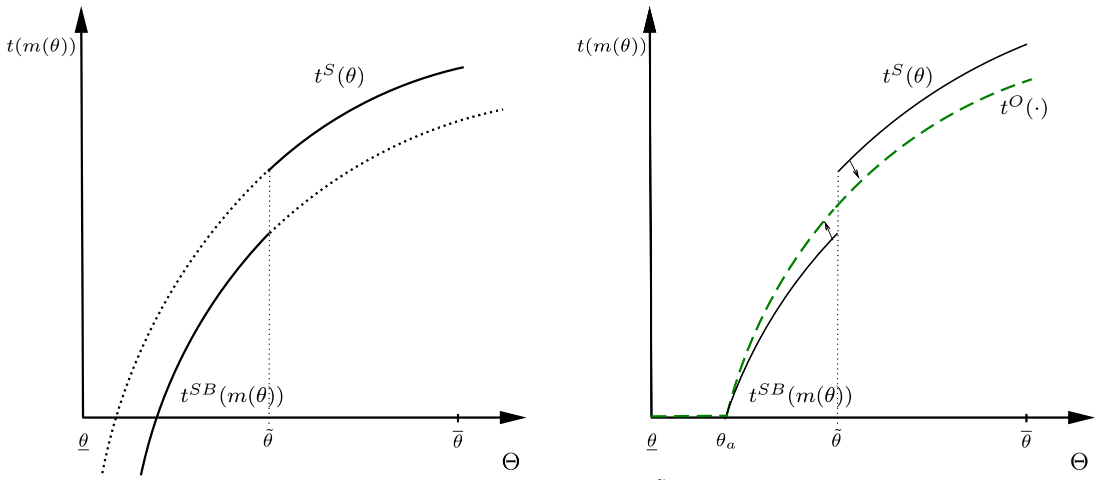

Figure 3: FOC (4.4) and (4.5) above and below $\tilde{\theta}$ and optimal transfer scheme

Suppose however that t^O were fixed and exactly as the standard transfer t^S on the portion above θ_a. Then rewriting the agent's moral hazard constraint (4.3) as

$$\int_{\underline{\theta}}^{\theta_a} v(h(1/\kappa))[1-p(m(x)-x)]dF_a(x|a) \quad + $$

$$\int_{\theta_a}^{\tilde{\theta}} v(t^S(m(x)))[1-p(m(x)-x)]dF_a(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t^S(x))dF_a(x|a) \quad = \quad c'(a)$$

where $h \equiv (v')^{-1}$ and κ is a constant, I can point to two effects. First, because the agent is presented with a constant $\int_{\underline{\theta}}^{\theta_a} v(h(1/\kappa))[1-p]dF > 0$, she has some incentives to reduce her effort. Second, $\int_{\theta_a}^{\tilde{\theta}} v(t^S(m(x)))[1-p]dF > \int_{\theta_a}^{\tilde{\theta}} v(t^S(x))dF$ (otherwise she would report truthfully), so she simultaneously has incentives to increase her action. Which of these dominates is not clear.

4.3 Equilibrium properties of the optimal contract

Now that a solution to Problem 1 has been derived, I explore some of its characteristics. It is already known that the optimal contract features *ex ante* rents, but I have remained silent as to the fixed component below θ_a. This is important because it directly affects the cost of the contract as well as the agent's *ex ante* incentives for effort and her *ex post* incentives to reveal information.

Proposition 4 *The optimal transfer t^O pays zero below θ_a.*

This is a fairly intuitive result. Anything below zero is not binding, as argued before. Anything above zero is too costly for two reasons. One, it induces a lower action from the agent because it insures her against failure. Two, it does not improve on information revelation (the threshold $\tilde{\theta}$). To see why, take $\tilde{\theta}$ interior and start from $t^O = 0$ below θ_a (in this case $\theta_a < \tilde{\theta}$ necessarily by (3.2)). Decreasing the threshold from $\tilde{\theta}$ to the next type down, say $\hat{\theta}$ costs some $\gamma > 0$ to be offered for *all* states. But that change from $\tilde{\theta}$ to $\hat{\theta}$ has zero measure. A consequence of Proposition 5 is that:

Corollary 2 *Completely truthful revelation (Case 1) can never occur in equilibrium.*

Given the penalties specified, the optimal contract that offers the agent some incentives to exert costly effort cannot simultaneously induce her to be completely truthful. Furthermore, because the optimal transfer function is concave, misreporting always occurs "at the bottom" (see Case 2). Indeed, the agent whose private information is the worst is the one with the strongest incentives to misreport when facing a concave transfer, and also with the lowest cost of misreporting.

5 The relationship between audit and transfers

From the agent's best response (4.3) one readily sees that a better audit (higher α) decreases effort.[12] Thus audit and transfer could be construed as strategic substitutes, since a higher action is associated (at least weakly) with a higher transfer. However they move in the *same* direction.

Proposition 5 *Transfer t^O and audit investment α^O co-vary positively.*

Optimally, high-power contracts are necessarily accompanied with a large enough investment in the audit technology. Conversely, it is because the audit is sufficiently precise that the

[12]Fix t and let a^* solve the agent's first-order condition (4.3) and differentiate with respect to α, $\frac{da^*}{d\alpha} \equiv \frac{da^*}{dp}\frac{dp}{d\alpha} < 0$.

contract is high-powered. Increasing t in isolation in response the the moral hazard problem is destructive; it requires a simultaneous increase in audit. This claim may be slightly counterintuitive but may be explained as follows. As α increases, the agent's choice a^* decreases because the ability to misreport enhances the marginal benefit of a high action, but the audit curtails that. So the principal's expected cost of a given action decreases and in response he increases the transfer (in each state). This is what leads to a tightening of the moral hazard problem in (4.9).

From a practical standpoint, Proposition 5 together with Condition (3.2) suggest it may not be the lack of audit that is the culprit in high-profile scandals such as Enron or Lehman Brothers. There is little doubt that firms of that nature are subject to audit. Rather the audit may not have been sufficient given the incentives offered. It is well documented that Enron executives engaged in information manipulation in spite of being audited. I am willing to add it was because the incentives were so powerful.

The next result highlights what makes t and α co-vary. The primitives of the problem are: *(i)* the properties of the distribution $F(\theta|a)$, *(ii)* the agent's risk-aversion, *(iii)* the cost of effort and *(iv)* the principal's payoff function (here trivially linear in the state).

Proposition 6 *Transfer t^O and audit investment α^O both:-*

1. *decrease in the dispersion of the distribution (in the sense of SOSD);*
2. *decrease in the agent's risk-aversion;*
3. *decrease as the cost of effort $(c(a))$ increases;*
4. *increase in the principal's payoffs.*

6 Discussion

6.1 Other penalties

The paper purposefully departs from optimal penalties because these necessarily lead truthful revelation (unless they conflict with a limited liability constraint, which is essentially equivalent to the present model). Here I discuss two potential modifications to the model in this respect.

6.1.1 Harsher penalties

The model could allow for penalties $-l < 0$. Then the first-order condition (3.1) would become $v't'(1-p) - p'(m-\theta)\left[v(t(m)) - v(-l)\right] = 0$ and clearly *(i)* there would be less exaggeration in equilibrium and *(ii)* for some l large enough, $m(\theta) = \theta\ \forall\theta$ (no misreporting). In the latter case, one would revert to model closer to that of Mookherjee and Png [26].[13] If l were not large enough, the problem would remain as here, albeit muted. The only significant difference is that the threshold θ_a would be such that f_a/f would be negative.

6.1.2 Penalties conditioned on offense

The Maximal Punishment Principle (see Baron and Besanko [3], now MPP) asserts that the penalty should be as severe as possible, and thus swiftly rules out conditioning it on the offense (e.g. small deviations from the state θ could be met with fines that commensurate). Setting the MPP aside, suppose that the principal instead uses some fine $\varphi \equiv \varphi(\widehat{m}-\theta)$ where $\varphi(0) = 0$. The agent expected utility then becomes $U = (1-p)v(t(\widehat{m})) + pv(t(\widehat{m}) - \varphi(\widehat{m}-\theta))$ and one can see that the truth-telling condition (3.2) turns into

$$v't'(\theta) = p'(0)\left[v(t(m)) - v(t(m) - \varphi(m-\theta))\right]|_{m=\theta} + p(m-\theta;\cdot)v'\varphi'|_{m=\theta}$$

i.e. $v't'(\theta) = 0$. In other words, driving a wedge between the transfers when the agent reports truthfully and does not, is essential. That is, $\varphi(\cdot)$ must be discontinuous at 0. How large a

[13] Noting that here truthful revelation would obtain immediately from the exogenous penalty.

wedge (discontinuity) is discussed above at some length. The MPP applies in this model as in many others because the audit generates no false negatives.

6.2 Audit: modeling choice

According to most accounting standards (e.g. US GAAP or the AASB in Australia), an audit seeks to provide a *reasonable* assurance that statements are free from material errors. As a result, a sampling procedure is usually adopted by financial auditors, who can verify the details of the transaction(s).[14] Statistical sampling is also followed by ISO-accredited companies for the purpose of quality assurance.[15] But in either case, the audit is *always* performed. The technology $p(\cdot;\alpha)$ I have chosen displays exactly these two characteristics.

Furthermore, absent additional (possibly unbounded, as in Mookherjee and Png [26]) punishments or rewards, the construction of Border and Sobel [4] or Mookherjee and Png [26] *cannot* deliver separation, let alone truthful revelation. To see why, observe that the audit technology can be rewritten $p(\omega;\alpha)$ and interpreted as a probability of running the audit, given some message ω – as in those papers. Then truth-telling requires $v(t(\theta)) = \max_{\omega\in\Theta} v(t(\omega))\,[1-p(\omega)]$, i.e. $v't'(\theta) = 0$; hence the need for fines or rewards in [4, 26].

6.3 Participation fee and binding constraint

The agent receives an *ex ante* rent in this model; the participation constraint fails to bind. This could be addressed with an *ex ante* participation fee, say ϕ. Then a contract entails a tariff $T = (t(\cdot),\phi)$ and the agent's expected utility reads $U = (1 -$

[14] "If controls are assessed as appropriate and operating as expected then lower levels of substantive testing is expected. [...] appropriate sampling (either statistically -in total or stratified - or judgementally when a small number of items make up much of the volume) is performed and transactions and account balances verified. The steps involved include tracing transactions from the general ledger back to supporting documents or from initiating documents through to the ledger to ensure that they are appropriately included." Mark Pickering, Auditor at Deloitte Touche Tohmatsu, 1986-91

[15] ISO: International Organization for Standardization.

$p)v(t(\widehat{m}) - \phi) + pv(-\phi)$ where $v(-\phi) < 0$. The truth-telling condition (3.2) becomes $v't'(\theta) = p'(0)\,[v(t(m) - \phi) - v(-\phi)]$. Because $v(t(\cdot)$ is concave, $v(t - \phi) - v(-\phi) > v(t)$ for each t, so for a fixed transfer function the truth-telling condition holds for a larger set of states θ. That is, $-\phi$ acts like $-l$ (see the first paragraph of this discussion). When ϕ is not too large, the information revelation problem remains as in the main text.

The main purpose of the fee ϕ is to render the participation constraint binding; suppose such a fee does exist. When $\mu > 0$ however the optimal transfer function still retains the same shape. The reason is that ϕ is paid *ex ante*, so *ex post* the agent still faces a gamble $\{p, 1-p\}$ over utilities $\{v(-\phi), v(t(m) - \phi)\}$ versus taking some really bad $v(t(\theta) - \phi)$.

If the participation constraint is made to bind the agent no longer receives an *ex ante* rent but an *ex post* information rent $U(t^O, \theta) = [1-p(m(\theta)-\theta)]v(t^O(m(\theta)))-v(t^O(\theta)) > 0$, $\forall \theta < \tilde{\theta}$ that is decreasing in the state θ.

6.4 Commitment

Commitment to the contract is an assumption that is both standard and very strong, even more so in this model where the principal commits himself to accept a(n) (obvious) lie (above $\overline{\theta}$) and to still compensate the agent according to her message. Implicit in this model is an extension of this commitment to any external auditor the principal may hire. To implement this, the principal may simply not inform the auditor about the support Θ. Otherwise one reverts to the limitation of direct mechanisms (see Roger [29]).

Commitment to running the audit is immaterial because the marginal cost of audit is naught. If for some reason the principal is unable to commit (to the audit, specifically), the game becomes one of cheap talk *à la* Crawford and Sobel [7] at the stage of information revelation. Ignoring the possibility of babbling equilibrium, auditing becomes no longer essential but may assist in improving information.

7 Conclusion

When a principal cannot observe the outcome of his agent's action in a moral hazard framework and needs to elicit this information from that very agent, he faces a problem of *ex post* adverse selection as well. With limited instruments, this introduces a fundamental tension between *ex ante* incentive, for which a contingent transfer is necessary, and *ex post* incentives, best addressed with a state-independent transfer. Type separation (not necessarily truthful revelation) requires the use of an *ex post* audit and penalties.

The *ex post* adverse selection problem is costly to the principal in three ways: first, the agent is able to exaggerate her actual performance and thereby may receive an inflated transfer. The principal's response introduces a first set of distortions. Second, because penalties are weak, they act as an implicit limited liability constraint. As a result the participation constraint cannot bind (there are rents) and the contract resembles an option. Last, the very fact that the contract entails a region with constant transfer implies that complete truthful revelation can never arise in equilibrium. There may be partial truthful revelation below a threshold; that is, the agent misreports her information in the worse states because the incentive is the strongest and the cost the lowest.

A key result of this paper is that the audit investment and the level of transfer co-vary. That is, the stronger the incentives offered to the agent, the more she must be audited to be kept in check. If thinking of real-life events (bankruptcies) of the past decade, it seems that this relationship may have been forgotten at times.

8 Appendix: Proofs

8.1 Preliminaries

I begin with a series of Lemmata that address the potential lack of smoothness of the agent's expected utility function U, and others that the will be useful throughout.

Lemma 3 *The function U is a.e. differentiable over Θ.*

Proof: By application of the Theorem of Lebesgue to a monotonically increasing function; i.e. by (3.2), U is monotonically increasing. ■

Then naturally:

Lemma 4 *Suppose a solution $m(t;\theta)$ of FOC (3.1) exists, then*

1. *this solution is unique;*

2. *$m(\theta)$ is a.e. differentiable and*

3. $\frac{dm}{d\theta} > 0$

Proof: Directly from the sorting condition $\frac{\partial^2 U}{\partial t \partial \theta} = v'p' > 0$, we know that condition (3.1) admits a unique maximiser when it binds. That $m(\theta;t)$ is increasing in θ is immediate from observing that the agent's optimisation problem is supermodular. I will need more that this statement though. Continuity of the solution $m(t;\theta)$ follows from the Theorem of the Maximum. To show that $m(\theta,t)$ is monotonically increasing, re-arrange (3.1) as $v't'/v = p'/1-p$, i.e. $d\ln(v(t(m)))/dm = -d\ln(1-p)/dm$. Take some $\theta' > \theta$ and suppose $m(\theta') \leq m(\theta)$. Then $p'(m(\theta')-\theta')/1-p\prime(m(\theta')-\theta') < p'(m(\theta)-\theta)/1-p\prime(m(\theta)-\theta)$, so that $d\ln(v(t(m(\theta'))))/dm < d\ln(v(t(m(\theta))))/dm$. Therefore $v(t(m(\theta'))) > v(t(m(\theta)))$ and since $v(\cdot)$ and $t(\cdot)$ are monotone increasing, $m(\theta') > m(\theta)$, a contradiction. The same can be shown if taking some $\theta' < \theta$ and supposing that $m(\theta') \geq m(\theta)$. It follows that $m(\theta,t)$ is a.e. differentiable, by application of the Theorem of Lebesgue, except at most for a finite set of points. Differentiate (3.1) with respect to θ and rearrange. ■

In spite of Lemma 3, there may still exist problematic discontinuities, especially at $\tilde{\theta}$, and this point is one of particular interest.

Lemma 5 *The function U is continuous and differentiable at $\tilde{\theta}$ when $\tilde{\theta} \in (\underline{\theta}, \overline{\theta})$*

Proof: I show that U cannot be discontinuous at $\tilde{\theta}$ and that by Condition (3.2) it must be also differentiable. The proof is written for U concave but also applies with obvious adjustments when it is convex. Suppose $v(t(\cdot))$ is at least weakly concave; since only upward deviations are of concern, the trouble is that we may have $v(t(\tilde{\theta})) < [1 - p(m(\tilde{\theta} - \varepsilon) - (\tilde{\theta} - \varepsilon))]v(t(m(\tilde{\theta} - \varepsilon)))$ for $\varepsilon > 0$, $\varepsilon \to 0$. Suppose so, then truth-telling cannot be an optimal response at $\tilde{\theta}$. So there must exist some value $\theta_0 < \tilde{\theta}$ (possibly $\underline{\theta}$) such that $v(t(\tilde{\theta})) \geq [1 - p(m(\theta) - \theta)]v(t(m(\theta)))$ for $\theta \in [\theta_0, \tilde{\theta})$. Let $\theta \to \tilde{\theta}$, this is exactly the definition of continuity. Now notice that

$$v' t'(\tilde{\theta}) = v(\tilde{\theta}) p'(0; \alpha) \Leftrightarrow \frac{\partial}{\partial \theta} v(t(\theta))|_{\tilde{\theta}} = \frac{\partial}{\partial \theta} [1 - p(m(\theta) - \theta)] v(t(m(\theta)))|_{\tilde{\theta}}$$

or $\frac{\partial}{\partial \theta} U|_R = \frac{\partial}{\partial \theta} U|_L$ at $\tilde{\theta}$. So U is differentiable. Condition (3.2) is a pasting condition at $\tilde{\theta}$. ■

Lemma 6 *The mapping* $m : \Theta \mapsto \mathcal{M}$ *is piece-wise weakly convex in* θ.

Proof: Take first $\tilde{\theta} \in (\underline{\theta}, \overline{\theta})$. $m(\theta)$ is increasing and a.e. differentiable by application of Lemma 1, with $m(\underline{\theta}) > \underline{\theta}$ for any $\tilde{\theta} > \underline{\theta}$. Because U is continuous and differentiable, $\lim_{\theta \uparrow \tilde{\theta}} m(\theta) = \theta$. Suppose now that $m(\theta) - \theta$ were increasing; then $dm(\theta)/d\theta > 1$ and $\lim_{\theta \uparrow \tilde{\theta}} m(\theta) \neq \theta$; so $m(\theta) - \theta$ must be decreasing, and consequently, $dm(\theta)/d\theta < 1$. Therefore $m(\theta)$ is convex when $\tilde{\theta} \in (\underline{\theta}, \overline{\theta})$. Now extend $\tilde{\theta}$ to $\overline{\theta}$ to obtain Case 3. ■

8.2 Proofs

Proof of Lemma 1: By pointwise optimization of Problem 1. Below $\tilde{\theta}$, $m(\theta) > \theta$, so the transfer $t^{SB} \equiv t(m(\theta))$, while above $\tilde{\theta}$, $t^S \equiv t(\theta)$. Notice that $\theta_a \leq \tilde{\theta}$, otherwise there exists an interval $[\tilde{\theta}, \theta_a]$ where t^O is constant and the agent reports truthfully. But this cannot be optimal by (3.1). ■

Proof of Proposition 1: The existence, sufficiency and uniqueness of such contract is shown in Jewitt, Kadan and Swinkels [17] (in particular, they show the multipliers μ, λ exist and are non-negative). To construct the contract, fix some action a^O and take the first-order condition. We know $\mu = 0$ necessarily, so below θ_a the transfer must be such that $1/v'$ remains non-negative. ■

Proof of Lemma 2: Fix some a. Integrate $1/v'$ over Θ:

$$\mathbb{E}_\theta \left[\frac{1}{v'(t^O)} \right] = \kappa \int_{\underline{\theta}}^{\overline{\theta}} dF(x|a) + \lambda \int_{\theta_a}^{\overline{\theta}} \frac{f_a}{f} dF(x|a) = \kappa + \lambda \int_{\theta_a}^{\overline{\theta}} f_a(x|a) dx.$$

where $\kappa \geq 0$. That is,

$$0 < \mathbb{E}_\theta \left[\frac{1}{v'(t^O)} \right] - \frac{1}{v'(t^O(\theta))} |_{\theta \leq \theta_a} = \lambda \int_{\theta_a}^{\overline{\theta}} f_a(x|a) dx.$$

(unless $v' = \infty$ for some t and that t is a constant). For any increasing t^O on some measure of Θ, the inequality must hold as $1/v'$ is increasing. Because $f_a/f \geq 0$ on $\left[\theta_a, \overline{\theta}\right]$ and strictly for at least a positive measure, $\lambda > 0$ necessarily. ■

Proof of Proposition 2: Fix a. Rewrite the first-order condition as $v'(t^O) = (\kappa + \lambda f_a/f)^{-1}$; let $h \equiv (v')^{-1}$. The function $h(\cdot)$ is continuous because v' is also continuous, so $t^O \equiv h\left([\kappa + \lambda f_a/f]^{-1}\right)$ is a continuous function. To show continuity at θ_a, recall that $\lambda \frac{f_a}{f}|_{\theta_a} = 0$ and f_a/f is continuous in θ, so continuity at θ_a follows. For the second part of the Proposition, restrict attention to $\theta \geq \theta_a$ and define $\tau(\theta) \equiv t^O \circ m(\theta)$. Then rewrite the FOC as $v'(\tau) - \left(\kappa + \lambda \frac{f_a}{f}\right)^{-1} = 0$, where $\tau(\theta)$ is a.e. differentiable; differentiate w.r.t. θ to find $v''\tau' + \lambda \frac{d}{d\theta}\left(\frac{f_a}{f}\right) / \left(\kappa + \lambda \frac{f_a}{f}\right)^2 = 0$. This verifies $\tau' > 0$ and therefore $t' > 0$ as required since $\frac{dm}{d\theta} > 0$. Re-arrange this expression and redefine the variables

$$\tau' = -\lambda \underbrace{\frac{1}{v''}}_{Y} \underbrace{\frac{\frac{d}{d\theta}\left(\frac{f_a}{f}\right)}{\left(\kappa + \lambda \frac{f_a}{f}\right)^2}}_{X}$$

Then $\tau'' \geq 0 \Leftrightarrow \left(\frac{dY}{d\theta} X + \frac{dX}{d\theta} Y\right) \leq 0$. With $Y < 0$, rewrite the second condition as

$$\frac{dY}{d\theta} X \leq -\frac{dX}{d\theta} Y \Leftrightarrow \frac{d}{d\theta} \ln -Y \leq \frac{d}{d\theta} \ln X,$$

$$\frac{d}{d\theta} \ln -\frac{1}{v''} \leq \frac{d}{d\theta} \ln \left(\frac{\frac{d}{d\theta}\left(\frac{f_a}{f}\right)}{\left(\kappa + \lambda \frac{f_a}{f}\right)^2} \right)$$

Since the ratio $\frac{f_a}{f}$ is increasing concave, the RHS is negative. It is immediate to verify by differentiation that the LHS is positive, so the necessary and sufficient condition cannot hold. Hence $\tau'' < 0$ (where it is differentiable), that is, the effective transfer $\tau(\theta)$ is concave in the type. To show it is concave in the *message*, call on Lemma 6 and observe that τ is the composition of the function $t(\cdot)$ and the convex function $m(\theta)$. Therefore $t(\cdot)$ must be concave in m. For the last item, observe that at $\tilde{\theta}$, $m(\tilde{\theta}) = \tilde{\theta}$ by (3.2) – the agent is truthful. Thus, under $t^O(\cdot)$:-

$$\begin{aligned} v(t^O(\tilde{\theta})) &= [1 - p(m(\tilde{\theta}) - \tilde{\theta})]v(t^O(m(\tilde{\theta}))) = v(t^O(m(\tilde{\theta}))) \\ \Leftrightarrow \quad t^O(\tilde{\theta}) &= t^O(m(\tilde{\theta})) \end{aligned} \tag{8.1}$$

directly from (3.2). From Lemma 1, $t^O(m(\theta)) = t^{SB}(m(\theta))$ for $\theta \leq \tilde{\theta}$ and $t^O(\theta) = t^S(\theta)$ for $\theta > \tilde{\theta}$. Both these transfer functions are continuous on their respective domains. Thus by (8.1) I have shown that $\lim_{\theta \uparrow \tilde{\theta}} t(m(\theta)) = t^O(m(\tilde{\theta})) = t^O(\tilde{\theta}) = \lim_{\theta \downarrow \tilde{\theta}} t(\theta)$, which is the definition of continuity. Last, the right-derivative of t^O at $\tilde{\theta}$ can be denoted $\frac{dt^O}{d\theta}|_{\tilde{\theta}}$, while the left-derivative is $\frac{dt^O}{dm}\frac{dm}{d\theta}|_{\tilde{\theta}}$, where $dm/d\theta|_{\tilde{\theta}} = 1$ since $m(\theta) = \theta$ at this point. Using this one more time, $\frac{dt^O}{dm}\frac{dm}{d\theta}|_{\tilde{\theta}} = \frac{dt^O}{d\theta}|_{\tilde{\theta}}$; i.e. the left- and right-derivative are identical at $\tilde{\theta}$, which defines differentiability. ■

Proof of Corollary 1: Take any two $\theta_1 < \theta_2$ and suppose that truthful revelation holds at θ_1, i.e. $v't'(\theta_1) \leq p'(0)v(t(\theta_1)$. Because t^O is everywhere non-decreasing and concave (and so is $v(\cdot)$), it must therefore be that $v't'(\theta_2) \leq v't'(\theta_1) \leq p'(0)v(t(\theta_1) \leq p'(0)v(t(\theta_2)$. Therefore the agent also reveals herself truthfully at θ_2; she does not jump away from truth-telling. ■

Proof of Proposition 3: Construct the Lagrangian with the objective function and the constraints (4.1)-(4.3). Apply the Envelop Theorem to the first constraint. Because $\tilde{\theta} \equiv \tilde{\theta}(\alpha, t)$, Leibnitz rule gives an additional term (e.g. $p(m(\tilde{\theta}) - \tilde{\theta}; \alpha)t(m(\tilde{\theta}))f(\tilde{\theta}|a)\frac{d\tilde{\theta}}{d\alpha}$). But it is naught at $\tilde{\theta}$, where $m(\tilde{\theta}) = \tilde{\theta}$. This gives the first-order conditions found in Lemma 1, as well as (4.9). When $\tilde{\theta} = \underline{\theta}$, this latter condition is meaningless. In this case the level of investment is determined by (3.2) at $\underline{\theta}$, i.e. (4.8). ■

Proof of Proposition 4: Any amount lower than zero is not binding. Take t^O to be zero below θ_{a^O}. Then necessarily by application of (3.2), $\tilde{\theta} > \theta_{a^O}$. All things otherwise equal, having $\tilde{\theta}$ interior is costly to the principal in that the expected transfer is higher (otherwise the agent would not misreport) and so is the agent's optimal action. So the principal may have incentives to lower $\tilde{\theta}$. The smallest possible change, $d\theta$, requires a fixed $\gamma > 0$ to be paid for *all* types (not just below θ_{a^O}). So the increase in expected cost is $\gamma > 0$, and because $d\theta$ has measure zero, it alters neither the agent's moral hazard constraint (4.3) nor her information revelation problem (4.1). Calling on continuity completes the argument for any measure $\int d\theta$. ■

Proof of Proposition 5: Let a^* solve the agent's moral hazard constraint (4.3). Differentiate (4.3) and with respect to α:

$$\begin{aligned} 0 &= -\int_{\underline{\theta}}^{\tilde{\theta}} v p_\alpha dF_a(x|a) \\ &+ \left[\int_{\underline{\theta}}^{\tilde{\theta}} v(t(m(x)))[1-p(m(x)-x)]dF_{aa}(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t(x))dF_{aa}(x|a) - c''(a) \right] \frac{da^*}{d\alpha} \end{aligned} \quad (8.2)$$

Since the term in the brackets is the agent's second-order condition, it is negative. Therefore $\frac{da^*}{d\alpha} < 0$. Next, take the first-order condition (4.4) (or (4.5), as necessary), multiply by f and v' and differentiate with respect to a:

$$f_a - v'\lambda f_{aa}(\cdot|a) - v'' dt/da \lambda f_a(\cdot|a) = 0.$$

Divide by v' and integrate from θ_a to $\overline{\theta}$, where SOC is $v''\lambda f_a(\cdot|a) < 0$:

$$\int_{\theta_a}^{\overline{\theta}} \frac{1}{v'} dF_a - \lambda \int_{\theta_a}^{\overline{\theta}} dF_{aa} - \int_{\theta_a}^{\overline{\theta}} \frac{SOC}{v'} \frac{dt}{da} dx = 0,$$

whence $\frac{dt}{da} < 0$ (from the perspective of the principal). Because $a = a^*$, combining these steps gives $\frac{dt}{d\alpha} > 0$. ■

Proof of Proposition 6: The following will be useful in several instances. Let a^* solve the agent's moral hazard constraint (4.3). Differentiate (4.3) with respect to t:

$$\begin{aligned} 0 &= \int_{\underline{\theta}}^{\tilde{\theta}} v'[1-p]dF_a(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v' dF_a(x|a) \\ &+ \left[\int_{\underline{\theta}}^{\tilde{\theta}} v(t(m(x)))[1-p(m(x)-x)]dF_{aa}(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t(x))dF_{aa}(x|a) - c''(a) \right] \frac{da^*}{dt} \end{aligned} \quad (8.3)$$

Since the term in the brackets is the agent's second-order condition, it is negative. Therefore $\frac{da^*}{dt} > 0$. To prove item *(i)*, consider two distributions $F^1(\theta|a)$ and $F^2(\theta|a)$, where F^2 is a mean-preserving spread of F^1 (see Rothschild and Stiglitz [31]). Fix t; because F^1 dominates

F^2 in the second order sense, it follows from (4.3) that at a^*

$$\int_{\underline{\theta}}^{\tilde{\theta}} v[1-p]dF_a^2 + \int_{\tilde{\theta}}^{\overline{\theta}} vdF_a^2 < \int_{\underline{\theta}}^{\tilde{\theta}} v[1-p]dF_a^1 + \int_{\tilde{\theta}}^{\overline{\theta}} vdF_a^1 \tag{8.4}$$

by application of the envelop theorem (to the messages). Now define the following variable $\theta_2 = \theta_1 + \epsilon$, where $\theta_2 \sim F^2$ and $\theta_1 \sim F^1$ (so θ_2 is more risky than θ_1, and (8.4) follows). Consider again (4.3), as under F^1, and differentiate with respect to ϵ at $\epsilon = 0$:

$$\begin{aligned} &\left[\int_{\underline{\theta}}^{\tilde{\theta}} v(t(m(x)))[1-p(m(x)-x)]dF_{aa}^1(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t(x))dF_{aa}^1(x|a) - c''(a)\right] \frac{da}{d\epsilon} \\ &\qquad + \frac{d}{d\epsilon}\left[\int_{\underline{\theta}}^{\tilde{\theta}} v[1-p]dF_a^1 + \int_{\tilde{\theta}}^{\overline{\theta}} vdF_a^1\right] = 0 \end{aligned}$$

By (8.4) the last term is negative, so from (8.3) $\frac{da}{d\epsilon} < 0$. Letting $\frac{da}{d\epsilon} \equiv \frac{da}{dt}\frac{dt}{d\epsilon}$, $\frac{dt}{d\epsilon} < 0$ as claimed. To show (ii), consider a family of utility functions $v(t;r)$ parametrized by r; risk aversion (i.e. the concavity of $v(\cdot;\cdot)$) increases in r. Suppose for simplicity that $v(t;r)$ is continuous and differentiable in r (as well as t). For a fixed action a, we know that

$$\frac{d}{dr}\left[\int_{\underline{\theta}}^{\tilde{\theta}} v(t;r)[1-p]dF(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t;r)dF(x|a)\right] < 0$$

using the envelop theorem again. That is, equivalently, for any two $r_2 > r_1$, $\int_{\underline{\theta}}^{\tilde{\theta}} v(t;r_2)[1-p]dF(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t;r_2)dF(x|a) < \int_{\underline{\theta}}^{\tilde{\theta}} v(t;r_1)[1-p]dF(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t;r_1)dF(x|a)$. It then follows from (4.3) that $a^*(r_2) < a^*(r_1)$; equivalently, differentiating (4.3)

$$\begin{aligned} 0 &= \frac{d}{dr}\left[\int_{\underline{\theta}}^{\tilde{\theta}} v(t;r)[1-p]dF_a(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t;r)dF_a(x|a)\right] \\ &+ \frac{da}{dr}\left[\int_{\underline{\theta}}^{\tilde{\theta}} v(t;r)[1-p]dF_{aa}(x|a) + \int_{\tilde{\theta}}^{\overline{\theta}} v(t;r)dF_{aa}(x|a) - c''(a)\right] \end{aligned} \tag{8.5}$$

Because the first term of (8.5) is negative it follows that $\frac{da}{dr} < 0$ as well. Making use of the fact that $\frac{da}{dt} > 0$ completes the argument. To prove (iii), consider two cost functions $c_1(a), c_2(a)$ such that $\forall a, c_2 > c_1$. Because $c_i', c_i'', c_i''' > 0, c_2 > c_1\ \forall a$ implies $c_2' > c_1'\ \forall a$. Fix t, from (4.3) we have that $a^*(c_2) < a^*(c_1)$. By (8.3) therefore $t(\theta, c_2) < t(\theta, c_1)\ \forall\theta$ (with obvious notation). For the last item, suppose the principal's payoff is some increasing function $\pi(\theta)$. From (4.7) it follows that a^O increases, and from (8.3) so does the transfer t. To complete the proof, apply Proposition 5. ■

References

[1] Araujo A. and H. Moreira (2001) "A general Lagrangian approach for non-concave moral hazard problems. "*Journal of Mathematical Economics*, Vol. 35, pp. 17-39

[2] Arya, A., Jonathan Glover and Shyam Sunder (1998) "Earnings management and the Revelation Principle. "*Review of Accounting Studies*, Vol. 3, pp. 7-34

[3] Baron D. and David Besanko "Regulation, Asymmetric Information, and Auditing." *The RAND Journal of Economics*, Vol. 15, No. 4, pp. 447-470

[4] Border, K. and J. Sobel (1987) "Samuarai accounting: a Theory of Auditing and Plunder" *The Review of Economic Studies*, Vol. 54, pp. 525-540

[5] Bushman, R. and Chandra Kanodia (1996) "A Note on Strategic Sampling in Agencies." *Management Science*, Vol.42, pp. 151-156

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

[7] Crawford, V. and Joel Sobel (1982) "Strategic Information Transmission." *Econometrica*, Vol. 50, pp. 1431-1451

[8] Demski, J. and Ronald Dye (1999) "Risk, Returns and Moral Hazard" *Journal of Accounting Research*, Vol. 37, No. 1, pp. 27-55

[9] Demski, J. and Hans Frimor (1999) "Performance measure garbling under renegotiation in multi-period agencies." *Journal of Accounting Research*, Vol. 37, Supplement, pp. 187-214

[10] Dye, R. and Sanjay Sridhar "Moral Hazard Severity and Contract Design" *The RAND Journal of Economics*, Vol. 36, No. 1, pp. 78-92

[11] Gromb, D. and David Martimort (2007), "Collusion and the organization of delegated expertise" *Journal of Economic Theory*, Vol. 137 (1), pp. 271-299

[12] Grossman, S. and Oliver Hart (1983), “An Analysis of the Principal-Agent Problem” *Econometrica*, vol. 51(1), pages 7-45.

[13] Holmström, B. (1979) “Moral hazard and observability.” *The Bell Journal of Economics*, Vol. 10, pp. 74-91

[14] Holmström, B. and Paul Milgrom (1987) “Aggregation and Linearity in the Provision of Intertemporal incentives.” *Econometrica*, Vol. 55, pp. 303-328

[15] Holmström, B. and Paul Milgrom (1991) “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Description.” *Journal of Law, Economics, & Organization*, Vol. 7, pp. 24-52

[16] Jewitt, I. (1988) “Justifying the first-order approach to principal-agent problems.” *Econometrica*, Vol. 56, pp. 1177-1190

[17] Jewitt, I. Ohad Kadan and Jeroen Swinkels “Moral hazard with bounded payments”, *Journal of Economic Theory* Vol. 143, pp. 59-82

[18] Kanodia, C. (1985) “Stochastic Monitoring and Moral Hazard.” *Journal of Accounting Research*, Vol.23, pp. 175-194

[19] Kartik, N (2009) “Strategic communication with lying costs”, *Review of Economic Studies*, 76(4), pp. 1359-1395.

[20] Kedia, S. and Thomas Philippon (2006) “The economics of fraudulent accounting”, *NYU working paper*

[21] Khalil, F. (1997) “Auditing without commitment. ” *Rand Journal of Economics*, Vol. 28, pp. 629-640

[22] Krähmer D. and Roland Strausz “Optimal procurement contract with pre-project planning.”, *Review of Economic Studies*, 78(3), pp. 1015-1041.

[23] Laffont J.-J. and David Martimort (2002) “The Theory of Incentive – the Principal-Agent model.” *Princeton University Press*

[24] Levitt, S. and Christopher Snyder (1997) “Is no News Bad News? Information Transmission and the Role of ”Early Warning” in the Principal-Agent Model.” *Rand Journal of Economics*, Vol. 28 (4), pp. 641-661

[25] Malcolmson, J. (2009) “Principal and expert agent.” *The B.E. Journal of Theoretical Economics*, Vol. 9 (1), Article 17.

[26] Mookherjee D. and Ivan Png (1989) “Optimal Auditing, Insurance, and Redistribution.” *The Quarterly Journal of Economics*, vol. 104(2), pp. 399-415

[27] Reinganum J. and Louis Wilde (1985) “Income Tax Compliance in a Principal-Agent Framework.” *The Journal of Public Economics*, vol. 26, pp. 1-18.

[28] Reinhart, C. and Kenneth Rogoff “From financial crash to debt crisis.” *American Economic Review*, vol. 101 (5) pp. 1677-1706.

[29] Roger, G. (2011) “Moral hazard with soft information.” *Working paper, UNSW*

[30] Rogerson, W. (1985) “The first-order approach to principal-agent problems.” *Econometrica*, Vol. 53, pp. 1357-1368.

[31] Rothschild, M. and Joseph Stiglitz (1970) “Increasing Risk I: a Definition. ” *Journal of Economic Theory*, Vol. 2, pp. 315-329