Submission Number: PET11-11-00094

Moral hazard with discrete soft information

Guillaume F Roger
The University of New South Wales

Abstract

I study a simple model of moral hazard with soft information. The risk-averse agent takes an action and she alone observes the stochastic outcome; hence the principal faces a problem of ex post adverse selection. Some measure of truthful revelation is necessary to implement any effort, for which an audit is required. And for effort to be induced the agent must be offered a rent. There exists a truth-telling equilibrium, and one with partial pooling, with effort in either case. To accommodate ex post information revelation, the principal must distort the transfer schedule, as compared to the standard moral hazard problem. Then effort is implemented for a smaller set of parameters than in the standard problem.

I regard this paper as exploratory. It is the foundation of a richer model on which I am still working. This paper is very simple yet very informative.

Submitted: February 27, 2011.

Moral hazard with discrete soft information

Guillaume Roger *

The University of New South Wales

January 2011

Abstract

I study a simple model of moral hazard with soft information. The risk-averse agent takes an action and she *alone* observes the stochastic outcome; hence the principal faces a problem of *ex post* adverse selection. Some measure of truthful revelation is *necessary* to implement any effort, for which an audit is required. And for effort to be induced the agent must be offered a rent. There exists a truth-telling equilibrium, and one with partial pooling, with effort in either case. To accommodate *ex post* information revelation, the principal must distort the transfer schedule, as compared to the standard moral hazard problem. Then effort is implemented for a smaller set of parameters than in the standard problem.

1 Introduction

In standard moral hazard problems the outcome of the agent's action is observable by the principal and may therefore (imperfectly) substitute itself for the non-observability of said action. Then a complete contract may be conditioned on the outcome. That is a convenient model, but not necessarily one that fits many relevant situations. Performance may be difficult to observe, or its observation may be considerably delayed. Sometimes it is not observed at all: for example, an accounting report is *not* a direct observation of the state of an enterprise. In this paper, attention is paid to the case where *neither* the action, *nor* the outcome are observable by the principal. The information is said to be *soft* in that it is subject to manipulation on the part of the agent:

*Financial support from the Australian School of Business at UNSW is gratefully acknowledged.

the principal receives *no* signal of performance whatsoever. 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. In the optimal taxation problem, the agent may undertake some investment (education) to enhance her productivity, and then be asked to reveal the latter to the tax authority.

Bar for the role of soft information, the model mirrors that of a standard moral hazard framework. A risk-neutral principal delegates production to a risk-averse agent protected by limited liability. The agent's action a governs the distribution of a stochastic outcome drawn from a discrete space, which she *alone* observes. That information must therefore be elicited *ex post*. In this construct, the principal is exposed to *ex ante* moral hazard and also faces a problem of adverse selection *ex post*. Instead of a standard incentive contract made of a prescribed action and a transfer conditioned on the outcome, the contract entails a prescribed action and a revelation mechanism. Because the principal otherwise observes nothing, this revelation mechanism must include an audit.[1] Although these problems interact, they can be handled in sequence because the action is sunk at the stage of information revelation.

This simple model delivers some important insights. First, absent some measure of *type separation*, the principal can only offer a trivial contract, in which the agent exerts no effort. Since type separation is equivalent to truthful revelation in this model, at least a subset of the *ex post* incentive constraints must hold for the principal to induce effort.[2] If not, the agent can freely lie and so pools at the top of the message space (which coincides with the type space). Because the principal is committed, he must pay the high transfer no matter what the true state is. Therefore the agent has no incentive to ever expend effort. Thus in any equilibrium where the principal induces effort, *some* separation must arise. In the present model, it also implies that at least *some* information must be (truthfully) revealed. Second, *any* information revelation generates an *ex ante* rent to the agent. In other words, satisfying any of the *ex post* incentive constraints and the moral hazard constraint, requires the (*ex ante*) participation constraint to be slack. The fundamental driving force

[1] Without audit it would be impossible to specify a non-trivial incentive-compatible mechanism at the information revelation stage, and therefore impossible to implement any effort in the first place.

[2] Throughout this article I will refer to "incentive constraints" as those addressing the adverse selection problem and "moral hazard" constraint as that dealing with the hidden action problem.

behind these two observations is the conflict between *ex post* incentives for information revelation, best addressed with transfers invariant in the message, and *ex ante* effort incentives, which need a sloped compensation schedule. The one implication of these two observations is that effort is more costly to the principal, and will therefore be implemented for a smaller set of parameters than in the standard problem.

Ignoring the trivial outcome of no effort, two mutually exclusive equilibria arise in this model. In the first one, the principal induces effort and the transfers are sufficient to elicit truthful information revelation (i.e. separation). However these transfers are distorted, as compared to the standard moral hazard problem, to accommodate the *ex post* incentive constraints. The compensation schedule is flatter, and the agent receives a (*ex ante*) rent. In the other equilibrium, the principal also induces effort, but the transfers offered are such that one *ex post* incentive constraint fails. Consequently there is partial pooling of types, i.e. incomplete information revelation. The transfers are distorted in the same direction – so the schedule is flatter as well – but one of them is never paid out. The agent also receives a rent. Given that effort is exerted, the principal's gross expected returns are the same in either equilibrium, so which of these equilibria the principal chooses to implement depends only on the rent left to the agent. That rent is characterised in terms of the primitives.

The works closest to this are Gromb and Martimort (2007), Green and Laffont (1986), and Levitt and Snyder (1997). The present model departs from all by adopting soft information in a very strong sense: the principal *never* observes any outcome.[3] Gromb and Martimort (2007) use the same sequence of events as here, however they study the incentives of expert(s) to search and report information about *others* (an exogenous project), not themselves. To overcome the adverse selection problem, their incentive contract must be made state dependent although they do not exert any influence on it. In contrast, here a share of the agent's compensation must be made state independent to induce information revelation. Levitt and Snyder (1997) 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. Here, information can only be observed, and reported, by the agent. The audit restores partial

[3]Gromb and Martimort's expert(s) receive(s) a soft signal, but whether a project is eventually successful is publicly (*ex post*) observable.

observability, and is therefore essential as compared to models of costly state-verification (e.g. Khalil (1997)), where it only assists in relaxing the incentive constraint of the agent. Cremer, Khalil and Rochet (1998a,b) allow for an agent to gather costly information about her own type before contracting. To emphasize the point, in these papers, information is still *exogenously* given although *ex ante* unknown to the agent. Here the agent has no *ex ante* private information, which only emerges *ex post*. Green and Laffont (1986) study the principal-agent problem with "partially verifiable information" in the sense that the agent's message is constrained to lie in an arbitrary subset $M(\theta)$ of the type space Θ, which varies with the true state in a publicly known fashion. $M(.)$-implementable mechanisms exist and need not elicit truth-telling.

2 Model

A principal delegates a task to an agent. At cost $\psi(a) \geq 0$ the agent undertakes an action $a \in \{\underline{a}, \overline{a}\}$ (with obvious ranking), with $\psi(\overline{a}) = \psi > \psi(\underline{a}) = 0$. This action yields a stochastic outcome $\theta \in \Theta \equiv \{\theta_1, \theta_2, \theta_3\}$, where $\theta_3 - \theta_2 = \theta_2 - \theta_1 = \Delta\theta > 0$. Let $\overline{\pi}_i \equiv \Pr(\theta_i | a = \overline{a})$ and $\underline{\pi}_i \equiv \Pr(\theta_i | a = \underline{a})$ denote the probabilities of each outcome conditional on the agent's action, with $\overline{\pi}_i > \underline{\pi}_i$. The agent's net utility is given by $u(t, a) \equiv v(t_i) - \psi(a)$, where $v(.)$ is a concave function with $v(0) = 0$. Thus the model is completely defined by the agent's preferences $v(.)$, the technology ψ and the parameters $\{\Theta, \pi_i(\theta|a)\}$. Here the agent *alone* observes the outcome θ. By application of the Revelation Principle, she reports a message $m \in \Theta$ to the principal, whereupon she receives the transfer t_i. The limited liability constraint applies throughout, so that $\forall i,\ t_i \geq 0$. The principal can commit to the contract and his net payoff reads $S(t; \theta) = \theta_i - t_i$. If the true state θ were observable by the principal, this construct would be a moot point and would collapse to the textbook moral hazard problem. The timing is almost standard:

1. The principal offers a contract $\mathcal{C} = \langle a, D \rangle$ consisting of an action a and a revelation mechanism $D = (\Theta, t(m))$ made of a message space and a transfer
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 only by the agent.
4. The agent report a message $m \in \Theta$

5. Transfers are implemented and payoffs are realised.

3 Restoring *ex post* observability

After she has taken some action a (now sunk), the agent maximises her utility $u(t, a)$ by choice of a message m. A mechanism is truthful if and only is the following constraints are satisfied

$$v(t_1) \geq v(t_2);\ v(t_2) \geq v(t_3);\ v(t_1) \geq v(t_3)$$

and

$$v(t_3) \geq v(t_2);\ v(t_2) \geq v(t_1);\ v(t_3) \geq v(t_1)$$

whence the first claim of this paper is immediate.

Proposition 1 *Absent any other instrument, the only (truthful) revelation mechanism requires* $t_1 = t_2 = t_3$. *Therefore* $a = \underline{a}$ *and* $t_i = 0\ \forall i$.

This result obtains because the principal entirely lacks *ex post* observability. Absent that, the principal is unable to address the fundamental tension between *ex ante* effort incentive and *ex post* information revelation. Thus a necessary element of any optimal contract is to restore at least some *ex post* observability. A natural avenue is the introduction of an *ex post* audit, which is costless (and therefore always) run in this model, but imperfect. With some probability $p(m - \theta)$, the agent's deception is uncovered and she receives zero. The function $p : \Theta \cup \{0\} \mapsto [0, 1]$ is increasing, symmetric and such that $p(0) = 0,\ p(2x) \geq 2p(x)$. With this, the agent has *ex post* expected utility

$$\mathbb{E}[u(t; p(.))] \equiv (1 - p(.))v(t(m)),$$

which she seeks to maximise by choice of the message $m \in \Theta$. The set of *ex post* incentive constraints now takes the form:

$$v(t_1) \geq (1 - p(\Delta\theta))v(t_2) \tag{3.1}$$

$$v(t_1) \geq (1 - p(2\Delta\theta))v(t_3) \tag{3.2}$$

$$v(t_2) \geq (1 - p(\Delta\theta))v(t_3) \tag{3.3}$$

$$v(t_2) \geq (1 - p(\Delta\theta))v(t_1) \tag{3.4}$$

$$v(t_3) \geq (1 - p(2\Delta\theta))v(t_1) \tag{3.5}$$

$$v(t_3) \geq (1 - p(\Delta\theta))v(t_2) \tag{3.6}$$

These constraints do not yield the standard implementability condition, as can be verified by adding them up pairwise. For example, add (3.1) and (3.4) to find $1 \geq (1 - p(\Delta\theta))$, which is trivially true and uninformative as to the shape of the transfer function. The system (3.1)-(3.6) forms the basis of the next claim.

Proposition 2 *There exist transfers $t_3 \geq t_2 \geq t_1$ such that constraints (3.1)-(3.6) hold. Whenever the local constraints (3.1) and (3.3) are satisfied, the global constraint (3.2) is necessarily slack. Whenever the global constraint (3.2) binds at least one of the local constraints fails.*

This existence result remains silent as to optimality and does not imply that transfers satisfying (3.1)-(3.6) solve the principal's problem. In particular, truthful revelation needs not be optimal.

4 The optimal contract

There always exists a trivial contract, in which the low action is sought from the agent. When $a = \underline{a}$, the principal needs only set $t_i = 0 \ \forall i$ – which incidentally elicits truth-telling *ex post*. I am interested in equilibria where effort is implemented.

Two cases of interest arise. In the first one, truthful revelation is elicited *ex post*, which may come at a cost to the principal. In the second case, the principal may not seek to satisfy the *ex post* incentive constraint because it is too costly. Let $\varphi \equiv v^{-1}(.)$ denote the inverse function of the agent's utility, and $v_i = v(t_i)$ for some t_i.

4.1 Truth-telling equilibrium

A truth-telling equilibrium is one where all the *ex post* incentive constraints (3.1)-(3.3) are satisfied. Following Proposition 2, only (3.1) and (3.3) need bind. The principal seeks to solve

Problem 1

$$\max_{v_i \geq 0} \sum_i \overline{\pi}_i \left[\theta_i - \varphi(v_i)\right]$$

s.t. (3.1)-(3.3) and

$$\sum_i \Delta\pi_i v_i \geq \psi \tag{4.1}$$

$$\sum_i \overline{\pi}_i v_i \geq \psi \tag{4.2}$$

The last two inequalities are the usual moral hazard and participation constraints. Attach multipliers γ_1, γ_2 to constraints (3.1) and (3.3) and λ and μ to each of (4.1) and (4.2), respectively. The necessary and sufficient first-order conditions of Problem 1 are

$$\varphi'(v_1) - \frac{\gamma_1}{\overline{\pi}_1} = \mu + \lambda \frac{\Delta\pi_1}{\overline{\pi}_1} \tag{4.3}$$

$$\varphi'(v_2) - \frac{\gamma_2 - \gamma_1(1-p)}{\overline{\pi}_2} = \mu + \lambda \frac{\Delta\pi_2}{\overline{\pi}_2} \tag{4.4}$$

$$\varphi'(v_3) + \frac{\gamma_2(1-p)}{\overline{\pi}_3} = \mu + \lambda \frac{\Delta\pi_3}{\overline{\pi}_3} \tag{4.5}$$

The MLRP ensures these conditions are not vacuous.

Lemma 1 *Suppose $\mu, \lambda > 0$ (as in the standard moral hazard problem), then at least two of (3.1)-(3.3) must be violated.*

Hence there cannot be an equilibrium in which the standard solution of the moral hazard problem also accommodates the *ex post* information revelation problem. Further, for the constraints (3.1)-(3.3) to hold, at least one of (4.1) or (4.2) must be slack, or violated. More precisely,

Lemma 2 *Suppose $\mu > 0$ and (3.1), (3.3) are satisfied, then the moral hazard constraint (4.1) is violated.*

So there is no solution to Problem 1, in which the participation constraint (4.2) binds, while the other constraints are satisfied. That is, either there is no truthful revelation *ex post* (by Lemma 1), or no effort can be induced, without affording the agent a rent (be Lemma 2. Therefore I can restrict attention to the set of utilities v_i such that (4.2) is slack. So, set $\mu = 0$ and define $\overline{\Pi} = \overline{\pi}_3 + (1-p)\overline{\pi}_2 + (1-p)^2\overline{\pi}_1$ and $\underline{\Pi} = \underline{\pi}_3 + (1-p)\underline{\pi}_2 + (1-p)^2\underline{\pi}_1$. With this in hand,

Proposition 3 *The lowest-cost truth-telling equilibrium in which the agent is induced to exert effort entails*

$$v_3^T = \frac{\psi}{\overline{\Pi} - \underline{\Pi}}$$

determined by a binding moral hazard constraint (4.1), and $v_1^T, v_2^T > 0$ determined by (3.1) and (3.3), both binding. The agent receives an ex ante *rent $U^T = \psi \frac{\underline{\Pi}}{\overline{\Pi} - \underline{\Pi}} > 0$*

This rent is is excess of the standard risk-premium the principal must pay (to partially insure the agent). To make the point more salient, the transfers are given by $t_3 = \varphi\left(\psi/(\overline{\Pi} - \underline{\Pi})\right)$, $t_2 = \varphi\left((1-p)\psi/(\overline{\Pi} - \underline{\Pi})\right)$ and $t_1 = \varphi\left((1-p)^2\psi/(\overline{\Pi} - \underline{\Pi})\right)$. The next claim immediately follows.

Proposition 4 *The schedule is flatter (than under the standard moral hazard problem).* v_1 *solving (4.3) exceeds the standard level and* v_3 *solving (4.5) is lower than the standard level.* v_2 *is ambiguous.*

This result owes to the fundamental tension between *ex post* incentive compatibility, best satisfied with constant transfers, and *ex ante* effort incentives, best addressed with a compensation conditioned on performance. The distortions tilt the schedule and are such that (3.1) and (3.3) are just binding. Of course the implication of this distortion is the rent $U^T > 0$, so it is costly to the principal. Finally,

Corollary 1 *The principal induces costly effort if and only if*

$$\sum_i \Delta\pi_i\theta_i \geq \psi\frac{\overline{\Pi}}{\overline{\Pi}-\underline{\Pi}} > \psi$$

the proof of which is obvious and therefore omitted. Finally we have

Proposition 5 *The high action is implemented for a narrower set of parameters than under the standard moral hazard problem.*

To see that recall simply that in the standard case, the cost of effort is given by the binding participation constraint (4.2).

4.2 No truthful revelation

Because combining *ex ante* effort incentives and *ex post* truthful revelation is costly, the principal may choose an alternative. In this case truthful revelation may purposefully not be sought, which may make him better off. In such an equilibrium, transfers are such that at least one of the *ex post* incentive constraints is violated.

Even when the agent sends a message that is not truthful, the principal does not update his beliefs as to the true state of the world because he has committed to the contract.[4] It then follows that if all incentive constraints (3.1)-(3.3) fail, the principal will necessarily offer a zero-effort contract. Indeed, the agent can only report θ_3, whereupon the principal must pay t_3. But then

[4]The principal has no further move in the game, so updating is a moot point. In particular, there is no renegotiation.

there is no incentive for the agent to exert any effort, so the principal offers only $\underline{t}_i = 0$.[5] Therefore, some measure of type separation is a *necessary* condition for the principal to want to induce action $\overline{a}$. One must also note that since some incentive constraint will fail (by design), (3.2) can no longer be ignored. However it is still true that there cannot be an equilibrium in which only (3.2) is violated, because (3.1) and (3.3) imply (3.2). Conversely, if (3.1) and (3.3) fail, it does not imply that (3.2) does. The principal's program is

Problem 2

$$\max_{v_i \geq 0} \sum_i \overline{\pi}_i \theta_i - \sum_i \rho_i \varphi(v_i)$$

s.t. (3.1)-(3.3) and

$$\sum_i \Delta \rho_i v_i \geq \psi \tag{4.6}$$

$$\sum_i \overline{\rho}_i v_i \geq \psi \tag{4.7}$$

where ρ_i denotes the probability of receiving a report i when the agent pools states (since some incentive constraint fails). The exact definition of ρ_i depends on the choice of pooling, that is, on which of the incentive constraints fail(s). At face value there are many combinations to consider; fortunately the next two lemmata are very useful to reduce the set of cases to investigate.

Lemma 3 *The principal does not offer a contract in which the agent exerts effort such that (3.2) and (3.3) are violated.*

In this case the agent pools her message at θ_3 and no separation obtains.

Lemma 4 *The principal does not offer a contract in which the agent exerts effort and any of*

1. *(3.1) and (3.2) or;*
2. *(3.1) and (3.3) or;*
3. *only (3.3);*

are violated.

[5] Incidentally, this elicits truthful reporting.

Thus the only viable case when the principal's contract is such that it does not induce the agent to truthfully reveal her information *ex post* requires (3.1) to be violated. And here too the participation constraint must be left slack in order to satisfy the moral hazard constraint.

Lemma 5 *There is no equilibrium such that the moral hazard constraint (4.6) is satisfied, the participation constraint (4.7) binds, and only (3.1) is violated*

So we know that in *any* equilibrium the agent receives an *ex ante* rent. With only one case to study and a slack constraint (4.7), Problem 2 becomes

Problem 3

$$\max_{v_i \geq 0} \sum_i \overline{\pi}_i \theta_i - [\overline{\pi}_3 \varphi(v_3) + (1 - \overline{\pi}_3)\varphi(v_2)]$$

s.t. (3.2),(3.3) and

$$\Delta\pi_3(v_3 - v_2) \geq \psi \tag{4.8}$$

$$\overline{\pi}_3 v_3 + (1 - \overline{\pi}_3)v_2 > \psi \tag{4.9}$$

Attach multipliers γ_2, γ_3 to (3.3), (3.2), and the usual λ to (7.5), the first-order conditions read

$$(1 - \overline{\pi}_3)\varphi'(v_2) - \gamma_2 = -\lambda\Delta\pi_3 \tag{4.10}$$

$$\overline{\pi}_3\varphi'(v_3) + [(1 - p(\Delta\theta)\gamma_2 + (1 - p(2\Delta\theta))\gamma_3] = \lambda\Delta\pi_3 \tag{4.11}$$

And we have

Proposition 6 *The least-cost non-truthful equilibrium in which the agent is induced to exert effort entails*

$$\begin{aligned} v_3^{NT} &= \frac{\psi}{\Delta\pi_3 p(\Delta\theta)} \\ v_2^{NT} &= (1 - p(\Delta\theta))v_3^{NT} \end{aligned}$$

determined by a binding moral hazard constraint (7.5), and and a binding (3.3). The agent receives an ex ante *rent* $U^{NT} = \frac{\psi}{\Delta\pi_3 p}(1 - p(1 - \underline{\pi}_3)) > 0$.

Again, directly from the first-order conditions:

Proposition 7 *The compensation schedule is flatter then in the standard moral hazard problem.*

And to complete the analysis, I also establish:

Corollary 2 *The high action is implemented only if*

$$\sum_i \Delta\pi_i\theta_i \geq \frac{\psi}{\Delta\pi_3 p(\Delta\theta)}[1 - p(\Delta\theta)(1-\overline{\pi}_3)] > \psi,$$

that is, for a smaller set of parameters than in the standard case.

Corollary 3 *Whether the principal chooses to offer a truth-telling contract depends on whether* $\frac{\overline{\Pi}}{\overline{\Pi}-\underline{\Pi}} \leq (>) \frac{1-p(\Delta\theta)(1-\overline{\pi}_3)}{\Delta\pi_3 p(\Delta\theta)}$.

4.3 Comparative statics

Altering any of the technology ψ or information structure $\pi(.|a)$ produces the same effects as in the standard moral hazard model. Of more interest are comparative statics with respect to the audit technology and the type space itself.

Corollary 4

1. $\frac{\partial U^T}{\partial p} < 0$
2. $\frac{\partial v_3^T}{\partial p} > 0$
3. $\frac{\partial U^{NT}}{\partial p} < 0$
4. $\frac{\partial v_3^{NT}}{\partial p} > 0$

Improving the audit technology tilts back the compensation schedule toward a steeper slope: the agent's utility in the good state increases. But it also eases the incentive constraints (3.1)-(3.3). The net effect is a decrease the agent's expected rent.

However one must note that, *in this model*, even a perfect audit technology does not enable the principal to implement the standard moral hazard schedule. Consider (3.1) and (3.3) and suppose $p(\Delta\theta) = 1$. Then truthful revelation still requires $v_1 \geq 0$ and $v_2 \geq 0$, and since $v_3 > 0$, the participation constraint still fails to bind. Recall that in the standard model, one must have at least $v_1 < 0$ to have a binding participation constraint. Of course, if the audit technology were perfect and costless, the principal should simply ignore the agent's message.

Last, it is immediate that $sign\ dp/d\Delta\theta = sign\ d\Delta\theta$ so that Corollary 4 carries over. Increasing the distance between types renders misreporting more hazardous here. This may be said to be an

artificial result of the properties of the function $p(.)$, but it also suggests that with audit, small lies only may be worthwhile.

5 Discussion

5.1 Monitoring

Should the principal audit the agent's report of an *outcome* (θ), or should he somehow gather information about action (a) – that is, monitor the agent? In the latter case, the agent is paid according to her action, not the outcome. Then the information revelation problem is moot and the risk-neutral principal bears all the risk. Implicitly in this paper it is presumed that monitoring is either too costly or outright impossible – as suggested by some of the examples.

5.2 Relation to M-implementability (Green and Laffont (1986))

These authors study the implementability of a social choice function when the agent may report a message from a set $M(\theta) \subset \Theta$, where $M(.)$ is *exogenous* and publicly known. They provide a necessary and sufficient condition – called the nested range condition (NRC) – for the agent to report her information truthfully. The NRC does not hold in our model, although it corresponds to a game of of "unidirectional distortions with an ordered space" (to use their words) – example a(2) in Green and Laffont.

Indeed, the NRC requires $M(\theta_3) = \{\theta_3\}$, $M(\theta_2) = \{\theta_2, \theta_3\}$, $M(\theta_1) = \{\theta_1, \theta_2, \theta_3\}$. In contrast, Constraints (3.2), (3.3) holding and (3.1) failing imply $M(\theta_3) = \{\theta_3\}$, $M(\theta_2) = \{\theta_2\}$, $M(\theta_1) = \{\theta_2\}$, whence $\theta_3 \notin M(\theta_1)$. This violates the definition of the NRC. Notice further that the sets $M(\theta_i)$ in this paper are endogenous, unlike in Green and Laffont (1986).

5.3 Separation versus truth-telling

The model analysed in this paper delivers two important results. First, *some* information revelation is necessary for costly effort to be implemented; second, eliciting information revelation (even partial) requires an *ex ante* rent to be left to the agent. That truthful revelation emerges as an equilibrium likely is an artefact of the discrete nature of the type space, as suggested by the discussion of Section 4.3. Indeed, in a separate paper, I show that truth-telling is *never* optimal

when types are continuous, except of course at the upper bound of the type space (and when messages are limited to the set of types). This is because continuous spaces allow for arbitrarily small lies. Thus what is important for the provision of effort incentive is not truthful revelation, but *type separation* (as emphasized in the introduction). To see why, observe first that when types (completely) pool in this model, one of them nonetheless reports truthfully. So, that is not lying per se that deters *ex ante* incentives, but pooling. Second, if all types face incentives such that they all report the same message, then it is immediate that the agent has no incentive to expend any effort. That is, pooling stifles effort.

Type separation needs not equate *truthful* information revelation. More generally, there may be other mechanisms with richer message spaces, that allow for complete type separation without it being truthful revelation. Exploring these is left to future research.

6 Conclusion

When the principal to a contract fraught with moral hazard also fails to observe any of the outcomes, he faces adverse selection *ex post*. This private information must be elicited from the agent through a distortion of the compensation structure. To satisfy information revelation *ex post*, the principal must leave an *ex ante* rent to the agent. To do so, he presents her with a flatter transfer scheme. This is a low(er)-power contract than in the standard moral hazard problem. This additional distortion is socially costly in that the high action can be implemented for a strictly smaller set of parameters than in the standard case. These results obtain because of the fundamental tension between effort provision and information revelation, which require different instruments.

7 Appendix: Proofs

Proof of Proposition 2: Suppose (3.1) and (3.3) are satisfied (either strictly or with some slack), then $v(t_1) \geq (1-p(\Delta\theta))^2 v(t_3)$. Since $(1-p(\Delta\theta))^2 > 1-p(2\Delta\theta)$, Condition (3.2) is necessarily slack. To show existence, take (3.1) and (3.3) binding. Then we have $v(t_3) > (1-p)v(t_3) = v(t_2) > (1-p)^2 v(t_3) = v_1$. For the last statement, take (3.2) binding. Then $v(t_1) = (1-p(2\Delta\theta))v(t_3) \geq (1-p(\Delta\theta))v(t_2)$ by (3.1) (or $\geq (1-p(\Delta\theta))v(t_1)$ by (3.3)). Either way it follows that $1-p(2\Delta\theta))v(t_3) \geq (1-p(\Delta\theta))^2 v(t_3)$ for both (3.1) and (3.3) to hold. This contradicts $(1-p(\Delta\theta))^2 > 1-p(2\Delta\theta)$. ■

Proof of Lemma 1: $\mu, \lambda > 0 \Leftrightarrow \sum_i \overline{\pi}_i v_i = \psi = \sum_i \Delta\pi_i v_i \Leftrightarrow \sum_i \underline{\pi}_i v_i = 0$. Since $v_3 \geq v_2 \geq v_1$ by MLRP and $\sum_i \overline{\pi}_i v_i = \psi > 0$, we must have $v_3 > 0 > v_1$. Therefore (3.2) cannot hold. Further, since $v_2 \geq v_1$, (3.1) must also be violated regardless of whether $v_2 \geq 0$ or $v_2 < 0$. Clearly in the later case (3.3) also fails to hold. ■

Proof of Lemma 2: $\mu > 0 \Leftrightarrow \sum_i \overline{\pi}_i v_i = \psi$ and when (4.1) holds, $\sum_i \overline{\pi}_i v_i \geq \psi + \sum_i \underline{\pi}_i v_i$. But but by (3.1)-(3.3) and MLRP, $\sum_i \underline{\pi}_i v_i > 0$. Then (4.1) implies

$$\begin{aligned} \sum_i \overline{\pi}_i v_i &\geq \psi + \sum_i \underline{\pi}_i v_i \\ \psi &\geq \psi + \sum_i \underline{\pi}_i v_i \\ 0 &\geq 0 + \sum_i \underline{\pi}_i v_i > 0 \end{aligned}$$

which is an obvious contradiction. So (4.1) must be violated. ■

Proof of Proposition 3: Set $\mu = 0$ and sum (4.3)-(4.5) to find

$$-\mathbb{E}_\Theta[\varphi'(v_i)] + (\gamma_1 + \gamma_2)p = 0$$

so that at least one of γ_1, γ_2 is strictly positive. Take the moral hazard constraint (4.1) and the *ex post* incentive constraint (3.1) and (3.3) binding to obtain v_3. Compute the cost of inducing effort as $C = \psi \frac{\overline{\Pi}}{\overline{\Pi}-\underline{\Pi}}$ and the rent R by subtracting the cost of effort ψ. To show this must be the lowest-cost contract, observe that v_3 must also solve

$$\varphi'(v_3) = \lambda \frac{\Delta\pi_3}{\overline{\pi}_3} - \frac{\gamma_2(1-p)}{\overline{\pi}_3} \tag{7.1}$$

Since $\varphi'(.)$ is increasing, $\gamma_2 = 0 \Rightarrow v_3 > \frac{\psi}{\overline{\Pi}-\underline{\Pi}}$ and $v_2 > (1-p)v_3 > (1-p)\frac{\psi}{\overline{\Pi}-\underline{\Pi}}$. It then follows that $v_1 \geq (1-p)v_2 > (1-p)^2 v_3 > (1-p)^2\frac{\psi}{\overline{\Pi}-\underline{\Pi}}$. A similar argument can be constructed for the case $\gamma_1 = 0$ by using (4.4) instead. ■

Proof of Proposition 4: In the standard problem (4.3)-(4.5) read

$$\varphi'(v_1) = \mu + \lambda\frac{\Delta\pi_1}{\overline{\pi}_1} \tag{7.2}$$

$$\varphi'(v_2) = \mu + \lambda\frac{\Delta\pi_2}{\overline{\pi}_2} \tag{7.3}$$

$$\varphi'(v_3) = \mu + \lambda\frac{\Delta\pi_3}{\overline{\pi}_3} \tag{7.4}$$

Increase v_i $\forall i$ by some arbitrarily small amount $\varepsilon > 0$, so that $\mu = 0$ but the cost of the contract comes within ε of the optimum. Then compare each of (7.2)-(7.4) to (4.3)-(4.5), recalling that $\varphi(.)$ is increasing convex. ■

Proof of Proposition 5: In the standard case, the high action is induced whenever $\sum_i \Delta\pi_i\theta_i \geq \psi$ since constraint (4.2) binds, while here $\sum_i \Delta\pi_i\theta_i \geq \psi\frac{\overline{\Pi}}{\overline{\Pi}-\underline{\Pi}} > \psi$ is required. Fix $\pi_i(\theta_i|a)$, this implies that at least some payoffs θ_i must increase. Fix Θ_i, it implies that at least some $\Delta\pi$ must increase. ■

Proof of Lemma 3: Suppose these constraints fail for some contract, then the agent only ever reports θ_3 and receives $\varphi(v_3)$ with probability 1. But then there is no incentive for effort. ■

Proof of Lemma 4:

1. The first case is not so obvious because it allows for $v_1 < 0 < v_2 < v_3$. But then the agent pools at θ_3 when observing θ_1. So Problem 2 becomes

   Problem 4

   $$\max_{v_i \geq 0} \sum_i \overline{\pi}_i\theta_i - [\overline{\pi}_2\varphi(v_2) + (1-\overline{\pi}_2)\varphi(v_3)]$$

   s.t. (3.3) and

   $$\Delta\pi_2(v_2 - v_3) \geq \psi \tag{7.5}$$

   $$\overline{\pi}_2 v_2 + (1-\overline{\pi}_2)v_3 \geq \psi \tag{7.6}$$

   Any effort on the part of the agent requires $v_3 \geq v_2$, which immediately violates the moral hazard constraint. The principal will never offer such a contract.

2. In the second case, v_1 again is "off equilibrium" because the agent reports θ_2 if observing θ_1 and otherwise pools at θ_3. The moral hazard constraint (7.5) rewrites $\Delta\pi_1(v_2 - v_3) \geq \psi > 0$, which is a contradiction again. The principal will never offer such a contract.

3. In the last instance, v_2 is also "off equilibrium" in the sense that the agent will always pool at θ_3 instead of reporting θ_2. Then the moral hazard constraint (7.5) becomes $\Delta\pi_1(v_1 - v_3) \geq \psi > 0$, which can never be satisfied. The principal will never offer such a contract either.

■

Proof of Lemma 5: Suppose (3.1) fails but (3.2) and (3.3) hold. θ_1 is never reported to the principal. By (4.7) and MLRP, $v_3 > v_2 > 0$. Regardless of the exact definition of ρ_i, take (4.7) binding then (4.6) is necessarily violated. ■

Proof of Proposition 6: Adding the first-order conditions, one finds

$$-\mathbb{E}_\Theta[\varphi'(v_i)] + p(\Delta\theta)\gamma_2 - (1 - p(2\Delta\theta))\gamma_3 = 0$$

whence $\gamma_2 > 0$ necessarily, and $v_2^{NT} = (1-p)v_3^{NT}$. γ_3 can be anything: since θ_1 is never reported, t_1 is never paid so any transfer satisfying (3.2) but not (3.1) will do. Combining the binding moral hazard constraint with $v_2^{NT} = (1-p)v_3^{NT}$ gives v_3^{NT}. To find the rent, first compute the principal cost of inducing effort

$$\begin{aligned} C^{NT} &= \overline{\pi}_3 v_3^{NT} + (1 - \overline{\pi}_3)v_2 \\ &= \overline{\pi}_3 \frac{\psi}{\Delta\pi_3 p(\Delta\theta)} + (1 - \overline{\pi}_3)(1-p)\frac{\psi}{\Delta\pi_3 p(\Delta\theta)} \\ &= \frac{\psi}{\Delta\pi_3 p(\Delta\theta)}[1 - p(\Delta\theta)(1 - \overline{\pi}_3)] \end{aligned}$$

and subtract the agent's effort cost ψ. ■

Proof of Proposition 7: As in the proof of Proposition 4 ■

References

[1] Crémer, F. Khalil and J.-C. Rochet(1998a), "Strategic Information Gathering before a Contract Is Offered. " *Journal of Econonmic Theory*

[2] Crémer, F. Khalil and J.-C. Rochet(1998b), "Contracts and Productive Information Gathering." *Games and Econonmic Behaviour* Vol. 25 (2) pp. 174-193.

[3] Green, J. and Jean-Jacques Laffont (1986) "Partially verifiable information and mechanism design." *The Review of Economic Studies*, Vol. LIII, pp. 447-456

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

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

[6] Laffont J.-J. and Jean Tirole (1986) "Using cost observation to regulate firms." *Journal of Political Economy*, Vol. 94, pp. 614-641

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

[8] 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